



Nested-solution facility location models

Ronald G. McGarvey, Andreas Thorsen

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1007/s11590-021-01759-4>

Nested-Solution Facility Location Models

Ronald G. McGarvey · Andreas Thorsen

Received: date / Accepted: date

Abstract Classical facility location models can generate solutions that do not maintain consistency in the set of utilized facilities as the number of utilized facilities is varied. We introduce the concept of nested facility locations, in which the solution utilizing p facilities is a subset of the solution utilizing q facilities, for all $i \leq p < q \leq j$, given some lower limit i and upper limit j on r , the number of facilities that will be utilized in the future. This approach is demonstrated with application to the p -median model, with computational testing showing these new models achieve reductions in both average regret and worst-case regret when $r \neq p$ facilities are actually utilized.

Keywords Facility location · Optimization · Practicality of model solutions

1 Introduction

Selecting the locations for a set of facilities to be opened in order to service a set of customers is one of the most widely studied and applied problems in all of operations research. Since the foundational work of Hakimi in [10], [11], thousands of authors have published works examining facility location problems, with [10] currently cited by more than 3,000 papers, according to Google Scholar.

However, as with any model, when an operations research analyst attempts to use results from a facility location problem formulation to influence decision making, the recommendations must be presented in a manner that is

R.G. McGarvey
Department of Industrial and Manufacturing Systems Engineering, and Truman School of Public Affairs, E3437 Laffer Hall, University of Missouri, Columbia, MO, 65211, USA
Tel.: +1 573-882-9564
E-mail: mcgarveyr@missouri.edu

A. Thorsen
Jake Jabs College of Business and Entrepreneurship, Montana State University, P.O. Box 173040, Bozeman, MT, 59717-3040, USA

comprehensible to the decision maker. One aspect of the solutions obtained by traditional facility location models that can be difficult to communicate to decision makers is a potential lack of consistency in the set of utilized facilities as the number of utilized facilities is varied. For example, consider a stylized problem in which each demand point is assigned service at its nearest opened facility, where the objective is to minimize the sum of the distances between each demand point and its assigned facility. Assume that demand exists at five points located along a unit line segment at locations 0, 0.25, 0.5, 0.75 and 1. Suppose that one new facility could be placed at any of these demand points; in this case it is apparent that the optimal solution is to place the facility at the line segment's midpoint. If two facilities were to be utilized, there are alternative optimal solutions (e.g., place the facilities at 0 and 0.75, or place the facilities at 0.25 and 0.75), but there is no optimal solution that places a facility at 0.5, demonstrating an inconsistency in the model's recommendations as the number of facilities increases from one to two.

The ability of the analyst to "sell" the modeling solution can be hindered by the extent to which this inconsistency property, while demonstrably valid from a mathematical perspective, is nevertheless at odds with a decision maker's intuition. The challenge imposed by this lack of consistency on the implementation of model results is most pronounced when the decision maker is uncertain about the number of facilities to be utilized in the future. In response, we introduce the concept of *nested* facility locations, defined as follows. A set of facility location solutions are said to be nested if the solution utilizing p facilities is a subset of the solution utilizing q facilities, for all $i \leq p < q \leq j$, given some lower limit i and upper limit j on r , the number of facilities that will be utilized in the future. Observe that such a set of nested solutions cannot exhibit the inconsistency discussed above. **The nested concept is consistent with a situation in which an initial decision is made to prepare a set of j locations for potential future use, the decision maker then engages with various constituencies to discuss the $j - i + 1$ alternative solutions (and does not wish to demonstrate this inconsistency during those discussions), and finally the actual number of facilities to be opened is then determined.**

These concepts, applicable to situations in which the future number of facilities is uncertain, have parallels to the broader class of models for optimization under uncertainty. More recently, a body of research has arisen examining linkages between combinatorial optimization (such as the facility location models of [10]) and machine learning methods, e.g., [12], which explicitly considered facility location problems as an application area in which machine learning methods could be used to address uncertainties in input parameters.

Previous works have considered facility location models when the number of opened facilities is uncertain. The p -median problem with an uncertain number of future facilities was first examined by [7]. These authors assumed that an initial set of p facilities must be opened, and that it will subsequently become necessary to open an additional set of between 0 and q facilities, assuming a minimax regret objective. While their model ensures that the initial p facilities selected are included in every future set of $p, p + 1, \dots, p + q$ facilities, these

authors do not enforce that the optimal set of $p + i$ facilities are included in the optimal set of $p + j$ facilities where $0 < i < j \leq q$; that is, the solutions are not nested. Later, [5] considered this same problem, minimizing the expected regret, with a slightly modified mathematical representation, and provided heuristic solution methods that were found to be effective for large problem instances, albeit with a limitation that $q = 1$. In [16], the authors considered a problem in which both the number of future facilities is uncertain and future demand is uncertain, identifying a set of initial facilities to open and a set of future relocations that include potentially opening new facilities and closing facilities opened in the initial time period. Clearly, as this formulation allows for facilities to be closed, the nested property is not enforced here.

The main contributions of this paper are the introduction of the concept of nested facility locations and the formulation of the nested p -median problem using two objective functions, minimizing expected regret and minimizing maximum regret. The use of a regret-based objective function is common in optimization problems (and facility location problems specifically) in which the underlying data are assumed to be uncertain and the decisions made at the outset constrain the recourse decisions available to the decision maker in the future; see e.g., [1], [7], [13], [14]. We develop two heuristic approaches, based on Lagrangian relaxation, to efficiently solve the model under each objective for large problem instances. Our computational results demonstrate both the novelty of the model compared to a similar approach by [7], and also the improved performance of our approach over a classic p -median model under an uncertain number of facilities.

2 Mathematical Model

Consider the following p -median problem in which a set of facilities are opened from a set of candidate facility locations, and each member of a set of demand points is assigned to receive service from a single opened facility, such that the total distance between demand points and their assigned facilities is minimized.

