Transportation Research Part B 64 (2014) 54–72

Contents lists available at ScienceDirect

Transportation Research Part B

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

Reliable *p*-median facility location problem: two-stage robust models and algorithms

Yu An [a], Bo Zeng [a,*], Yu Zhang [b], Long Zhao [a]

[a] Department of Industrial and Management Systems Engineering, University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620, USA
[b] Department of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620, USA

ARTICLE INFO

Article history:
Received 14 January 2013
Received in revised form 26 February 2014
Accepted 27 February 2014

Keywords:
Facility location problem
Reliable distribution network
Two-stage robust optimization
Disruption-caused demand change
Capacity

ABSTRACT

In this paper, we propose a set of two-stage robust optimization models to design reliable *p*-median facility location networks subject to disruptions. We analyze their structural properties, and implement the *column-and-constraint generation* method with customized enhancement strategies, which is more effective than Benders cutting plane method. Numerical experiments are performed on real data and management insights on system design are presented. In particular, our study demonstrates the strong modeling capability of two-stage robust optimization scheme by including two practical issues, i.e., demand changes due to disruptions and facility capacities, which receive little attention in reliable distribution network design research. Results show the significant influence of the demand change on the network configuration.

1. Introduction

The determination of facility locations and client assignments are among the most crucial issues in designing an efficient distribution network. To address these issues, various facility location models have been formulated and studied for decades, including those based on *p*-median and fixed-charge facility location formulations and their extensions (Daskin, 1995; Drezner, 1995; Revelle et al., 2008; Melo et al., 2009). The applications of those facility location models can be found in various industries, including manufacturing, retail, and healthcare (Barahona and Jensen, 1998; Teo and Shu, 2004; Jia et al., 2007). Although it is expected by designers that the distribution network works reliably, the system itself and/or its working environment could be seriously affected by various disruptions. For example, some facilities may be disabled by natural disasters, labor strikes, or terrorism threats. Since the material or information flows are generated, processed, and distributed by facilities, facility disruptions could significantly deteriorate the performance of the whole network and result in enormous economic losses (see the descriptions in Snyder et al. (2012) and references therein). In addition, in a disruptive situation, the whole system may need to deal with a demand pattern which is totally different from that in the normal disruption-free situation (Ergun et al., 2010). Ignoring those issues may lead to a less reliable configuration of the distribution network that is not efficient in mitigating disruptions.

To consider disruptions in system design for better reliability, several recent studies, including Snyder and Daskin (2005), Berman et al. (2007), Chen et al. (2011), Cui et al. (2010), Li and Ouyang (2010), Lim et al. (2010), Li et al. (2013), and Peng et al. (2011), propose to proactively consider disruptions and the incurred cost of countermeasures in the system design stage. The countermeasures, i.e., mitigation or recourse operations, are to reassign clients to survived facilities such that they

* Corresponding author. Tel.: +1 813 974 5588; fax: +1 813 974 5953.
E-mail address: bzeng@usf.edu (B. Zeng).

http://dx.doi.org/10.1016/j.trb.2014.02.005

can be served and the impact of disruptions can be minimized. Hence, the objective of system design is to minimize the (weighted) overall cost, including the operation cost in the normal situation when all facilities function properly, and the cost of mitigation in disruptive situations. To analytically represent this new design scheme, based on the explicit probabilistic information, several compact (nonlinear) mixed integer programs or scenario-based two-stage stochastic programming formulations are developed and customized exact or approximation algorithms are designed to solve real instances. (Snyder and Daskin, 2005; Chen et al., 2011; Cui et al., 2010; Li and Ouyang, 2010; Lim et al., 2010; Shen et al., 2011; Peng et al., 2011).

Nevertheless, in many situations, either accurate method does not exist, or data are not sufficient to exactly characterize probability distributions, or data are contaminated to provide precise information. Under such situations, probabilistic models, e.g., the aforementioned two types of models, could be inappropriate or lead to infeasible solutions. To address this challenge, robust optimization (RO) method, which simply assumes an uncertainty set to capture random data, is developed to provide solutions that are robust to any perturbations within the uncertainty set. To model the situation where some decisions can be made and implemented after the uncertainty is revealed, robust optimization is extended to include the second stage recourse decisions so that the available information can be fully utilized to produce a less conservative solution. After their introduction, original robust optimization method and its two-stage extension have been applied in many operational and engineering areas (Ben-Tal et al., 2009; Bertsimas et al., 2011), such as facility location problems with random demands (Atamturk and Zhang, 2007; Baron et al., 2011; Gabrel et al., 2014). In fact, comparing with demand uncertainty, disruptions are often less likely to be described by accurate probabilistic information. For example, earthquakes in California or hurricanes in Florida could cause facilities or client sites in those regions to be disrupted. However, it is very difficult to estimate the number of earthquakes or hurricanes in next 10 years based on historical/statistical data. Hence, in this paper, we apply the concept of uncertainty set to capture the random disruptions and employ robust optimization method to study reliable facility location problems.

Specifically, we adopt two-stage robust optimization approach to investigate the reliable p-median problem, where location decisions are made before (*here-and-now*) and recourse (mitigation) decisions are made after disruptions being revealed (*wait-and-see*). We mention that such a modeling framework exactly captures the decision making sequence in real operations. In particular, due to its strong modeling capability, we are able to extend our study to consider facility capacities and demand changes due to disruptions. The former situation is very challenging for probabilistic models while the latter has not been analytically investigated in existing literature. We further implement two solution algorithms, i.e., *Benders decomposition* and *column-and-constraint generation* methods. The latter one is enhanced by a few improvement strategies based on structural properties. A set of numerical experiments are performed to generate insights on the algorithm performance and the network design.

The rest of the paper is organized as follows. Section 2 reviews relevant literature on probabilistic models and two-stage robust optimization models. Section 3 introduces two-stage robust optimization reliable p-median models and analyzes their properties. Section 4 describes our solution algorithms. Numerical results and insights on system design are presented in Section 5, followed by Section 6, where the paper is concluded and future research directions are discussed.

2. Literature review

In this section, we briefly review two types of relevant studies on the facility location problem: probability based reliable facility location models and (two-stage) robust optimization formulations. Results on classical and deterministic facility location problems can be found in Daskin (1995) and Drezner (1995). For problems with uncertain demands and costs, readers are referred to a comprehensive review in Snyder (2006).

The research by Drezner (1987) is probably the first one studying facility location problem with unreliable facilities while Snyder and Daskin (2005) present the first reliable facility location models with inclusion of mitigation/recourse operations and costs. They implement Lagrangian relaxation algorithms within a branch-and-bound scheme to solve the resulting linear mixed integer programs for real instances. Chen et al. (2011) consider a combined facility location and inventory management system subject to facility failures. The authors develop an exact polynomial-time algorithm to handle the nonlinearity introduced by inventory costs and apply Lagrangian relaxation as the solution method. By relaxing the assumption that all sites share the same failure rate, Cui et al. (2010) build a nonlinear mixed integer program and develop both Lagrangian relaxation and continuum approximation (CA) methods for this challenging problem. To reduce the complexity of the nonlinear form, Lim et al. (2010) study a simplified model where clients are assigned to (unreliable) facilities and reliable backup facilities if needed. Shen et al. (2011) present both scenario-based stochastic programming and a nonlinear mixed integer programming model and show that they are generally equivalent. Also, a constant-ratio approximation algorithm for the case where all failure rates are identical is proposed. Li and Ouyang (2010) study a problem with correlated probabilistic disruptions and solve their model by CA method. Indeed, because CA technique could be useful to derive analytical insights, this approximation approach has been adopted to study the reliable facility location model in a competitive market environment by Wang and Ouyang (2013) and to investigate the effect of misestimation of (a single) failure rate on the network configuration by Lim et al. (2013). In addition to considering facility failures, Li et al. (2013) study the problem with a fortification strategy where the unreliable facilities can be fortified by hardening operations under a budget. Recently, this line of research is extended to investigate more general reliable network design problems. Peng et al. (2011) consider a reliable multiple-echelon logistics network design problem where disruptions can happen in multiple echelons. An et al. (2011) study reliable hub-and-spoke network design problems in which hubs could be disrupted and affected flows will be rerouted through survived operational hubs. From those aforementioned studies, we observe that (i) either complicated nonlinear mixed integer programs or large-scale scenario-based stochastic programs are necessary to build the model. When professional solvers are not efficient to deal with those models, customized algorithms, either analytical or heuristic ones, are developed and (ii) some practical situations are not sufficiently investigated. For example, very limited research is done for capacitated models except Peng et al. (2011) and Lim et al. (2013), and no exact algorithm has been developed. Although it is noted that disruptions could cause different demand patterns (Ergun et al., 2010), the impact of such type of demand changes has not been captured or included in the study of facility location problem.

Different from nonlinear mixed integer programs or scenario-based stochastic programs that are developed based on precise probabilistic information, robust optimization based location models, including those developed with two-stage robust optimization method, assume a probability-free uncertainty set and seek to determine locations that are robust to any perturbations in that uncertainty set. Baron et al. (2011) build a multi-period capacitated fixed charge (single-stage) robust location model and investigate the impact of different uncertainty sets on facility locations. Gülpınar et al. (2012) propose to use tractable (single-stage) robust optimization method to approximately solve stochastic facility location problem with a chance constraint. Atamturk and Zhang (2007), Gabrel et al. (2014), and Zeng and Zhao (2013) develop two-stage (tri-level) robust optimization formulations for location–transportation problems where locations and capacities are determined in the first stage and transportation decisions are adjusted after demand is realized. Different solution algorithms are proposed by them respectively, including an approximation algorithm (Atamturk and Zhang, 2007), Benders cutting plane algorithm (Gabrel et al., 2014), and the *column-and-constraint generation* (C&CG) algorithm (Zeng and Zhao, 2013). The *column-and-constraint generation* algorithm demonstrates a superior computational performance over Benders cutting plane method in the two-stage facility location problem subject to random demands.

We also note a survivable network design problem presented by Smith et al. (2007) which is formulated as a tri-level model where enemy's attack on arcs plays a role similar to site disruptions in the presented two-stage robust facility location models. It is different from our research by considering a commodity flow based network and employing Benders cutting plane as the solution method.

To make this paper focused, we restrict this study to p-median problem and leave the study of two-stage RO formulations for another classical model, i.e., the fixed-charge facility location problem, as a future research direction. Research presented in this paper makes the following contributions to the literature.

(i) To the best of our knowledge, no research has been done to apply two-stage RO to formulate reliable facility location problems with consideration of disruptions. Hence, this paper presents the first set of reliable facility location formulations using two-stage robust optimization tools.
(ii) Because of the modeling advantages of two-stage RO, we consider real features that have received very limited or no attention. They are finite capacities of facilities and demand changes due to disruptions.
(iii) In addition to some analytical study on these models, we customize and implement solution algorithms to perform numerical experiments. We also present management insights based on the numerical results from instances with real data.

3. Two-stage robust *p*-median reliable models

In this section, we present our formulations on two-stage RO reliable p-median facility location problem. We first consider uncapacitated robust models and then extend our work to consider capacitated cases. Existing research generally ignores the demand changes due to disruptions. We show that, by using the two-stage robust optimization framework, demand changes can be easily incorporated. We also derive structural properties of these models.

3.1. Robust uncapacitated p-median facility location models

Different from stochastic programming models that explicitly consider all possible uncertain scenarios (two-stage) robust optimization models use an uncertainty set to describe the concerned possible scenarios without depending on probability information. In the context of disruption description, we employ a cardinality constrained uncertainty set, which is probably the most used set to describe discrete uncertainties (see Atamturk and Zhang (2007) and Bertsimas and Sim (2004) for examples). Specifically, assuming that all sites in set J are homogeneous and considering all possible scenarios with up to k simultaneous disruptions, the uncertainty set, i.e., the disruption set in this paper, can be represented as

$$A = \left\{ \mathbf{z} \in \{0,1\}^{|J|} : \sum_{j \in J} z_j \leqslant k \right\}, \tag{1}$$

where z_j is the indicator variable for site j, i.e., $z_j = 1$ if site j is disrupted and $z_j = 0$ otherwise. Note that, although there may exist an exponential number of disruptive scenarios, this formulation provides an implicit but compact algebraic format to capture all of them. In the remainder of this paper, unless explicitly mentioned, we employ this disruption set to perform our study.

