Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
European Journal of Operational Research 159 (2004) 574–585
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Discrete Optimization

A framework for demand point and solution space aggregation analysis for location models

R.L. Francis [a,*], T.J. Lowe [b], A. Tamir [c], H. Emir-Farinas [a]

[a] *Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA*
[b] *Tippie College of Business, University of Iowa, Iowa City, IA 52242, USA*
[c] *Department of Statistics & Operations Research, School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel*

Received 11 January 2002; accepted 25 September 2002
Available online 8 October 2003

Abstract

Many location problems can be formulated as minimizing some location objective function subject to upper bounds on other location constraint functions. When such functions are subadditive and nondecreasing in the distances (a common occurrence), worst-case demand point aggregation error bounds are known. We show how to solve a relaxation and a restriction of the aggregated problem in such a way as to obtain lower and upper bounds on the optimal value of the original problem. We consider some applications to covering and related problems.

Keywords: Location; Aggregation error; Covering; Demand points; K-centers

1. Introduction

Location problems often involve finding locations of new facilities that provide services of some kind to existing facilities, also called demand points (DPs). When such problems occur in urban contexts, each private residence can be a DP. Thus there can be too many DPs to be modeled individually, and aggregation of the DPs becomes necessary; indeed, sometimes only aggregated data is available. This aggregation creates a more tractable model, but also introduces model error. It is naturally of interest to examine how much error is introduced, and the effect of the level of aggregation upon the error. Essentially the modeler is faced with a tradeoff: less model accuracy for more model tractability, or vice versa.

Hillsman and Rhoda (1978) were perhaps the first to study errors associated with DP aggregation for the K-median problem. Their classification of aggregation errors was further studied by Current and Schilling (1987), Erkut and Bozkaya (1999) and Zhao and Batta (1999). Plastria (2001), in an effort to reduce aggregation error, further refined the Hillsman–Rhoda error classification, and made a strong case for using the centroid of

* Corresponding author. Tel.: +1-352-392-6727; fax: +1-352-392-3537.
E-mail addresses: francis@ise.ufl.edu (R.L. Francis), tlowe@blue.weeg.uiowa.edu (T.J. Lowe), atamir@post.tau.ac.il (A. Tamir), hulya@ufl.edu (H. Emir-Farinas).

0377-2217/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0377-2217(03)00433-8

each aggregation group as the representative point for the group in the reduced K-median problem. Current and Schilling (1990) considered errors due to aggregation for the (discrete) covering problem, a problem we consider in some detail in this paper.

Methods for reducing DP aggregation error have been proposed by Current and Schilling (1987), Bowerman et al. (1999), and Hodgson and Neuman (1993). General background discussions on DP aggregation and resulting errors can be found in Francis et al. (1999), Zhao and Batta (1999, 2000) and Plastria (2001).

Zhao and Batta (2000) point out that there is another kind of aggregation error; it is perhaps less obvious but often occurs. Due to budgetary constraints, only a subset of the set of all possible feasible solutions may be considered. For example, potential location sites of interest may be enumerated in a list. With more resources to consider sites of interest, additional sites might be added to the list. This approach is usually viewed as solving a restriction of the actual problem, but it can also be viewed as a "solution space" aggregation, since a large collection of potential sites is, in effect, aggregated into a smaller collection. Zhao and Batta (1999, 2000) also give a general background discussion of DP aggregation, as do Francis et al. (1999, 2002).

A useful idea in aggregation analysis is that of an error bound. Suppose X denotes a choice of new facility locations, and $f(X)$ denotes a location model objective without DP aggregation. Let $f'(X)$ denote the resulting approximating location model after some form of aggregation. An *error bound* is a number, say eb, so that $|f'(X) - f(X)| \leqslant \text{eb}$ for all X of interest. Such DP aggregation error bounds originated with Zemel (1985) and Francis and Lowe (1992), and are now known for a large class of network location models (Francis et al., 2000), including the K-center, K-median and other models. Sometimes error bounds can be "two-sided", that is, $f'(X) \leqslant f(X) + \text{eb}_1$ for all X, and $f(X) \leqslant f'(X) + \text{eb}_2$ for all X. Such error bounds are worst-case error measures, since they hold for all X. However, if they are small compared to function values, they can be useful. Computational experience to date (Rayco et al., 1997, 1999) indicates they can be especially helpful for problems with worst-case structure, such as K-center models, which are known to be closely related to covering models (Kolen and Tamir, 1990).

In this paper we consider aggregation for some constrained location minimization models. The objective function to be minimized is a location function, and there are also constraints that impose upper bounds on costs of other location functions. Starting with an original (unaggregated) model, say (Pr), we do aggregation to obtain an approximating model, say (Pr′). We find a restriction of (Pr′) say ($\text{Pr}'_{\text{restr}}$) and a relaxation of (Pr′) say (Pr'_{relx}), that also "serve" as a restriction and relaxation respectively of (Pr). Assuming all three approximating problems can be solved to optimality, we can thus obtain bounds on minimal values of both (Pr) and (Pr′). When the bounds are (nearly) equal we can conclude that an optimal solution to (Pr′) is (nearly) optimal to (Pr). Our work generalizes earlier aggregation results (Francis and Lowe, 1992) that were specific to covering models.