Sets

- I : set of demand points
- J : set of candidate facility locations

Decision variables

- x_j : equals one if facility j is opened, equals zero otherwise
- y_{ij} : equals one demand point i is assigned to be served by facility j , equals zero otherwise
- z : total distance between demand points and the facilities to which they are assigned

Data parameters

- ϕ : number of facilities to be opened

- γ_{ij} : distance between demand point i and facility j

$$\min z = \sum_i \sum_j \gamma_{ij} y_{ij} \quad (1)$$

$$\sum_j x_j = \phi \quad (2)$$

$$\sum_j y_{ij} = 1 \quad \forall i \quad (3)$$

$$y_{ij} \leq x_j \quad \forall i, \forall j \quad (4)$$

$$x_j \in \{0, 1\} \quad \forall j \quad (5)$$

$$y_{ij} \geq 0 \quad \forall i, \forall j \quad (6)$$

Constraint (6) is sufficient to limit variable y_{ij} to take only binary values, based on the model structure [6]. Suppose this p -median model (1)–(6) were solved across a range of values for ϕ , denoted as follows:

Sets

- H : set of potential numbers of facilities to be opened

Data parameters

- $\hat{\phi}_h$: number of facilities to be opened under set element $h \in H$, ordered such that $\hat{\phi}_{h+1} = \hat{\phi}_h + 1 \quad \forall h$

Let $\hat{\alpha}_h$ denote the optimal objective function value obtained when this p -median model (1)–(6) is solved with $\hat{\phi}_h$ facilities to be opened.

The nested p -median problem can then be formulated as follows:

Decision variables

- \hat{x}_{jh} : equals one if facility j is opened when $\hat{\phi}_h$ facilities are opened, equals zero otherwise
- \hat{y}_{ijh} : equals one if demand point i is assigned to be served by facility j when $\hat{\phi}_h$ facilities are opened, equals zero otherwise
- \hat{z}_h : total distance between demand points and the facilities to which they are assigned when $\hat{\phi}_h$ facilities are opened
- \hat{w} maximum relative penalty for the nested solution objective function, across all elements of set H

$$\min \sum_h \hat{z}_h \quad (7)$$

$$\min \hat{w} \quad (8)$$

$$\sum_j \hat{x}_{jh} = \hat{\phi}_h \quad \forall h \quad (9)$$

$$\sum_j \hat{y}_{ijh} = 1 \quad \forall i, \forall h \quad (10)$$

$$\hat{y}_{ijh} \leq \hat{x}_{jh} \quad \forall i, \forall j, \forall h \quad (11)$$

$$\hat{x}_{jh} \in \{0, 1\} \quad \forall j, \forall h \quad (12)$$

$$\hat{y}_{ijh} \geq 0 \quad \forall i, \forall j, \forall h \quad (13)$$

$$\hat{z}_h = \sum_i \sum_j \gamma_{ij} \hat{y}_{ijh} \quad \forall h \quad (14)$$

$$\hat{x}_{jh} \geq \hat{x}_{j,(h-1)} \quad \forall j, \forall h > 1 \quad (15)$$

$$\hat{w} \geq \frac{\hat{z}_h - \hat{\alpha}_h}{\hat{\alpha}_h} \quad \forall h \quad (16)$$

Minimizing expected regret is equivalent to (for the p -median problem) minimizing expected distance across the elements of set H [13]. Therefore, model (7), (9)–(15), denoted *NestedPmedExpectedRegret (NPER)*, minimizes the expected regret of the nested model. Alternatively, formulation (8)–(16), denoted *NestedPmedMinimaxRegret (NPMR)* solves a minimax relative regret objective for the nested model.

The p -median problem is known to be NP-hard [9]. For NPER, if $|H| = 1$, NPER reduces to the classical p -median problem, and thus NPER is also NP-hard [14]. Because NPMR is minimax relative regret model for the p -median, NPMR is also NP-hard [1].

3 Initial Computational Results: Illustrating Regret

In this section, we conduct an initial numerical study to demonstrate the regret calculations, evaluating the performance of the NPER and NPMR models using a 30-instance set of benchmark problems from the online repository [15]. Each instance defines an ordered list of 100 locations and provides a matrix of the distances between each pair of locations. For our computational testing, we assume for each instance that the first 20 locations in the ordered list are candidate facility locations and that the other 80 locations are demand points. Assume an uncertain number of facilities are to be opened, ranging from a minimum of one facility to a maximum of four facilities. Models NPER and NPMR were both coded in GAMS 33.2 and solved using CPLEX 12.10.

3.1 Contrasting results with NOFUN

To help illustrate the difference between these nested models and the “Number Of Facility Locations Uncertain” (NOFUN) formulation of [7], consider one instance (number 27) from this test problem set. Models NPER and NOFUN (with an expected regret objective function) were each solved for this instance. Figure 1 presents an illustration of this test problem, graphed in the (x,y) -plane. The text boxes appearing inside the graph indicate for which values of ϕ_h a facility is opened at the corresponding location. For example, the black box located nearest to the lower-left corner of the figure indicates a facility location that was utilized in the NOFUN solution for $\hat{\phi}_h = 4$, but which was not opened for any other value of $\hat{\phi}_h$. Observe that the NPER solution utilizes a total of four locations, as required by the nested property. The NOFUN solution, however, utilizes a total of six different locations. Both model solutions are identical for $\hat{\phi}_h = 1$ and $\hat{\phi}_h = 2$. Moreover, both models exhibit the nested property for values of $\hat{\phi}_h$ between one and three (although the third facility opened is at a different location for each model). However, observe that the NOFUN solution for $\hat{\phi}_h = 4$ is a significant departure from the NOFUN solution for $\hat{\phi}_h = 3$, with only one location in common between the two solutions.