Next, we develop our two-stage RO reliable p-median facility location models. Let I be the set of client sites (clients for short) and $J \subseteq I$ be the set of potential facility sites. Following the convention of previous research (Snyder and Daskin, 2005; Chen et al., 2011; Cui et al., 2010), we assume that $I = J$. Each client $i \in I$ has a demand d_i and the unit cost of serving i by the facility at $j \in J$ is $c_{ij} \geqslant 0$ with $c_{ii} = 0$. We use $\mathbf{y}$ and $\mathbf{x}$ to denote the first stage (the normal situation without disruptions) decision variables: $y_j = 1$ means that a facility is located at j, $y_j = 0$ otherwise; $x_{ij} \in [0,1]$ represents the portion of i's demand served by j in the normal situation. Note that the first stage decision variables are to be fixed before any disruptive scenario $\mathbf{z}$ in set A is realized. In a disruptive scenario, as in Snyder and Daskin (2005) and Cui et al. (2010), a disrupted facility cannot serve any client. However, system reliability can be achieved by implementing recourse or mitigation operations such as re-assigning clients to survived facilities. So, we introduce $\mathbf{w}$ and $\mathbf{q}$ to represent the second stage recourse operation decisions in a disruptive scenario, where $w_{ij} \in [0,1]$ represents the portion of demand d_i served by the survived facility at j and $q_i \in [0,1]$ represents the unsatisfied portion. Each unit of unsatisfied demand of d_i will incur a penalty M.

To the best of our knowledge, all existing formulations on facility location problem assume that all sites keep generating regular demands in spite of disruptions. However, under some situations, disruptions will introduce new demand patterns. On the one hand, demands of non-essential or luxury products often vanish in natural disaster-caused disruptions. On the other hand, some daily necessities and protection-based items, such as medicines and batteries, will significantly increase (Ergun et al., 2010). To capture such phenomenon, we introduce a parameter θ to reflect the client demand change due to a disruption. Then, in a disruptive scenario, the demand of client i is set to $(1-\theta z_i)d_i$, which depends on the site disruption status z_i and θ. Clearly, by setting θ to a positive value (subject to $\leqslant 1$) or to a negative value, we can model the disruption-caused demand reduction or increase, respectively. Hence, we consider $\theta \in (-\infty, 1]$ in the remainder of this paper. Next, we present the two-stage RO reliable p-median facility location model with up to k simultaneous disruptions.

<u>RO-PMP$_\theta$</u>

$$V_\theta(p,k,\rho) = \min_{\mathbf{x},\mathbf{y}} \quad (1-\rho)\sum_i\sum_j c_{ij}d_ix_{ij} + \rho \max_{\mathbf{z}\in A} \min_{(\mathbf{w},\mathbf{q})\in S(\mathbf{y},\mathbf{z})} \left(\sum_i\sum_j c_{ij}(1-\theta z_i)d_iw_{ij} + \sum_i M(1-\theta z_i)d_iq_i \right), \tag{2}$$

$$\text{s.t.} \quad x_{ij} \leqslant y_j, \quad \forall i,j, \tag{3}$$

$$\sum_j x_{ij} = 1, \quad \forall i, \tag{4}$$

$$\sum_j y_j = p, \tag{5}$$

$$x_{ij} \geqslant 0, \quad \forall i,j; \; y_j \in \{0,1\}, \quad \forall j, \tag{6}$$

where

$$S(\mathbf{y},\mathbf{z}) = \{w_{ij} \leqslant 1 - z_j, \quad \forall i,j, \tag{7}$$

$$w_{ij} \leqslant y_j, \quad \forall i,j, \tag{8}$$

$$\sum_j w_{ij} + q_i = 1, \quad \forall i, \tag{9}$$

$$w_{ij} \geqslant 0, \quad \forall i,j; \; q_i \geqslant 0, \quad \forall i\}. \tag{10}$$

The objective function in (2) seeks to minimize the weighted sum of the operation costs in the normal disruption-free situation and in the worst disruptive scenarios in A. The weight $\rho \in [0,1]$ is a parameter reflecting the system designer's attitude towards the disruption cost. Clearly, a larger ρ indicates that the designer is more conservative and willing to configure the system in a way such that less recourse/mitigation operation costs will incur in disruptive situations. Constraints in (3)–(5) are from the classical p-median model and simply mean that a client can be assigned to a facility only if the facility is built, the entire demand of a client has to be served, and the total number of facilities is p, respectively.

The *max* operator identifies the disruptive scenario(s) in A yielding the largest operation cost, given the location $\mathbf{y}$. The second *min* seeks the least costly mitigation solution while the set $S(\mathbf{y},\mathbf{z})$ defines the possible recourse operations. That is, given the definition of y_j and z_j, constraints (7) and (8) ensure that in any disruptive scenario, client i's demand can only be assigned to established *and* survived facilities. Then, constraints in (9) represent that the portion $(1-q_i)$ of i's demand has to be served and the rest will be lost and penalized. In this paper, our research focuses on the nontrivial cases where $k \leqslant p-1$. Otherwise, there will be no mitigation operations in any worst disruption scenario and the problem reduces to the p-median formulation.

Note from (2)–(10) that the two-stage RO is a very adaptable modeling framework. By setting $\theta = 0$, we can compactly formulate the regular robust facility location problem without demand change, similar to existing studies on reliable facility location models (Snyder and Daskin, 2005; Cui et al., 2010; Peng et al., 2011). We can also consider the more involved situations by simply letting $\theta \neq 0$. As we mention, no existing work on the reliable facility location problem studies the impact of demand changes due to disruptions on the system design. One possible reason is that, if the demand change factor is considered, classical probabilistic models have to evaluate all possible scenarios while different scenarios will have different coefficients in their objective functions, which makes it very challenging to have a compact and tractable formulation.

Nevertheless, the two-stage robust optimization scheme provides us a convenient modeling framework to address this issue. Indeed, even for the most sophisticated situation where demand changes are site dependent, our model can easily capture it by introducing site specific θ_i into its objective function (2). We leave research on this line as a future direction to keep our paper focused.

Although this robust formulation is a complicated tri-level optimization problem, we can analyze the impact of θ by deriving some structural properties.

Remark 1.

(*i*) The "closest" principle always holds when assigning demands to facilities. Specifically, in the normal disruption-free situation, demand d_i from client i is served by a facility that is closest to client i. In a disruptive situation, d_i is either served by a survived facility that is closest to i, or abandoned if the unit service cost from that facility is more than M.

(*ii*) Let $v(\mathbf{y},\mathbf{z})$ be the optimal value of RO-PMP$_\theta$ for a given $\mathbf{y}$ and $\mathbf{z}$, including costs from both normal and disruptive situation $\mathbf{z}$. Because of (*i*), we have $v(\mathbf{y},\mathbf{z})$ is a linear non-increasing function with respect to $\theta \in (-\infty, 1]$. Furthermore, $v(\mathbf{y}) = \max_{\mathbf{z}} v(\mathbf{y},\mathbf{z})$ is a piecewise linear non-increasing function with respect to $\theta \in (-\infty, 1]$. Different pieces correspond to different $\mathbf{z}$.

(*iii*) $V_\theta(p,k,\rho) = \min_{\mathbf{y}} v(\mathbf{y})$ is a non-increasing quasi-convex function with respect to $\theta \in (-\infty, 1]$.

Based on Remark 1, when $\theta \in [0,1]$, i.e., a disruption either does not affect demands or causes demand reduction, the worst case disruptions can be further characterized.

Lemma 1. *When M is sufficiently large, i.e., $M \geqslant \max_{i,j} c_{i,j}$, consider given facility location $\mathbf{y}^*$ and the disruption set A. If $\theta \in [0,1]$, the worst case disruptions and therefore demand reductions happen only at facility sites, i.e., those with $y_j^* = 1$.*

Proof. We prove it by contradiction. Consider a worst case disruptive situation $\mathbf{z}^1$ where a disruption happens at site j_0, on which there is no facility, i.e., $z_{j_0}^1 = 1$ and $y_{j_0}^* = 0$. Let C^1 be the operation cost under this disruptive situation.

As $p - k \geqslant 1$, there exists a facility, say j_1 with $y_{j_1}^* = 1$, survived in the disruptive situation $\mathbf{z}^1$. Consider two disruptive situations: $\mathbf{z}'$ where $z'_{j_0} = 0$, and $z'_j = z_j^1$ for $j \neq j_0$, and $\mathbf{z}^2$ where $z_{j_0}^2 = 0, z_{j_1}^2 = 1$ and $z_j^2 = z_j^1$ for $j \neq j_0$ and $j \neq j_1$. Denote the operation cost under $\mathbf{z}'$ by C', and that under $\mathbf{z}^2$ by C^2.

First, it is clear that $C' \geqslant C^1$, because demand from client j_0 must be served by some facility in $\mathbf{z}'$ while this demand is decreased in $\mathbf{z}^1$ and incurs less cost. Note that the equality could be achieved only if $\theta = 0$.

Second, under the disruptive situation $\mathbf{z}^2$, because the facility at j_1 is not available, all its served clients' demands, including the demand from j_1, will be served by other survived facilities, which will be further and more costly. Also, because $c_{j_1 j_1} = 0$, the demand from site j_1 will not incur any service cost in $\mathbf{z}'$. So, we have $C^2 \geqslant C'$. Note that the equality could be achieved only if $\theta = 1$.

Because $C^2 \geqslant C' \geqslant C^1$ and equalities cannot be achieved simultaneously (given θ only takes a single value), we have the desired contradiction. □

However, the case with $\theta < 0$, i.e., a disruption will cause demand increase, is more complicated. Indeed, as demonstrated in the following example, for a given facility location solution $\mathbf{y}^*$, the worst case disruption may happen at a non-facility site.

Example 1. In the uncapacitated 4-site network ($\theta = -1$) in Fig. 1(a), two facilities are built on site 2 and 4. The unit service costs are symmetric and they are $c_{12} = c_{34} = c_{13} = c_{24} = 1$, and $c_{14} = c_{23} = 1.41$. The demands are $d_1 = d_3 = 100, d_2 = d_4 = 10$, and the penalty M is set to 15. Consider $k = 1$. We observe in Fig. 1(b) that a disruption at site 2 (or site 4, respectively) will disable its facility, cause a demand increase, and incur a recourse cost of 261. Nevertheless, any disruption at a client site, such as site 1 in Fig. 1(c), will cause a larger demand increase and incur a higher recourse cost of 300. Therefore, the worst case disruption could happen at a non-facility site.

Note from this example that we cannot properly evaluate the reliability of a distribution network by just considering facility disruptions, if demands increase along with disruptions. It further highlights the importance to adopt the presented robust p-median facility location model to analytically design the distribution network to hedge against disruptions.

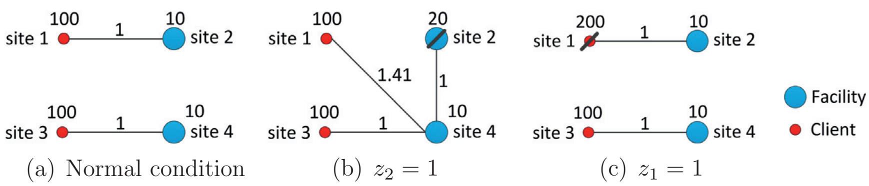


Fig. 1. A 4-site network with $\theta = -1$.

Because the "closest" principle holds in both normal and disruptive situations, the following result is valid when demands must be met by existing facilities. Let $V_\theta(p,k,\rho|\mathbb{C})$ denote its optimal value under some condition $\mathbb{C}$. Then, we can evaluate the function $V_\theta(p,k,\rho)$ by its input parameters.

Lemma 2. *For a given facility location* $\mathbf{y}_0$*, let* $C_r(\mathbf{y}_0)$ *and* $C_{\mathbf{z}^*}(\mathbf{y}_0)$ *be the operating costs in the normal situation and a worst disruptive situation* $\mathbf{z}^*$ *in A, respectively. When M is sufficiently large, we have* $C_{\mathbf{z}^*}(\mathbf{y}_0) \geqslant C_r(\mathbf{y}_0)$.