An overview of our paper is as follows. In Section 2 we present notation, the basic location models (Pr) and (Pr′), and assume the existence of certain vectors of two-sided error bounds. The underlying location context of interest involves locating facilities on a network, using shortest path distances between points. Our abstraction of this context assumes a metric space. In Section 3 we present Theorem 1, our basic bounding result. In Section 4 we consider a class of so-called SAND location models (Francis et al., 2000), essentially ones that are subadditive and nondecreasing in distances. For such models, Theorems 2 and 3 provide formulas to compute the two-sided error bounds. At the end of the section we specialize these results to covering and K-center models. In Section 5 we present some computational experience for covering problems. The experience indicates our results can be useful for covering and related models. A short conclusion section ends our paper.

2. Notation and definitions

In a typical location model we have a set of DPs (existing facilities), embedded in some metric space such as an (undirected) transport network; we need to select a subset of points X in the space where servers (new facilities) will be located. Usually, there are constraints on X, depending possibly on cost considerations, and proximity of the servers to the DPs. The goal is to choose a feasible subset X, optimizing some given utility function, say $f(X)$, defined on the collection of all feasible subsets X. In this section we introduce a quite general and formal model that will capture the above characteristics and components of a location problem.

Let M be a metric space, with $d(x,y)$ denoting the distance between any pair of points $x,y \in M$. We denote by 2^M the power set of M, the collection of all subsets of M. Given is a domain $S \subset 2^M$, with real functions $f(X)$, $h_1(X),\dots,h_k(X)$, from S to $\Re^1$, and a subdomain $S' \subset S$. We also have (approximating) real functions $f'(X)$, $h'_1(X),\dots,$ $h'_k(X)$ from S to $\Re^1$, and a vector $\tau = (\tau_1,\dots,\tau_k) \in \Re^k$. We consider the following general optimization problems:

$$\Pr(\tau) : v(\tau) = \min\{f(X) : h(X) \leqslant \tau, X \in S\},$$

$$\Pr'(\tau) : v'(\tau) = \min\{f'(X) : h'(X) \leqslant \tau, X \in S'\},$$

where $h(X) = (h_1(X),\dots,h_k(X))$, $h'(X) = (h'_1(X),$ $\dots,h'_k(X))$. We assume the above two minima exist, as well as subsequent ones. Note the problem statements also define two functions, v and v'.

Motivated by DP aggregation in location models, our main interest is in a subclass of the above defined as follows.

Given are two multi-subsets of M, a DP set $P = \{p_1,\dots,p_m\}$ and an aggregate demand point (ADP) set $P' = \{p'_1,\dots,p'_m\}$. The DPs are distinct, but the ADPs are not; hence the need for a multi-subset. For each i, p'_i is the point that p_i is aggregated into. For any finite subset X of M, and $y \in M$ define $D(X,y) = \min_{x\in X} d(x,y)$. Let $D(X,P)$ be the vector in $\Re^m_+$ defined by $D(X,P) = (D(X,p'_1),\dots,D(X,p_m))$. Let $D(X,P')$ be defined by $D(X,P') = (D(X,p'_1),\dots,D(X,p'_m))$. We shall also use the notation $T(P,P') = (d(p_1,p'_1),\dots,$ $d(p_m,p'_m))$ in Section 4.

For $j = 0,1,2,\dots,k$, let g_j be a real function from $\Re^m_+$ to $\Re^1$. Suppose that $f(X) = g_0(D(X,P))$, $f'(X) = g_0(D(X,P'))$, and $h_j(X) = g_j(D(X,P))$, $h'_j(X) = g_j(D(X,P'))$, for $j = 1,\dots,k$. (We assume that each $X \in S$ is a finite subset of M.) We refer to the foregoing subclass of models designated with prime symbols as *location aggregation models.*

As an example, consider the K-median problem with the additional constraints requiring that each DP in P will be served within a distance which is not greater than some given radius t. In this case we define $\tau = (t,\dots,t)$, $S = \{X : X \subset M, |X| = K\}$, $k = m$, $g_0(u_1,\dots,u_m) = \sum_{i=1}^m u_i$, and $g_j(u_1,\dots,$ $u_m) = u_j$ for $j = 1,\dots,m$.

We make the following assumptions with respect to the general optimization problems:

Assumption I. There exist $\alpha_1 \in \Re^1_+$ and $\beta_1 \in \Re^k_+$ such that for any $X \in S$, there exist $Y_X \in S'$, and

(i) $f'(Y_X) \leqslant f(X) + \alpha_1$,
(ii) $h'(Y_X) \leqslant h(X) + \beta_1$.

Assumption II. There exist $\alpha_2 \in \Re^1_+$ and $\beta_2 \in \Re^k_+$ such that for any $X \in S$,

(i) $f(X) \leqslant f'(X) + \alpha_2$,
(ii) $h(X) \leqslant h'(X) + \beta_2$.

(In some cases we will simply take $Y_X = X$.)

3. General bounding result

Fig. 1 gives an overview of this section. The upper portion of the figure shows the original problem $\Pr(\tau)$ and the three aggregated problems $\Pr'(\tau)$, $\Pr'(\tau+\beta_1)$, and $\Pr'(\tau-\beta_2)$. The results of this section are the three consequences, proven in Theorem 1. Consequence 1 uses $\Pr(\tau)$, $\Pr'(\tau+\beta_1)$, and Assumption I. Consequence 2 uses both assumptions and the three aggregated problems. Consequence 3 uses $\Pr(\tau)$, $\Pr'(\tau-\beta_2)$, and Assumption II. The three consequences provide lower and upper bounds on both $v(\tau)$ and $v'(\tau)$.