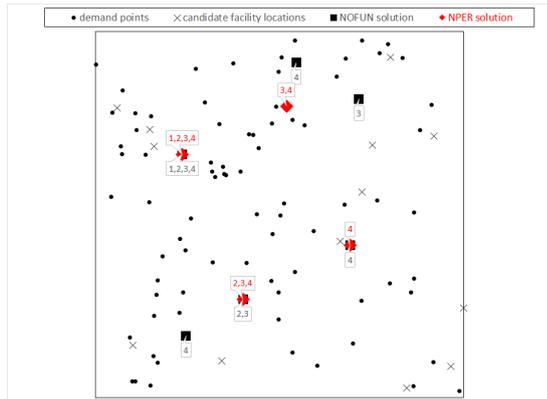


Fig. 1 Facility locations opened in NPER and NOFUN solutions, problem instance 27

3.2 Calculating regret

As an illustration of how we calculate regret, consider again instance number 27 from the test problem data set. Suppose that the p -median model (denoted PMED) were solved for some value of ϕ . This PMED model solution could then be utilized as the basis for a set of nested facility opening recommendations. For example, suppose that the 2-median solution for a particular test instance

opened facilities at locations j_1 and j_2 . Model NPER could then be run, forcing these two locations to define the solution for $\hat{\phi}_h = 2$, thereby generating an optimized nested facility solution based on this 2-median solution; denote the solutions obtained from this two-step approach as *Variations on the p -median solution* ($VpMS$). This $VpMS$ presents the best possible performance that could be achieved, were a p -median solution implemented, but it later become necessary to add or remove facilities from this solution.

Figure 2 illustrates this test problem instance, with the text boxes appearing inside the graph again indicating when a facility is opened at the corresponding location. Here, the PMED solutions are presented for $\phi \in \{1, 2, 3\}$, and are denoted by black boxes. The V2MS, denoted by red diamonds, is also presented for $\hat{\phi}_h \in \{1, 2, 3\}$. Observe that the V2MS presents a nested solution that is identical to the PMED solution for $\phi = 2$ (by design).

Consider now the different solutions corresponding to one opened facility. Due to the nested property, the V2MS is limited to selecting one of the two facilities utilized in the 2-median solution; this 1-facility solution from the V2MS has objective function value equal to 235,518. Contrast this to the optimal 1-median solution, which has objective function value ($\hat{\alpha}_1$) equal to 218,369, and utilizes a facility that does not appear in the 2-median solution. Here, the V2MS has a relative regret equal to 7.9% when one facility is opened. Similarly, the optimal 3-median solution utilizes a facility that does not appear in either the 1-median or 2-median solutions, and is thus able to outperform the nested 3-facility solution obtained from the V2MS. Table 1 presents the regret calculations for this test problem instance, evaluating the V2MS relative to the PMED solution when 1, 2, 3 and 4 facilities are to be utilized.

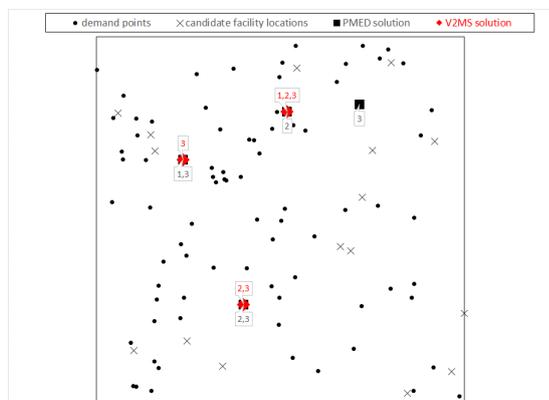


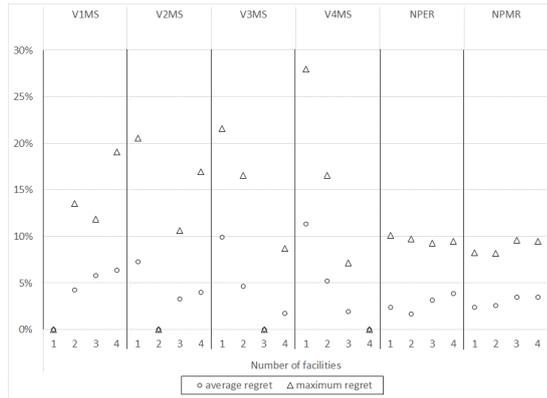
Fig. 2 Facility locations opened in V2MS and PMED solutions, problem instance 27

Table 1 Regret calculations for problem instance 27

ϕ	PMED objective	V2MS objective	Regret
1	218,369	235,518	7.9%
2	155,883	155,883	0%
3	125,693	126,314	0.5%
4	105,416	110,166	4.5%

3.3 Comparing solution quality

Models NPER and NPMR were solved across all 30 test instances, with set $H = \{1, 2, 3, 4\}$. To provide a basis for comparison of solution quality, model PMED was also solved for 1, 2, 3 and 4 facilities for each test instance. Note that this required, in total, 30 runs of NPER, 30 runs of NPMR, and 120 runs of PMED. Each of these PMED model solutions was then utilized as the basis for a set of nested facility opening recommendations, denoted $VpMS$ as discussed above, and thus requiring an additional 120 runs of NPER. Figure 3 presents the average regret and maximum regret, across the 30 test instances, for each model solution, varying the actual number of facilities to be opened on the horizontal axis.

**Fig. 3** Average regret and maximum regret across solutions

Observe that the $VpMS$ solution has zero regret when p facilities are to be opened, by definition. However, should a p -median solution be implemented, but it subsequently become necessary to increase or decrease the number of facilities, the regret can be quite large. For the V1MS solution, the average regret ranges between 4.3% (for two opened facilities) and 6.4% (for four opened facilities), with corresponding maximum regret values of 13.5% and 19.0%. For the V2MS, V3MS and V4MS solutions, the greatest values for both average regret and maximum regret occur at $\phi_h=1$ opened facility (with average regret

values of 7.3%, 9.9% and 11.3%, respectively, and maximum regret values of 20.6%, 21.6% and 28.0%, respectively).

