University of California Transportation Center
UCTC-FR-2010-02 (identical to UCTC-2010-02)

Reliable Facility Location Design under the Risk of Disruptions

Tingting Cui, University of California, Berkeley,
Yanfeng Ouyang, University of Illinois at Urbana-Champaign, and
Zuo-Jun Max Shen, University of California, Berkeley
February 2010

Submitted to *Operations Research*

Reliable Facility Location Design under the Risk of Disruptions

Tingting Cui
Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720
tingting@ieor.berkeley.edu

Yanfeng Ouyang
Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801
yfouyang@illinois.edu

Zuo-Jun Max Shen
Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720
shen@ieor.berkeley.edu

Reliable facility location models consider unexpected failures with site-dependent probabilities, as well as possible customer reassignment. This paper proposes a compact mixed integer program (MIP) formulation and a continuum approximation (CA) model to study the reliable uncapacitated fixed charge location problem (RUFL) which seeks to minimize initial setup costs and expected transportation costs in normal and failure scenarios.

The MIP determines the optimal facility locations as well as the optimal customer assignments, and the MIP is solved using a custom-designed Lagrangian Relaxation (LR) algorithm. The CA model predicts the total system cost without details about facility locations and customer assignments, and it provides a fast heuristic to find near-optimum solutions. Our computational results show that the LR algorithm is efficient for mid-sized RUFL problems and that the CA solutions are close to optimal in most of the test instances. For large-scale problems, the CA method is a good alternative to the LR algorithm that avoids prohibitively long running times.

Key words: facility location, reliability, mixed integer program, Lagrangian relaxation, heuristics, continuum approximation

1. Introduction

The classic uncapacitated fixed charge location problem (UFL) selects facility locations and customer assignments in order to balance the trade-off between initial setup costs and day-to-day transportation costs. However, some of the constructed facilities may become unavailable due to disruptions caused by natural disasters, terrorist attacks or labor strikes. When a facility failure

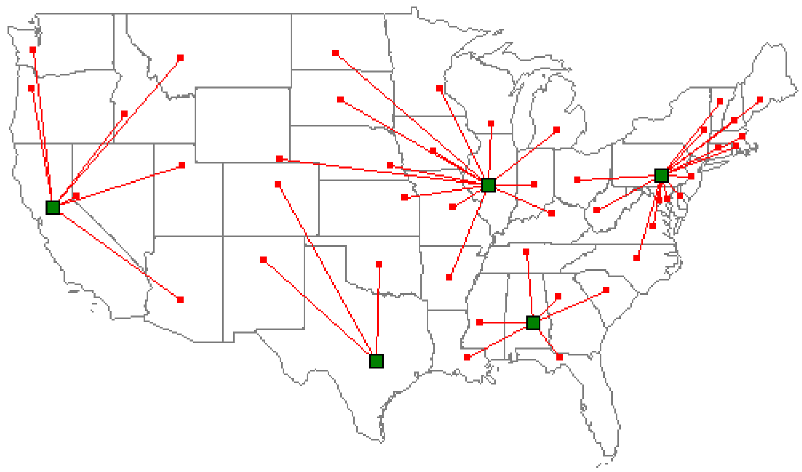


Figure 1 UFL solution to 49-data set

occurs, customers may have to be reassigned from their original facilities to others that require higher transportation costs. In this paper we present facility location models that minimize normal construction and transportation costs as well as hedge against facility failures within the system.

The reliable location model was first introduced by Snyder and Daskin (37) to handle facility disruption. Their motivating example is as follows. Consider a supply network that serves 49 cities, consisting of all state capitals of the continental United States and Washington, DC. Demands are proportional to the 1990 state populations and the fixed costs are proportional to the median house prices. The optimal UFL solution for this problem is shown in Figure 1. This solution has a fixed cost of \$348,000 and a transportation cost of \$509,000 (at \$0.00001 per mile per unit of demand). However, if the facility in Sacramento, CA failed, customers from the entire west-coast region would have to get service from the facilities in Springfield, IL and Austin TX, which would increase the transportation cost to \$1,081,000 (112%). Table 1 lists the "failure cost", the transportation cost associated with each facility failure.

Location	Failure Cost	% Increase
Sacramento, CA	1,081,229	112%
Harrisburg, PA	917,332	80%
Springfield, IL	696,947	37%
Montgomery, AL	639,631	26%
Austin, TX	636,858	25%
Transp. cost w/o failures	508,858	0%

Table 1 Failure costs of UFL solution

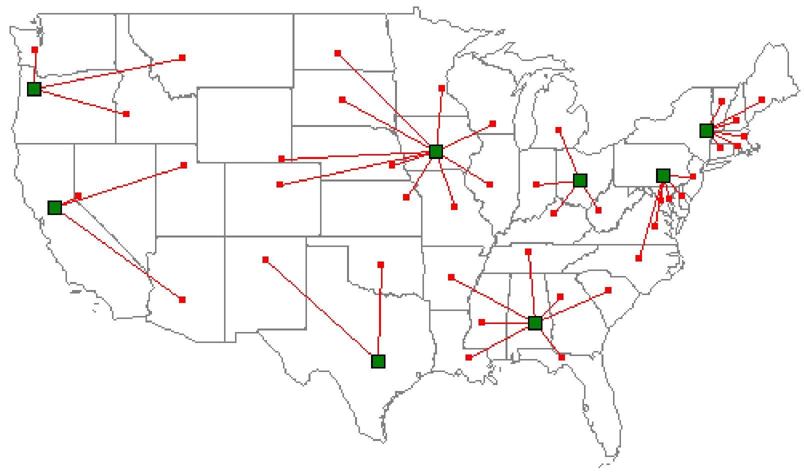


Figure 2 A more reliable solution

Snyder and Daskin (37) suggested that locating facilities in the capitals of CA, NY, TX, PA, OH, AL, OR, and IA (Figure 2) is a more reliable solution. In this solution, the maximum failure cost is reduced to $500,216, less than the smallest failure cost in Table 1. However, three additional facilities are used in this solution resulting in a total location and day-to-day transportation cost of $919,298 - a 7.25% increase from the UFL optimal solution.

Realistically, no company would accept a supply network with high normal operating costs just to hedge against very rare facility disruptions. In order to balance the trade-off between normal operating costs and failure costs, the network structure should depend on how likely the candidate sites may get disrupted, as well as their closeness to the potential customers. In Snyder and Daskin (37), all facility locations are assumed to have identical failure probabilities, which might not be very representative of practical situations. Let us illustrate how site-dependent failure probabilities impact the choice of facility locations. Specifically, suppose that the facilities are vulnerable to hurricane related disasters. Facilities located in the Gulf coast area (TX, LA, MS, AL and FL) all have a 10% chance of disruption, while other potential sites have a much lower failure probability of 0.1%. It is cost efficient here to hedge against disruption by locating facilities in the capitals of CA, PA, IL, GA and OK (Figure 3). In this solution, the two facilities along the Gulf coast (TX and AL) are moved to adjacent "safer" locations. Although the failure costs of CA and PA are high, we choose not to build "backup" facilities for them because their probability of disruption

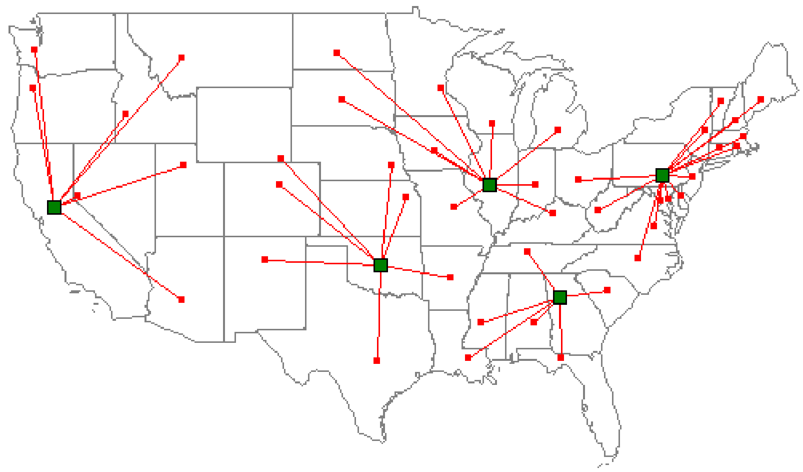


Figure 3 A cost efficient solution

is so small. The expected failure cost in this solution is about \$4,000, compared to \$130,344 in the UFL optimal solution in Figure 1, and the location and day-to-day transportation costs are increased by only 3.6%. Table 2 compares the normal operating costs and the expected failure costs of the three solutions.

Location	Failure Cost	Failure Probability	Expected Cost
Solution 1			
Sacramento, CA	1,081,229	0.001	1,081
Harrisburg, PA	917,332	0.001	917
Springfield, IL	696,947	0.001	697
Montgomery, AL	639,631	0.1	63,963
Austin, TX	636,858	0.1	63,686
Expected failure cost			130,344
Normal operating cost			857,128
Solution 2			
Location	Failure Cost	Failure Probability	Expected Cost
Sacramento, CA	500,216	0.001	500
Albany, NY	419,087	0.001	419
Austin, TX	476,374	0.1	47,637
Harrisburg, PA	409,383	0.001	409
Columbus, OH	434,172	0.001	434
Montgomery, AL	474,640	0.1	47,464
Salem, OR	389,484	0.001	389
Des Moines, IA	452,305	0.001	452
Expected failure cost			97,706
Normal operating cost			919,298
Solution 3			
Location	Failure Cost	Failure Probability	Expected Cost
Sacramento, CA	1,058,226	0.001	1,058
Harrisburg, PA	908,672	0.001	909
Springfield, IL	681,786	0.001	682
Atlanta, GA	679,022	0.001	679
Oklahoma City, OK	660,985	0.001	661
Expected failure cost			3,989
Normal operating cost			888,009

Table 2 Comparisons of the normal operating costs and the expected failure costs

In this paper, we seek to design supply networks that are both reliable and cost efficient. We minimize the expected transportation costs in both the regular and the failure scenarios (plus the fixed construction costs) to balance the trade-off between normal and emergency operating costs. The failure of each facility site is assumed to be independent and the probability is taken as a prior. Unlike in Snyder and Daskin (37), the failure probabilities are allowed to be site-dependent. The facility location decisions and customer assignments are made at the first stage, before any failures occur. Each customer can be assigned to up to $R \geq 1$ facilities to hedge against failures. After any disruptions occur, each customer is served by her closest assigned operating facility; if all her assigned facilities have failed then a penalty cost is charged. We feel that it is reasonable to restrict each customer's facility assignments to a pre-determined subset of all open facilities. In reality a customer may not be able to get service from all facilities due to system compatibility, limited capacity, or simply excessive transportation costs. Our computational results indicate that the choice of R has no significant impact on the network structure of the optimal solutions.

In the sequel we present two distinct models to address the reliable facility location problem – one discrete and the other continuous. Our discrete model is a linear mixed integer program (MIP) that computes the optimal facility locations and customer assignments. Unlike most scenario-based stochastic programming formulations that require exponentially many variable and constraints, our MIP formulation is polynomial in the number of candidate sites. The MIP is solved efficiently using a custom-designed Lagrangian Relaxation (LR) algorithm. However, due to the complexity of the underlying problem, the computational cost of this method for large problem instances can be excessive, and very few insights can be drawn from the numerical solutions. Thus, we develop a continuum approximation (CA) model which predicts the total system cost without the details of facility locations and customer assignments. Managerial insights (e.g., how the solution varies as key parameters change) can be drawn directly from the CA model. In addition, the CA approach can be used as a heuristic to find near-optimal solutions.

The remainder of this paper is organized as follows. We review literature on facility location in Section 2. In Section 3, we formulate the reliable facility location problem as a linear mixed-integer program (MIP) and provide a Lagrangian relaxation algorithm. Section 4 introduces the continuum approximation (CA) model. Computational results for both the discrete and the CA models are discussed in Section 5. Section 6 concludes the paper and discusses future research.

2. Literature Review

The extensive literature on facility location dates back to its original formulation in 1909 and the Weber problem (39). Traditionally, facility location problems are modeled as discrete optimization problems and solved with mathematical programming techniques. Daskin (15) and Drezner (16) provide good introductions to and surveys of this topic.

Recently, reliability issues in supply chain design are of particular interest. Most of the existing literature focuses on facility congestions from stochastic demand. Daskin (13, 14), Ball and Lin (2), ReVelle and Hogan (35), and Batta *et. al.* (3) all attempted to increase the system availability through redundant coverage.

Focus on system failures due to facility disruptions in supply chain design is gaining attention recently (33, 34). In the traditional locational analysis literature, Snyder and Daskin (37) propose an implicit formulation of the stochastic P-median and fixed-charge problems based on level assignments, where the candidate sites are subject to random disruptions with equal probability. Work by Zhan *et. al.* (40) and Berman et al. (4) relax the assumption of uniform failure probabilities. Zhan *et. al.* formulate the stochastic fixed-charged problem as a nonlinear mixed integer program and provides several heuristic solution algorithms. Berman et al. focus on an asymptotic property of the problem. They prove that the solution to the stochastic P-median problem coincides with the deterministic problem as the failure probabilities approach zero. They also propose heuristics with bounds on the worst-case performance.