Proposition 1. *When M is sufficiently large, the function* $V_\theta(p,k,\rho)$ *is (i) non-increasing with respect to p; (ii) non-decreasing with respect to k; and (iii) non-decreasing with respect to* ρ.

Proof. Statements in (i) and (ii) are straightforward. We give the proof for the statement (iii). Consider $\rho_1 \leqslant \rho_2$ and their corresponding optimal facility locations $\mathbf{y}_1$ and $\mathbf{y}_2$.

Clearly, as $\mathbf{y}_2$ may not be optimal when $\rho = \rho_1$, we have

$$V_\theta(p,k,\rho_1) = V_\theta(p,k,\rho_1|\mathbf{y}=\mathbf{y}_1) \leqslant V_\theta(p,k,\rho_1|\mathbf{y}=\mathbf{y}_2).$$

Given that $\rho_1 \leqslant \rho_2$, it follows from Lemma 2 that

$$V_\theta(p,k,\rho_1|\mathbf{y}=\mathbf{y}_2) \leqslant V_\theta(p,k,\rho_2|\mathbf{y}=\mathbf{y}_2).$$

Therefore, we have

$$V_\theta(p,k,\rho_1) = V_\theta(p,k,\rho_1|\mathbf{y}=\mathbf{y}_1) \leqslant V_\theta(p,k,\rho_2|\mathbf{y}=\mathbf{y}_2) = V_\theta(p,k,\rho_2). \quad \square$$

We mention that results in Proposition 1 could be useful in algorithm design and implementation validation. For example, to deal with complicated instances with large p,k and ρ, strong lower or upper bounds can be obtained from instances with small p,k and ρ, which are likely to be computationally simpler.

Next, we extend our study to the capacitated facility location problem, whose reliable models have received little research attention.

3.2. The robust capacitated facility location model

The capacitated p-median facility location (CPMP) problem is an extension of the classical facility location model. Besides the same objective function and decision variables as in the classical uncapacitated facility location problem, it assumes that each potential facility has a capacity, i.e., an upper bound on the amount of demand that it can serve (Sridharan, 1995). Let K_j denote the capacity of site j. The two-stage robust capacitated p-median facility location problem is shown as follows:

<u>RO-CPMP$_\theta$</u>

$$V_\theta^C(p,k,\rho) = \min_{\mathbf{x},\mathbf{y}} \quad (1-\rho)\sum_i\sum_j c_{ij}d_ix_{ij} + \rho \max_{\mathbf{z}\in A} \min_{(\mathbf{w},\mathbf{q})\in S^C(\mathbf{y},\mathbf{z})} \left(\sum_i\sum_j c_{ij}(1-\theta z_i)d_iw_{ij} + \sum_i M(1-\theta z_i)d_iq_i\right), \tag{11}$$

$$\text{s.t.} \quad (3)\text{–}(6),$$

$$\sum_i d_ix_{ij} \leqslant K_jy_j, \quad \forall j, \tag{12}$$

with

$$S^C(\mathbf{y},\mathbf{z}) = \left\{(7)\text{–}(10) \quad \sum_i (1-\theta z_i)d_iw_{ij} \leqslant K_jy_j, \quad \forall j\right\}. \tag{13}$$

Constraints (12) ensure that the total demand served by facility j does not exceed its capacity K_j. Constraints (13) impose the similar requirement on the survived facility j.

Because of the facility capacity constraints, RO-CPMP$_\theta$ is less trackable than the uncapacitated one. Nevertheless, under several mild assumptions, some properties can be derived.

We assume that (i) $K_j \geqslant d_j$ for all j, and (ii) the service costs satisfy the triangular inequality. The first assumption indicates that in both normal and disruptive situations, it is feasible to serve the whole demand of a (survived) facility site by the facility on it. The second assumption shows that it always leads to less cost to serve the demand of a (survived) facility site by the facility on it. Note that when $\theta < 0$, because demands increase along with disruptions, the total capacity from existing facilities may not be sufficient and therefore the penalty due to unmet demands will be incurred. When $\theta \in [0,1]$, a result similar to Lemma 1 can be easily derived.

Lemma 3. *Under assumptions (i) and (ii), when M is sufficiently large, consider a given facility solution* $\mathbf{y}^*$ *and a disruption set A. We have that if* $\theta \in [0,1]$*, the worst case disruptions and therefore demand reductions happen only at facility sites, i.e., those with* $y_j^* = 1$.

Also, similar to Lemma 2 and Proposition 1, we derive the following results to evaluate $V_\theta^C(p,k,\rho)$ by its input parameters.

Lemma 4. *When M is sufficiently large, for a given facility solution* $\mathbf{y}_0$ *and a worst disruptive situation* $\mathbf{z}^*$*, we have* $C_{\mathbf{z}^*}(\mathbf{y}_0) \geqslant C_r(\mathbf{y}_0)$.

Proposition 2. *When M is sufficiently large, the function* $V_\theta^C(p,k,\rho)$ *is (i) non-increasing with respect to p; (ii) non-decreasing with respect to k; and (iii) non-decreasing with respect to* ρ.

4. Solution algorithms

Two-stage RO models in general are very difficult to solve (Ben-Tal et al., 2004). When the second stage mitigation problem is a linear program (LP), as in each of the models we introduce so far, Benders decomposition method can be employed to seek optimal solutions (Bertsimas et al., 2013; Jiang et al., 2011). However, Benders method is not efficient in dealing with real size instances. A different solution method, the *column-and-constraint generation* algorithm, denoted by C&CG algorithm, was developed in Zeng and Zhao (2013) recently, which shows a superior performance over Benders method in solving practical problems. In this paper, we adopt C&CG method as the primary solution method to solve the proposed RO models. We first provide details of a customized C&CG method for our robust models and then present a set of enhancement strategies. We also briefly discuss the implementation of Benders decomposition method.

4.1. Implementation of C&CG algorithm

We select RO-PMP$_\theta$ to describe the development of the customized C&CG algorithm. Because the capacitated robust model is of a similar structure, C&CG can be implemented with minor modifications.

C&CG algorithm is implemented within a two level master-sub problem framework. In the subproblem, for a given solution $(\mathbf{x}^*, \mathbf{y}^*)$ to the first stage decision problem, we solve the remaining *max*−*min* problem to identify the worst scenario. As the unsatisfied demand will be penalized in any disruptive situation, the second stage mitigation problem is always feasible. Hence, we can take the dual and obtain a *max*−*max* problem, which is actually a maximization problem. Specifically, let **u**, **v**, and **s** be the dual variables of the constraints (7)–(9) respectively. The resulting nonlinear maximization formula of subproblem is as follows:

<u>NL-SubP</u>

$$\max_{\mathbf{z},\mathbf{u},\mathbf{v},\mathbf{s}} \sum_i \sum_j (1-z_j)u_{ij} + \sum_i \sum_j y_j^* v_{ij} + \sum_i s_i, \tag{14}$$

$$\text{s.t.} \quad u_{ij} + v_{ij} + s_i \leqslant c_{ij} d_i (1-\theta z_i), \quad \forall i,j, \tag{15}$$

$$s_i \leqslant M d_i (1-\theta z_i), \quad \forall i, \tag{16}$$

$$\sum_{j \in J} z_j \leqslant k, \tag{17}$$

$$u_{ij} \leqslant 0, \quad \forall i,j; \;\; v_{ij} \leqslant 0, \quad \forall i,j; s_i \textit{ free}, \quad \forall i; z_j \in \{0,1\}, \quad \forall j. \tag{18}$$

As the nonlinear terms are the products of a continuous variable and a binary variable, we can linearize this formulation by replacing them with a set of new variables, i.e., $U_{ij} = u_{ij} z_j$, and using *big-M* method. We denote the *big-M* as $\mathbb{M}$ to differentiate it from the penalty coefficient *M*.

As a result, the linearized subproblem is:

<u>SubP</u>

$$\mathcal{Q} = \max_{\mathbf{z},\mathbf{u},\mathbf{v},\mathbf{s},\mathbf{U}} \sum_i \sum_j \left(u_{ij} - U_{ij} + y_j^* v_{ij} \right) + \sum_i s_i, \tag{19}$$

$$\text{s.t.} \quad u_{ij} + v_{ij} + s_i \leqslant c_{ij} d_i (1-\theta z_i), \quad \forall i,j, \tag{20}$$

$$s_i \leqslant M d_i (1-\theta z_i), \quad \forall i, \tag{21}$$

$$U_{ij} \geqslant u_{ij}, \quad \forall i,j, \tag{22}$$

$$U_{ij} \geqslant -\mathbb{M} z_j, \quad \forall i,j, \tag{23}$$

$$U_{ij} \leqslant u_{ij} + \mathbb{M}(1-z_j), \quad \forall i,j, \tag{24}$$

$$\sum_{j \in J} z_j \leqslant k, \tag{25}$$

$$\begin{aligned} & u_{ij} \leqslant 0, \quad \forall i,j; U_{ij} \leqslant 0, \quad \forall i,j; \;\; v_{ij} \leqslant 0, \quad \forall i,j; \\ & s_i \textit{ free}, \quad \forall i; \;\; z_j \in \{0,1\}, \quad \forall j. \end{aligned} \tag{26}$$

Note that the linearized subproblem (SubP), which is a MIP problem, can be solved by a professional MIP solver. Next, we describe the details of the *column-and-constraint generation algorithm* along with the formulation of the *master problem*, which will be solved iteratively. In each iteration n, a significant scenario $\mathbf{z}^n$ will be identified through solving SubP. Then, a set of recourse variables $(\mathbf{w}^n, \mathbf{q}^n)$ and corresponding constraints in the forms of (28)–(31) associated with this particular scenario will be created and added to the *master problem*, whose complete set of variables are listed in (32). Let *UB* and *LB* be the upper and lower bounds respectively, *Gap* be the relative gap between *UB* and *LB*, n be the iteration index and ϵ be the optimality tolerance.

4.1.1. Column-and-constraint generation algorithm

1. Set $LB = -\infty, UB = \infty$, and $n = 0$.
2. Solve the following *master problem* (MP) and obtain an optimal solution $(\mathbf{x}^n, \mathbf{y}^n, \eta^n)$ and set *LB* to the optimal value of the MP.
MP

$$\min \quad (1-\rho)\sum_i\sum_j c_{ij}d_i x_{ij} + \rho\eta, \tag{27}$$

$$\text{s.t.} \quad (3)\text{–}(6),$$

$$\eta \geqslant \sum_i\sum_j c_{ij}(1-\theta z_i^l)d_i w_{ij}^l + \sum_i M(1-\theta z_i^l)d_i q_i^l, \quad \forall l = 1,2,\ldots,n, \tag{28}$$

$$\sum_j w_{ij}^l + q_i^l = 1, \quad \forall i, l = 1,2,\ldots,n, \tag{29}$$

$$w_{ij}^l \leqslant 1 - z_j^l, \quad \forall i,j,l = 1,2,\ldots,n, \tag{30}$$

$$w_{ij}^l \leqslant y_j, \quad \forall i,j,l = 1,2,\ldots,n, \tag{31}$$

$$x_{ij} \geqslant 0, \quad \forall i,j; \ y_j \in \{0,1\}, \quad \forall j; \eta \ \mathit{free};$$

$$w_{ij}^l \geqslant 0, \quad \forall i,j,l = 1,2,\ldots,n; q_i^l \geqslant 0, \quad \forall i,l = 1,2,\ldots,n. \tag{32}$$

3. Solve SubP with respect to $(\mathbf{x}^n, \mathbf{y}^n)$ and derive an optimal solution $(\mathbf{z}^n, \mathbf{u}^n, \mathbf{v}^n, \mathbf{s}^n)$ and its optimal value $\mathcal{Q}^n$. Update

$$UB = \min\left\{UB, (1-\rho)\sum_i\sum_j c_{ij}d_i x_{ij}^n + \rho\mathcal{Q}^n\right\}.$$

4. If $Gap = \frac{UB-LB}{LB} \leqslant \epsilon$, an ϵ-optimal solution is found, terminate. Otherwise, create recourse variables $(\mathbf{w}^n, \mathbf{q}^n)$ and corresponding constraints associated with the identified $\mathbf{z}^n$ and add them to MP. Update $n = n + 1$. Go to Step 2. □

It has been proven in Zeng and Zhao (2013) that C&CG algorithm converges to an optimal solution in finite iterations. Different from C&CG method, after solving SubP, Benders decomposition method will iteratively supply a single cutting plane in the following form to its master problem that only carries the first stage decision variables $(\mathbf{x}, \mathbf{y})$

$$\eta \geqslant \sum_i\sum_j \left(1 - z_j^n\right)u_{ij}^n + \sum_i\sum_j v_{ij}^n y_j + \sum_i s_i^n.$$

Comparing these two types of algorithms, Zeng and Zhao (2013) theoretically show that (i) C&CG method is of a much less computational complexity. In our study, it depends on the cardinality of the disruption set. However, for Benders decomposition method, its computational complexity depends on the product of cardinality of the disruption set and the number of extreme points of the dual for the recourse problem. (ii) For C&CG method, its generated constraints are always stronger than those generated by Benders decomposition method. Because of its proven computational advantage and solution capability, C&CG method is recently employed to solve robust optimization problems in different applications (Zugno and Conejo, 2013; Hervet et al., 2013).