The NPER and NPMR solutions both significantly reduce the potential exposure to large values of average regret and maximum regret, relative to the Vp MMS solutions. The NPER average regret varies between a minimum value of 1.7% (for two opened facilities) and a maximum of 3.9% (for four opened facilities), with maximum regret values ranging between 9.3% and 10.1% (for three and one opened facilities, respectively). Similarly, the NPMR average regret varies between 2.4% and 3.5% (for one and three opened facilities, respectively), while its maximum regret values range between 8.2% and 9.6% (for two and three opened facilities, respectively).

4 Solution approach

In this section we develop a heuristic approach based on Lagrangian relaxation (LR) to solve NPER and NPMR. A commonly used LR approach to solve a standard p -median problem involves relaxing constraints (3) [8]. This type of constraint appears in NPER and NPMR in constraints (10). The additional complicating constraints in NPER and NPMR are the nesting constraints (15). For our LR algorithm, we replace the original nesting constraints (15) from the NPER and NPMR models with the following set of constraints:

$$\hat{x}_{j,h} \geq \hat{x}_{j,h-h'} \quad \forall j, \forall h > 1, \forall h' < h \quad (17)$$

The constraint set (17) includes constraints (15) as well as other redundant constraints, however we find empirically that relaxing constraints (17) enhances the lower bound.

4.1 Solving NPER using Lagrangian relaxation

We define λ_{ih} and $\nu_{j,h,h'}$ as the Lagrange multiplier vectors associated with constraints (10) and (17), respectively. The Lagrangian relaxation $L(\lambda, \nu)$ of NPER with respect to constraints (10) and (17) is given by

$$\begin{aligned} \min_{\hat{x}, \hat{y}} \quad & \sum_h \sum_i \sum_j \gamma_{ij} \hat{y}_{ijh} + \sum_i \sum_h \lambda_{ih} (1 - \sum_j \hat{y}_{ijh}) \\ & + \sum_j \sum_{h>1} \sum_{h'<h} \nu_{j,h,h'} (\hat{x}_{j,h-h'} - \hat{x}_{j,h}), \end{aligned} \quad (18)$$

subject to constraints (9), (11)-(13).

Rearranging terms to isolate the \hat{x} and \hat{y} variables, $L(\lambda, \nu)$ is rewritten as

$$\begin{aligned}
\min_{\hat{x}, \hat{y}} & \sum_i \sum_j \sum_h (\gamma_{ij} - \lambda_{ih}) \hat{y}_{ijh} + \sum_i \sum_h \lambda_{ih} + \sum_j \left\{ \hat{x}_{j,1} \sum_{h=2}^{|H|} \nu_{j,h,h-1} \right\} \\
& + \sum_j \left\{ \sum_{h=2}^{|H|-1} \hat{x}_{jh} \left\{ \left(\sum_{h'=h+1}^{|H|} \nu_{j,h',h'-h} \right) - \left(\sum_{h'=1}^{h-1} \nu_{j,h,h'} \right) \right\} \right\} \\
& - \sum_j \left\{ \hat{x}_{j,|H|} \left\{ \sum_{h'=1}^{|H|-1} \nu_{j,|H|,h'} \right\} \right\}
\end{aligned} \tag{19}$$

subject to constraints (9), (11)-(13).

4.1.1 Calculating the NPER lower bound

Define ρ_{jh} as the contribution of the (facility j , scenario h)-pair to the objective function of $L(\lambda, \nu)$. Then, ρ_{jh} can be constructed based on (19) as:

$$\rho_{jh} = \begin{cases} \Delta_{jh} + \sum_{h=2}^{|H|} \nu_{j,h,h-1} & \text{if } h = 1 \\ \Delta_{jh} + \left(\sum_{h'=h+1}^{|H|} \nu_{j,h',h'-h} \right) - \left(\sum_{h'=1}^{h-1} \nu_{j,h,h'} \right) & \text{if } 1 < h < |H| \\ \Delta_{jh} - \sum_{h'=1}^{|H|-1} \nu_{j,|H|,h'} & \text{if } h = |H| \end{cases} \tag{20}$$

where $\Delta_{jh} = \sum_i \min(0, \gamma_{ij} - \lambda_{ih})$.

For a given set of multipliers λ and ν , the optimal solution for this problem can be obtained using the following approach. For each scenario h , denote the $\hat{\phi}_h$ smallest values of ρ_{jh} as $P_h^* = \{\hat{\rho}_{j_1}, \hat{\rho}_{j_2}, \dots, \hat{\rho}_{j_{\hat{\phi}_h}}\}$. Then, for each h , set $\hat{x}_{jh} = 1$ for $j \in P_h^*$. Then set $\hat{y}_{ijh} = \hat{x}_{jh}$ if $\gamma_{ij} - \lambda_{ih} < 0$, otherwise set \hat{y}_{ijh} to zero. Then calculate the objective function value of $L(\lambda, \nu)$, denoted as $Z_L(\lambda, \nu)$.

Note that since we relax constraints (17), the optimal solution to $L(\lambda, \nu)$ may not have a nested solution. This motivated our decision to formulate the LR using the nesting constraints (17), which include many redundant instances, rather than the more-compact nesting constraints (15). In this way, we would expect to encounter more frequent nesting in the lower bound solution in later iterations of the algorithm (which we observed to occur in our computational testing).

4.1.2 Calculating the NPER upper bound

We can find an upper bound to NPER by generating a feasible solution using part of the lower bound solution to $L(\lambda, \nu)$. The lower bound solution utilizes $\tilde{\Phi} \geq \hat{\phi}_{|H|}$ facilities. We apply Algorithm 1, described below, to find $\hat{\phi}_{|H|}$ nested facilities out of the $\tilde{\Phi}$ facilities chosen in the lower bound solution. The idea behind Algorithm 1 is to use information about a facility's contribution to

$L(\lambda, \nu)$ as a proxy for its value in an upper bound solution. After values of \hat{x}_{jh} are determined using Algorithm 1, it is simple to set the corresponding values for \hat{y}_{ijh} by setting $\hat{y}_{ijh} = 1$ for the closest utilized facility j (i.e., $\hat{x}_{jh} = 1$). We may then calculate the objective function, denoted as Z_U .