All of the above literature is based on discrete optimization. Most of the discrete location models are NP-hard and thus it is difficult to obtain good solutions for large problem instances within a limited time frame. This fact motivates research on the continuum approximation (CA) method as an alternative to solving large-scale facility location problems. Building on the earlier work in Newell (28, 29) and Daganzo (7, 8), Daganzo and Newell (9) propose a CA approach for the traditional facility location problem. While conditions are slowly-varying, the cost of serving the demand near a facility location is formulated as a function of a continuous facility density (number of facilities per unit area) that can be efficiently optimized in a point-wise way. Note that the inverse of facility density is the influence area size (area per facility). The optimization yields the desired facility density and influence area size near each candidate location, which informs the design of discrete facility locations. It is shown in various contexts that the CA approach gives good approximate solutions to large-scale logistics problems by focusing on key physical issues such as the facility size and demand distribution (21, 22, 23, 5, 6, 10, 12). See Langevin et al. (26) and Daganzo (11) for reviews of the CA model. Ouyang and Daganzo (31) and Ouyang (32) propose methods to efficiently transform output from the CA model into discrete design strategies. The former reference also analytically validates the CA method for the traditional facility location problem. Recently, Lim et al. (27) propose a reliability CA model for facility location problems with uniform customer density. For simplification, a specific type of failure-proof facility is assumed to exist; a customer is always re-assigned to a failure-proof facility after its nearest regular facility fails, regardless of other (and nearer) regular facilities. We relax these rather strong assumptions in our work.

3. The Discrete Model

In this section we formulate the discrete model that minimizes the sum of the normal operating cost and the expected failure cost. We first show how this problem can be formulated as an MIP and then develop a Lagrangian relaxation algorithm to efficiently solve the problem.

3.1. Formulation

Define I to be the set of customers, indexed by i, and J to be the set of candidate facility locations, indexed by j. For the ease of notation, we also use I and J to indicate the cardinalities of the sets. Each customer $i \in I$ has a demand rate of λ_i. The cost to ship a unit of demand from facility $j \in J$ to customer $i \in I$ is denoted by d_{ij}. Associated with each facility $j \in J$ are the fixed location cost

f_j and the probability of failure $0 \le q_j < 1$. The events of facility disruptions are assumed to be independent.

Each customer is assigned to up to $R \ge 1$ facilities, and can be serviced by these and only these facilities. There is a cost ϕ_i associated with each customer $i \in I$ that represents the penalty cost of not serving the customer per unit of missed demand. This cost may be incurred even if some of her assigned facilities are still online, given that ϕ_i is less than the cost of serving i via any of these facilities. This rule is modeled using an "emergency" facility, indexed by $j = J$, that has fixed cost $f_J = 0$, failure probability $q_J = 0$ and transportation cost $d_{iJ} = \phi_i$ for customer $i \in I$.

The variables used in this model are the location variables (X), the assignment variables (Y) and the probability variables (P):

$$X_j = \begin{cases} 1, & \text{if a facility } j \text{ is open} \\ 0, & \text{otherwise} \end{cases}$$
$$Y_{ijr} = \begin{cases} 1, & \text{if facility } j \text{ is assigned to customer } i \text{ at level } r \\ 0, & \text{otherwise} \end{cases}$$
$$P_{ijr} = \text{probability that facility } j \text{ serves customer } i \text{ at level } r.$$

We employ the modeling techniques introduced by Snyder and Daskin (37) for assigning customers to facilities at multiple levels. A "level-r" assignment for a customer $i \in I$ will serve her if and only if all of her assigned facilities at levels $0, \cdots, r-1$ have failed. At optimality, each customer $i \in I$ should have exactly R assignments, unless i is assigned to the emergency facility at certain level $s < R$. If a customer i is indeed assigned to exactly R regular facilities at levels $0, \cdots, R-1$, she must also be assigned to the emergency facility J at level R to capture the possibility that all of the R regular facilities may fail. Finally, P_{ijr} is the probability that facility j serves customer i at level r, given her other assigned facilities at levels 0 to $r-1$.

The reliability UFL problem (**RUFL**) is formulated as:

$$\text{(RUFL)} \quad \text{Min} \sum_{j=0}^{J-1} f_j X_j + \sum_{i=0}^{I-1} \sum_{j=0}^{J} \sum_{r=0}^{R} \lambda_i d_{ij} P_{ijr} Y_{ijr} \tag{1a}$$

$$\text{s.t.} \sum_{j=0}^{J-1} Y_{ijr} + \sum_{s=0}^{r-1} Y_{iJs} = 1 \quad \forall\ 0 \le i \le I-1,\ 0 \le r \le R \tag{1b}$$

$$\sum_{r=0}^{R-1} Y_{ijr} \le X_j \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J-1 \tag{1c}$$

$$\sum_{r=0}^{R} Y_{iJr} = 1 \quad \forall\, 0 \le i \le I-1 \tag{1d}$$

$$P_{ij0} = 1 - q_j \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J \tag{1e}$$

$$P_{ijr} = (1-q_j) \sum_{k=0}^{J-1} \frac{q_k}{1-q_k} P_{i,k,r-1} Y_{i,k,r-1} \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J,\ 1 \le r \le R \tag{1f}$$

$$X_j,\ Y_{ijr} \in \{0,1\} \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J,\ 0 \le r \le R. \tag{1g}$$

The objective function (1a) is the sum of the fixed costs and the expected transportation costs. Constraints (1b) enforce that for each customer i and each level r, either i is assigned to a regular facility at level r or she is assigned to the emergency facility J at certain level $s < r$ (taking $\sum_{s=0}^{r-1} Y_{iJs} = 0$ if $r = 0$). Constraints (1c) limit customer assignments to only the open facilities, while constraints (1d) require each customer to be assigned to the emergency facility at a certain level. (1e)-(1f) are the "transitional probability" equations. P_{ijr}, the probability that facility j serves customer i at level r, is just the probability that j remains open if $r = 0$. For $1 \le r \le R$, P_{ijr} is equal to $\frac{q_k(1-q_j)}{1-q_k} P_{i,k,r-1}$ given that facility k serves customer i at level $r-1$. Note that constraints (1b) imply that $Y_{i,k,r-1}$ can equal 1 for at most one $k \in J$, which guarantees correctness of the transitional probabilities.

Formulation (1a)-(1g) is nonlinear. However, the only nonlinear terms are $P_{ijr} Y_{ijr},\ 0 \le i \le I-1,\ 0 \le j \le J,\ 0 \le r \le R$, each being a product of a continuous variable and a binary variable. We apply the linearization technique introduced by Sherali and Alameddine (36) by replacing each $P_{ijr} Y_{ijr}$ with a new variable W_{ijr}. For each $0 \le i \le I-1$, $0 \le j \le J$ and $0 \le r \le R$ a set of new constraints is added to the formulation to enforce $W_{ijr} = P_{ijr} Y_{ijr}$:

$$W_{ijr} \le P_{ijr} \tag{2a}$$

$$W_{ijr} \le Y_{ijr} \tag{2b}$$

$$W_{ijr} \ge 0 \tag{2c}$$

$$W_{ijr} \ge P_{ijr} + Y_{ijr} - 1. \tag{2d}$$

The linearized formulation (**LRUFL**) is stated below:

$$\text{(LRUFL)} \quad \text{Min} \sum_{j=0}^{J-1} f_j X_j + \sum_{i=0}^{I-1} \sum_{j=0}^{J} \sum_{r=0}^{R} \lambda_i d_{ij} W_{ijr} \tag{3a}$$

$$\text{s.t. } (1b)-(1d) \tag{3b}$$

$$P_{ijr} = (1-q_j) \sum_{k=0}^{J-1} \frac{q_k}{1-q_k} W_{i,k,r-1} \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J,\ 1 \le r \le R \tag{3c}$$

$$(2a)-(2d) \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J,\ 1 \le r \le R \tag{3d}$$

$$X_j,\ Y_{ijr} \in \{0,1\} \quad \forall\, 0 \le i \le I-1,\ 0 \le j \le J,\ 0 \le r \le R. \tag{3e}$$

Unlike scenario based stochastic programming problems that have exponentially many variables and constraints, our formulation is compact and polynomial in size. Proposition 1 shows the equivalence of (LRUFL) to the scenario based formulation.

Proposition 1. *If $R = J$, then formulation* (1a)-(1g) *is equivalent to the stochastic programming formulation that covers all failure scenarios.*

In general, (LRUFL) is not equivalent to the scenario based formulation if $R < J$. However, our computational results show that the choice of R has little impact on the optimal facility locations. Similar to classic facility location problems, we do not enforce a customer to be served by her closest open facility in our formulation. It is proved in Snyder and Daskin (37) that the optimal solution always assigns a customer to open facilities level by level in increasing order of distance, given that all facilities are equally likely to fail. The following proposition extends this result to the case where the facility failure probabilities are different across sites.

Proposition 2. *In any optimal solution* $(\mathbf{X}, \mathbf{Y}, \mathbf{P})$ *of (RUFL), if $Y_{ijr} = 1$ and $Y_{ik,r+1} = 1$, then $d_{ij} \le d_{ik}$, for all $0 \le i \le I-1,\ 0 \le j \le J$, and $0 \le r \le R$.*

Proposition 2 tells us that for a given subset of facilities assigned to a customer, the optimal assignment levels only depend on the distances from the customer to these facilities. However, if more than R facilities are constructed, it may be sub-optimal to assign each customer to her R closest facilities. As the following example shows, it may be optimal to assign a customer to a facility that is farther away but less likely to fail.

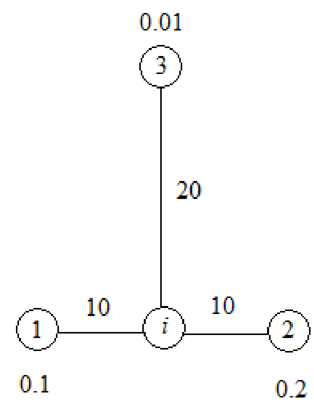


Figure 4 Fraction of a supply network

Example 1. Consider a fraction of a supply network depicted in Figure 4. Three facilities are constructed around customer i. The distances from i to the facilities are $d_{i1} = d_{i2} = 10$, and $d_{i3} = 20$. The failure probabilities of the three facilities are $q_1 = q_3 = 0.1$, and $q_2 = 0.2$. The demand rate at i is $\lambda_i = 1$ and the penalty for not serving a unit of demand is $\phi_i = 1000$. Suppose that each customer is only allowed one primary and one back-up facility ($R = 2$). If we assign customer i to the two closest facilities 1 and 2, then the expected transportation/penalty cost for this customer is 29.8. However, the optimal strategy is to assign i to facilities 1 and 3, which reduces the expected transportation cost to 11.98.

Example 1 implies that even with fixed facility locations, the customer assignment problem is combinatorial and requires more sophisticated solution methods. In this regard, our model is harder than that in Snyder and Daskin (37), in which the customer assignment problem with fixed facility locations can be easily solved. We discuss how to decompose the customer assignment problem using Lagrangian relaxation in section 3.2, and how to efficiently solve the individual customer assignment problem in section 3.3.

3.2. The Lagrangian Relaxation Algorithm

The linear mixed-integer program (LRUFL) can be solved using commercial software packages like ILOG CPLEX, but generally such an approach takes an excessively long time even for moderately sized problems. This fact motivates the development of a Lagrangian relaxation algorithm. Relaxing constraints (1c) with multipliers μ yields the following objective function:

$$\sum_{j=0}^{J-1}(f_j - \sum_{i=0}^{I-1}\mu_{ij})X_j + \sum_{i=0}^{I-1}\sum_{j=0}^{J}\sum_{r=0}^{R}\lambda_i d_{ij} W_{ijr} + \sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\sum_{r=0}^{R-1}\mu_{ij} Y_{ijr}.$$

For given value of μ, the optimal value of $\mathbf{X}$ can be found easily:

$$X_j = \begin{cases} 1 & \text{if } f_j - \sum_{i=0}^{I-1}\mu_{ij} < 0 \\ 0 & \text{otherwise.} \end{cases}$$

To find the optimal $\mathbf{Y}$, the customer assignment decision, note that the problem is separable in i. For given Lagrangian multipliers μ an individual customer's assignment problem is referred to as the *relaxed subproblem* (**RSP**). The complexity of RSP is demonstrated in Example 1, in which the simple heuristic leads to suboptimal solutions. We discuss efficient algorithms for RSP in Section 3.3.

We use standard subgradient optimization technique to update the Lagrangian multipliers μ, as described in Fisher (18). If the Lagrangian process fails to converge in a certain number of iterations, we use branch-and-bound to close the gap. As a benchmark, we tested our algorithm on the same data sets used by Snyder and Daskin (37). The computational results are discussed in Section 5.1.

3.3. The Relaxed Subproblem

Below is the MIP formulation of the relaxed subproblem with respect to customer i (RSP$_i$). For ease of notation, we omit the subscript i in Y_{ijr}, P_{ijr} and W_{ijr}.

$$(\text{RSP}_i) \quad \text{Min } \Phi_i = \sum_{j=0}^{J}\sum_{r=0}^{R}\lambda_i d_{ij} W_{jr} + \sum_{j=0}^{J-1}\sum_{r=0}^{R-1}\mu_{ij} Y_{jr} \tag{4a}$$

$$\text{s.t. } \sum_{j=0}^{J-1} Y_{jr} + \sum_{s=0}^{r-1} Y_{Js} = 1 \quad \forall\ 0 \le r \le R \tag{4b}$$

$$\sum_{r=0}^{R-1} Y_{jr} \le 1 \quad \forall\ 0 \le j \le J-1 \tag{4c}$$

$$\sum_{r=0}^{R} Y_{Jr} = 1 \tag{4d}$$

$$P_{j0} = 1 - q_j \quad \forall\ 0 \le j \le J \tag{4e}$$

$$P_{jr} = (1 - q_j) \sum_{k=0}^{J-1} \frac{q_k}{1 - q_k} W_{k,r-1} \quad \forall\ 0 \le j \le J,\ 1 \le r \le R \tag{4f}$$

$$Y_{jr} \in \{0, 1\} \quad \forall\ 0 \le j \le J,\ 0 \le r \le R \tag{4g}$$

$$(2a) - (2d). \tag{4h}$$

We propose two methods to solve the relaxed subproblem: one exact algorithm that finds the optimal customer assignment, and one fast approximate algorithm that provides an lower bound.