We are now ready to prove the main result relating problems $\Pr(\tau)$ and $\Pr'(\tau)$. (Note that this result is applicable to the general optimization problems and not only to the subclass defined above by the functions $g_0,g_1,\dots,g_k$.)

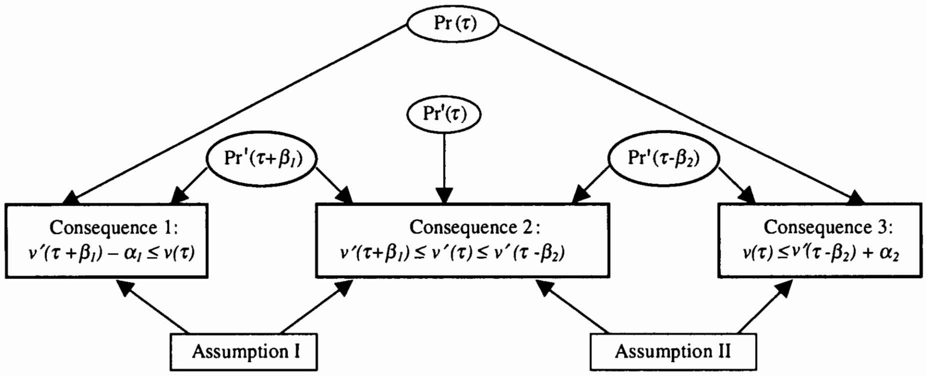

Fig. 1. Overview of section: consequences and information used to obtain them.

Theorem 1. *Suppose that the minimal objective function values* $v(\tau)$, $v'(\tau)$, $v'(\tau+\beta_1)$ *and* $v'(\tau-\beta_2)$ *are well defined. Then, under Assumptions I and II,*

$$v'(\tau+\beta_1)-\alpha_1 \leqslant v(\tau) \leqslant v'(\tau-\beta_2)+\alpha_2,$$

$$v'(\tau+\beta_1) \leqslant v'(\tau) \leqslant v'(\tau-\beta_2).$$

Proof. The inequalities $v'(\tau+\beta_1) \leqslant v'(\tau) \leqslant v'(\tau-\beta_2)$ follow directly from the definition of $\mathrm{Pr}'(\tau)$, and the nonnegativity of β_1 and β_2.

Next we prove $v(\tau) \leqslant v'(\tau-\beta_2)+\alpha_2$. Assuming all minima of interest exist, we have

$$\begin{aligned} v(\tau) &= \min\{f(X) : h(X) \leqslant \tau, X \in S\} \\ &\leqslant \min\{f(X) : h(X) \leqslant \tau, X \in S'\} \\ &\leqslant \min\{f'(X)+\alpha_2 : h(X) \leqslant \tau, X \in S'\} \\ &\leqslant \min\{f'(X)+\alpha_2 : h'(X) \leqslant \tau-\beta_2, X \in S'\} \\ &= \alpha_2+\min\{f'(X) : h'(X) \leqslant \tau-\beta_2, X \in S'\} \\ &= \alpha_2+v'(\tau-\beta_2). \end{aligned}$$

(Note that the first inequality follows from $S' \subset S$, the second follows from Assumption $\mathrm{II_i}$, and the third follows from Assumption $\mathrm{II_{ii}}$.)

To prove that $v'(\tau+\beta_1)-\alpha_1 \leqslant v(\tau)$, let $X^* \in S$ be optimal for $\mathrm{Pr}(\tau)$. Hence $f(X^*) = v(\tau)$ and $h(X^*) \leqslant \tau$. From Assumption I, there exists $Y_{X^*} \in S'$ such that

$$h'(Y_{X^*}) \leqslant h(X^*)+\beta_1,$$

$$f'(Y_{X^*}) \leqslant f(X^*)+\alpha_1.$$

Since $h(X^*) \leqslant \tau$, we obtain $h'(Y_{X^*}) \leqslant h(X^*)+\beta_1 \leqslant \tau+\beta_1$. The latter implies that Y_{X^*} is feasible for $\mathrm{Pr}'(\tau+\beta_1)$. In particular,

$$v'(\tau+\beta_1) \leqslant f'(Y_{X^*}).$$

Combining the latter inequality with the above upper bound on $f'(Y_{X^*})$, we conclude

$$\begin{aligned} &-\alpha_1+v'(\tau+\beta_1) \\ &\quad\leqslant -\alpha_1+f'(Y_{X^*}) \leqslant f(X^*) = v(\tau). \qquad \square \end{aligned}$$