Algorithm 1: Calculating the facility locations for the upper bound

```

Initialize by calculating  $\rho_{jh}$  for all  $j, h$ .
for  $h \leftarrow 1$  to  $|H|$  do
    Create a sorted list of  $\rho_{jh}$  denoted as  $S_h^* = \{S_j^1, S_j^2, \dots, S_j^{|J|}\}$ 
    for  $j \leftarrow 1$  to  $|J|$  do
        Denote the ordinality of facility  $j$  in  $S_h^*$  as  $\hat{S}_h^*(j)$ .
    end
end
for  $j \leftarrow 1$  to  $|J|$  do
    if Facility  $j$  is utilized in current lower bound solution then
        Set  $SumRank(j) = \sum_h \hat{S}_h^*(j)$ 
    else
        Set  $SumRank(j) = \infty$ 
    end
end
Create a sorted list of  $SumRank(j)$  denoted as
 $SR = \{SR_j^1, SR_j^2, \dots, SR_j^{|J|}\}$ 
for  $j \leftarrow 1$  to  $|J|$  do
    Denote the ordinality of facility  $j$  in  $SR$  as  $\hat{SR}(j)$ .
end
for  $h \leftarrow 1$  to  $|H|$  do
    for  $k \leftarrow 1$  to  $|J|$  do
        if  $\hat{SR}(j) \leq \hat{\phi}_h$  then
            Set  $\hat{x}_{j,h} = 1$ 
        else
            Set  $\hat{x}_{j,h} = 0$ 
        end
    end
end

```

4.1.3 Subgradient optimization method for NPER

A subgradient method (Algorithm 2) is used to calculate values for the Lagrange multipliers λ_{ih} and $\nu_{j,h,h'}$ at each iteration. We use subgradient smoothing to enhance convergence of the lower bound [2]. The parameters for the subgradient method are listed in Table 2 below, followed by Algorithm 2.

Table 2 Subgradient method parameters for solving NPER

n	iteration number
\bar{z}	Current best upper bound
\underline{z}	Current best lower bound
$\lambda_{ih}^n, \nu_{j,h,h'}^n$	Lagrange multiplier values at iteration n
θ_1, θ_2	Step-size coefficient for Lagrange multipliers λ_{ih} and $\nu_{j,h,h'}$, respectively
t_1, t_2	Step size for the Lagrange multipliers λ_{ih} and $\nu_{j,h,h'}$, respectively
$\Gamma_{\lambda_{ih}}^n, \Gamma_{\nu_{j,h,h'}}^n$	subgradients for the Lagrange multipliers λ_{ih} and $\nu_{j,h,h'}$, respectively, for iteration n
$\tilde{\Gamma}_{\lambda_{ih}}^n, \tilde{\Gamma}_{\nu_{j,h,h'}}^n$	smoothed subgradients for the Lagrange multipliers λ_{ih} and $\nu_{j,h,h'}$, respectively, for iteration n
σ	minimum duality gap termination criteria
σ^n	duality gap in iteration n
ζ	exponential smoothing constant
N	maximum number of iterations termination criteria

Algorithm 2: Subgradient optimization algorithm to solve NPER

Initialize by setting $\bar{z} = \infty$, $\underline{z} = -\infty$, $\lambda_{ih}^1 = 1$, $\nu_{j,h,h'}^1 = 0$, $n = 1$

while $\sigma^n > \sigma$ **and** $n < N$ **do**

Step 1. Solve $L(\lambda^n, \nu^n)$. Set $\underline{z} = \max\{\underline{z}, Z_L(\lambda^n, \nu^n)\}$

Step 2. Apply Algorithm 1 to calculate the upper bound solution,

$\hat{x}_{jh}^n, \hat{y}_{ijh}^n, \forall i, j, h$. Set $\bar{z} = \min\{\bar{z}, Z_U\}$.

Step 3. Calculate the duality gap: $\sigma^n = \frac{\bar{z} - \underline{z}}{\underline{z}}$

Step 4. Make the following calculations:

Step 4a. Calculate subgradients

$$\Gamma_{\lambda_{ih}}^n = 1 - \sum_j \hat{y}_{ijh}^n, \forall i, h$$

$$\Gamma_{\nu_{j,h,h'}}^n = \hat{x}_{j,h-h'}^n - \hat{x}_{jh}^n, \forall h > 1, h' < h$$

Step 4b. Calculate step sizes

$$t_1 = \theta_1(\bar{z} - \underline{z}) / \sum_{i,h} (\Gamma_{\lambda_{ih}}^n)^2, \forall i, h$$

$$t_2 = \theta_2(\bar{z} - \underline{z}) / \sum_{i,h} (\Gamma_{\nu_{j,h,h'}}^n)^2, \forall h > 1, h' < h$$

Step 4c. Calculate smoothed subgradients

if $n > 1$ **then**

$$\tilde{\Gamma}_{\lambda_{ih}}^n = (1 - \zeta)\tilde{\Gamma}_{\lambda_{ih}}^{n-1} + \zeta\Gamma_{\lambda_{ih}}^n, \forall i, h$$

$$\tilde{\Gamma}_{\nu_{j,h,h'}}^n = (1 - \zeta)\tilde{\Gamma}_{\nu_{j,h,h'}}^{n-1} + \zeta\Gamma_{\nu_{j,h,h'}}^n, \forall h > 1, h' < h$$

else

$$\tilde{\Gamma}_{\lambda_{ih}}^n = \Gamma_{\lambda_{ih}}^n, \forall i, h$$

$$\tilde{\Gamma}_{\nu_{j,h,h'}}^n = \Gamma_{\nu_{j,h,h'}}^n, \forall h > 1, h' < h$$

end

Step 4d. Calculate updated Lagrange multipliers

$$\lambda_{ih}^{n+1} = \lambda_{ih}^n + t_1 \tilde{\Gamma}_{\lambda_{ih}}^n, \forall i, h$$

$$\nu_{j,h,h'}^{n+1} = \nu_{j,h,h'}^n + t_2 \tilde{\Gamma}_{\nu_{j,h,h'}}^n, \forall h > 1, h' < h$$

Step 5: Set $t = t + 1$

end

4.2 Solving NPMR using Lagrangian relaxation

We re-organize constraint (16) as follows:

$$0 \geq \frac{\hat{z}_h}{\hat{\alpha}_h} - (1 + \hat{w}) \quad \forall h \quad (21)$$

We define λ_{ih} , $\nu_{j,h,h'}$, and δ_h as the Lagrange multipliers associated with constraints (10), (17), and (21), respectively.