4.2. Algorithm enhancement

In this section, we study how to improve the computational performance of the basic C&CG method on solving reliable p-median facility location problems. In particular, note that the numbers of variables and constraints in MP will quickly increase over iterations, which may demand excessive amount of computational time for large instances. Hence, we develop a few enhancement strategies to reduce the computational expenses.

Variable fixing. Variable fixing technique has been widely used within Lagrangian relaxation algorithms. It has been proven to be effective in reducing computation time for complicated network design problems (see Snyder and Daskin (2005) and Contreras et al. (2011) for examples). Now, we extend this idea to improve C&CG method.

For any $i \in I$, we have the following results:

Proposition 3. *Assume that UB is derived from a feasible solution to RO-PMP$_\theta$.*

(*i*) *if* $V_\theta(p,k,\rho|y_i=0) > UB$ *for some i, we have* $y_i = 1$ *in any optimal solution and* (*ii*) *if* $V_\theta(p,k,\rho|y_i=1) > UB$ *for some i, we have* $y_i = 0$ *in any optimal solution.*

Consequently, we can implement the following *variable fixing* steps within C&CG method.

Corollary 1. *Assume that* $\mathbf{y}^\circ$ *is the best facility location solution obtained and the corresponding objective function value is* UB°. *The following operations can be implemented without losing any optimal solution:*

(i) *if* $y_i^\circ = 1$ *and the optimal value of MP with an additional constraint* $y_i = 0$ *is strictly greater than* UB°, *fix* $y_i = 1$ *in MP;*
(ii) *if* $y_i^\circ = 0$ *and the optimal value of MP with an additional constraint* $y_i = 1$ *is strictly greater than* UB°, *fix* $y_i = 0$ *in MP.*

Note that once y_i is fixed, MP's feasible space will be reduced, which may lead to better solutions with less computational time. Although computing MP to optimality maybe difficult in practice, we can derive its lower bound within a time limit and use that lower bound to perform the aforementioned variable fixing operations.

Multiple scenario generation. Note that C&CG method generates, through solving SubP, a single scenario (i.e., its corresponding variables and constraints as in (28)–(32)) in each iteration. Because every scenario yields a valid lower bound, we actually can identify multiple significant scenarios, instead of a single optimal one. By supplying those scenarios to MP, we can further speed up the increase of the *LB*. Specifically, at each iteration *n*, given an optimal $\mathbf{z}^n$ that solves SubP, we create another disruptive scenario by modifying disrupted sites with the least demands to non-disrupted ones and changing non-disrupted sites with the largest demands to disrupted ones. Scenarios that replicate existing ones are eliminated in our implementation, although they will not affect the final solution.

We observe that variable fixing and multiple scenario generation typically work better for large-scale instances. For small-scale instances, because they will either incur extra computational time on probing variables or lead to larger MP with more scenarios, they may not show as good performance as the basic method does.

Good solutions of MP before convergence. Note that MP gradually evolves into a large-scale MIP that is computationally intensive. Indeed, when *Gap* of C&CG method is large, it is not necessary to derive an optimal solution of MP and a good feasible solution could be sufficient to generate significant disruptive scenarios. So, in the beginning iterations where *Gap* is large, we can set a relatively larger optimality tolerance for a good solution to MP. As *Gap* becomes smaller, a smaller optimality tolerance will be adopted for a better solution and a more precise lower bound. Also, according to Propositions 1 and 2, we can impose bound constraint on the objective function, which could also reduce the computational time of the branch-and-bound process in solving MP.

5. Numerical study and analysis

In this section, we first describe data and experimental setup. Then, we demonstrate the results of a set of numerical experiments and present our insights on various reliable *p*-median models.

All of our experiments are performed on the 49-site data set described in Snyder and Daskin (2005), which includes information of demands and site coordinates. We also consider a data set of 25 sites that are randomly selected from the 49-site data set as shown in Table 1.

In the study, c_{ij} is the Euclidean distance between site *i* and *j* obtained from site coordinates. For capacitated models, the capacity of each site is randomly generated between $[D/10, 3D/10]$ where *D* is the total demand of all sites. If the capacity is smaller than its demand, we set the value of capacity equal to the demand. For all problems with 25 or 49 sites, we test them with different parameter values, i.e., $\theta = -1$, 0, and 1, $\rho = 0.2, 0.4, p = 8, 10$, and $k = 1, 2$, and 3, totally 72 instances. For each instances, we also consider two cases where $M = 15$ and $M = \max_{i,j}\{c_{ij}\}$. The first value resembles a situation where an affected demand will be served by competitors if the service cost of using survived facilities is more than 15. The second value represents a situation where all demands must be served (if capacity is sufficient) in any disruptive scenario.

C&CG algorithm is our primary solution method and we apply it to solve all cases. For the comparison purpose, we also implement Benders decomposition method (BD) and benchmark it with C&CG on the RO-PMP$_1$ model with $|I| = 25$ and $M = 15$ to confirm the efficiency of C&CG algorithm over BD method. The optimality tolerance ϵ is set as 0.1% and time limit

Table 1
Cities in the 25-site data set.

No.	City	No.	City	No.	City	No.	City
0	Austin (TX)	7	St. Paul (MD)	14	Topeka (KS)	21	Pierre (SD)
1	Tallahassee (FL)	8	Baton Rouge (LA)	15	Charleston (WV)	22	Dover (DE)
2	Harrisburg (PA)	9	Frankfort (KY)	16	Salt Lake City (UT)	23	Washington (DC)
3	Columbus (OH)	10	Columbia (SC)	17	Lincoln (NE)	24	Montpelier (VT)
4	Richmond (VA)	11	Denver (CO)	18	Augusta (ME)		
5	Boston (MA)	12	Hartford (CT)	19	Boise City (ID)		
6	Annapolis (MD)	13	Des Moines (IA)	20	Helena (MT)		

Table 2
Computational performance of Benders decomposition on the RO-PMP$_1$ with $M = 15$.

$\|I\|$	p	k	$\rho = 0.2$				$\rho = 0.4$			
			Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	8	1	T	2321	1426.76	1.11	6994.2(m)	2308	1539.77	31.30
		2	T	2317	1525.92	7.54	T	2320	1721.44	38.53
		3	T	2350	1649.93	14.34	T	2351	1821.58	41.81
	10	1	6666.7(m)	2337	1024.11	6.61	5443.7(m)	2103	1067.20	33.19
		2	5274.8(m)	2101	1066.29	10.86	6887.9(m)	2377	1151.57	37.66
		3	6714.5(m)	2361	1119.92	14.56	6717.1(m)	2365	1292.44	44.47

Table 3
Computational results for uncapacitated instances with $M = 15$.

$\|I\|$	ρ	p	k	RO-PMP$_{-1}$				RO-PMP$_0$				RO-PMP$_1$			
				Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	0.2	8	1	0.8	2	1763.95		1.9	5	1558.09		1.9	5	1426.76	
			2	0.8	3	2233.97		2.8	6	1855.51		3.2	7	1525.92	
			3	40.2	19	2846.98		818.8	59	2292.82		11.3	12	1649.93	
		10	1	0.7	2	1364.19		0.7	2	1139.08		3.4	7	1024.11	
			2	2.1	6	1759.77		2.0	6	1374.09		5.9	9	1066.29	
			3	3.5	6	2088.41		6.4	10	1601.93		4.9	8	1119.92	
	0.4	8	1	0.7	2	2214.17		2.9	5	1802.45		2.0	7	1539.77	
			2	1.4	4	3086.90		16.3	14	2335.00		18.0	14	1721.44	
			3	184.4	32	4126.34		6627.6	97	3136.73		40.1	18	1821.58	
		10	1	0.7	2	1814.40		1.8	2	1364.19		3.2	7	1067.20	
			2	2.6	5	2503.45		3.9	7	1785.45		9.4	11	1151.57	
			3	4.3	8	3139.28		51.2	23	2235.93		25.8	15	1292.44	
49	0.2	8	1	4.8	4	6563.12		3.1	4	6145.54		2.2	3	5684.45	
			2	732.7	21	7502.22		48.2	10	6808.88		96.9	12	6020.47	
			3	T	32	8457.28	6.68	T	35	7426.71	5.64	6058.7	38	6364.45	
		10	1	2.1	3	5339.88		0.7	2	4950.32		9.1	6	4687.15	
			2	164.4	13	6090.26		126.3	12	5534.76		160.5	14	4982.95	
			3	T	30	7148.76	5.89	T	31	6111.83	4.27	737.5	20	5131.20	
	0.4	8	1	44.5	9	7405.65		4.5	4	6625.50		14.3	7	5898.97	
			2	T	37	8892.50	6.35	4208.4	30	8027.37		T	36	6528.61	0.23
			3	T	32	10625.10	15.87	T	28	9198.43	19.56	T	31	7215.81	8.11
		10	1	7.2	5	5988.67		3.7	4	5351.47		12.4	7	4796.70	
			2	T	27	7402.96	1.16	7165.2	31	6365.19		2153.1	24	5264.33	
			3	T	29	9337.32	17.71	T	28	7467.20	12.46	T	31	5697.75	4.52

7200 s. The master problems and the subproblems are solved by a mixed integer programming solver, CPLEX 12.5, at its default settings. All algorithms are implemented in C++ and tested on a Dell Optiplex 760 desktop computer (Intel Core 2 Duo CPU, 3.0 GHz, 3.25 GB of RAM) in Windows XP environment.

5.1. Algorithm performance

Table 2 presents the performance of BD methods on instances of RO-PMP$_1$. Tables 3–6 summarize the computational results of C&CG algorithm on the uncapacitated and capacitated reliable models with different M values. In those tables, the column Time (s) presents the computational time in seconds; the column Iter indicates the number of iterations; the column Obj shows the best objective value ever found; the column Gap (%) provides the relative gap in percentage if it is larger than ϵ. We use letter T in Time (s) column to indicate an instance which cannot be solved within the time limit. If the algorithm terminates because the memory is not sufficient for the solver to compute MP or SubP, we add "(m)" after the computation time in column Time (s).

Based on these tables, we observe that

(i) C&CG algorithm performs hundreds of times faster and takes much fewer iterations than the classical Benders decomposition method. This result confirms the observations made in Zeng and Zhao (2013) for the location–transportation network design problem. Actually, compared to results in Zeng and Zhao (2013), a more significant superiority of the enhanced C&CG algorithm is observed in solving reliable p-median problems. To further demonstrate the computational advantage of C&CG algorithm, Fig. 2 shows the convergence plots of two algorithms under the time limit of 15 s for a case ($\theta = 1, M = 15,\ |I| = 25, \rho = 0.2,\ p = 8,$ and $k = 3$). Obviously, these two methods present distinct

Table 4
Computational results for uncapacitated instances with $M = \max_{i,j}\{c_{ij}\}$.