3.3.1. An Exact Algorithm Following a similar argument to Proposition 2, given the subset of facilities that serve a certain customer, it is optimal to assign this customer to the facilities level by level in increasing order of the distances. Therefore the objective value of (RSP_i) only depends on the set of facilities that serve customer i. Define $\Phi_i(S)$ to be the minimum cost to serve customer i, using only facilities in S; i.e.,

$$\Phi_i(S) = \text{Min} \sum_{j=0}^{J} \sum_{r=0}^{R} h_i d_{ij} W_{jr} + \sum_{j \in S} \mu_{ij} \tag{5a}$$

$$\text{s.t. } (4b) - (4g) \tag{5b}$$

$$\sum_{r=0}^{R-1} Y_{jr} = 0 \quad \forall\ j \in \{1, \cdots, J-1\} \setminus S. \tag{5c}$$

It is clear that (RSP_i) is equivalent to the following minimization of a set function (MSF_i):

$$(\mathrm{MSF}_i) \quad \text{Min } \Phi_i(S) \tag{6a}$$

$$\text{s.t. } S \subseteq \{0, \cdots, J-1\} \tag{6b}$$

$$|S| \le R. \tag{6c}$$

We solve (MSF) using a special branch-and-bound algorithm, based on some unique properties of the set function Φ, as described in Proposition 3.

Proposition 3. *The set function Φ_i is supermodular, for all $i = 0, \cdots, I-1$.*

The minimization of a supermodular set function can be solved more efficiently, using the branch-and bound algorithm developed by Goldengorin et al. (20). The algorithm keeps track of A and

B, the set of facilities that have been forced in or out for each branch-and-bound node. The supermodularity of the objective function allows us to force out a facility if its addition to set A does not reduce the total cost. In an unconstrained problem, it is also possible to force in a facility if its deletion from $\{0, \cdots, J-1\} \setminus B$ increases the total cost. However, since MSF is subject to the cardinality constraint (6c), the second option does not apply here.

3.3.2. An Approximate Solution Although the exact algorithm in Section 3.3.1 takes advantage of special structure of the problem, its worst case complexity is still exponential. In this section we provide a fast approximate algorithm that finds lower bounds for the Lagrangian procedure.

In our approximate solution, we replace the variable probability P_{jr} with fixed numbers. Let $j_0, j_1, \cdots, j_{J-1}$ be an ordering of the facilities such that $q_{j_0} \le q_{j_1} \le \cdots \le q_{J-1}$. Define

$$\begin{aligned} \alpha_r &= (1 - q_{j_r}) \prod_{\ell=0}^{r-1} q_{j_\ell} \\ \beta_r &= \prod_{\ell=0}^{r-1} q_{j_\ell}. \end{aligned}$$

We define a reformulation of the relaxed subproblem (RRSP) by replacing P_{jr} with α_r if $0 \le j \le J-1$, and replacing P_{Jr} with β_r:

$$(\text{RRSP}_i) \quad \text{Min} \sum_{j=0}^{J-1} \sum_{r=0}^{R-1} (\lambda_i d_{ij} \alpha_r + \mu_{ij}) Y_{jr} + \sum_{r=0}^{R} \lambda_i d_{iJ} \beta_r Y_{Jr} \tag{7a}$$

$$\text{s.t.} \sum_{j=0}^{J-1} Y_{jr} + \sum_{s=0}^{r-1} Y_{Js} = 1 \quad \forall\ 0 \le r \le R \tag{7b}$$

$$\sum_{r=0}^{R-1} Y_{jr} \le 1 \quad \forall\ 0 \le j \le J-1 \tag{7c}$$

$$\sum_{r=0}^{R} Y_{Jr} = 1 \tag{7d}$$

$$Y_{jr} \in \{0, 1\} \quad \forall\ 0 \le j \le J,\ 0 \le r \le R. \tag{7e}$$

The following proposition states that we can solve (RRSP) for a lower bound of (RSP).

PROPOSITION 4. *The (RRSP) formulation* (7a)-(7e) *yields a lower bound to the relaxed subproblem* (4a)-(4h).

We note that the (RRSP) formulation (7a)-(7e) leads to a combinatorial assignment problem, which can be solved in strongly polynomial time using the Hungarian algorithm (25). In our numerical tests, we use both the exact and the approximate algorithm to get the best combination of speed and accuracy.

Although our compact MIP formulation and the Lagrangian relaxation algorithm are significant improvements over scenario based stochastic programming formulations, the worst case complexity is still exponential, due to the NP-hardness of the underlying problem. Furthermore, because only numerical results are available from the discrete model, very few managerial insights can be drawn from the optimal solutions. In the next section, we overcome these difficulties by introducing the continuum approximation (CA) model.

4. The Continuum Approximation Model

In this section, we restrict our attention to the planar version of the reliable facility location problem, defined over a large set of customers in the continuous metric space $\mathcal{S} \subseteq \mathbb{R}^2$, where the demand rate λ, fixed cost f, failure probability q and the penalty cost ϕ are continuous functions of the location $x \in S$. All these spatial attributes are assumed to vary continuously and slowly in x. Suppose that the cost units are set so that the transportation cost for serving a unit demand at x by a facility at x_j is equal to the distance measured by the Euclidean metric, $\|x - x_j\|$. In addition, we assume that $\phi(x) \geq \max\{\|x - x_j\| : \forall x_j \in S\}$, for all $x \in S$. Under such assumption, a customer shall always be assigned to exactly R facilities if available.

4.1. Infinite Homogeneous Plane

We first consider the case where $\mathcal{S} = \mathbb{R}^2$, and all parameters, λ, ϕ, q, f, are constant everywhere. We will first identify optimal results for this simpler case and then use them as building blocks to design solution methods for more general cases.

It is clear that on a homogeneous plane, given any set of locations $\mathbf{x} = \{x_1, \cdots x_n\}$, a customer should always be assigned to the R nearest facilities. Otherwise we could reduce the cost by simply switching this customer over to a closer facility. Thus, any design $\mathbf{x}$ (subject to failure) determines the assignment of customer demand.

¿From the perspective of a generic facility j, it will serve every customer on the 2-d plane with a certain probability (depending on its failure probability and that of other facilities). The whole area $\mathcal{S}$ can be partitioned into non-overlapping subareas $\mathcal{R}_{j0}, \mathcal{R}_{j1}, \mathcal{R}_{j2}, \cdots$, such that $\mathcal{R}_{jk}, \forall k$, contains the subset of customers for whom facility j is the $(k+1)^{\text{th}}$ nearest facility. With this definition, for every j there is a non-overlapping partition if we ignore the boundaries of these subareas,

$$\bigcup_k \mathcal{R}_{jk} = \mathcal{S}, \text{ and } \mathcal{R}_{jk} \bigcap \mathcal{R}_{jk'} = \varnothing, \forall k' \neq k.$$

Since every customer will always go to the nearest available facility, the customer at $x \in \mathcal{R}_{jk}$ will go to facility j only after all of its k "nearest" facilities have failed, and if $k+1 \leq R$. Facility j will serve customers at x with the following service probability:

$$P(x, x_j|\mathbf{x}) = (1-q)q^k, \text{ if } x \in \mathcal{R}_{jk}, \tag{8}$$

which decreases with k.

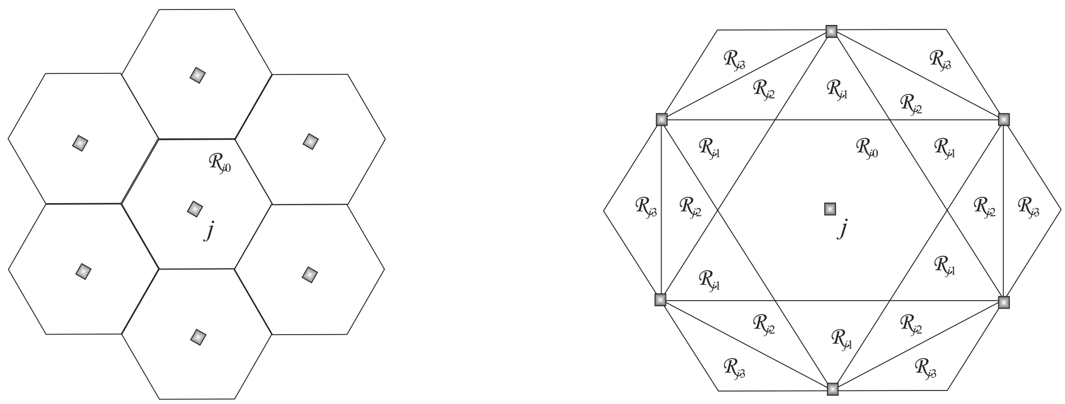
Particularly, the *initial service area* $\mathcal{R}_{j0}$ denotes the subarea of $\mathcal{S}$ served by facility j before any failure; i.e., $\mathcal{R}_{j0} := \{x : \|x - x_j\| \leq \|x - x_i\|, \forall i\} \subseteq \mathcal{S}$. Further denoting the set of initial service areas by $\mathbf{R} := \{\mathcal{R}_{10}, \mathcal{R}_{20}, \ldots, \mathcal{R}_{n0}\}$, they should form another area partition (ignoring boundaries):

$$\bigcup_j \mathcal{R}_{j0} = \mathcal{S} \text{ and } \mathcal{R}_{j0} \bigcap \mathcal{R}_{j'0} = \varnothing, \forall j' \neq j.$$

Proposition 5 shows that the optimal facility design on a homogeneous plane has the following special structure.

Proposition 5. *In an infinite homogeneous Euclidean plane, the optimal initial service areas should form a regular hexagon tessellation of the plane, while the facilities are at the centroids of the initial service areas; see Figure 5(a).*

With Proposition 5, we can estimate the expected cost incurred by one facility on an infinite homogeneous plane. First of all, the probability that a particular facility serves a customer diminishes approximately exponentially with the distance between them. This is because the number of



(a) (b)

Figure 5 Regular hexagon tessellation in an infinite homogeneous 2-d Euclidean plane: (a) Initial service areas; (b) Service subarea partition for facility j

facilities closer to the customer (i.e., k), is approximately proportional to the square of the distance, while the service probability in (8) decreases exponentially with k. From the facility's perspective, the number of available customers grows only polynomially with the distance. Hence, the expected service cost incurred to one facility on an infinite homogenous plane is bounded from above even when $R \to \infty$.

The regular hexagonal tessellation design in Figure 5(a) obviously leads to the service subarea partition in Figure 5(b). An arbitrary facility j has an initial service area size $A := |\mathcal{R}_{j0}|$ and may fail with a probability of q. For this facility to serve customers that only go to R nearest facilities, we define the following useful term:

$$L := \int_{x \in \mathcal{S}} \|x - x_j\| P(x, x_j|\mathbf{x}) dx = \sum_{k=0}^{R-1} \int_{x \in \mathcal{R}_{jk}} \|x - x_j\| (1-q) q^k dx, \tag{9}$$

where the second equality holds from (8).

Note that λL can be taken as the expected transportation cost of a generic facility to serve all its potential customers, while the average traveled distance of a customer is given by $L/(RA)$. Apparently, $L < \infty$ (since $q < 1$) and its value should only depend on three factors, A, R and q.

By dimensional analysis and the Buckingham-Π Theorem (24), the dimensionless quantities, $L/A^{\frac{3}{2}}$, R, and q, must be interdependent; *i.e.*, there must exist a unique function G such that

$$L/A^{\frac{3}{2}} = G(R, q). \tag{10}$$

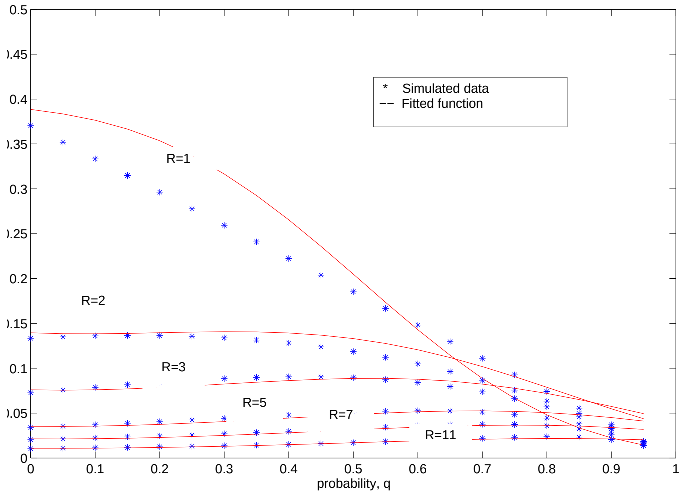


Figure 6 Simulated and fitted $L/(RA)^{3/2}$ for Euclidean metric

We note that $G(R,q)$ can be interpreted as the expected transportation cost for a generic facility to serve all its potential customers when $\lambda = 1$ and $A = 1$. The exact functional form of G is unknown; however it only depends on the distance metric and can be estimated by a simulation.

For purpose of our analysis, we hypothesize that $\ln(L/A^{3/2})$ can be approximated by a linear function of a list of polynomial terms of R and q. For the Euclidean metric, least squares regression with the simulated data in Figure 6 (with $1 \le R \le 11, 0 \le q \le 0.95$) yields

$$G(R,q) \approx \exp(-0.930 - 0.223q + 4.133q^2 - 2.906q^3 - 1.542\pi q^2/R), \tag{11}$$

The R-square value for the above regression equals 0.96, indicating a very good fit, especially for $R \ge 2$ and $q \le 0.5$ (the realistic range of parameters for the reliability problem). In the numerical example, we will use (11) to approximate $G(R,q)$. It should be noted, however, that (11) is by no means the only way to estimate $G(R,q)$; rather, it is a plausible and simple choice. The CA approach presented in this paper can still be applied with any alternatives of (11).