4. Application to SAND location aggregation models

In this section we apply the above general theorem to constraint, objective and solution domain aggregation for so-called SAND location models (Francis et al., 2000). To introduce such models we first define a SAND function. Let g be any real function from $\Re^m_+$ to $\Re^1$. If, for any $U, V \in \Re^m_+$ the function g satisfies $g(U+V) \leqslant g(U)+g(V)$, then g is called *subadditive*. If, for any $U, V \in \Re^m_+$ with $U \leqslant V$ the function g satisfies $g(U) \leqslant g(V)$, then g is called *nondecreasing*. If g is both subadditive and nondecreasing we say it is *SAND*. Clearly, if g is SAND, then for any $U, V, W \in \Re^m_+$ with $U \leqslant V+W$ we have $g(U) \leqslant g(V)+g(W)$. Note, since $0 \leqslant 0+V$ for any $V \in \Re^m_+$, that $g(0) \leqslant g(0)+g(V)$ implies $0 \leqslant g(V)$; our SAND functions are always nonnegative valued. If g is subadditive and $g(0)=0$, $V \leqslant W$ implies $V \leqslant 0+W$, so $g(V) \leqslant g(0)+g(W) = g(W)$ and it is unnecessary to assume g is nondecreasing. Well-known SAND functions (Francis et al., 2000) include the sum function, $g(U) = u_1 + \cdots + u_m$, and the max function, $g(U) = \max\{u_1, \ldots, u_m\}$.

Let S be a domain and let $S' \subset S$ be a subdomain. Let r be a nonnegative real. We say that S' is an *r-cover* of S if for any $X \in S$, there exists $Y_X, Y_X \in S'$, such that $D(Y_x, x) \leqslant r$ for any $x \in X$. Without loss of generality, we assume $|Y_X| \leqslant |X|$.

The r-cover definition implies that for each $x \in X$, there exists $y_x \in Y_X$ with $d(x, y_x) \leqslant r$. As a motivating example, suppose M comes from an undirected travel network using shortest path distances. Let r be half the longest arc length in the network. For each $x \in X$, we can take y_x to be a closest vertex to x, and then have $d(y_x, x) \leqslant r$. We then set $Y_X = \{y_x : x \in X\}$.

To obtain an overview of this section consider Table 1. Table 1 is a summary of notation (lines 1–3), along with the results of Theorems 2 and 3, which establish the formulas in lines 4 and 5 for the α's and β's. Lines 6 and 7 summarize the relationships of the α's and β's to the functions for the original and approximating (aggregated) problems.

Lemma 1. *Suppose that S' is an r-cover of S. For any $X \in S$, and $i = 1, \ldots, m$,*

$$D(Y_X, p'_i) \leqslant D(X, p_i) + r + d(p_i, p'_i).$$

Proof. Let $x \in X$ satisfy $D(X, p_i) = d(x, p_i)$, and let $y_x \in Y_X \in S'$ satisfy $d(x, y_x) \leqslant r$. Then, $D(Y_X, p'_i) \leqslant d(y_x, p'_i) \leqslant d(y_x, x) + d(x, p_i) + d(p_i, p'_i) \leqslant r + d(x, p_i) + d(p_i, p'_i) = r + D(X, p_i) + d(p_i, p'_i)$. □

Lemma 2. *Let $e \in \Re^m_+$ be defined by $e = (1, \ldots, 1)$. Let g be a SAND function from $\Re^m_+$ to $\Re^1$, and suppose that S' is an r-cover of S. Then for any $X \in S$,*

$$g(D(Y_X, P')) \leqslant g(D(X, P)) + g(re + T(P, P')).$$

Proof. From Lemma 1 we have $D(Y_X, P') \leqslant D(X, P) + re + T(P, P')$. The result follows directly from the fact that g is a *SAND* function. □

Remark 1. If we let $S' = S$, then S' is a 0-cover of S. In particular, for any $X \in S$ we have $Y_x = X$. In this case, we obtain

$$g(D(X, P')) \leqslant g(D(X, P)) + g(T(P, P')),$$

$$g(D(X, P)) \leqslant g(D(X, P')) + g(T(P, P')).$$

Theorem 2. *Let g_0 be a SAND function from $\Re^m_+$ to $\Re^1$.*

(a) *Then $f(X) = g_0(D(X, P))$ and $f'(X) = g_0(D(X, P'))$ satisfy Assumption II_i with*

$$\alpha_2 = g_0(T(P, P')).$$

(b) *Moreover, if S' is an r-cover of S, then $f(X) = g_0(D(X, P))$ and $f'(Y_X) = g_0(D(Y_X, P'))$ satisfy Assumption I_i with*

$$\alpha_1 = g_0(re + T(P, P')).$$

Proof. Part (a) follows from the Remark above, while (b) follows from Lemma 2. □

Table 1
Notation, and summary of Theorems 2 and 3

	Objective information	Constraint information
1	$D(X,P) = (D(X,p_i))$	$T(P,P') = (d(p_1,p'_1),\ldots,d(p_m,p'_m))$
2	$f(X) = g_0(D(X,P))$	$h(X) = (g_1(D(X,P)),\ldots,g_k(D(X,P)))$
3	$f'(X) = g_0(D(X,P'))$	$h'(X) = (g_1(D(X,P')),\ldots,g_k(D(X,P')))$
4	$\alpha_1 = g_0(re + T(P,P'))$	$\beta_1 = (g_1(re + T(P,P')),\ldots,g_k(re + T(P,P')))$
5	$\alpha_2 = g_0(T(P,P'))$	$\beta_2 = (g_1(T(P,P')),\ldots,g_k(T(P,P')))$
6	$f(X) \leqslant f'(X) + \alpha_2$, all X	$h(X) \leqslant h'(X) + \beta_2$, all X
7	$f'(Y_X) \leqslant f(X) + \alpha_1$, $X \in S$, $Y_X \in S'$	$h'(Y_X) \leqslant h(X) + \beta_1$, $X \in S$, $Y_X \in S'$

S' is an r-cover of S; for every $X \in S$, $\exists Y_X$ in S' such that $D(Y_X, x) \leqslant r$ for any $x \in X$. All g functions are SAND.