The Lagrangian relaxation $L(\lambda, \nu, \delta)$ of NPMR is given by

$$\begin{aligned} \min_{\hat{x}, \hat{y}, \hat{w}} \quad & \hat{w} + \sum_i \sum_h \lambda_{ih} (1 - \sum_j \hat{y}_{ijh}) \\ & + \sum_j \sum_{h>1} \sum_{h'<h} \nu_{j,h,h'} (\hat{x}_{j,h-h'} - \hat{x}_{j,h}) \\ & + \sum_h \delta_h \left(\frac{\sum_i \sum_j \gamma_{ij} \hat{y}_{ijh}}{\hat{\alpha}_h} - 1 - \hat{w} \right) \end{aligned} \quad (22)$$

subject to constraints (9), (11)-(13).

Rearranging terms to isolate the \hat{x} , \hat{y} , and \hat{w} variables, $L(\lambda, \nu, \delta)$ is rewritten as

$$\begin{aligned} \min_{\hat{x}, \hat{y}, \hat{w}} \quad & \sum_i \sum_j \sum_h \left(\frac{\gamma_{ij}}{\hat{\alpha}_h} \delta_h - \lambda_{ih} \right) \hat{y}_{ijh} + \sum_i \sum_h \lambda_{ih} + \sum_j \left\{ \hat{x}_{j,1} \sum_{h=2}^{|H|} \nu_{j,h,h-1} \right\} \\ & + \sum_j \left\{ \sum_{h=2}^{|H|-1} \hat{x}_{j,h} \left\{ \left(\sum_{h'=h+1}^{|H|} \nu_{j,h',h'-h} \right) - \left(\sum_{h'=1}^{h-1} \nu_{j,h,h'} \right) \right\} \right\} \\ & - \sum_j \left\{ \hat{x}_{j,|H|} \left\{ \sum_{h'=1}^{|H|-1} \nu_{j,|H|,h'} \right\} \right\} \\ & + \hat{w} \left(1 - \sum_h \delta_h \right) \end{aligned} \quad (23)$$

subject to constraints (9), (11)-(13).

4.2.1 Calculating the lower and upper bounds for NPMR

Define ρ_{jh} as the contribution of the (facility j , scenario h)-pair to the objective function of $L(\lambda, \nu, \delta)$. Then, ρ_{jh} can be constructed based on (23) as:

$$\rho_{jh} = \begin{cases} \Pi_{jh} + \sum_{h=2}^{|H|} \nu_{j,h,h-1} & \text{if } h = 1 \\ \Pi_{jh} + \left(\sum_{h'=h+1}^{|H|} \nu_{j,h',h'-h} \right) - \left(\sum_{h'=1}^{h-1} \nu_{j,h,h'} \right) & \text{if } 1 < h < |H| \\ \Pi_{jh} - \sum_{h'=1}^{|H|-1} \nu_{j,|H|,h'} & \text{if } h = |H| \end{cases} \quad (24)$$

Table 3 Additional subgradient method parameters to solve NPMR

δ_h^n	Lagrange multiplier values at iteration n
θ_3	Step-size coefficient for Lagrange multipliers δ_h
t_3	Step size for the Lagrange multipliers δ_h
$\Gamma_{\delta_h}^n$	subgradients for the Lagrange multipliers δ_h , for iteration n
$\tilde{\Gamma}_{\delta_h}^n$	smoothed subgradients for the Lagrange multipliers δ_h , for iteration n

where $\Pi_{jh} = \sum_i \min(0, (\frac{\gamma_{ij}}{\alpha_h} \delta_h - \lambda_{ih}))$. For a given set of multipliers λ , ν , and δ , the optimal solution and objective function value for this problem can be obtained using the same approach developed for NPER, with one change. That is, the above formulation will have the following characteristic: In each iteration, if $1 - \sum \delta_h > 0$, then it is optimal to set $\hat{w} = 0$. Otherwise, if $1 - \sum \delta_h < 0$, it is optimal to set $\hat{w} = \infty$. Denote the optimal objective function value of the lower bound problem as $Z_L(\lambda, \nu, \delta)$. We then calculate the upper bound for NPMR using the same technique developed for NPER in Algorithm 1, with objective function value Z_U equal to the maximum relative regret value obtained from equation (16).

4.2.2 Subgradient optimization method for NPMR

The parameters for our subgradient method to solve NPMR include the parameters used in the NPER subgradient method (listed in Table 2) as well as those listed in Table 3. The details of the method are described in Algorithm 3.

Algorithm 3: Subgradient optimization algorithm to solve NPMR

Initialize by setting $\bar{z} = \infty$, $z = -\infty$, $\lambda_{ih}^1 = 1$, $\nu_{j,h,h'}^1 = 0$, $\delta_h = 0$, $n = 1$

while $\sigma^n > \sigma$ and $n < N$ **do**

 Step 1. Solve $L(\lambda^n, \nu^n, \delta^n)$. Set $z = \max\{z, Z_L(\lambda^n, \nu^n)\}$

 Step 2. Apply Algorithm 1 to calculate the upper bound solution, $\hat{x}_{jh}^n, \hat{y}_{ijh}^n, \hat{w}_h^n, \forall i, j, h$. Set $\bar{z} = \min\{\bar{z}, Z_U\}$.