$\|I\|$	ρ	p	k	RO-PMP$_{-1}$				RO-PMP$_0$				RO-PMP$_1$			
				Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	0.2	8	1	1.9	2	1763.95		1.8	5	1558.09		1.3	5	1426.76	
			2	0.6	3	2233.97		2.3	6	1855.51		3.7	7	1525.92	
			3	177.6	32	2914.07		2287.6	71	2396.15		39.4	19	1734.64	
		10	1	1.8	2	1364.19		0.8	2	1139.08		6.6	9	1026.99	
			2	3.5	6	1759.77		2.7	6	1374.09		7.2	10	1083.22	
			3	3.8	6	2088.41		14.5	12	1616.81		5.6	8	1119.92	
	0.4	8	1	0.6	2	2214.17		2.8	5	1802.45		3.8	7	1539.77	
			2	1.4	4	3086.90		18.8	14	2335.00		25.1	15	1738.11	
			3	545.6	46	4197.12		T	88	3228.20	3.03	258.4	33	2010.03	
		10	1	1.7	2	1814.40		0.7	2	1364.19		7.7	10	1094.38	
			2	2.9	5	2503.45		4.6	7	1785.45		16.7	13	1183.81	
			3	6.0	8	3139.28		79.8	26	2252.60		35.1	17	1292.44	
49	0.2	8	1	5.5	4	6567.18		4.7	4	6149.60		1.8	3	5686.38	
			2	T	30	8132.07	3.18	T	35	7599.45	6.43	77.7	10	6026.61	
			3	T	29	9561.88	13.09	T	32	8329.70	11.02	5518.2	38	6432.67	
		10	1	2.5	3	5339.88		1.9	2	4950.32		10.3	6	4687.15	
			2	192.5	13	6090.26		138.1	12	5534.76		261.2	16	4992.81	
			3	T	32	7288.48	8.28	T	36	6520.93	11.16	1117.1	22	5131.20	
	0.4	8	1	55.1	9	7405.65		5.7	4	6633.61		5.8	4	5902.83	
			2	T	29	9974.33	10.97	T	34	9294.07	15.56	5316.4	30	6583.29	
			3	T	25	12920.90	25.70	T	27	10829.30	23.77	T	31	7384.12	9.61
		10	1	10.5	5	5988.67		4.7	4	5351.47		14.6	7	4796.70	
			2	T	26	7330.81	0.91	7112.8	31	6365.19		T	35	5316.31	0.13
			3	T	29	9673.43	19.77	T	32	8384.91	23.48	T	30	5792.46	7.16

Table 5
Computational results for capacitated instances with $M = 15$.

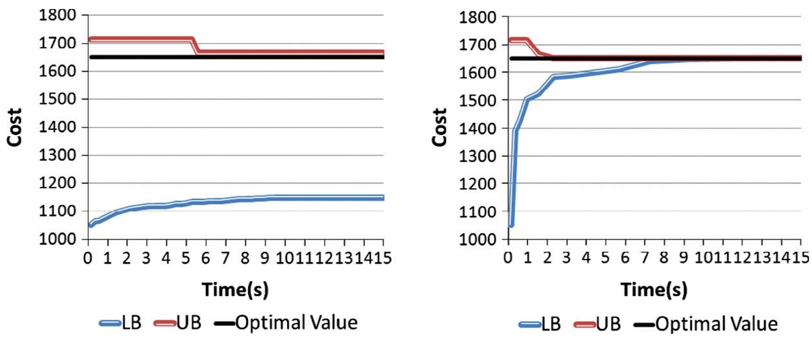
$\|I\|$	ρ	p	k	RO-CPMP$_{-1}$				RO-CPMP$_0$				RO-CPMP$_1$			
				Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	0.2	8	1	0.9	2	2069.34		1.7	3	1783.39		2.8	4	1550.58	
			2	2.2	3	2686.70		4.8	6	2126.51		7.5	7	1638.56	
			3	4.6	5	3370.00		10.4	9	2466.17		6.7	7	1725.95	
		10	1	0.5	2	1474.23		0.6	2	1148.48		1.4	4	1036.93	
			2	1.2	2	1923.53		1.2	3	1383.48		4.6	6	1094.58	
			3	2.5	3	2405.02		2.8	3	1642.20		7.7	8	1180.54	
	0.4	8	1	1.3	2	2620.20		1.1	3	2024.04		3.8	5	1623.29	
			2	5.3	5	3809.90		25.9	11	2791.94		11.5	9	1799.26	
			3	40.6	13	5021.64		543.0	33	3495.95		79.9	18	2015.52	
		10	1	1.4	2	1981.33		0.7	2	1373.58		4.3	6	1083.45	
			2	1.8	2	2869.74		1.7	3	1843.59		12.1	10	1224.17	
			3	3.9	4	3832.70		14.8	10	2307.08		181.7	25	1403.52	
49	0.2	8	1	20.3	5	6946.01		13.2	4	6432.77		24.8	5	6180.25	
			2	7150.5	24	7970.07		T	26	7290.78	6.30	1165.5	19	6525.89	
			3	T	21	8977.90	8.49	T	21	8095.26	12.93	7173.6	33	6937.76	
		10	1	11.4	4	5616.50		7.6	4	5045.10		19.4	5	4778.10	
			2	3968.1	22	6493.59		3335.1	26	5752.22		442.7	15	5097.14	
			3	T	20	7361.09	8.55	T	24	6187.69	3.65	3476.3	23	5250.64	
	0.4	8	1	164.6	9	7730.58		28.3	6	6939.44		87.2	9	6426.95	
			2	T	22	9591.76	12.70	T	22	8582.31	15.45	T	25	7113.72	2.20
			3	T	16	11616.60	26.69	T	19	10144.40	25.84	T	22	7937.46	11.82
		10	1	111.7	8	6248.44		23.8	6	5434.73		39.1	7	4881.02	
			2	T	21	7838.19	6.14	T	28	6888.70	8.38	T	27	5552.00	1.79
			3	T	17	9496.51	18.43	T	22	7753.12	12.64	T	23	5885.55	6.35

patterns in Fig. 2. In Fig. 2(a) Benders decomposition method cannot reduce the gap between upper bound and lower bound. In particular, it fails to improve lower bound. In Fig. 2(b), however, upper and lower bounds of C&CG method quickly converge to the optimal value in a short time.

(ii) The computation complexity of C&CG algorithm increases with the problem size $|I|$ and k, as well as the weight coefficient ρ. In all four tables, the most challenging instances are those with largest $|I|, k$, and ρ. Note that all instances with $k = 1$ are easy to compute. Most small size instances with $|I| = 25$ can be solved to optimality or with a small optimality gap while some instances with $|I| = 49$ are difficult. A closer analysis shows that SubP is easier to compute

Table 6
Computational results for capacitated instances with $M = \max_{i,j}\{c_{ij}\}$.

$\lvert I \rvert$	ρ	p	k	RO-CPMP$_{-1}$				RO-CPMP$_0$				RO-CPMP$_1$			
				Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	0.2	8	1	2.1	2	2069.34		2.5	3	1783.39		3.2	5	1550.58	
			2	9.8	7	2729.23		5.7	6	2126.51		23.7	12	1752.15	
			3	27.4	9	3944.68		243.4	20	2772.17		83.6	17	1880.17	
		10	1	1.1	2	1474.23		0.8	2	1148.48		4.3	6	1037.56	
			2	2.6	2	1923.53		2.4	3	1383.48		5.8	6	1094.58	
			3	8.5	5	2486.07		11.2	8	1726.22		24.7	12	1226.52	
	0.4	8	1	1.7	2	2620.20		2.7	3	2024.04		5.1	6	1656.87	
			2	22.2	9	3884.64		38.1	12	2816.64		243.3	27	2024.80	
			3	440.8	21	5948.72		2387.6	37	3863.25		981.5	38	2263.44	
		10	1	1.6	2	1981.33		0.6	2	1373.58		9.1	8	1103.78	
			2	2.4	2	2869.74		2.2	3	1843.59		14.5	10	1224.17	
			3	12.7	7	3920.41		91.3	17	2425.86		1248.9	46	1520.56	
49	0.2	8	1	28.8	5	7008.87		82.7	8	6517.80		57.7	7	6197.10	
			2	T	19	8953.19	12.73	T	24	8175.68	6.71	5238.2	24	6585.34	
			3	T	17	11586.60	32.24	T	21	10005.10	30.47	T	22	7963.78	15.91
		10	1	16.5	4	5616.50		8.3	4	5045.10		33.4	6	4794.66	
			2	T	21	6671.10	3.45	T	29	5949.65	3.65	473.9	15	5097.14	
			3	T	20	8396.78	22.37	T	21	7215.95	18.27	T	26	5292.58	1.01
	0.4	8	1	414.1	10	7731.58		133.6	9	7086.20		276.4	11	6475.30	
			2	T	18	11470.30	24.19	T	21	10205.30	19.71	T	23	7232.63	5.55
			3	T	17	17546.90	49.89	T	19	13976.10	47.42	T	20	10561.40	34.83
		10	1	117.6	8	6248.44		26.4	6	5434.73		59.6	8	4914.15	
			2	T	20	8153.59	9.54	T	28	7209.75	14.43	T	27	5552.00	1.79
			3	T	19	10833.00	30.05	T	23	9732.88	34.73	T	22	5969.43	9.35



(a) Benders decomposition method (b) C&CG method

Fig. 2. Convergence plots within 15 s.

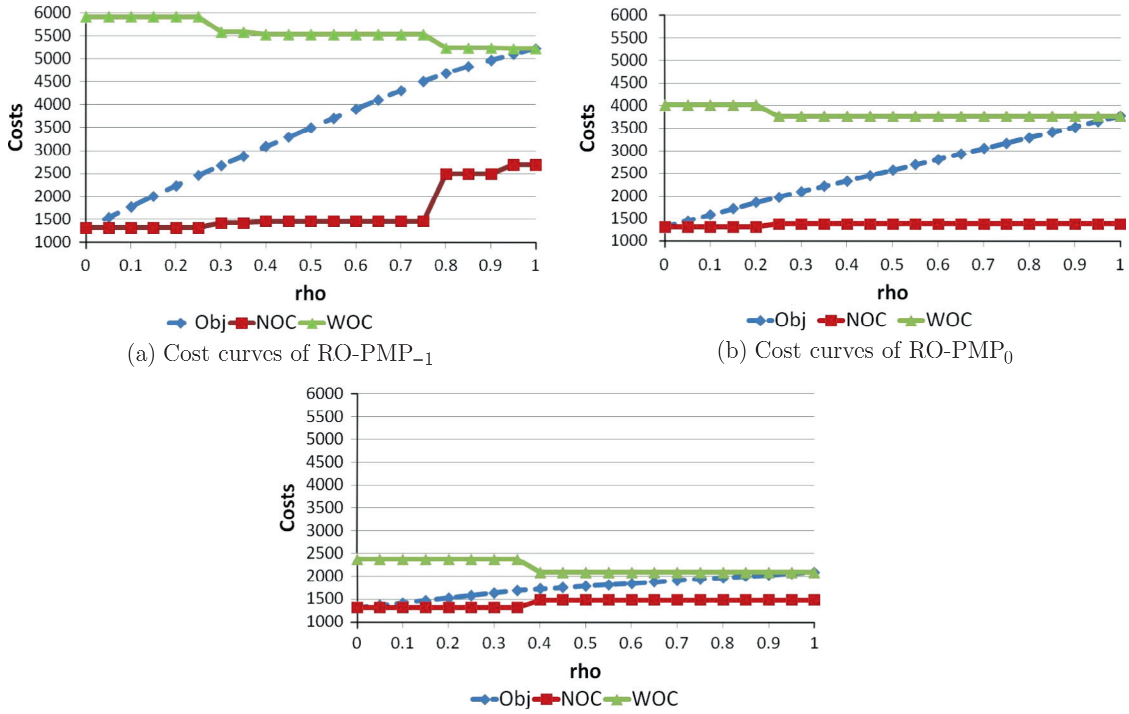
and the actual bottleneck is to solve MP, which will grow into a large MIP problem over iterations. As CPLEX, a general-purpose MIP solver, is currently called to solve MP, one possible direction of future research is to develop a specialized algorithm that takes advantage of the structure of MP for a faster computation.