Theorem 3. *For* $j = 1, \ldots, k$, *let* g_j *be a SAND function from* $\Re^m_+$ *to* $\Re^1$. *Let* $h(X) = (h_1(X), \ldots, h_k(X))$, *and* $h'(X) = (h'_1(X), \ldots, h'_k(X))$, *where* $h_j(X) = g_j(D(X,P))$ *and* $h'_j(X) = g_j(D(X, P'))$, $j = 1, \ldots, k$.

(a) *The functions* $h(X)$ *and* $h'(X)$ *satisfy Assumption* II_{ii} *with*

$$\beta_2 = (g_1(T(P,P')), \ldots, g_k(T(P,P')).$$

(b) *Moreover, if* S' *is an* r*-cover of* S, *then the functions* $h(X)$ *and* $h'(Y_X)$ *satisfy Assumption* I_{ii} *with*

$$\beta_1 = (g_1(re + T(P,P')), \ldots, g_k(re + T(P,P'))).$$

Proof. The proof is the same as for Theorem 2. □

Finally, we conclude that if the functions $g_0, g_1, \ldots, g_k$ are SAND functions, and S' is an r-cover of S, then Theorem 1 is applicable to the respective location aggregation model, where α_1, α_2, β_1 and β_2 are defined in Theorems 2 and 3.

In the remainder of this section we specialize the above results to a class of location problems; covering problems, and constrained K-center problems. We shall see it is sometimes possible to exploit extra problem structure to find smaller entries in the α and β vectors than those given by Theorems 2 and 3.

The basic covering problem appears in line 1 of Table 2, with related aggregated problems in lines 2–4. Note that S' is an r-cover of S, $h(X) = (D(X, P_1), \ldots, D(X, p_m))$, $h'(X) = (D(X, p'_1), \ldots, D(X, p'_m))$, $\tau = (t_1, \ldots, t_m)$. Theorem 3 gives $\beta_2 = (d(p_1, p'_1), \ldots, d(p_m, p'_m))$, $\beta_1 = re + \beta_2$. Because the original and approximating covering problems have the same objective and do not depend (directly) on P, $f(X) = |X| = f'(X)$. Thus $f(X) \leqslant f'(X) + 0$ for all X, so Assumption II_i is satisfied with $\alpha_2 = 0$. Now consider Assumption I_i. Since S' is an r-cover of S, for any $X \in S$ there exists $Y_X \in S'$ such that $|Y_X| \leqslant |X|$, which means $f(Y_X) \leqslant f(X) + 0$ for all X. Hence Assumption I_i is met with $\alpha_1 = 0$. We note the basic inequalities (see the Remark below) used to obtain the relaxation and restriction are $-d(p_i, p'_i) \leqslant D(X, p_i) - D(X, p'_i) \leqslant d(p_i, p'_i)$, which hold for all X and i. We also use Theorem 3 to obtain the relaxation.

By considering Table 2 we obtain the following.

Qualitative Insight 1. $\Pr'(\tau)$ should be a good aggregation if $t_i \gg r + d(p_i, p'_i)$ for all i.

The insight follows from noticing that if the other terms on the right-hand-sides of the constraints for the problems in lines 2–4 are small compared to the t_i, then all the right-hand-side values will be essentially the same, indicating the associated problem minimal objective function values will all be in some "small" interval, say $[a, b]$. By Theorem 1, the minimal objective function value of the problem of line 1 will also be in $[a, b]$. Thus $\Pr'(\tau)$ should be a good approximation to $\Pr(\tau)$.

There is, of course, no guarantee that a feasible solution to $\Pr'(\tau)$ will be feasible to $\Pr(\tau)$. However, a direct consequence of Assumption II is the following.

Remark 2

(a) *Almost feasibility:* If X is a feasible solution to $\Pr'(\tau)$, then X is "almost" feasible to $\Pr(\tau)$, that is, $D(X, P) \leqslant \tau + \beta_2$, $X \in S$.

(b) *Feasibility:* If X is a feasible solution to $\Pr'(\tau)$ and $D(X, P') \leqslant \tau - \beta_2$, then X is feasible to $\Pr(\tau)$.