Following from (10), we define the average per unit area cost of the system to be

$$z(A) := (f + \lambda L)/A + \lambda\phi q^R = f/A + \lambda G(R,q)\sqrt{A} + \lambda\phi q^R, \tag{12}$$

where f/A is the fixed cost per unit area (since each initial service area consists of a single facility), and $\lambda\phi q^R$ is the expected penalty cost per unit area for failure to serve the customers.

The optimal size of the initial service area, and thus the minimum system cost can be obtained by minimizing the average cost per unit area; i.e.,

$$\min_{A>0} z(A).$$

4.2. Heterogeneous Plane

To handle more realistic cases, we allow the parameters λ, ϕ, q, f to be varying functions of the location x in a bounded area $\mathcal{S}$ in this section. Similar to the homogeneous case, we look for a continuous function, $A(x) \in \mathbb{R}_+, x \in \mathcal{S}$, that approximates the initial service area size of a facility near x.

We assume that $\mathcal{S}$ is far larger than $A(x)$; i.e. *approximately infinite*; and all parameters, $f(x)$, $\lambda(x)$, $q(x)$ and $\phi(x)$ are slowly varying; i.e. *approximately constant*. When the parameters are approximately constant over a region comparable to the size of several influence areas, the influence area size $A(x)$ should also be approximately constant on that scale. In this case, we define $z(A(x),x)$ to be the cost of serving a unit area near x when the influence area size is approximately $A(x)$; i.e.,

$$z(A(x),x) := f(x)/A(x) + \lambda(x)G(R,q(x))\sqrt{A(x)} + \lambda(x)\phi(x)q(x)^R. \tag{13}$$

The total system cost can be approximated by integrating $Z(A(x),x)$ over the service area $\mathcal{S}$:

$$\int_{x\in\mathcal{S}} z(A(x),x)dx. \tag{14}$$

The optimal system cost can be approximated by solving a point-wise optimization problem for each $x \in \mathcal{S}$:

$$\min_{A(x)>0} z(A(x),x). \tag{15}$$

4.3. Feasible Discrete Location Design

Formula (14) yields an estimate of the total system cost without providing a discrete facility design. However, the optimal initial service area sizes, $A^*(x), \forall x \in \mathcal{S}$, can be used as guidelines to obtain feasible discrete location designs.

The optimal number of initial service areas, n^*, is approximately given by

$$n^* := \int_{\mathcal{S}} [A^*(x)]^{-1} dx.$$

The disk model by Ouyang and Daganzo (31) searches for a set of n non-overlapping disks, each having a round shape (i.e., approximating hexagons) and a proper size, that cover most of $\mathcal{S}$. A disk centered at x will have size $\alpha A^*(x)$, where the scaling parameter α is slightly smaller than 1 to ensure that the round disks can jointly cover most of $\mathcal{S}$ without leaving the region.

The disks move within $\mathcal{S}$ in search of a non-overlapping distribution pattern. To automate the sliding procedure, repulsive forces acting on the centers of the disks are imposed on any overlapping disks and on any disks that lie outside of $\mathcal{S}$. The disks then move under these forces in small steps, and the disk sizes and forces are updated simultaneously. Ouyang and Daganzo (31) and Ouyang (32) provide detailed discussions on how to choose step sizes, how to introduce necessary random perturbations, and how to decrease α incrementally until all forces vanish (i.e., when a desired non-overlapping pattern is found). Then, the disk centers will be used as the facility locations and the customer demands will be assigned accordingly. This procedure will give a near-optimal feasible solution to the planar problem.

4.4. Remarks on the CA Model

The point-wise optimization problem (15) can be solved in closed-form to provide managerial insights. Section 5.3 shows a simple example of sensitivity analysis based on the CA formula.

It should be noted that the implementation framework of the CA model does not rely directly on the specific function form of $G(R,q)$. Other forms of $G(R,q)$ (e.g., those providing clearer managerial insights or better regression fit) can be easily applied.

There are several potential sources of inaccuracy in the CA model. First, the CA model is expected to perform well for large-scale systems with slow-varying conditions. This is because we have assumed constant conditions in a fairly large area (with size $\approx RA$). If, in certain cases, system parameter values change rapidly with x, the CA model may not yield very accurate results.

However, much like the traditional CA method, the error here is likely to cancel out across different customers due to the law of large numbers. Section 5 will use numerical examples to show that the continuum approximation model yields near-optimal results despite violations to the assumption on slow-varying conditions.

Second, we ignore the boundary of $\mathcal{S}$ while developing $P(\cdot)$. Such simplification, however, is not likely to introduce severe errors because the probability that a facility serves a particular customer diminishes geometrically (rapidly for small q) with the distance between them. Therefore, the influence of the boundary for large-scale problems (i.e., $n \gg 1$) is likely to be small. Also, Gersho (19) showed that even for finite (but large) two-dimensional planes, the influence will only be significant for initial service areas directly touching the periphery of $\mathcal{S}$, and hence small.

Nevertheless, the influence of the boundary will cause remarkable errors in certain ill-posed situations. Note that in the derivations in Section 4.1, we assume that $n \geq R$. In extreme cases (e.g., facility set-up costs f are very high compared with customer penalties ϕ), the ideal facility density shall be very low (i.e., $A \to \infty$). When the customer area $|\mathcal{S}|$ is finite, it is possible that the number of total facilities $n < R$ (or even $n \to 0$). In such cases, it is not possible to force every customer to consider R facilities—the assumption used to derive cost formula (9) is violated. Hence, sometimes it is not reasonable to specify that every customer shall consider at most R facilities. Alternatively, we may postulate that a customer, if served at all, shall only be served by a facility within a *maximum service distance*. Appendix B provides discussions on such formulations.

5. Computational Results

We conducted a series of computational experiments to test the performance of the discrete and the CA model. We also demonstrate the use of CA for sensitivity analysis.

5.1. Discrete Model Results

The discrete model was tested on two types of networks - the “real” network based on the US map with 49 or 88 nodes and the “random” network generated on a unit square region with 50 or 100 nodes (the data set was kindly provided by L. Snyder and is available from his website (38)). The failure probabilities q_j in the real networks are calculated using $q_j = 0.1e^{-D_j/400}$, in which D_j is the great cycle distance (in miles) between location j and New Orleans, LA. In the random networks, q_j are randomly generated from a uniform distribution between 0 and 0.2. For each data set, we test our algorithm for $R = 2$, 3 and 4. The Lagrangian relaxation/branch-and-bound procedure is executed to a tolerance of 0.5%, or up to 3600 seconds (60 minutes) in CPU time. The algorithm was coded in C++ and tested on an Intel Pentium 4 3.20GHz processor with 1.0 GB RAM under Linux. Parameter values for the Lagrangian relaxation algorithm can be found in Table 3, and the algorithm performance is summarized in Table 4.

Parameter	Value
Optimal tolerance	0.005
Maximum number of approximate iterations at root node	1000
Maximum number of exact iterations at root node	500
Maximum number of approximate iterations at child nodes	200
Maximum number of exact iterations at child nodes	100
Initial value for μ_{ij}	optimal dual of LP relaxation

Table 3 Parameter values for the Lagrangian relaxation

Nodes	Re-assignment level	Root LB	Root UB	Root gap (%)	Overall UB	Overall gap (%)	CPU time (sec.)
49	2	878,856	880,098	0.141	880,098	0.141	9
49	3	870,229	874,423	0.482	874,423	0.482	19
49	4	870,143	874,323	0.481	874,323	0.481	50
88	2	1,230,120	1,234,860	0.385	1,234,860	0.385	66
88	3	1,217,650	1,223,610	0.489	1,223,610	0.489	458
88	4	1,217,270	1,223,290	0.494	1,223,290	0.494	883
50	2	6,336	6,363	0.415	6,363	0.415	1
50	3	6,331	6,363	0.496	6,363	0.496	3
50	4	6,331	6,363	0.496	6,363	0.496	4
100	2	11,922	11,981	0.494	11,981	0.494	22
100	3	11,921	11,980	0.499	11,980	0.499	35
100	4	11,911	11,970	0.489	11,970	0.489	21

Table 4 LR Algorithm Performance

We notice that the maximum re-assignment level R does not affect the optimal facility locations in all of our test instances, although a higher R in general helps to reduce the optimal cost. Figure 7 and 8 illustrate the optimal facility locations for the 49-node and the 88-node problem respectively. Table 5 and 6 list the percentage of covered demand, fixed cost, and failure probability at each optimal facility location in the two problem instances. In both cases, the optimal solutions avoid highly risky areas such as LA and MS. In areas with moderate risk, clusters of facilities are formed to hedge against possible disruptions. In areas with low risk (OR, CA and AZ), facilities are located relatively sparsely.

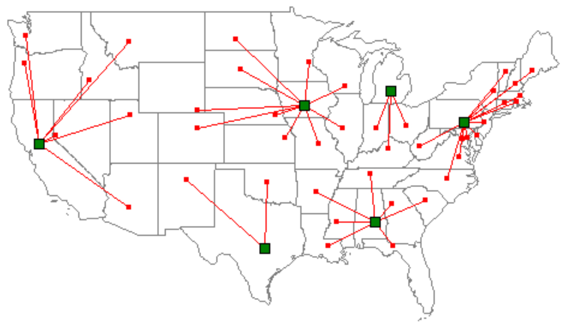


Figure 7 Optimal Solution to the 49-Node Problem

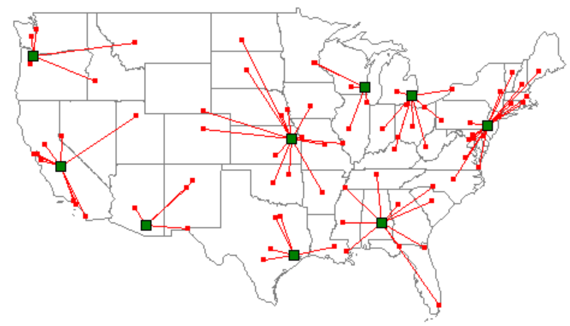


Figure 8 Optimal Solution to the 88-Node Problem

Location	Demand Covered	Fixed Cost	Failure Probability
Sacramento, CA	19%	115,800	0.001
Austin, TX	9%	72,600	0.043
Harrisburg, PA	29%	38,400	0.012
Lansing, MI	12%	48,400	0.013
Montgomery, AL	17%	62,200	0.053
Des Moines, IA	15%	49,500	0.014

Table 5 Optimal Locations for the 49-Node Problem

Location	Demand Covered	Fixed Cost	Failure Probability
Houston, TX	10%	58,000	0.043
Philadelphia, PA	27%	49,400	0.012
Detroit, MI	10%	25,600	0.013
Milwaukee, WI	10%	53,500	0.014
Portland, OR	3%	59,200	0.001
Tucson, AZ	5%	66,800	0.001
Fresno, CA	17%	80,300	0.001
Montgomery, AL	9%	62,200	0.053
Topeka, KS	8%	48,800	0.024

Table 6 Optimal Locations for the 88-Node Problem

Our algorithm appears to have performed efficiently on the random test instances. However, the algorithm convergence is slow for some of the real test instances. Due to the computational complexity of finding exact solutions for the relaxed subproblems (RSP), we can only afford to run the exact algorithm for a very limited number of iterations (100 at each B&B node as compared to 2000 in Snyder and Daskin (37)), The approximate algorithm for RSP is fast, but the bound it provides can be lax in some circumstances. For fixed-charge location problems, it is generally more efficient to relax the assignment constraint (1b) instead of the linking constraint (1c). However, in our case relaxing (1c) allows us to decompose the customer assignment problem, a key step in the algorithm development. To improve the efficiency of the algorithm, we need to find ways to solve the relaxed subproblems more quickly and better decomposiition mechanisms with tighter bounds. These goals will be pursued in our future research.

5.2. CA Model Resutls

To test the performance of the CA approach, we consider a $[0,1]\times[0,1]$ unit square, where customer demands are distributed according to a density function $\lambda(x)$. A facility built at location x incurs a cost of $f(x)$ and may fail with probability $q(x)$. As a benchmark, we also construct and solve analogous discrete test instances by partitioning the unit square into $7\times 7=49$ identical square cells; the center of each cell represents a candidate facility location as well as the consolidation point of the customer demand from that cell.

We group our test instances into two categories: the homogeneous case and the heterogeneous case. In the homogeneous case, all system parameters are constant over space; i.e., $\lambda(x)=\lambda, f(x)=f$, and $q(x)=q\ \forall x$. We generate 16 test instances with key parameters taking values from $q\in\{0.05,0.10,0.15,0.20\}$, $\lambda\in\{50000,100000,150000,500000\}$. The fixed cost is $f=1000$ for all 16 instances.

In the heterogeneous case, we let the key parameters be continuous functions that can vary across space, defined as follows:

$$\lambda(x)=\lambda(1+\Delta_\lambda\cos(\pi x_{[2]})),\ f(x)=fe^{-\|x\|},\ q(x)=q[1+\Delta_q\cos(\pi\|x\|)],\forall x,$$

where $\|x\|$ is the Euclidean distance from x to the origin, and $x_{[2]}$ is the second coordinate of x. Note that q and λ control the average magnitude of failure probabilities and demand densities, while Δ_q and Δ_λ control the variability of these parameters. We generate 20 test instances in total, with q and Δ_q drawn from $q\in\{0.1,0.2\}$ and $\Delta_q\in\{0.1,0.2,0.3,0.4,0.5\}$, and Δ_λ taking values from $\Delta_\lambda\in\{0.0,1.0\}$. The average demand density is set to be $\lambda=100000$ and the average fixed cost is set to $f=1000$ for all 20 instances.