 Step 3. Calculate the duality gap: $\sigma^n = \frac{\bar{z} - z}{z}$

 Step 4. Make the following calculations:

 Step 4a. Calculate subgradients

$$\Gamma_{\lambda_{ih}}^n = 1 - \sum_j \hat{y}_{ijh}^n, \forall i, h$$

$$\Gamma_{\nu_{j,h,h'}}^n = \hat{x}_{j,h-h'}^n - \hat{x}_{jh}^n, \forall h > 1, h' < h$$

$$\Gamma_{\delta_h}^n = \sum_i \sum_j (\gamma_{ij} \hat{y}_{ijh}^n / \hat{\alpha}_h - 1 - \hat{w})$$

 Step 4b. Calculate step sizes

$$t_1 = \theta_1(\bar{z} - z) / \sum_{i,h} (\Gamma_{\lambda_{ih}}^n)^2, \forall i, h$$

$$t_2 = \theta_2(\bar{z} - z) / \sum_{i,h} (\Gamma_{\nu_{j,h,h'}}^n)^2, \forall h > 1, h' < h$$

$$t_3 = \theta_3(\bar{z} - z) / \sum_{i,h} (\Gamma_{\delta_h}^n)^2, \forall h$$

 Step 4c. Calculate smoothed subgradients

if $n > 1$ **then**

$$\tilde{\Gamma}_{\lambda_{ih}}^n = (1 - \zeta)\tilde{\Gamma}_{\lambda_{ih}}^{n-1} + \zeta\Gamma_{\lambda_{ih}}^n, \forall i, h$$

$$\tilde{\Gamma}_{\nu_{j,h,h'}}^n = (1 - \zeta)\tilde{\Gamma}_{\nu_{j,h,h'}}^{n-1} + \zeta\Gamma_{\nu_{j,h,h'}}^n, \forall h > 1, h' < h$$

$$\tilde{\Gamma}_{\delta_h}^n = (1 - \zeta)\tilde{\Gamma}_{\delta_h}^{n-1} + \zeta\Gamma_{\delta_h}^n, \forall h$$

else

$$\tilde{\Gamma}_{\lambda_{ih}}^n = \Gamma_{\lambda_{ih}}^n, \forall i, h$$

$$\tilde{\Gamma}_{\nu_{j,h,h'}}^n = \Gamma_{\nu_{j,h,h'}}^n, \forall h > 1, h' < h$$

$$\tilde{\Gamma}_{\delta_h}^n = \Gamma_{\delta_h}^n, \forall h$$

end

 Step 4d. Calculate updated Lagrange multipliers

$$\lambda_{ih}^{n+1} = \lambda_{ih}^n + t_1 \tilde{\Gamma}_{\lambda_{ih}}^n, \forall i, h$$

$$\nu_{j,h,h'}^{n+1} = \nu_{j,h,h'}^n + t_2 \tilde{\Gamma}_{\nu_{j,h,h'}}^n, \forall h > 1, h' < h$$

$$\delta_h^{n+1} = \delta_h^n + t_3 \tilde{\Gamma}_{\delta_h}^n, \forall h$$

 Step 5: Set $t = t + 1$

end

5 Computational Results

We tested the LR-based solution algorithm for NPER (denoted LR-ER) and NPMR (denoted LR-MR) on a set of test problems from the online repository [4]. These problems are drawn from the p -median instances examined in [3]; in our computational testing we utilize the problem instances containing 400 or more nodes (denoted $pmed16$ through $pmed40$ at [4]). We assume that all nodes are both demand points and potential facility locations, thus $I = J$. Models NPER and NPMR were again both coded in GAMS 33.2 and solved

Table 4 NPER Computational Experience

Problem size $ I = J $	# instances	OTS-ER			LR-ER		
		Gap	Runtime	UB	Gap	Runtime	\bar{z}
400	5	0%	413	43,384	0.36%		43,425
500	5	20%	367	51,450	0.50%		46,119
600	5	60%	599	57,614	0.38%	600	47,312
700	4	100%	600	91,829	2.21%		51,274
800	3	100%	600	75,389	2.61%		55,267
900	3	100%	600	85,225	2.16%		54,772

using CPLEX 12.10, here using nesting constraints (17) rather than (15), for consistency with LR-ER and LR-MR. The computational results obtained with this off-the-shelf optimization software will be denoted OTS-ER (for NPER) and OTS-MR (for NPMR).

In our computational testing we initialized the LR parameters $\theta_1 = \theta_2 = \theta_3 = 0.75$ and $\zeta = 0.8$. Values for step-size coefficients θ_1, θ_2 were reduced by 10% after 30 unimproved iterations. The Lagrange multipliers were initialized as $\lambda_{ih}^1 = \nu_{j,h,h'}^1 = \delta_h^1 = 2$, and the termination criteria were set at $\sigma = 10^{-6}$ and $N = 1000$. For all instances tested, we assumed set $H = \{1, 2, 3, 4\}$.

The results of the computational testing for NPER appear in Table 4. Each value presented in this table for OTS-ER and LR-ER represents the average value, across all instances, for each problem size. A maximum run time of 600 seconds was utilized with GAMS/CPLEX, thus the results for LR-ER present the model's status at 600 seconds of run time, to allow for comparison. Column UB presents the best feasible objective function value obtained by GAMS/CPLEX. Observe that for the smallest problems ($|I| = |J| = 400$), OTS-ER was able to obtain a provably optimal solution in all instances, with an average run time of 413 seconds, outperforming LR-ER. For the problem size $|I| = |J| = 500$, OTS-ER was able to obtain a provably optimal solution for 4 of the 5 instances, however in the remaining instance OTS-ER terminated after 600 seconds with an optimality gap of 100%. Similarly, for the problem size $|I| = |J| = 600$, OTS-ER was able to obtain a provably optimal solution for 1 of the 5 instances, in 3 of the remaining instance OTS-ER terminated after 600 seconds with an optimality gap of 100%. For these two problem sizes, LR-ER was able to obtain solution with average optimality gaps of 0.50% and 0.38%, respectively, for $|I| = |J| = 500$ and 600. For all instances of the largest problem sizes ($|I| = |J| = 700, 800$ and 900), OTS-ER terminated after 600 seconds with an optimality gap of 100%. Observe that for these large problem sizes, the best feasible solutions obtained by OTS-ER had objective function values that were 79%, 36% and 56% greater, on average, than the best feasible solutions obtained by LR-ER for $|I| = |J| = 700, 800$ and 900, respectively.