(iii) Including additional features does not incur significant computational expense. Compared with models without capacity restrictions or demand changes, capacitated ones are slightly harder while models with demand changes could be easier. Hence, our two-stage RO formulations of reliable p-median problems are computationally robust to additional features or restrictions.

(iv) Although for many instances the optimal objective values are the same for the different M values, the large penalty coefficient M generally negatively impacts the computational performance, which is more significant for instances in capacitated models. One explanation is that large penalty coefficient M forces demand that was served by a disrupted facility to be served by survived ones, instead of being simply treated as unmet demand. As a result, the optimization complexity increases.

5.2. Impact of the reliability

In this section, we investigate the effect of including the worst disruptive scenarios on the system configuration and operations. Specifically, for different ρ values, after deriving an optimal solution, we compute the corresponding normal



(a) Cost curves of RO-PMP$_{-1}$

(b) Cost curves of RO-PMP$_0$

(c) Cost curves of RO-PMP$_1$

Fig. 3. Effect of ρ value on the robust uncapacitated facility location models.

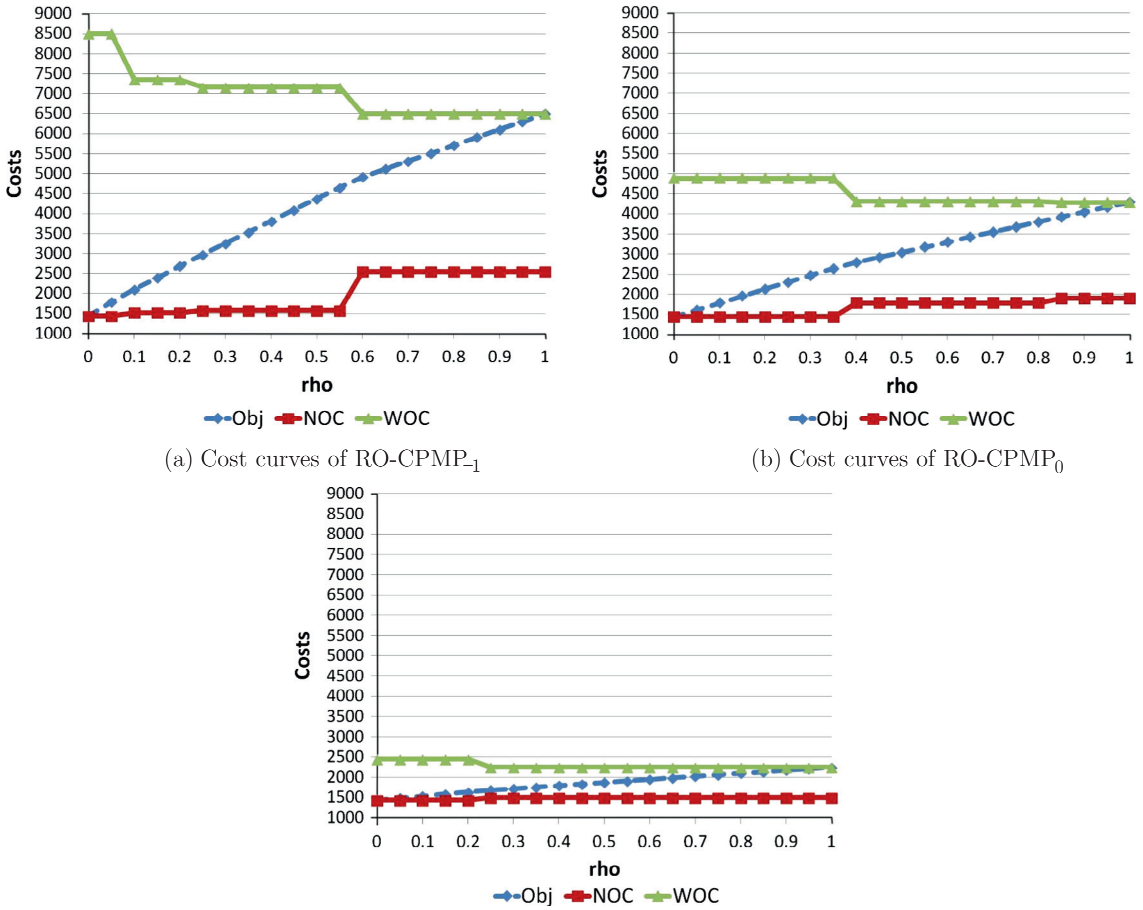
operation cost (*NOC*) and worst case operation cost (*WOC*). Then, we plot those operation costs with respect to ρ. It is obvious that the classical p-median model can be obtained by setting ρ to 0 and the formulation to minimize only the worst case cost can be obtained by setting ρ to 1. Figs. 3 and 4 present our results for 25-site models with $M = 15$. Note that the weighted objective function values are also included.

Clearly, the two cost functions demonstrate a monotone property, or a "staircase pattern", over ρ. Within a single stair the optimal system configurations are the same for different ρ values. Actually, the small number of stairs implies that the optimal system configuration based on two-stage RO is not very sensitive to different ρ. However, when ρ keeps increasing, *WOC* would decrease while *NOC* would increase. Sometimes a slight increase in *NOC* will lead to a significant decrease in *WOC*. Such a phenomenon is also observed in the stochastic programming based reliable facility location models (Snyder and Daskin, 2005). Overall, a desired trade-off between *NOC* and *WOC* can be achieved by selecting a configuration of an appropriate stair. Another observation is that the *WOC*s and hence the objective values of the models with demand changes are quite different from those of the classical models ($\theta = 0$). It is reasonable since, for example, in RO-PMP$_1$/RO-CPMP$_1$, under a disruptive scenario, the demands of disrupted clients will disappear and will not incur any cost, which counteracts the cost increase due to failed facilities.

5.3. Comparison of stochastic programming and RO models

As mentioned in Section 2, stochastic programming (SP) has been used as the primary tool to develop models and corresponding algorithms to study the reliable facility location problems. To evaluate the solution quality of two-stage RO models, we compare the solutions from our model RO-PMP$_0$ and the SP model presented in Snyder and Daskin (2005) using the same data set and penalty coefficient (i.e., the uncapacitated instance with $|I| = 25, M = 15$ and $\theta = 0$). For the RO model, we first set the weight coefficient $\rho = 0.2$ and $1 - \rho = 0.8$ for the worst case and normal situation operation costs, respectively. We set the failure probability of each site to 0.01 in the SP model. It renders the probability of the normal disruption-free situation to 0.78, which roughly matches the weight of normal situation in the RO model. We compare two measures, i.e., *NOC* and *WOC*, for solutions of both models to evaluate their performances. For the SP model, WOC^{SP} is obtained by inserting its optimal facility locations to the RO model. Numerical results are provided in Table 7. We also consider a more conservative situation in which ρ is set to 0.4 in RO.

We note in Table 7 that, when $\rho = 0.2$, the qualities of SP and RO solutions are basically close in normal situations. An SP solution might have an equal or a little bit less *NOC*, while its *WOC* could be significantly more, compared to an RO solution. When $\rho = 0.4$, because more consideration is placed on the worst case performance in RO, RO solutions might not be in favor



(a) Cost curves of RO-CPMP$_{-1}$

(b) Cost curves of RO-CPMP$_0$

(c) Cost curves of RO-CPMP$_1$

Fig. 4. Effect of ρ value on the robust capacitated facility location models.

Table 7
Comparison of SP and RO models.

ρ	p	k	SP		RO		$\Delta_{NOC}(\%) = \frac{NOC^{SP} - NOC^{RO}}{NOC^{RO}}$	$\Delta_{WOC}(\%) = \frac{WOC^{SP} - WOC^{RO}}{WOC^{RO}}$
			NOC^{SP}	WOC^{SP}	NOC^{RO}	WOC^{RO}		
0.2	8	2	1313.74	4022.6	1313.74	4022.6	0.00	0.00
		3	1313.74	6732.09	1417.77	5793.01	−7.34	16.21
	10	2	913.976	3214.53	913.976	3214.53	0.00	0.00
		3	913.976	4651.45	967.934	4137.91	−5.57	12.41
0.4	8	2	1313.74	4022.6	1380.83	3766.26	−4.86	6.81
		3	1313.74	6732.09	1546.41	5522.21	−15.05	21.91
	10	2	913.976	3214.53	981.016	2992.1	−6.83	7.43
		3	913.976	4651.45	967.934	4137.91	−5.57	12.41

of *NOC* while they have less *WOC*. Consequently, we observe more clearer differences in *NOC* and *WOC* between SP and RO solutions among all cases. Nevertheless, the difference in *NOC* for these two models is not drastic. Those results suggest that, SP and two-stage RO models are comparable, even when a relatively large weight is assigned to *WOC* in RO. Hence, our presented two-stage RO models are not overly conservative. Those results also indicate that, instead of relying on accurate probabilistic information to build an SP model, two-stage RO provides a dependable modeling alternative for practical usage that requires much less information support.

Table 8
Comparison between RO-PMP$_{-1}$, RO-PMP$_1$, and RO-PMP$_0$ ($|I| = 25$).

ρ	p	k	RO-PMP$_{-1}$ OL	RO-PMP$_1$ OL	RO-PMP$_0$ OL	$\Delta^{-1}_{NOC}(\%)$	$\Delta^{-1}_{WOC}(\%)$	$\Delta^{1}_{NOC}(\%)$	$\Delta^{1}_{WOC}(\%)$
0.2	8	1	0 1 2 3 5 8 11 13	0 1 2 3 5 8 11 13	0 1 2 3 5 8 11 13	0.00	0.00	0.00	0.00
		2	0 1 2 3 5 8 11 13	0 1 2 3 5 8 11 13	0 1 2 3 5 8 11 13	0.00	0.00	0.00	0.00
		3	0 1 3 5 8 11 13 23	0 1 2 3 4 5 7 11	0 1 2 3 4 5 11 13	5.56	9.54	−4.09	13.82
	10	1	0 1 2 3 4 5 8 11 13 16	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 8 11 13 16	0.00	0.00	−6.83	23.62
		2	0 1 3 5 6 7 8 10 11 14	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 8 11 13 16	−9.01	15.36	−6.83	25.35
		3	0 1 3 5 7 8 10 11 14 23	0 1 2 3 4 5 8 11 13 16	0 1 2 3 4 5 8 10 11 13	−4.53	12.49	5.90	9.69
0.4	8	1	0 1 2 3 5 8 11 13	0 1 2 3 5 8 11 13	0 1 2 3 5 8 11 13	0.00	0.00	0.00	0.00
		2	0 1 3 5 6 8 11 14	0 1 2 3 4 5 7 11	0 1 2 3 8 11 12 13	−5.35	8.19	−6.59	29.26
		3	0 1 8 9 11 12 14 23	0 1 2 3 4 5 7 11	0 1 3 5 6 10 11 17	−12.20	10.87	4.61	27.68
	10	1	0 1 2 3 4 5 8 11 13 16	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 8 11 13 16	0.00	0.00	−6.83	23.62
		2	0 1 3 5 7 8 10 11 14 23	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 7 8 11 14	−3.24	5.60	0.00	0.00
		3	0 1 3 7 8 10 11 12 14 23	0 1 2 3 4 5 7 8 11 16	0 1 2 3 4 5 8 10 11 13	−10.45	15.38	−0.67	20.50

Table 9
Comparison between RO-CPMP$_{-1}$, RO-CPMP$_1$, and RO-CPMP$_0$ models ($|I| = 25$).