Table 3 first lists the constraints of the three aggregated problems with duplicate ADPs in column 2. The process of aggregation makes some constraints redundant. Thus column 3 shows the

Table 2
The family of covering problems

	Covering problem	Objective	Constraints of problem
1	Original: $\Pr(\tau)$	$\min \lvert X\rvert$	$D(X, p_i) \leqslant t_i$, all i; $X \in S$
2	Aggregation: $\Pr'(\tau)$	$\min \lvert X\rvert$	$D(X, p'_i) \leqslant t_i$, all i; $X \in S'$
3	Restriction: $\Pr'(\tau - \beta_2)$	$\min \lvert X\rvert$	$D(X, p'_i) \leqslant t_i - d(p_i, p'_i)$, all i; $X \in S'$
4	Relaxation: $\Pr'(\tau + \beta_1)$	$\min \lvert X\rvert$	$D(X, p'_i) \leqslant r + t_i + d(p_i, p'_i)$, all i; $X \in S'$

Table 3
Covering constraints for aggregated covering problems; S' is an r-cover of S; $I_j = \{i : p_i = q_j\}$, $j = 1, \ldots, n$

Problem	Covering constraints with		
	ADPs $p'_1, \ldots, p'_m$	Distinct ADPs: $q_1, \ldots, q_n$; $\tau_j^- = \min\{t_i - d(p_i, q_j) : i \in I_j\}$; $\tau_j = \min\{t_i : i \in I_j\}$; $\tau_j^+ = r + \min\{t_i + d(p_i, q_j) : i \in I_j\}$	All $t_i = t$ also $\tau_j^- = t - \max\{d(p_i, q_j) : i \in I_j\}$; $\tau_j = t$; $\tau_j^+ = r + t + \min\{d(p_i, q_j) : i \in I_j\}$
Restriction $\Pr'(\tau - \beta_2)$	$D(X, p'_i) \leqslant t_i - d(p_i, p'_i)$, $X \in S'$	$D(X, q_j) \leqslant \tau_j^-$, $X \in S'$	$D(X, q_j) \leqslant \tau_j^-$, $X \in S'$
Aggregation $\Pr'(\tau)$	$D(X, p'_i) \leqslant t_i$, $X \in S'$	$D(X, q_j) \leqslant \tau_j$, $X \in S'$	$D(X, q_j) \leqslant t$, $X \in S'$
Relaxation $\Pr'(\tau + \beta_1)$	$D(X, p'_i) \leqslant t_i + r + d(p_i, p'_i)$, $X \in S'$	$D(X, q_j) \leqslant \tau_j^+$, $X \in S'$	$D(X, q_j) \leqslant \tau_j^+$, $X \in S'$

constraints again with distinct ADPs. Finally column 4 shows a simplification of the constraints when the covering radii are identical. Constraint aggregation for the covering problem can reduce substantially the number of constraints. With n distinct ADPs, there are exactly n constraints for the aggregate problem, reduced from m for the original problem. Denote by $Q = \{q_1, \ldots, q_n\}$ the set of all distinct ADPs. Let I_j denote the set of all indices of DPs aggregated into q_j, that is, $I_j = \{i : p'_i = q_j\}$, $j = 1, \ldots, n$.

In Section 5, we provide examples separately for the case $r = 0$. Note $r = 0$ if $S' = S$. A consideration of Table 3, along with reasons similar to those for Qualitative Insight 1, leads to the following note and qualitative insight.

Extra structure note. If all $t_i = t$, and there is some $i \in I_j$ with $p_i = q_j$, then $\min\{d(p_i, q_j) : i \in I_j\} = 0$ and $\tau_j^+ = r + t$.

For reasons similar to those given for Qualitative Insight 1, we also have

Qualitative Insight 2. If all $t_i = t$, then Q and S' should be a good aggregation if, for all $j = 1, \ldots, n$,

$$t \gg \max\{d(p_i, q_j) : i \in I_j\},$$

$$t \gg r + \min\{d(p_i, q_j) : i \in I_j\}.$$

***K*-Center problem.** Suppose, instead of the above covering problem, we have a K-center objective $f(X) = \max\{w_i D(X, p_i) : i = 1, \ldots, m\}$, with the same distance constraints and given information of the above covering problem, and distinct ADPs $q_1, \ldots, q_n$. Define $\omega_j \equiv \max\{w_i : i \in I_j\}$ for $j = 1, \ldots, n$, and $\mathrm{eb} \equiv \max\{w_i d(p_i, q_j) : i \in I_j, j = 1, \ldots, n\}$. We then have $f'(X) = \max\{\omega_j D(X, q_j) : j = 1, \ldots, n\}$. We can take β_1 and β_2 to be the same as for the covering problem. Assuming S' is an r-cover of S, Theorem 2 applies to this problem with $\alpha_2 = \mathrm{eb}$ and $\alpha_1 = \max\{w_i(r + d(p_i, p'_i)) : i = 1, \ldots, m\}$. Note, if all $w_j = 1$ then $\alpha_1 = r + \alpha_2$.

We can obtain a smaller value of α_1 with an extra assumption. Suppose each ADP is some DP, so that $f'(X) \leqslant f(X)$ for all X. Note that Lemma 1 applies for any choice of the p'_i, say $p'_i = a_i$, $i = 1, \ldots, m$. Hence, with; $A = \{a_1, \ldots, a_m\}$, $D(Y_X, A) \leqslant D(X, P) + re + T(P, A)$. If $A = P$ then $D(Y_X, P) \leqslant D(X, P) + re$. Thus by applying any SAND function g, we can conclude that $g(D(Y_X, P)) \leqslant g(D(X, P)) + g(re)$. When g is the weighted max-function of this problem we also know that $f'(X) \leqslant f(X)$ for all X, and $f(Y_X) = g(D(Y_X, P))$, $f(X) = g(D(X, P))$. Thus $f'(Y_X) \leqslant f(Y_X) \leqslant f(X) + g(re)$. Thus (Assumption I) $\alpha_1 = g(re) = \max\{w_i r : i = 1, \ldots, m\} = r \max\{w_i : i = 1, \ldots, m\}$.