The results of the computational testing for NPMR appear in Table 5. Each value presented in this table for OTS-MR and LR-MR represents the

Table 5 NPMR Computational Experience

Problem size $ I = J $	# instances	OTS-MR		LR-MR	
		Runtime	UB	Runtime	\bar{z}
400	5	3600	60.9%	600	10.1%
500	5		53.6%		11.9%
600	5		37.1%		12.3%
700	4		65.6%		10.8%
800	3		42.1%		11.5%
900	3		51.6%		10.6%

average value, across all instances, for each problem size. Due to its difficulty in obtaining a solution, the maximum run time was increased to 3600 seconds for GAMS/CPLEX. However, the results for LR-MR again present the model's status at 600 seconds of run time, to emphasize its performance advantage. For all 25 instances tested, GAMS/CPLEX returned a solution with an optimality gap of 100% (and an objective function lower bound of -10^{-4}), thus we do not present the optimality gap in this table. Observe that the best feasible solutions obtained by OTS-MR are considerably worse than the solutions obtained by LR-MR (despite OTS-MR being provided six times the run time allowed to LR-MR), with OTS-MR having objective function values that were 6.0, 4.5, 3.0, 6.1, 3.7 and 4.9 times the best feasible solutions obtained by LR-MR, on average, for $|I| = |J| = 400, 500, 600, 700, 800$ and 900 , respectively.

6 Conclusion

This paper has introduced the concept of *nested* facility locations, in which the solution utilizing p facilities is a subset of the solution utilizing q facilities, for all $i \leq p < q \leq j$, given some lower limit i and upper limit j on r , the number of facilities that will be utilized in the future. Such an approach overcomes a challenge in the solutions to traditional facility location models, which can generate solutions that do not maintain consistency in the set of utilized facilities as the number of utilized facilities is varied. The nested facility location concept is demonstrated with application to the classic p -median model, with computational testing showing these new models achieve reductions in worst-case regret when $r \neq p$ facilities are actually utilized, relative to the best possible performance that could be achieved, were a p -median solution implemented, but it later become necessary to add or remove facilities from this solution.

We develop two heuristic approaches, based on Lagrangian relaxation (denoted LR-ER and LR-MR, respectively), to efficiently solve the model under each objective for large problem instances. Computational testing on a set of large problem instances found LR-ER and LR-MR to perform well, significantly outperforming the commercial solver GAMS/CPLEX on the largest problem instances.

Future extensions to this research could extend the nested concept to other facility location models, such as the p -center or maximum covering models. The nested concept could potentially be applied to facility location models that consider a budget on future facility opening and closing decisions, as in [16]. It would also be possible to extend these concepts to continuous facility location models, in which the facility locations are selected as coordinates in a planar region. Another potential extension would incorporate the nested concept to situations in which additional aspects of the model were assumed to be uncertain, such as the distances between demand points and facilities. The nested facility location concept could also be extended to stochastic optimization approaches, in which uncertainties are characterized by probability distributions.

Acknowledgements Research reported in this manuscript was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM104417. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Averbakh, I.: Minmax regret solutions for minimax optimization problems with uncertainty. *Operations Research Letters* **27**, 57–65 (2000)
2. Baker, B.M., Sheasby, J.: Accelerating the convergence of subgradient optimisation. *European Journal of Operational Research* **117**(1), 136–144 (1999)
3. Beasley, J.E.: A note on solving large p -median problems. *European Journal of Operational Research* **21**, 270–273 (1985)
4. Beasley, J.E.: <http://people.brunel.ac.uk/~mastjjb/jeb/orlib/pmedinfo.html> (2004)
5. Berman, O., Drezner, Z.: The p -median problem under uncertainty. *European Journal of Operational Research* **189**, 19–30 (2008)
6. Current, J., Daskin, M.S., Schilling, D.: Discrete network location models. In: Z. Drezner, H.W. Hamacher (eds.) *Facility Location: Applications and Theory*, pp. 81–118. Springer-Verlag, New York (2002)
7. Current, J., Ratick, S., ReVelle, C.: Dynamic facility location when the total number of facilities is uncertain: A decision analysis approach. *European Journal of Operational Research* **110**, 597–609 (1997)
8. Daskin, M.S., Maass, K.L.: *The p -Median Problem*, pp. 21–45. Springer International Publishing, Cham (2015)
9. Garey, M.R., Johnson, D.S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company (1979)
10. Hakimi, S.L.: Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research* **12**(3), 450–459 (1964)
11. Hakimi, S.L.: Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research* **13**(3), 462–475 (1965)
12. Lodi, A., Mossina, L., Rachelson, E.: Learning to handle parameter perturbations in combinatorial optimization: An application to facility location. *EURO Journal on Transportation and Logistics* **9**(4), 100023 (2020)
13. Snyder, L.V.: Facility location under uncertainty: a review. *IIE Transactions* **38**, 537–554 (2006)
14. Snyder, L.V., Daskin, M.S.: Stochastic p -robust location problems. *IIE Transactions* **38**(11), 971–985 (2006)
15. Sobolev Institute of Mathematics: http://www.math.nsc.ru/AP/benchmarks/UFLP/Eng1/uf1p_eucl_eng.html (2018)

-
16. Sonmez, A.D., Lim, G.J.: A decomposition approach for facility location and relocation problem with uncertain number of future facilities. *European Journal of Operational Research* **218**, 327–338 (2012)