The penalty cost is fixed at $\phi(x) = \sqrt{2}$, and the reassignment level is set to $R = 2$ for all 36 test instances in both categories.

For each test instance, we list the CA predicted cost Z_{CA}, the feasible discrete solution found by the disk algorithm Z_{CA}^D, and the optimal solution found by the Lagrangian algorithm Z_{LR}^D. The superscript 'D' here stands for discrete customer demand (i.e., assuming the continuous demand distribution as an approximation). The number of facilities constructed in the CA and the discrete model solutions are also listed respectively under n_{CA}^* and n_{LR}^*. For comparison, we calculate ε^D, the gap between Z_{CA}^D and Z_{LR}^D for each test instance. The results are summarized in Table 7 for the homogeneous cases and in Table 8 for the heterogeneous cases. We are also interested in comparing the model results assuming continuous demand (i.e., considering the aggregated discrete demand as an approximation), and those results are listed in Appendix C.

Table 7 CA cost estimates, feasible solutions, and LR solutions for the homogeneous cases

q	$\lambda(10^4)$	Z_{CA}	Z_{CA}^D	Z_{LR}^D	n_{CA}^*	n_{LR}^*	ε^D (%)
0.05	5	13908.5	14694.1	14281.1	5	5	2.89
0.1	5	14430.9	15600.1	15134.1	5	5	3.08
0.15	5	15345.4	16762.3	16251.5	5	5	3.14
0.2	5	16632.0	18180.6	17633.4	5	5	3.1
0.05	10	22151.2	22895.4	22607.4	7	7	1.27
0.1	10	23199.4	24470.8	24281.5	7	8	0.78
0.15	10	25015.7	26605.3	24281.5	7	8	0.74
0.2	10	27568.7	29298.8	28954.8	7	9	1.19
0.05	15	28704.7	30538.9	29460.2	10	10	3.66
0.1	15	29309.9	32702.9	31807.3	10	10	2.82
0.15	15	30667.5	35790.7	34988.7	10	11	2.29
0.2	15	32880.6	39546.5	38945.2	10	10	1.54
0.05	50	65504.6	65586.0	54164.0	21	49	21.09
0.1	50	70771.4	72551.7	62506.1	21	49	16.07
0.15	50	79752.2	83885.3	74026.1	21	49	13.32
0.2	50	92354.9	95861.2	88724.3	21	49	8.04

Our test results show that the CA method is a promising tool for finding near optimal solutions. Even under the discrete demand distribution (i.e., considering the continuous demand distribution as an approximation), the optimality gap is below 4% in most test instances. Particularly, even when the demand distribution is significantly variable across space ($\lambda(x)$ varying from 0 to 2λ), the gaps is mostly within $4 - 7\%$.

We note that under very high demand density $\lambda = 500,000$, the discrepancy between the CA and the LR solutions is more significant in terms of both the number of constructed facilities and

Table 8 CA cost estimates, feasible solutions, and LR solutions for the heterogeneous case

q	Δ_q	Δ_λ	Z_{CA}	Z_{CA}^D	Z_{LR}^D		n_{CA}^* n_{LR}^*	ε^D (%)
0.1	0.1	0	18235	19563	18971.1	12	15	3.12
0.1	0.2	0	18115.3	19405.5	18726.6	12	14	3.63
0.1	0.3	0	18012.8	19245.2	18631.9	12	15	3.29
0.1	0.4	0	17927.5	19464.7	18366.3	12	15	5.98
0.1	0.5	0	17859.4	18807.9	18349.6	12	14	2.5
0.2	0.1	0	22158.7	23827.9	23074.8	12	14	3.26
0.2	0.2	0	21668.7	23126.6	22665.4	12	15	2.03
0.2	0.3	0	21243.6	22621.5	22076.7	12	17	2.47
0.2	0.4	0	20884.0	22044.2	21426.4	12	16	2.88
0.2	0.5	0	20590.4	21462.3	20978.2	12	14	2.31
0.1	0.1	1	16667.8	18573.0	17670.0	11	14	5.11
0.1	0.2	1	16646.7	18505.9	17529.8	11	14	5.57
0.1	0.3	1	16634.4	18411.5	17435.3	11	16	5.6
0.1	0.4	1	16630.7	18354.0	17293.5	11	13	6.13
0.1	0.5	1	16635.4	18541.6	17177.7	11	13	7.94
0.2	0.1	1	18864.4	22932.2	22029.6	11	13	4.1
0.2	0.2	1	18740.0	22548.1	21770.8	11	13	3.57
0.2	0.3	1	18659.9	22284.3	21369.5	11	14	4.28
0.2	0.4	1	18623.1	22157.3	21174.1	11	16	4.64
0.2	0.5	1	18628.6	21936.2	20731.1	11	14	5.81

the total system cost. This discovery is not surprising. With high demand density, the difference of demand distribution used in the CA and the discrete model is magnified. In general, the marginal benefit of building an additional facility is lower in the CA model than in the discrete model, since demand is distributed throughout the area in the former, and aggregated at the cell centers in the latter. Hence, the CA model tends to build fewer facilities when the customer demand is extremely dense. We anticipate that the discrepancies would decrease if the input to the discrete model has finer resolution (i.e., more demand aggregation nodes).

To verify our hypothesis, we further divide the unit square to 10×10 identical cells and aggregate demand to the 100 cell centers. We test all 4 instances with high demand density $\lambda = 500,000$ in the 100-node network. The results are listed in Table 10.

Table 9 CA cost estimate, feasible solutions, and LR solutions in the 100-node network

q	$\lambda(10^4)$	Z_{CA}	Z_{CA}^D	Z_{LR}	n_{CA}^*	n_{LR}^*	ε (%)
0.05	50	65504.6	68263.3	66173.5	21	22	3.16
0.10	50	70771.4	74491.9	73528.5	21	24	1.31
0.15	50	79752.2	84207.7	83114.6	21	23	1.32
0.20	50	92354.9	97045.5	95837.4	22	25	1.26

The results in Table 10 indicate that the CA solutions are more consistent with the discrete model solutions with more demand aggregation nodes. This implies that the CA model should match the discrete model best when there are a large number of candidate locations (when the discrete model would incur the most computational difficulty). In this sense, the LR and the CA methods can serve as the complement of each other.

5.3. Sensitivity Analysis with CA

The system cost predicted by the CA model is continuous in all parameters, and is thus a useful tool for sensitivity analysis. In this section, we demonstrate how to use CA to study the impact of the key parameters on the structure of the optimal system design. In particular, we are interested in knowing how the degree of demand aggregation affects the system cost. In other words, all other things being equal, is it preferable to have evenly distributed demand or aggregated demand?

The CA model suggests that the total cost is determined by (14). It is easy to verify that $Z(A(x),x)$ is modular in A and that the point-wise optimal initial service area can be determined by

$$A^*(x) = \left(\frac{2f(x)}{\lambda(x)G(R,q)}\right)^{\frac{2}{3}}.$$

Plugging $A^*(x)$ back in (13) gives us the cost "density" near point x

$$z(x) \equiv z(A^*(x),x) = (2^{-\frac{2}{3}} + 2^{\frac{1}{3}}) f(x)^{\frac{1}{3}} \lambda(x)^{\frac{2}{3}} G^{\frac{2}{3}}(R, q(x)) + \phi(x)\lambda(x)q(x)^R. \tag{16}$$

Clearly, $z(x)$ is concave in λ. From Jensen's inequality, we know that the total cost decreases as the degree of demand aggregation increases.

To verify our findings, we designed numerical tests using the LR algorithm. The key parameters are determined by

$$\lambda(x) = \lambda(1 + \Delta_\lambda \cos(\pi x_{[2]})),\ f(x) = 1000,\ q(x) = 0.2, \phi(x) = \sqrt{2}\ \forall x.$$

We generated 30 test instances, 10 each for three different levels of average demand λ at 50000, 100000 or 150000. The demand variation Δ_λ ranges from 0 to 0.9. Like the previous tests, we aggregate demand to 49 discrete points. Each test instance is solved by the Lagrangian algorithm, and then the percentage change in the optimal cost is calculated, using the case $\Delta_\lambda = 0$ as the benchmark. The test results are illustrated in Figure 9.

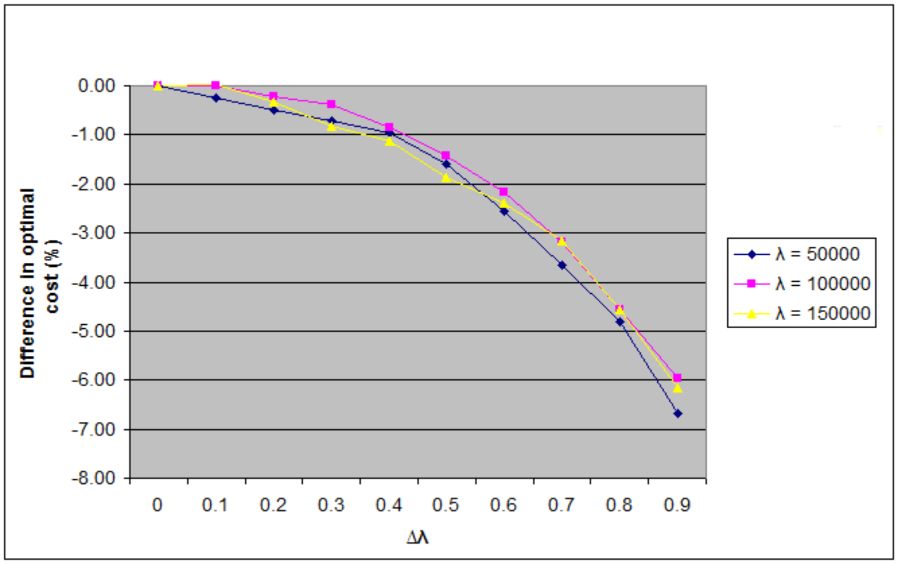


Figure 9 Optimal Cost v.s. demand variability

Clearly, the test results from the discrete model verify the predictions made by the CA model, with the total cost decreasing by up to 7% as the demand variation increases from 0 to 0.9. This result implies that it is beneficial to aggregate demand. In reality, this principle is commonly implemented through the use of warehouses and distribution centers which serve as points for demand aggregation.

6. Conclusions

Supply chains are vulnerable to disruptions caused by natural disasters, terrorist attacks or man-made defections. The consequences of disruptions are often disastrous despite their rare occurrence. However, the emergency cost can be significantly reduced through a proactive approach during the design phase. We present two distinct models to find facility location solutions that are both reliable and cost efficient. Our discrete model is a mixed integer linear program. With our custom-designed Lagrangian relaxation algorithm, it efficiently computes the global optimal solutions for small or medium sized problems. Our continuum approximation model omits details of the facility locations and customer assignments, but provides managerial insights and serves as a valuable tool for sensitivity analysis. It can also be used as an efficient heuristic to find near-optimal solutions for large problem instances.

Our findings also bring up new questions for future research. First, we plan to introduce capacity limits into the model, as opposed to the uncapaciated case in this study. Although exogenous facility capacities will not significantly increase the complexity of the models and the solution algorithms, it is also possible to let the system endogenously determine the capacity level for each facility, at a certain reservation cost. Second, only static decision rules are considered in this study, ignoring the duration and the frequency of the facility disruptions. Incorporating these factors into our model will allow us to examine optimal decision rules in a dynamic environment. Regarding the CA approach, we are particularly interested in developing effective procedures to map discrete data into the continuous setting so that the CA model can be used to solve existing discrete problems in the literature. We are also interested in seeking better understanding or estimates of function $G(R, q)$ so as to potentially improve the performance of the CA model. Finally, we would explore other applications of the CA model, especially in the field of integrated supply chain design.

References

[1] Ball, W.W.R. and Coxeter, H.S.M. (1987) *Mathematical Recreations and Essays (13th Edition).* New York: Dover.

[2] Ball, M.O. and Lin, F.L. (1993) A reliability model applied to emergency service vehicle location. *Operations Research* 41(1): 18-36.

[3] Batta, R., Dolan, J.M. and Krishnamurthy, N.N. (1989) The maximal expected covering location problem - revisited. *Transportaion Science* 23(4): 277-287.

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

[5] Campbell, J. F. (1993a) Continuous and discrete demand hub location problems. *Transportation Research Part B*, 27: 473-482.

[6] Campbell, J. F. (1993b) One-to-many distribution with transshipments: An analytic model. *Transportation Science*, 27: 330-340.

[7] Daganzo, C.F. (1984a) The length of tours in zones of different shapes. *Transportation Research Part B*, 18(2), 135-146.

[8] Daganzo, C.F. (1984b) The distance traveled to visit N points with a maximum of C stops per point: A manual tour-building strategy and case study. *Transportation Science*, 18, 331-350.

[9] Daganzo, C.F. and Newell, G.F. (1986) Configuration of physical distribution networks. *Networks* 16: 113-132.

[10] Daganzo, C.F. and Erera, A.L. (1999) On planning and design of logistics systems for uncertain environments. In: *New Trends in Distribution Logistics*. Springer, Berlin, Germany, 3-21.

[11] Daganzo, C.F. (1992) *Logistics Systems Analysis (1st Edition)*, Springer, Berlin, Germany.

[12] Dasci, A. and Verter, V. (2001) A continuous model for production-distribution system design. *European Journal of Operational Research* 129: 287-298.

[13] Daskin, M.S. (1982) Application of an expected covering model to emergency medical service system design. *Decision Science* 13(3): 416-439.

[14] Daskin, M.S. (1983) A maximum expected covering location model: Formulation, properties and heuristic solution. *Transportation Science* 17(1): 48-70.

[15] Daskin, M.S. (1995) *Network and Discrete Location: Models, Algorithms, and Applications*. John Wiley, New York, 1995.

[16] Drezner, Z., editor. (1995) *Facility Location: A Survey of Applications and Methods*. Springer, New York, 1995.