5. Computational experience

In this section we present some computational experimentation, illustrating Theorem 1 for the covering location model with rectilinear distances, for several large-scale data sets, assuming $r = 0$. The approach employs several different aggregation methods, which we outline only; see Emir-Farinas and Francis (2004) for more details. We observe from the experimentation that solving only the aggregated problem often underestimates the actual minimal objective function value. Also, it is possible, with enough ADPs (a relatively small number compared to the number of DPs), to find a restriction and relaxation of the aggregated problem that together provide an optimal solution to the original problem.

Some notation is useful. Given any ADP set Q, for each $q_j \in Q$, let I_j again denote the set of indices of DPs aggregated into ADP q_j (those closest to q_j). For each ADP q_j, define $\gamma_j = \min\{d(q_j, p_i) : p_i \in I_j\}$ and $\delta_j = \max\{d(q_j, p_i) : p_i \in I_j\}$. With reference to Table 3, we then have $\tau_j^- = t - \delta_j$, $\tau_j^+ = t + \gamma_j$ for all q_j. With reference to Table 1, because there is no aggregation of DPs in the objective function, we have $\alpha_1 = \alpha_2 = 0$. Further, the vectors β_1 and β_2 are given by $\beta_1 = (\tau_j^+)$, $\beta_2 = (\tau_j^-)$.

For any given ADP set Q, let LB and UB denote the respective optimal objective function values of the following aggregated problems:

Relaxation:

$$\text{LB} = \min |X| \ \text{s.t.}\, D(X, q_j) \leqslant t + \gamma_j, \quad q_j \in Q;$$

Restriction:

$$\text{UB} = \min |X| \ \text{s.t.}\, D(X, q_j) \leqslant t - \delta_j, \quad q_j \in Q.$$

Theorem 1 gives $\text{LB} \leqslant v(\tau)$, $v'(\tau) \leqslant \text{UB}$, where $v(\tau)$, $v'(\tau)$ are the optimal objective function values of the original and aggregated covering problems respectively.

We considered three aggregation methods, which we refer to as pick the farthest (PTF), random, and independent projection algorithm (IPA). PTF is based on the 2-approximation algorithm of Dyer and Frieze (1985), and has previously been used by Daskin et al. (1989) for aggregation for various covering problems. PTF constructs an ADP set Q as follows. The first element of Q is any randomly chosen DP. Then, if Q denotes the current ADP set, Q is augmented with a DP p_i for which $D(Q, p_i)$ is largest (p_i is the farthest DP from Q). Termination occurs when $|Q|$ is sufficiently large. The complexity of the method is $O(|Q|m)$. The Random method, used for comparison purposes and for its simplicity, makes a random selection of ADPs from among the DPs.

IPA is designed specifically for rectilinear distances. For any given planar covering problem with rectilinear distances we use a well-known linear transformation $T(x, y) = (x + y, -x + y)$ (Francis et al., 1992, p. 231) to the DPs to obtain an equivalent problem with Tchebyshev distances. The transformed DPs are then projected onto the x and the y axes. These two sets of projected points on the axes provide the input data for solving two covering problems on the real line to optimality, using a covering radius much smaller than the value t of interest. The solution to each covering problem on the real line is a collection of intervals. The Cartesian product of the two sets of intervals forms a collection of cells in the plane. For each cell that has at least one DP a (Tchebyshev) 1-center problem is solved to find the ADP for all DPs in that cell. Applying the inverse transformation $T^{-1}(u, v) = \frac{1}{2}(u - v, u + v)$ to this collection of ADPs provides the set Q of ADPs for the original covering problem with rectilinear distances. Due to ranking the DPs for the covering problems on the real line, the complexity of IPA is $O(m \log m)$. Several factors must be considered in making a choice of the covering radius ρ for each covering problem on the line. If we want the distance between each DP and its ADP to be at most ε (say), we choose $\rho \leqslant \varepsilon$. If we want approximately n ADPs, and A denotes the total area of the set containing all the DPs, we rely on a formula from Francis and Rayco (1996) and choose ρ approximately equal to $\sqrt{(A/(2n))}$. To have an aggregation of high quality, we also want $\rho \ll t$, the covering radius of the original problem. Some degree of trial and error may be necessary to find a good choice of ρ.

We used two data sets. A random problem had 50,000 DPs uniformly distributed in a square of dimensions 1000 by 1000, and a covering radius of $t = 250$. A second problem with real data, for Palm Beach County, Florida, had 69,960 DPs and a radius of 50,000 feet (15,240.2 m). The DPs are power transformers for a utility network, and closely follow the street network. The dark regions

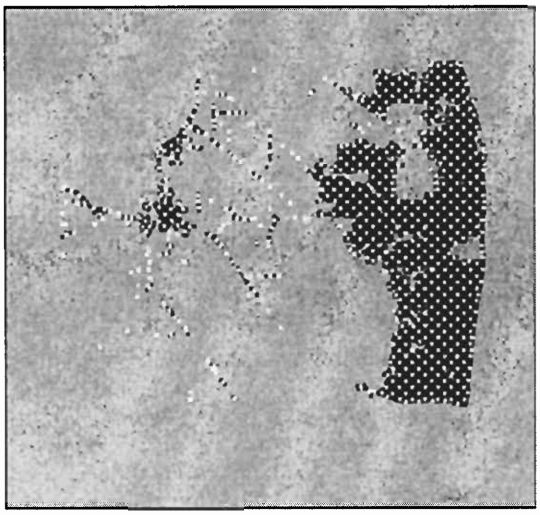

Fig. 2. DPs, and 703 ADPs located according to IPA for Palm Beach Co.

in Fig. 2 illustrate the DPs for the latter data set. The light points in Fig. 2 are the ADPs, found with IPA.