ρ	p	k	RO-CPMP$_{-1}$ OL	RO-CPMP$_1$ OL	RO-CPMP$_0$ OL	$\Delta^{-1}_{NOC}(\%)$	$\Delta^{-1}_{WOC}(\%)$	$\Delta^{1}_{NOC}(\%)$	$\Delta^{1}_{WOC}(\%)$
0.2	8	1	0 1 3 4 5 8 11 13	0 1 2 3 4 5 11 13	0 1 3 4 5 8 11 13	0.00	0.00	5.72	7.73
		2	0 1 3 4 5 8 11 13	0 1 2 3 4 5 11 13	0 1 2 3 4 5 11 13	−5.41	15.53	0.00	0.00
		3	0 1 3 4 8 11 12 13	0 1 2 3 4 5 11 13	0 1 2 3 4 5 11 13	−8.73	15.37	0.00	0.00
	10	1	0 1 2 3 4 5 8 11 13 16	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 8 11 13 16	0.00	0.00	−6.77	22.18
		2	0 1 2 3 4 5 8 10 11 13	0 1 2 3 4 5 8 11 13 16	0 1 2 3 4 5 8 11 13 16	−5.52	9.77	0.00	0.00
		3	0 1 2 3 4 5 8 10 11 13	0 1 2 3 4 5 8 11 13 16	0 1 2 3 4 5 8 10 11 13	0.00	0.00	5.84	−1.14
0.4	8	1	0 1 3 4 5 8 11 13	0 1 2 3 4 5 7 11	0 1 3 4 8 11 12 13	3.64	1.30	5.14	28.70
		2	0 1 3 4 8 11 12 13	0 1 2 3 4 5 7 11	0 1 2 3 4 5 8 17	13.15	1.62	18.97	34.97
		3	0 1 4 8 9 11 12 13	0 1 2 3 4 5 11 13	0 1 2 3 4 5 11 13	−19.30	23.20	0.00	0.00
	10	1	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 7 8 11 14	0 1 2 3 4 5 8 11 13 16	−6.77	6.05	−6.77	22.19
		2	0 1 2 3 4 5 8 10 11 13	0 1 2 3 4 5 7 8 11 16	0 1 2 3 4 5 8 11 13 16	−5.52	9.77	−6.15	12.30
		3	0 1 2 3 4 5 8 10 11 13	0 1 2 3 4 5 7 8 10 11	0 1 2 3 4 5 8 10 11 13	0.00	0.00	−5.83	11.88

5.4. Effect of demand changes

As we mentioned earlier, demand changes due to disruptions have not been included or investigated in any existing reliable facility location models. So, it remains unknown that how demand changes will affect system design and operations, or how approximate the results we have if we ignore the demand change factor when it does exist. To explore the impact of demand changes, Tables 8 and 9 present optimal configurations of the models with $\theta = -1, 1$, and 0 (M = 15). The column *OL* represents the optimal locations of facilities. We then insert the optimal locations derived from $RO - PMP_0/RO - CPMP_0$ into $RO - PMP_{-1}$ and $RO - PMP_1/RO - CPMP_{-1}$ and $RO - CPMP_1$, and compute the associated operation costs in normal and the worst disruptive situations. The relative changes of *NOC*s and *WOC*s for RO-PMP$_{-1}$ model with respect to its own optimal *NOC*s and *WOC*s are presented in the column $\Delta^{-1}_{NOC}(\%)$ and $\Delta^{-1}_{WOC}(\%)$, respectively. Similar results for RO-PMP$_1$ are in columns $\Delta^{1}_{NOC}(\%)$ and $\Delta^{1}_{WOC}(\%)$. For a case, if the optimal facility locations are different among the three models, we will highlight the locations by underlines.

We observe that the demand change plays a significant role in determining system configuration. In all 24 instances, including uncapacitated and capacitated ones, there are 21 instances whose optimal facility locations are different for different θ values. In fact, when we put more weight on the worst disruptive situations or consider facility capacities, the impact of demand changes becomes more significant. For example, when $\rho = 0.4$, there are 11 instances (out of 12) on which those models yield different solutions. Furthermore, the different facility locations present different performances in both normal and the worst disruptive situations. From the columns $\Delta^{-1}_{NOC}(\%)$ and $\Delta^{1}_{NOC}(\%)$, we note that the system configuration derived without demand changes could incur more or less cost in the normal situation, which can hardly be predicted beforehand. In fact, the difference can be as significant as -19.30% or 18.97%, definitely a nontrivial value. Nevertheless, in the worst disruptive situation, it is generally observed that the system configuration derived without demand changes will incur much more operation cost, which can be up to 34.97% in a capacitated instance. Therefore, we can conclude that the demand change factor, if it exists in the practice, should not be ignored in system design, especially when the weight coefficient ρ is large and capacity needs to be considered.

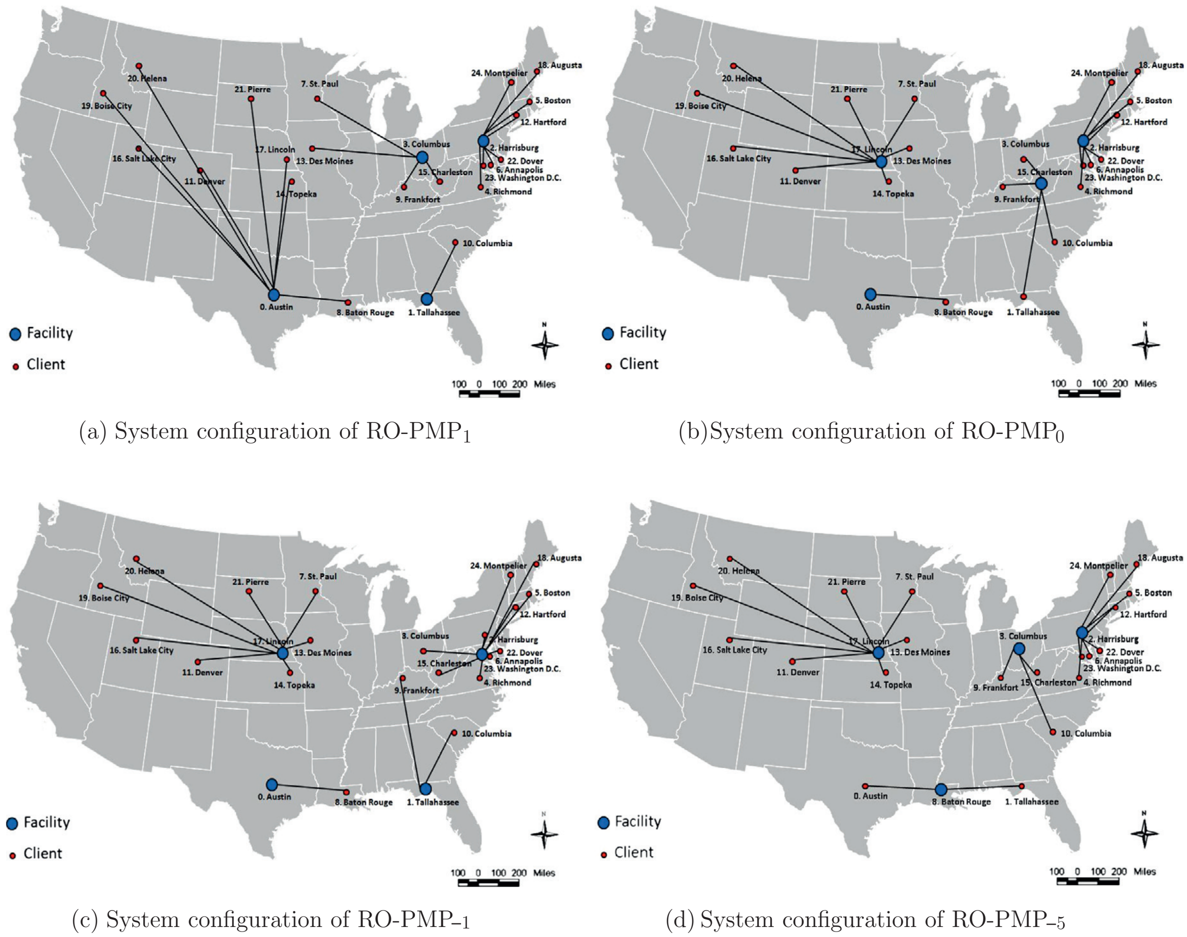


Fig. 5. Comparison of optimal configurations of uncapacitated models with different θ values.

To develop insights on system configuration, we plot optimal facility locations and normal situation assignments of a small-scale instance in Fig. 5, where $p = 4, k = 1$ and $|I| = 25$ with $\rho = 0.2$ and $M = \max_{i,j}\{c_{ij}\}$. We also include a relatively extreme case where $\theta = -5$, which indicates a four-time demand increase of a disrupted site.

We mention that sites are numbered according to the descending order of demand quantities. In Fig. 5, we observe a clear trend on selecting facility locations. When $\theta = 1$, four facilities are constructed on the sites of the smallest indices, i.e., those with largest demands. With θ becoming smaller, facilities are generally built on sites with smaller demands. One explanation is that: on the one hand, with θ becomes smaller, more demands should be served in disruptive situations; on the other hand, if sites with large demands are selected for facilities, although it saves cost in the normal situation, disruptions on those facility sites will incur very high service costs as demands will be served by facilities far away. Therefore, if demands increase under disruptions, to balance the cost and the risk, facilities should be built on sites with smaller demands to avoid the large cost in the worst scenarios.

5.5. A correlated disruption set

The disruption set with a simple cardinality restriction in Eq. (1) indicates that all sites are of the identical failure possibility and there is little correlation among them. In this section, given the adaptability of our modeling framework, we investigate a different disruption set that carries some correlations. Specifically, we partition J into a few subsets and assume that sites in each subset are temporally or spatially correlated. Hence, the number of disruptions in each subset can be better estimated. Also, we have an overall budget constraint to restrict the total number of disruptions. The disruption set takes the following form:

$$A_1 = \left\{ \mathbf{z} \in \{0,1\}^{|J|} : \sum_{j \in J_1} z_j \leqslant k_1, \sum_{j \in J_2} z_j \leqslant k_2, \ldots, \sum_{j \in J_L} z_j \leqslant k_L, \sum_{l=1}^{L} \sum_{j \in J_l} a_j^l z_j \leqslant b \right\}. \tag{33}$$

Table 10
Computational results for uncapacitated instances with a correlated uncertainty set.

$\|I\|$	ρ	p	b	RO-PMP$_{-1}$				RO-PMP$_0$				RO-PMP$_1$			
				Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	0.2	8	15	1.1	2	1763.95		1.5	5	1558.09		2.4	5	1426.76	
			30	0.9	3	2233.97		3.6	7	1820.23		4.5	7	1525.92	
			40	9.8	11	2455.70		19.5	15	1989.57		6.4	10	1592.05	
		10	15	1.6	2	1364.19		0.4	2	1139.08		2.7	7	1024.11	
			30	2.7	6	1759.77		2.9	6	1374.09		4.8	8	1060.02	
			40	17.1	14	2003.68		11.8	12	1490.87		5.5	9	1088.26	
	0.4	8	15	0.8	2	2214.17		1.7	5	1802.45		3.4	7	1539.77	
			30	1.5	4	3086.90		12.2	12	2306.74		15.6	13	1674.14	
			40	22.6	16	3352.64		27.1	16	2458.94		18.9	14	1724.86	
		10	15	0.4	2	1814.40		0.6	2	1364.19		3.2	7	1067.20	
			30	2.5	5	2503.45		4.8	7	1785.45		6.7	9	1139.02	
			40	11.6	12	2883.19		43.4	19	2055.38		27.3	16	1221.07	
49	0.2	8	15	4.7	4	6563.12		3.1	4	6145.54		1.3	3	5684.45	
			30	73.5	11	7325.44		30.6	8	6631.47		203.4	16	6007.48	
			40	19.8	7	7447.91		10.6	5	6643.61		165.4	14	6046.56	
		10	15	2.1	3	5339.88		0.6	2	4950.32		9.6	6	4687.15	
			30	38.5	9	5865.65		8.4	5	5307.93		46.7	10	4887.14	
			40	266.4	15	6251.45		152.8	13	5559.63		995.9	23	5033.25	
	0.4	8	15	44.9	9	7405.65		4.9	4	6625.50		14.2	7	5898.97	
			30	570.3	17	8649.65		125.4	13	7393.31		972.1	21	6284.88	
			40	177.4	13	8954.28		116.1	11	7393.31		1064.6	20	6284.88	
		10	15	7.9	5	5988.67		3.2	4	5351.47		11.7	7	4796.70	
			30	1620.0	20	7158.15		68.8	11	5947.01		1302.4	22	5200.39	
			40	94.1	12	7575.79		81.2	11	6192.15		T	34	5357.96	1.38

Table 11
Computational results for capacitated instances with a correlated uncertainty set.