[17] Fejes Toth, L. (1959) Sur la representation d'une population infinie par un nombre fini d'dlements. *Acta Math*, Academy of Science of Hungary, 10: 299-304.

[18] Fisher, M.L. (1981) The Lagrangian relaxation method for solving integer programming problems. *Management Science* 27: 1-18.

[19] Gersho, A. (1979) Asymptotically optimal block quantization, *IEEE Transactions on Information Theory* IT-25(4): 373-380.

[20] B. Goldengorin, G. Sierksma, G. A. Tijssen and Tso, M. (1999) The data-correcting algorithm for the minimization of supermodular functions. *Management Science* 45: 1539-1551.

[21] Hall, R.W. (1984) Travel distance through transportation terminals on a rectangular grid. *Journal of the Operational Research Society* 35(12): 1067-1078.

[22] Hall, R.W. (1986) Discrete models/continuous models. *OMEGA - International Journal of Management Science* 14(3): 213-220.

[23] Hall, R.W. (1989) Configuration of an overnight package air network. *Transportation Research A* 23A(2): 139-149

[24] Johnson, W. (1944) *Math. and Physical Principles of Engineering Analysis*, McGraw Hill.

[25] Kuhn, H.W. (1955) The Hungarian Method for the assignment problem. *Naval Research Logistics Quarterly* 2: 83C97.

[26] Langevin, A., Mbaraga, P. and Campbell, J.F. (1996) Continuous approximation models in freight distribution: An overview. *Transportation Research Part B* 30: 163-188.

[27] Lim, M., Daskin, M., Chopra, S., and Bassamboo, A. (2007) Managing risks of facility disruptions, *Working Paper*, Northwestern University.

[28] Newell, G.F. (1971) Dispatching policies for a transportation route, *Transportation Science* 5: 91-105.

[29] Newell, G.F. (1973) Scheduling, location, transportation and continuum mechanics: some simple approximations to optimization problems, *SIAM J. Appl. Math.* 25(3): 346-360.

[30] Okabe, A., Boots, B. and Sugihara, K. (1992) *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley, Chichester, UK.

[31] Ouyang, Y. and Daganzo, C.F. (2006) Discretization and validation of the continuum approximation scheme for terminal system design, *Transportation Science* 40(1): 89-98.

[32] Ouyang, Y. (2007) Design of vehicle routing zones for large-scale distribution systems. *Transportation Research Part B* 41(10): 1079-1093.

[33] Qi, L. and Shen, Z.-J.M. (2007) A supply chain design model with unreliable supply. *Naval Research Logistics* 54: 829-844.

[34] Qi, L., Shen, Z.-J. M., and Snyder, L. V. (2008) The effect of supply disruptions on supply chain network design. Submitted for publication.

[35] Revelle, C. and Hogan, K. (1989) The maximum availability location problem. *Transportation Science* 23(3): 192-200.

[36] Sherali, H.D. and Alameddine, A. (1992) A new reformulation-linearization technique for bilinear programming problems. *Journal of Global Optimization* 2(4): 379-410.

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

[38] http://www.lehigh.edu/ lvs2

[39] Weber, A. (1957) *Theory of the Location of Industries.* University of Chicago Press, IL, 1957.

[40] Zhan, R.L., Shen, Z.-J.M., Daskin, M.S. (2007) System Reliability with Location-Specific Failure Probabilities. Working paper, University of California, Berkeley.

Appendix A: Proof of the Propositions

A.1. Proof of Proposition 1

Let $\Omega = \{0,1\}^J$ be the set of failure scenarios. For each $\omega \in \Omega$, let p_ω be the probability that scenario ω will occur, also let $\delta_{j\omega}$ be the binary parameter indicating whether or not facility j is operational in scenario ω. The scenario based stochastic program (SSP) is formulated as follows

$$\text{(SSP)} \quad \text{Min} \sum_{j=0}^{J-1} f_j X_j + \sum_{i=0}^{I-1} \sum_{j=0}^{J} \sum_{\omega \in \Omega} \lambda_i d_{ij} p_\omega Y_{ij\omega} \tag{17a}$$

$$\text{s.t.} \sum_{j=0}^{J} Y_{ij\omega} = 1 \quad \forall\ 0 \le i \le I-1,\ \omega \in \Omega \tag{17b}$$

$$\sum_{i=0}^{I-1} Y_{ij\omega} \le \delta_{j\omega} X_j \quad \forall\ 0 \le j \le J-1,\ \omega \in \Omega \tag{17c}$$

$$X_j,\ Y_{ij\omega} \in \{0,1\} \quad \forall\ 0 \le i \le I-1,\ 0 \le j \le J,\ \omega \in \Omega, \tag{17d}$$

where X_j is the binary variable indicating whether or not a facility is build at location j, and $Y_{ij\omega}$ is equal to one if and only if customer i is served by facility j in scenario ω ($Y_{iJ\omega} = 1$ indicates that the penalty cost is incurred in scenario ω).

To verify that the scenario based formulation (17a)-(17d) is equivalent to the compact (RUFL) formulation (1a)-(1g), we first show how to map an optimal solution of (RUFL) to a feasible solution of (SSP). Let $(\mathbf{X}, \mathbf{Y}, \mathbf{P})$ be an optimal solution of (RUFL), we construct a solution $(\mathbf{X}', \mathbf{Y}')$ for (SSP) by letting $\mathbf{X}' = \mathbf{X}$. For each $0 \le i \le I-1$ and $0 \le r \le J$, let $j(i,r) \in \{0 \le j \le J:\ Y_{ijr} = 1\}$, *i.e.* $j(i,r)$ is the unique facility that serves customer i at level r. The customer assignment in each scenario is determined as follows (by convention, we let $\delta_{J\omega}=1$ for all $\omega\in\Omega$)

$$Y'_{ij\omega}=\begin{cases}1 & \text{if } j=j(i,r) \text{ for some } 0\le r\le J,\ \delta_{j\omega}=1, \text{and } \delta_{j(i,\ell)\omega}=0,\ \forall\ 0\le\ell\le r-1\\ 0 & \text{otherwise.}\end{cases}$$

By construction, $(\mathbf{X}',\mathbf{Y}')$ is feasible to (SSP). Next, we show that $(\mathbf{X}',\mathbf{Y}')$ achieves the same object value as $(\mathbf{X},\mathbf{Y},\mathbf{P})$. Let $\Phi(\mathbf{X},\mathbf{Y},\mathbf{P})$ and $\Psi(\mathbf{X}',\mathbf{Y}')$ be the objective function of (RUFL) and (SSP) respectively. Also, define $\Omega(i,r)=\{\omega\in\Omega:\ Y'_{ij(i,r)\omega}=1\}$, *i.e.* $\Omega(i,r)$ is the set of scenarios in which customer i is served by facility $j(i,r)$. It follows that

$$\begin{aligned}\Psi(\mathbf{X}',\mathbf{Y}') &= \sum_{j=0}^{J-1} f_j X'_j+\sum_{i=0}^{I-1}\sum_{j=0}^{J}\sum_{\omega\in\Omega}\lambda_i d_{ij} p_\omega Y'_{ij\omega}\\ &= \sum_{j=0}^{J-1} f_j X_j+\sum_{i=0}^{I-1}\sum_{r=0}^{J}\lambda_i d_{i,j(i,r)}\sum_{\omega\in\Omega(i,r)} p_\omega\\ &= \sum_{j=0}^{J-1} f_j X_j+\sum_{i=0}^{I-1}\sum_{r=0}^{J}\lambda_i d_{i,j(i,r)}(1-q_{j(i,r)})\prod_{\ell=0}^{r-1} q_{j(i,\ell)}\\ &= \sum_{j=0}^{J-1} f_j X_j+\sum_{i=0}^{I-1}\sum_{r=0}^{J}\sum_{j=0}^{J}\lambda_i d_{ij} P_{ijr} Y_{ijr}\\ &= \Phi(\mathbf{X},\mathbf{Y},\mathbf{P}),\end{aligned}$$

which implies that solving (SSP) yields a lower bound to (RUFL).

Conversely, given an optimal solution $(\mathbf{X},\mathbf{Y})$ to (SSP), we construct a feasible solution $(\mathbf{X}',\mathbf{Y}',\mathbf{P}')$ to (RUFL), also by letting $\mathbf{X}'=\mathbf{X}$. Without loss of generality, we assume that $Y_{ij\omega}=1$ if and only if $j=\min\{0\le k\le J:\ \delta_{k\omega}X_k=1,\ d_{ik}\le d_{ik'}\ \forall\ k'\ne k \text{ s.t. } \delta_{k'\omega}X_{k'}=1\}$ (by convention, we assume $X_J=1$), *i.e.* each customer is always served by her closest open facility, and if there are more than one facilities that are equally close, we break the tie by choosing the facility with the lowest index.

Let $N=\{0\le j\le J-1:\ X_j=1\}$ be the set of facilities that are constructed in the optimal solution to (SSP). For each customer i, let $j(i,0),\ j(i,1),\ \cdots,\ j(i,|N|)$ be an ordering of the facilities in $N\bigcup\{J\}$ such that for all $1\le r\le|N|$, $d_{i,j(i,r-1)}\le d_{i,j(i,r)}$, and if $d_{i,j(i,r-1)}=d_{i,j(i,r)}$, then

$j(i,r-1) < j(i,r)$. Also, define $\Omega(i,r) = \{\omega \in \Omega : \ \delta_{j(i,r)\omega} = 1, \text{ and } \delta_{j(i,\ell)\omega} = 0, \ \forall\ 0 \leq \ell \leq r-1\}$. We set the values of $\mathbf{Y}'$ and $\mathbf{P}'$ as follows

$$Y'_{ijr} = \begin{cases} 1 & \text{if } j = j(i,r) \text{ and } d_{ij} \leq d_{iJ} \\ 0 & \text{otherwise} \end{cases}$$
$$P'_{ijr} = \begin{cases} (1 - q_{j(i,r)}) \prod_{\ell=0}^{r-1} q_{j(i,\ell)} & \text{if } j = j(i,r) \text{ and } d_{ij} \leq d_{iJ} \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that by construction $(\mathbf{X}', \mathbf{Y}', \mathbf{P}')$ is feasible to (RUFL). The objective value associated with solution is

$$\begin{aligned}
\Phi(\mathbf{X}', \mathbf{Y}', \mathbf{P}') &= \sum_{j=0}^{J-1} f_j X'_j + \sum_{i=0}^{I-1} \sum_{r=0}^{J-1} \sum_{j=0}^{J-1} \lambda_i d_{ij} P'_{ijr} Y'_{ijr} \\
&= \sum_{j=0}^{J-1} f_j X_j + \sum_{i=0}^{I-1} \sum_{r=0}^{|N|} \lambda_i d_{i,j(i,r)} (1 - q_{j(i,r)}) \prod_{\ell=0}^{r-1} q_{j(i,\ell)} \\
&= \sum_{j=0}^{J-1} f_j X_j + \sum_{i=0}^{I-1} \sum_{r=0}^{J} \lambda_i d_{i,j(i,r)} \sum_{\omega \in \Omega(i,r)} p_\omega \\
&= \sum_{j=0}^{J-1} f_j X_j + \sum_{i=0}^{I-1} \sum_{j=0}^{J} \sum_{\omega \in \Omega} \lambda_i d_{ij} p_\omega Y_{ij\omega} \\
&= \Psi(\mathbf{X}, \mathbf{Y}).
\end{aligned}$$

Therefore, the optimal solution to (SSP) is also a lower bound to (RUFL). This completes our proof. Q.E.D.

A.2. Proof of Proposition 2

Suppose, for a contradiction, that $(\mathbf{X}, \mathbf{Y}, \mathbf{P})$ is optimal for (RUFL) where $Y_{ijr} = Y_{ik,r+1} = 1$ and $d_{ij} > d_{ik}$ for some $0 \leq i \leq I-1$, $0 \leq j \leq J$, and $0 \leq r \leq R$. We will show that by "swapping" j and k the objective value will decrease. Obviously $j \leq J-1$, otherwise j is the pseudo facility and customer i cannot be assigned to facility k as a backup. We consider two cases based on whether or not k is the pseudo facility.

If $k \leq J-1$ we construct a different solution $(\mathbf{X}', \mathbf{Y}', \mathbf{P}')$ as follows:

$$X' = X;$$
$$Y'_{h\ell s} = \begin{cases} 1 & \text{if } h = i, \ell = k, \ s = r \text{ or } h = i, \ell = j, \ s = r+1, \\ 0 & \text{if } h = i, \ell = j, \ s = r \text{ or } h = i, \ell = k, \ s = r+1, \\ Y_{h\ell s} & \text{otherwise;} \end{cases}$$

$$P'_{h\ell s} = \begin{cases} \frac{1-q_k}{1-q_j} P_{jr} & \text{if } h=i, \ell=k,\ s=r, \\ \frac{q_k(1-q_j)}{1-q_k} P'_{kr} = q_k P_{jr} & \text{if } h=i, \ell=j,\ s=r+1, \\ 0 & \text{if } h=i, \ell=j,\ s=r \text{ or } h=i, \ell=k,\ s=r+1, \\ P_{h\ell s} & \text{otherwise.} \end{cases}$$

By construction, $(\mathbf{X}', \mathbf{Y}', \mathbf{P}')$ is a feasible solution. Let $\Phi(\mathbf{X}, \mathbf{Y}, \mathbf{P})$ be the objective value associated with $(\mathbf{X}, \mathbf{Y}, \mathbf{P})$, it follows that:

$$\begin{aligned} \Phi(\mathbf{X}', \mathbf{Y}', \mathbf{P}') - \Phi(\mathbf{X}, \mathbf{Y}, \mathbf{P}) &= \lambda_i(P'_{kr} d_{ik} + P'_{j,r+1} d_{ij} - P_{jr} d_{ij} - P_{k,r+1} d_{ik}) \\ &= \lambda_i[d_{ik}(P'_{kr} - P_{k,r+1}) - d_{ij}(P_{jr} - P'_{j,r+1})] \\ &= \lambda_i\{d_{ik}[\frac{1-q_k}{1-q_j} P_{jr} - \frac{q_j(1-q_k)}{1-q_j} P_{jr}] - d_{ij}(P_{jr} - q_k P_{jr})\} \\ &= \lambda_i(1-q_k)(d_{ik} - d_{ij}) P_{jr} < 0. \end{aligned}$$

The case in which $k = J$ is similar, except that $Y'_{ij,r+1} = P'_{ij,r+1} = 0$, which reduces the cost even more. This implies a contradiction to that $(\mathbf{X}, \mathbf{Y}, \mathbf{P})$ is optimal. Q.E.D.

A.3. Proof of Proposition 3

Let $S \subseteq \{0, \cdots, J-1\}$ be a subset of candidate locations, and $u, v \in \{0, \cdots, J-1\} \setminus S$, we show that

$$\Phi_i(S \cup \{u, v\}) - \Phi_i(S \cup \{u\}) \geq \Phi_i(S \cup \{v\}) - \Phi_i(S). \tag{18}$$

Assume that $S = \{j_1,\ j_2,\ \cdots,\ j_n\}$ where $d_{ij_1} \leq d_{ij_2} \leq \cdots \leq d_{ij_n}$, i.e., we sort elements in S in nondecreasing order of their distance to customer i. Let

$$\begin{aligned} \bar{n} &= \inf\{1 \leq k \leq n:\ d_{ij_k} \leq \phi_i\} \\ s &= \inf\{1 \leq k \leq n:\ d_{ij_k} \leq d_{iu}\} \\ t &= \inf\{1 \leq k \leq n:\ d_{ij_k} \leq d_{iv}\}. \end{aligned}$$

In addition, define

$$P_k = \begin{cases} \prod_{\ell=1}^{k} q_{j_\ell} & 1 \leq k \leq \bar{n} \\ 1 & k = 0, \end{cases}$$

$$C_k = \begin{cases} P_{k-1}(1-q_{j_k}) d_{ij_k} & 1 \leq k \leq \bar{n} \\ P_{\bar{n}} \phi_i & k = \bar{n}+1. \end{cases}$$

Following a similar argument as in the proof of Proposition 2, we know that it is optimal to assign the facilities level by level in increasing order of distance, until the transportation cost exceeds the penalty cost, i.e.,

$$\begin{aligned}\Phi_i(S) &= \lambda_i \sum_{k=1}^{\bar{n}} P_{k-1}(1-q_{j_k})d_{ij_k} + P_{\bar{n}}\phi_i + \sum_{j\in S}\mu_{ij}\\ &= \lambda_i \sum_{k=1}^{\bar{n}+1} C_k + \sum_{j\in S}\mu_{ij}.\end{aligned}$$

Without loss of generality, we assume that d_{iu} and d_{iv} are less than the penalty cost ϕ_i, i.e. $s \le \bar{n}$ and $t \le \bar{n}$. It follows that

$$\begin{aligned}\Phi_i(S\cup\{v\}) - \Phi_i(S) &= \lambda_i[\sum_{k=1}^{t} C_k + P_t(1-q_v)d_{iv} + q_v \sum_{k=t+1}^{\bar{n}+1} C_k] + \sum_{j\in S\cup\{v\}}\mu_{ij} - \lambda_i\sum_{k=1}^{\bar{n}+1} C_k - \sum_{j\in S}\mu_{ij}\\ &= \lambda_i[P_t(1-q_v)d_{iv} - (1-q_v)\sum_{k=t+1}^{\bar{n}+1} C_k] + \mu_{iv}\\ &= \lambda_i(1-q_v)[P_t d_{iv} - \sum_{k=t+1}^{\bar{n}+1} C_k] + \mu_{iv}.\end{aligned}$$

Note that the first item in the last equation is negative, because

$$\begin{aligned}P_t d_{iv} - \sum_{k=t+1}^{\bar{n}+1} C_k &= P_t[d_{iv} - \sum_{k=t+1}^{\bar{n}} (\prod_{\ell=t+1}^{k-1} q_{j_\ell})(1-q_{j_k})d_{ij_k} - (\prod_{\ell=t+1}^{\bar{n}} q_{j_\ell})\phi_i]\\ &< P_t d_{iv}[1 - (\prod_{\ell=t+1}^{k-1} q_{j_\ell})(1-q_{j_k}) - \prod_{\ell=t+1}^{\bar{n}} q_{j_\ell}] = 0.\end{aligned}$$

To show that (18) holds, we consider the following two cases.

Case 1: $d_{iu} \le d_{iv}$. In this case, it follows that

$$\begin{aligned}\Phi_i(S\cup\{u,\ v\}) - \Phi_i(S\cup\{u\}) &= \lambda_i[\sum_{k=1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u\sum_{k=s+1}^{t} C_k + q_u P_t(1-q_t)d_{iv} + q_u q_v \sum_{k=t+1}^{\bar{n}+1} C_k\\ &\quad + \sum_{j\in S\cup\{u,\ v\}}\mu_{ij} - \lambda_i[\sum_{k=1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u\sum_{k=s+1}^{\bar{n}+1} C_k] - \sum_{j\in S\cup\{u\}}\mu_{ij}\\ &= \lambda_i[q_u P_t(1-q_v)d_{iv} - q_u(1-q_v)\sum_{k=t+1}^{\bar{n}+1} C_k] + u_{iv}\\ &= \lambda_i q_u(1-q_v)(P_t d_{iv} - \sum_{k=t+1}^{\bar{n}+1} C_k) + \mu_{iv}.\end{aligned}$$

Clearly (18) holds in this case, since $0 \le q_u \le 1$ and $P_t d_{iv} - \sum_{k=t+1}^{\bar{n}+1} C_k < 0$.

Case 2: $d_{iu} > d_{iv}$. In this case $t \le s$, and the following assertion holds:

$$\begin{aligned}
\Phi_i(S\cup\{u,\ v\}) - \Phi_i(S\cup\{u\}) &= \lambda_i[\sum_{k=1}^{t} C_k + P_t(1-q_v)d_{iv} + q_v \sum_{k=t+1}^{s} C_k + q_v P_s(1-q_u)d_{iu} + q_v q_u \sum_{k=s+1}^{\bar{n}+1} C_k] \\
&\quad + \sum_{j\in S\cup\{u,\ v\}} \mu_{ij} - \lambda_i[\sum_{k=1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u \sum_{k=s+1}^{\bar{n}+1} C_k] - \sum_{j\in S\cup\{u\}} \mu_{ij} \\
&= \lambda_i\{P_t(1-q_v)d_{iv} - (1-q_v)[\sum_{k=t+1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u \sum_{k=s+1}^{\bar{n}+1} C_k]\} + \mu_{iv} \\
&= \lambda_i(1-q_v)\{P_t d_{iv} - [\sum_{k=t+1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u \sum_{k=s+1}^{\bar{n}+1} C_k]\} + \mu_{iv}.
\end{aligned}$$

We claim that (18) holds in this case, because

$$\begin{aligned}
&[\sum_{k=t+1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u \sum_{k=s+1}^{\bar{n}+1} C_k] - \sum_{k=t+1}^{\bar{n}+1} C_k \\
\le\ &[\sum_{k=t+1}^{s} C_k + P_s(1-q_u)d_{iu} + q_u \sum_{k=s+1}^{\bar{n}+1} C_k] - \sum_{k=s+1}^{\bar{n}+1} C_k \\
=\ &(1-q_u)[P_s d_{iu} - \sum_{k=s+1}^{\bar{n}+1} C_k] < 0.
\end{aligned}$$

Q.E.D.

A.4. Proof of Proposition 4

First, we introduce an equivalent formulation of (RSP) by "spliting" the decision variables:

$$Y_{jr} = \begin{cases} 1 & \text{if the level } r \text{ facility for this customer has the same transportation distance as facility } j \\ 0 & \text{otherwise.} \end{cases}$$

$$Z_{jr} = \begin{cases} 1 & \text{if the level } r \text{ facility for customer } i \text{ has the same failure probability as facility } j \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that RSP is equivalent to the following problem:

$$\text{Min} \sum_{j=0}^{J}\sum_{r=0}^{R} \lambda_i d_{ij} W_{jr} + \sum_{j=0}^{J-1}\sum_{r=0}^{R-1} \mu_{ij} Y_{jr} \tag{19a}$$

$$\text{s.t. (4b)} - \text{(4d)} \tag{19b}$$

$$\sum_{j=0}^{J-1} Z_{jr} + \sum_{s=0}^{r-1} Z_{Js} = 1 \quad \forall\ 0 \le r \le R \tag{19c}$$

$$\sum_{r=0}^{R-1} Z_{jr} \le 1 \quad \forall\ 0 \le j \le J-1 \tag{19d}$$

$$\sum_{r=0}^{R} Z_{Jr} = 1 \tag{19e}$$

$$P_{j0} = 1 - q_j \quad \forall\ 0 \le j \le J \tag{19f}$$

$$P_{jr} = (1-q_j)\sum_{k=0}^{J-1} \frac{q_k}{1-q_k} W_{i,k,r-1} \quad \forall\ 0 \le j \le J,\ 1 \le r \le R \tag{19g}$$

$$W_{jr} \le P_{jr} \quad \forall\ 0 \le j \le J,\ 0 \le r \le R \tag{19h}$$

$$W_{jr} \le Z_{jr} \quad \forall\ 0 \le j \le J,\ 0 \le r \le R \tag{19i}$$

$$W_{jr} \ge 0 \quad \forall\ 0 \le j \le J,\ 0 \le r \le R \tag{19j}$$

$$W_{jr} \ge P_{jr} + Z_{jr} - 1 \quad \forall\ 0 \le j \le J,\ 0 \le r \le R \tag{19k}$$

$$Y_{jr},\ Z_{jr} \in \{0,1\} \quad \forall\ 0 \le j \le J,\ 0 \le r \le R \tag{19l}$$

$$Y_{jr} = Z_{jr} \quad \forall\ 0 \le j \le J,\ 0 \le r \le R. \tag{19m}$$

If we remove the last constraint (19m), the customer is allowed to choose an arbitrary combination of transportation cost and failure probability. Next, we show that the (RRSP) formulation (7a)-(7e) is equivalent to formulation (19a) - (19l), based on the following lemma.

LEMMA 1. *There exists an optimal solution* $(\mathbf{Y}^*, \mathbf{Z}^*, \mathbf{P}^*)$ *to formulation* (19a) - (19l)*, such that if* $Z^*_{jr} = 1,\ Z^*_{k,r+1} = 1$ *and* $r+1 \le R-1$*, then* $q_j \le q_k$*.*

Proof of Lemma 1. Suppose that $(\mathbf{Y}, \mathbf{Z}, \mathbf{P})$ is an optimal solution to formulation (19a) - (19l), such that $Z_{jr} = 1,\ Z_{k,r+1} = 1,\ j,k \le R-1$ and $q_j > q_k$. Let u and v be the facilities assigned to this customer at level r and $r+1$, *i.e.* $Y_{ur} = 1$ and $Y_{v,r+1} = 1$. We construct a new solution $(\mathbf{Y}', \mathbf{Z}', \mathbf{P}')$ as follows:

$$\mathbf{Y}' = \mathbf{Y};$$
$$Z'_{\ell s} = \begin{cases} 1 & \text{if } \ell = k,\ s = r \text{ or } h = i, \ell = j,\ s = r+1, \\ 0 & \text{if } \ell = j,\ s = r \text{ or } h = i, \ell = k,\ s = r+1, \\ Z_{\ell s} & \text{otherwise;} \end{cases}$$
$$P'_{\ell s} = \begin{cases} \frac{1-q_k}{1-q_j} P_{jr} & \text{if } \ell = k,\ s = r, \\ \frac{q_k(1-q_j)}{1-q_k} P'_{kr} = q_k P_{jr} & \text{if } \ell = j,\ s = r+1, \\ 0 & \text{if } \ell = j,\ s = r \text{ or } h = i, \ell = k,\ s = r+1, \\ P_{\ell s} & \text{otherwise.} \end{cases}$$

By construction, $(\mathbf{Y}', \mathbf{Z}', \mathbf{P}')$ is a feasible solution to formulation (19a) - (19l). Define $G(\mathbf{Y}, \mathbf{Z}, \mathbf{P})$ to be the objective value of formulation (19a) - (19l) associated with solution $(\mathbf{Y}, \mathbf{Z}, \mathbf{P})$. The following assertion holds:

$$\begin{aligned}
G(\mathbf{Y}', \mathbf{Z}', \mathbf{P}') - G(\mathbf{Y}, \mathbf{Z}, \mathbf{P}) &= \lambda_i (P'_{kr} d_{iu} + P'_{j,r+1} d_{iv} - P_{jr} d_{iu} - P_{k,r+1} d_{iv}) \\
&= \lambda_i [d_{iu}(P'_{kr} - P_{jr}) + d_{iv}(P'_{j,r+1} - P_{k,r+1})] \\
&= \lambda_i \{ d_{iu} [\frac{1-q_k}{1-q_j} P_{jr} - P_{jr}] - d_{iv}(q_k P_{jr} - \frac{q_j(1-q_k)}{1-q_j} P_{jr}) \} \\
&= \frac{q_j - q_k}{1-q_j} \lambda_i P_{jr} (d_{iu} - d_{iv}).
\end{aligned}$$

Following a similar argument as in the proof of Proposition 2, $d_{iu} \le d_{iv}$, implying $G(\mathbf{Y}', \mathbf{Z}', \mathbf{P}') \le G(\mathbf{Y}, \mathbf{Z}, \mathbf{P})$. Therefore, if an optimal solution does not satisfy the condition in Lemma 1, we can always construct an alternative optimal solution by swapping j and k. This completes the proof of Lemma 1.