$\|I\|$	ρ	p	b	RO-CPMP$_{-1}$				RO-CPMP$_0$				RO-CPMP$_1$			
				Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)	Time (s)	Iter	Obj	Gap (%)
25	0.2	8	15	1.3	2	2069.34		2.6	3	1783.39		2.3	4	1550.58	
			30	1.2	3	2686.70		3.7	5	2126.51		4.6	6	1597.10	
			40	9.8	8	3253.49		11.9	8	2331.96		6.9	7	1626.47	
		10	15	0.7	2	1474.23		1.9	2	1148.48		2.3	4	1036.93	
			30	2.5	2	1923.53		1.7	3	1383.48		4.7	6	1084.08	
			40	13.1	9	2259.29		15.3	10	1612.96		5.6	6	1119.55	
	0.4	8	15	1.5	2	2620.20		1.4	3	2024.04		3.8	5	1623.29	
			30	5.7	5	3809.90		32.6	12	2791.94		15.2	11	1737.18	
			40	54.6	14	4670.89		11.6	7	2895.55		14.1	10	1791.85	
		10	15	1.5	2	1981.33		0.8	2	1373.58		4.4	6	1083.45	
			30	1.5	2	2869.74		1.8	3	1843.59		20.7	12	1224.17	
			40	18.4	11	3398.16		58.4	17	2211.98		23.1	12	1291.03	
49	0.2	8	15	21.6	5	6946.01		7.6	4	6432.77		14.7	5	6180.25	
			30	7126.0	23	7919.65		T	27	7062.09	1.40	723.5	15	6470.81	
			40	T	20	8214.50	0.16	452.5	11	7062.09		572.3	15	6470.81	
		10	15	11.8	4	5616.50		7.4	4	5045.10		10.4	5	4778.10	
			30	757.7	14	6290.50		224.6	12	5466.82		295.4	14	5064.34	
			40	T	21	6775.00	2.08	T	21	5796.18	1.22	5917.2	26	5215.26	
	0.4	8	15	172.6	9	7730.58		27.3	6	6939.44		87.6	9	6426.95	
			30	T	20	9585.18	12.30	T	25	8255.52	10.02	1373.8	16	6791.59	
			40	T	17	10139.70	7.01	7037.1	22	7948.09		1961.9	17	6791.59	
		10	15	115.4	8	6248.44		21.6	6	5434.73		21.0	7	4881.02	
			30	T	18	7881.48	5.07	7141.2	23	6276.43		T	24	5466.90	1.41
			40	T	19	8813.58	9.55	T	20	6696.92	5.93	T	23	5519.86	1.84

In our numerical study, $L = 2$, sites are randomly assigned to J_1 and J_2, $k_1 = 2$ and $k_2 = 1$, a_j^1 takes the value of 10 for all $j \in J_1$ and $a_j^2 = 15$ for all $j \in J_2$. Experiments are performed with $n = 25, 49$, $\rho = 0.2, 0.4$, $p = 8, 10$, $b = 15, 30, 40$, and $M = 15$. We mention that it is rather a simple set just for the demonstration of the impact of correlation. The computational performance of C&CG algorithm is presented in Tables 10 and 11.

Note that with $b = 15, 30$, and 40 in A_1, the number of disruptions over the entire site set can be 1, 2 and 3 respectively, which resembles the set we study in (1) with $k = 1, 2$ and 3. Nevertheless, comparing Table 10 with Tables 3 and 11 with Table 5, we observe that: (i) the algorithm performance is generally better for the correlated set with less iterations; (ii) the objective function value, i.e., the weighted operation cost, is often smaller. Both points can be explained by the fact that the disruption set in (33) is a tighter and more precise description of all kinds of disruptive scenarios, if correlations exist; and A_1 is a subset of that defined in (1) with an appropriate k. So, with more structural information available, both the master and subproblems are easier or smaller to compute than those with the cardinality set (1). Clearly, the cardinality set (1) is an overestimation of disruptions if disruptions actually happen according to the pattern defined in (33). Hence, in terms of system design and operations, the cardinality set (1) could lead to a more conservative system configuration with a higher objective function value as it is overprotective towards some *unrealistic* disruptive situations. So, with a proper description of the correlation, a less conservative system design can be achieved with the desired reliability level.

6. Conclusion

In this paper, we propose two-stage robust optimization based models for the reliable p-median facility location problem. In particular, we demonstrate the strong modeling capability of two-stage robust optimization framework by considering two practical features, i.e., facility capacity and demand change due to site disruption, in a compact fashion, which otherwise would demand for large-scale scenario-based stochastic programming formulations and have received little attention. We study those models, and customize and implement exact computing algorithms, i.e., the column-and-constraint generation and Bender decomposition methods, to solve them. From a set of computational experiments, we note that (i) the column-and-constraint generation algorithm drastically outperforms the other method. Instances with up to 49 sites, including those with demand changes and capacities, can be solved exactly or with a reasonable gap; (ii) by assigning an appropriate weight to the operation cost from the worst disruptive situations in the objective function, a considerable decrease of that cost could be achieved by a small increase in the operation cost in the normal situation; (iii) demand changes due to disruptions should not be ignored in system design. Otherwise, a different network configuration could result in a huge increase of the operation cost in the worst disruptive situation; and (iv) a description of disruption correlations, even in a simple form, makes those models less challenging and leads to less conservative system designs with the desired reliability.

A clear direction of the future research is to explore the problem structure to enhance the column-and-constraint generation algorithm for larger scale instances. So, a customized procedure can be developed to replace the professional solver for a better performance. Moreover, in real practice, one disruption can cause a complicated demand pattern. It could vary at different phases of the disruption and is also highly dependent on population and economic conditions of different locations (Ergun et al., 2010). Hence, another interesting direction is to extend the recourse problem to a multi-period formulation so that it can capture dynamic and site-specific demand variations.

Acknowledgments

This research was partially supported by National Science Foundation through Grant CMMI-1235135. The authors are grateful to anonymous referees, editors, and Professor Hai Yang for their constructive comments and suggestions to improve the quality and clarity of the paper.

References

An, Y., Zhang, Y., Zeng, B., 2011. The Reliable Hub-and-spoke Design Problem: Models and Algorithms. Tech. Rep., under reivew, University of South Florida.
Atamturk, A., Zhang, M., 2007. Two-stage robust network flow and design under demand uncertainty. Operations Research 55 (4), 662–673.
Barahona, F., Jensen, D., 1998. Plant location with minimum inventory. Mathematical Programming 83 (1), 101–111.
Baron, O., Milner, J., Naseraldin, H., 2011. Facility location: a robust optimization approach. Production and Operations Management 20 (5), 772–785.
Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A., 2004. Adjustable robust solutions of uncertain linear programs. Mathematical Programming 99 (2), 351–376.
Ben-Tal, A., El Ghaoui, L., Nemirovski, A., 2009. Robust Optimization. Princeton Univ Press.
Berman, O., Krass, D., Menezes, M., 2007. Facility reliability issues in network p-median problems: strategic centralization and co-location effects. Operations Research 55 (2), 332–350.
Bertsimas, D., Sim, M., 2004. The price of robustness. Operations Research 52 (1), 35–53.
Bertsimas, D., Brown, D., Caramanis, C., 2011. Theory and applications of robust optimization. SIAM Review 53 (3), 464–501.
Bertsimas, D., Litvinov, E., Sun, X., Zhao, J., Zheng, T., 2013. Adaptive robust optimization for the security constrained unit commitment problem. IEEE Transactions on Power Systems 28 (1), 52–63.
Chen, Q., Li, X., Ouyang, Y., 2011. Joint inventory–location problem under the risk of probabilistic facility disruptions. Transportation Research Part B 45 (7), 991–1003.
Contreras, I., Cordeau, J.-F., Laporte, G., 2011. The dynamic uncapacitated hub location problem. Transportation Science 45 (1), 18–32.
Cui, T., Ouyang, Y., Shen, Z., 2010. Reliable facility location design under the risk of disruptions. Operations Research 58 (4), 998–1011.
Daskin, M., 1995. Network and Discrete Location: Models, Algorithms, and Applications. Wiley-Interscience.
Drezner, Z., 1987. Heuristic solution methods for two location problems with unreliable facilities. Journal of the Operational Research Society 38 (6), 509–514.
Drezner, E., 1995. Facility Location: A Survey of Applications and Methods. Springer.
Ergun, O., Karakus, G., Keskinocak, P., Swann, J., Villarreal, M., 2010. Operations research to improve disaster supply chain management. Wiley Encyclopedia of Operations Research and Management Science. http://dx.doi.org/10.1002/9780470400531.eorms0604.

Gabrel, V., Lacroix, M., Murat, C., Remli, N., 2014. Robust location transportation problems under uncertain demands. Discrete Applied Mathematics 164 (1), 100–111.

Gülpınar, N., Pachamanova, D., Çanakoğlu, E., 2012. Robust strategies for facility location under uncertainty. European Journal of Operational Research 225 (1), 21–35.

Hervet, C., Faye, A., Costa, M.-C., Chardy, M., Francfort, S., 2013. Solving the two-stage robust FTTH network design problem under demand uncertainty. Electronic Notes in Discrete Mathematics 41, 335–342.

Jia, H., Ordóñez, F., Dessouky, M., 2007. A modeling framework for facility location of medical services for large-scale emergencies. IIE Transactions 39 (1), 41–55.

Jiang, R., Zhang, M., Li, G., Guan, Y., 2011. Benders Decomposition for the Two-stage Security Constrained Robust Unit Commitment Problem. Tech. Rep., University of Florida.

Li, X., Ouyang, Y., 2010. A continuum approximation approach to reliable facility location design under correlated probabilistic disruptions. Transportation Research Part B 44 (4), 535–548.

Li, Q., Zeng, B., Savachkin, A., 2013. Reliable facility location design under disruptions. Computers and Operations Research 40 (4), 901–909.

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

Lim, M.K., Bassamboo, A., Chopra, S., Daskin, M.S., 2013. Facility location decisions with random disruptions and imperfect estimation. Manufacturing and Service Operations Management 15 (2), 239–249.

Melo, M., Nickel, S., Saldanha-Da-Gama, F., 2009. Facility location and supply chain management – a review. European Journal of Operational Research 196 (2), 401–412.

Peng, P., Snyder, L., Lim, A., Liu, Z., 2011. Reliable logistics networks design with facility disruptions. Transportation Research Part B 45 (8), 1190–1211.

Revelle, C., Eiselt, H., Daskin, M., 2008. A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research 184 (3), 817–848.

Shen, Z., Zhan, R., Zhang, J., 2011. The reliable facility location problem: formulations, heuristics, and approximation algorithms. INFORMS Journal on Computing 23 (3), 470–482.

Smith, J.C., Lim, C., Sudargho, F., 2007. Survivable network design under optimal and heuristic interdiction scenarios. Journal of Global Optimization 38 (2), 181–199.

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

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

Snyder, L., Atan, Z., Peng, P., Rong, Y., Schmitt, A., Sinsoysal, B., 2012. OR/MS Models for Supply Chain Disruptions: A Review (2012). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1689882> (accessed 01.19.13).

Sridharan, R., 1995. The capacitated plant location problem. European Journal of Operational Research 87 (2), 203–213.

Teo, C., Shu, J., 2004. Warehouse–retailer network design problem. Operations Research 52 (3), 396–408.

Wang, X., Ouyang, Y., 2013. A continuum approximation approach to competitive facility location design under facility disruption risks. Transportation Research Part B 50, 90–103.

Zeng, B., Zhao, L., 2013. Solving two-stage robust optimization problems using a column-and-constraint generation method. Operations Research Letters 41 (5), 457–461.

Zugno, M., Conejo, A.J., 2013. A Robust Optimization Approach to Energy and Reserve Dispatch in Electricity Markets. Tech. Rep., Technical University of Denmark.