Without loss of generality, we can fix $\mathbf{Z} = \mathbf{Z}^*$ and $\mathbf{P} = \mathbf{P}^*$ in formulation (19a) - (19l), which leads to the (RRSP) formulation (7a)-(7e). Since formulation (19a) - (19l) is a relaxation of (RSP), it follows that (RRSP) yields a lower bound for (RSP). Q.E.D.

A.5. Proof of Proposition 5

The following lemma gives a necessary optimality condition for facility location design and customer allocation.

Lemma 2. *The optimal facility locations should satisfy the following conditions:*

1. *the initial service areas* $\mathbf{R}$ *(i.e., initial customer allocation before any failure) should form a Voronoi tessellation;*

2. *the location of each facility should be the centroid of all customer demands weighted by this facility's service probability to the customers.*

Proof of Lemma 2. The first condition is obvious from the fact that for any given facility location design, every customer always goes to the nearest available facility. The second necessary condition can be proven by examining the cost objective with respect to an infinitesimal perturbation of one generic facility location, x_j, while holding $\mathcal{R}_{jk}, \forall j, k$, fixed. Let $\mathcal{F}(x_j)$ denote the expected service cost of a facility located at x_j to serve all its potential customers. Consider an arbitrary location perturbation Δ_x and a scalar $\epsilon > 0$. Consider an arbitrary location perturbation Δ_x and a scalar $\epsilon > 0$:

$$\begin{aligned}\mathcal{F}(x_j + \epsilon\Delta_x) - \mathcal{F}(x_j) &= \sum_{k=0}^{R-1} \int_{x \in R_{jk}} (1-q)q^k \lambda \left\{ \|x - x_j - \epsilon\Delta_x\| - \|x - x_j\| \right\} dx \\ &= \sum_{k=0}^{R-1} \int_{x \in R_{jk}} (1-q)q^k \lambda \left\{ \frac{\|x - x_j - \epsilon\Delta_x\|^2 - \|x - x_j\|^2}{\|x - x_j - \epsilon\Delta_x\| + \|x - x_j\|} \right\} dx.\end{aligned}$$

It is easy to show that the first-order condition $\lim_{\epsilon \to 0} \frac{1}{\epsilon}\{\mathcal{F}(x_j + \epsilon\Delta_x) - \mathcal{F}(x_j)\} = 0$ requires that the optimal facility location x_j satisfies

$$x_j = \frac{\sum_{k=0}^{R-1} \int_{x \in R_{jk}} (1-q)q^k \lambda x dx}{\sum_{k=0}^{R-1} \int_{x \in R_{jk}} (1-q)q^k \lambda dx} = \frac{\sum_{k=0}^{R-1} \int_{x \in R_{jk}} P(x, x_j|\mathbf{x}) x dx}{\sum_{k=0}^{R-1} \int_{x \in R_{jk}} P(x, x_j|\mathbf{x}) dx}.$$

Hence, the optimal facility location x_j is the centroid of all customer demands weighted by the corresponding service probability. This completes the proof of Lemma 2.

It is worth noting that the above proof does not require S to be homogeneous and infinite. Hence, Lemma 2 holds also for finite and heterogeneous S.

Since the plane is infinite and homogeneous, the facility locations and all service areas should be translationally and rotationally symmetric. The initial service area of every facility (which, as a Voronoi polygon, must be convex (30)) should have the facility location as its centroid. Hence, collectively they should form a centroidal Voronoi tessellation—which should then minimize the total customer initial access cost (before any failure) to the facilities. As pointed out by Gersho (19), Fejes Toth (17) proved that this cost is minimized under the Euclidean metric when the shape of the initial service areas are exactly congruent (i.e., of same shape and size) and form a regular hexagonal tessellation of the space. Gersho further proved that even in a finite 2-d plane, regular hexagonal tessellations should cover most of the space if the number of facilities is sufficiently large (19). This result leads to Proposition 5.   Q.E.D.

Appendix B: Alternative CA Formulation

Rather than specifying R, we may alternatively assume that a customer at x, if served at all, shall only be served by a facility within a *maximum service distance* $\theta(x)$. Note that $\theta(x)$ approximately corresponds to the reassignment level R by

$$\pi\theta^2(x) \approx RA(x).$$

Note that some models in the literature, e.g. (37), assume that $\phi(x) \equiv \theta(x), \forall x$ (i.e., the customers are willing to travel to a facility as long as the travel cost is no larger than the penalty cost). Defining $\phi(x)$ and $\theta(x)$ separately generalizes such assumptions.

For any facility at x_j there are two possible cases.

B.1. Case 1: $R > 1$ or $A(x) < \pi\theta^2(x)$

When the maximum service distance is sufficiently large, there are customer reassignments upon facility failure. Substituting $R = \pi\theta^2(x)/A(x)$ into (9), we have

$$z(A) := \frac{f}{A} + \phi\lambda q^{\frac{\pi\theta^2}{A}} + \lambda g(q)\sqrt{A}\left[e^{-1.542q^2/\theta^2}\right]^A, \tag{20}$$

where $g(q) := \exp(-0.930 - 0.223q + 4.133q^2 - 2.906q^3)$. Formula (14) can be minimized by finding the optimal function $A(x) \in [0, \pi\theta^2(x))$ at every point $x \in \mathcal{S}$.

B.2. Case 2: $0 < R < 1$ or $A(x) \geq \pi\theta^2(x)$

This case is extreme, where the maximum service distance is very small and hence no customer reassignment is possible. Within the influence area of facility j, customers with a distance to x_j larger than $\theta(x_j) \approx \theta(x)$ shall never be served regardless of facility failure scenarios. The customers within distance $\theta(x_j) \approx \theta(x)$ from x_j would be served if facility j does not fail (with probability $1 - q(x_j)$) and not served otherwise. Then, the expected costs (fixed charge, penalty, transportation), incurred in an area of size $A(x_j)$, become approximately

$$\begin{aligned} f(x_j) &+ \phi(x_j)\cdot\lambda(x_j)[A(x_j) - \pi\theta^2(x_j)] + \phi(x_j)q(x_j)\cdot\lambda(x_j)[\pi\theta^2(x_j)] \\ &+ \int_0^{\theta(x_j)} r[1 - q(x_j)]\lambda(x_j)2\pi r dr \end{aligned}$$

$$= f(x_j)+\phi(x_j)\lambda(x_j)A(x_j) + \left[\frac{2\theta(x_j)}{3} - \phi(x_j)\right][1-q(x_j)]\lambda(x_j)\pi\theta^2(x_j).$$

Therefore, the expected cost per unit area near $x \approx x_j$ can be approximated by

$$z(A(x),x) := \frac{1}{A(x)}\left[f(x) + \left(\frac{2\theta(x)}{3} - \phi(x)\right)[1-q(x)]\lambda(x)\pi\theta^2(x)\right] + \lambda(x)\phi(x).$$

Formula (14) can be minimized by finding the optimal function $A(x) \in [\pi\theta^2(x), \infty)$ at every point $x \in \mathcal{S}$ regarding the following functional:

$$z(A) := \frac{1}{A}\left[f + \left(\frac{2\theta}{3} - \phi\right)(1-q)\lambda\pi\theta^2\right] + \lambda\phi. \tag{21}$$

In summary, note that from (21),

$$z(\pi\theta^2) = \frac{f}{\pi\theta^2} + \lambda\left[\frac{2\theta}{3}(1-q) + \phi q\right], \text{ and } z(\infty) = \lambda\phi.$$

Obviously, $z(A)$ decreases monotonically with A if $f\left(\phi - \frac{2\theta}{3}\right)(1-q)\lambda\pi\theta^2$. The optimal solution to (20) on $[0, \pi\theta^2)$ shall be compared with $z(\infty) = \lambda\phi$. On the other hand, $z(A)$ increases monotonically with A if $f < \left(\phi - \frac{2\theta}{3}\right)(1-q)\lambda\pi\theta^2$. Case 2 will surely be suboptimal; the solution found in Case 1 will be the optimal solution.

Appendix C: Comparison of the CA and the Discrete Model under Continuous Demand

In Section 5.2, we have compared the CA and the discrete model under aggregated discrete customer demand (i.e., considering the continuous demand as an approximation). This appendix provides additional results on the cost difference of the two models under continuous demand distribution (i.e., considering the discrete demand as an approximation). In addition to the results shown in Section 5.2, we implement the CA model for each test instance through the following procedure:

(i) Compute the continuous solution $A^*(x)$ (15), the optimal number of facilities n^*_{CA}, and the predicted total cost Z_{CA} (without discrete facility locations);

(ii) Use the disk model described in Section 4.3 to translate $A^*(x)$ into a feasible planar solution (i.e., facility can be anywhere in the unit square) and compute the planar cost Z^P_{CA}; and

(iii) Round the planar facility locations to the nearest cell centers, and compute Z^C_{CA}, the exact total system cost for the CA solution under continuous customer demand.

For comparison, we solve the discrete version of the problem as follows:

(i) Apply the LR algorithm to obtain the optimal solution with discrete demand input;

(ii) Compute the cost for the LR solution to serve continuous customer demand Z^C_{LR}, where the superscript 'C' stands for continuous customer demand. While doing this, we simply enforce that each customer goes to the nearest existing facility.

The percentage difference between the CA model cost and the discrete model cost under continuous customer demand; i.e., $\varepsilon^C = \frac{Z^C_{CA}-Z^C_{LR}}{Z^C_{LR}}$, are summarized in Table 10 for the homogeneous instances and in Table 11 for the heterogeneous instances.

Table 10 Cost comparison of the CA and the discrete model under continuous demand: the homogeneous cases

q	$\lambda(10^4)$	Z_{CA}	Z^P_{CA}	Z^C_{CA}	n^*_{CA}	Z^C_{LR}	n^*_{LR}	ε^C (%)
0.05	5	13908.5	14687.2	14694.2	5	14521.9	5	1.19
0.10	5	14430.9	15546.3	15608.2	5	15705.8	5	-0.62
0.15	5	15345.4	16666.9	16777.4	5	16865.8	5	-0.52
0.20	5	16632.0	18047.2	18201.9	5	18318.1	5	-0.63
0.05	10	22151.2	23270.2	23397.5	7	23484.7	7	-0.37
0.10	10	23199.4	24797.6	24951.4	7	25104.2	8	-0.61
0.15	10	25015.7	26886.8	27063.9	7	27192.0	8	-0.47
0.20	10	27568.7	29538.5	29734.8	7	30093.5	9	-1.19
0.05	15	28704.7	30085.7	31002.8	10	30963.3	10	0.13
0.10	15	29309.9	32162.6	33004.7	10	33292.4	10	-0.86
0.15	15	30667.5	35484.1	36263.4	10	36339.7	11	-0.21
0.20	15	32880.6	39222.6	39877.9	10	40234.9	10	-0.89
0.05	50	65504.6	67197.0	71305.8	21	79225.5	49	-10.00
0.10	50	70771.4	73815.2	77062.8	21	85395.6	49	-9.76
0.15	50	79752.2	83703.2	87206.1	21	94838.5	49	-8.05
0.20	50	92354.9	96457.2	100348.9	21	107554.2	49	-6.70

We note that most ε^C is negative in most cases, indicating that the CA model is slightly better when the customer demand is continuous.

Table 11 Cost comparison of the CA and the discrete model under continuous demand: the heterogeneous case

q	Δ_q	Δ_λ	Z_{CA}	Z_{CA}^P	Z_{CA}^C	n_{CA}^*	Z_{LR}^D	n_{LR}^*	ε^C (%)
0.1	0.1	0.0	18235	19061.8	20027.3	12	19995.4	15	0.16
0.1	0.2	0.0	18115.3	18891.9	19868.8	12	19817.4	14	0.26
0.1	0.3	0.0	18012.8	18753.3	19721.1	12	19875.7	15	-0.78
0.1	0.4	0.0	17927.5	18734.3	19943.9	12	19493.9	15	2.31
0.1	0.5	0.0	17859.4	18499.1	19500.8	12	19453.6	14	0.24
0.2	0.1	0.0	22158.7	23489.3	24376.4	12	24046.2	14	1.37
0.2	0.2	0.0	21668.7	22726.7	23573.6	12	23747.5	15	-0.73
0.2	0.3	0.0	21243.6	22190.3	23061.7	12	23231.2	17	-0.73
0.2	0.4	0.0	20884	21693.4	22593.4	12	22540.0	16	0.24
0.2	0.5	0.0	20590.4	21368.8	22139.0	12	22266.6	14	-0.57
0.1	0.1	1.0	16667.8	18337.7	19286.0	11	19362.9	14	-0.4
0.1	0.2	1.0	16646.7	18241.1	19137.8	11	19203.7	14	-0.34
0.1	0.3	1.0	16634.4	18136.5	19050.5	11	19314.8	16	-1.37
0.1	0.4	1.0	16630.7	18105.1	18971.4	11	18843.0	13	0.68
0.1	0.5	1.0	16635.4	18198.9	19278.5	11	18852.6	13	2.26
0.2	0.1	1.0	18864.4	22717.6	23525.2	11	23326.4	13	0.85
0.2	0.2	1.0	18740	22387.6	23258.8	11	23270.6	13	-0.05
0.2	0.3	1.0	18659.9	22169.5	22810.1	11	23027.1	14	-0.94
0.2	0.4	1.0	18623.1	22075.7	23253.0	11	23075.8	16	0.77
0.2	0.5	1.0	18628.6	21768.2	22486.7	11	22686.9	14	-0.88