All the aggregated covering problems were solved as continuous rectilinear covering problems, using well-known reduction techniques (see Garfinkel and Nemhauser, 1972) with CPLEX (Version 7.0), and a well-known finite dominating set (FDS) principle. The FDS principle, introduced by Hooker et al. (1991) for location problems on networks, narrows the search for an optimal solution to a predetermined finite set of points. For the continuous planar location-covering problem with rectilinear distances, one version of the FDS principle is as follows. Consider the set of all ℓ_1 (rectilinear) neighborhoods, squares each with a center at some ADP and radius being the corresponding covering upper bound; each edge of each square makes a 45° angle with an axis. It can be shown that a FDS in this case consists of the set S of (a) all centers of isolated squares, together with (b) edge intersections of all pairs of nondisjoint squares. (We actually used a superset of S, by including all four vertices of the rectangle formed by the intersection of each pair of nondisjoint squares.) Note that the total number of points in S to be considered is bounded above by $|Q| + 2|Q|(|Q| - 1)/2 = |Q|^2$.

Tables 4 and 5 show the upper and lower bounds on $v(\tau)$ for the uniform and Palm Beach County data sets respectively. In the tables, N/A means $\delta_j > t$ for some ADP q_j. For IPA, the number of ADPs is chosen to be close to the nominal number of interest. For both random and PTF, the aggregated problem is identical to the lower bounding problem, since the ADPs are a subset of the DPs, and therefore all $\gamma_j = 0$. For IPA, we show the optimal value of the aggregated problem, $v'(\tau)$; it is the same as the corresponding lower bound value except for one instance.

Note that IPA and PTF usually give the best upper bounds (UBs) and lower bounds (LBs) respectively. Regardless of the aggregation level and the scheme, LBs and UBs are bounds on the same $v(\tau)$. Therefore, the maximum (minimum) of the LBs (UBs) is also a LB (an UB) on $v(\tau)$. We highlight the best lower bounds and upper bounds in the tables.

Table 4
Lower and upper bounds on $v(\tau)$ for the uniform DP data set

No. of ADPs	IPA			No. of ADPs	PTF		Random	
	LB	$v'(\tau)$	UB		LB	UB	LB	UB
53	8	8	15	50	9	28	7	N/A
151	11	11	17	150	10	18	9	20
255	10	10	16	250	11	17	10	19
349	10	10	14	350	12	16	10	18
442	10	10	**13**	450	12	15	11	17
543	12	12	14	550	12	15	11	17
661	12	12	14	650	12	15	12	17
757	12	12	14	750	12	15	11	16
863	11	12	14	850	**13**	14	11	17

Table 5
Lower and upper bounds on $v(\tau)$ for the Palm Beach County data set

No. of ADPs	IPA			No. of ADPs	PTF		Random	
	LB	$v'(\tau)$	UB		LB	UB	LB	UB
100	11	11	16	100	12	20	6	N/A
199	12	12	16	200	13	17	6	N/A
306	12	12	15	300	13	16	6	N/A
399	12	12	15	400	13	15	8	N/A
498	13	13	15	500	13	16	8	N/A
599	**14**	14	15	600	13	15	8	N/A
703	13	13	**14**	700	13	15	7	N/A
807	13	13	**14**	800	13	15	8	N/A
899	13	13	**14**	900	13	15	8	N/A

For the uniform problem, we conclude $v(\tau) = 13$. Even though the aggregate problem providing the LB has the same number of centers as the aggregate problem providing the UB, it may not be feasible. By contrast, the aggregated problem with 442 ADPs located with IPA and providing the UB gives an optimal solution to the original problem.

Similarly, the best LB and UB are 14 for the Palm Beach County problem. The aggregated problems with 703, 807 and 899 ADPs located according to IPA all give an optimal solution to the original unaggregated covering problem. Note there is no aggregation level shown in Table 5 with LB = 14 = UB. However, we did find such a level, using IPA with 2047 ADPs.

Fig. 2 shows both the DPs and the ADPs for the Palm Beach County data set, based on IPA with 703 ADPs. Fig. 3 shows the solution to the corresponding restriction, which gives a provably optimal solution to the original problem.

Figs. 4 and 5 give histograms of the gammas and deltas for the problem of Fig. 2 found using IPA. Note the variety of values in the two histograms, and the opposite skewness tendencies of the distributions. By perturbing the radius t of the aggregated problem ($\tau_j^- = t - \delta_j$, $\tau_j^+ = t + \gamma_j$) we obtain both the restricted and relaxed problem. The opposite skewness tendencies mean that the restriction is more of a perturbation than the relaxation.

In conclusion, it is clear that the quality of the solutions obtained by this bounding approach is sensitive to the number of ADPs, and that there is

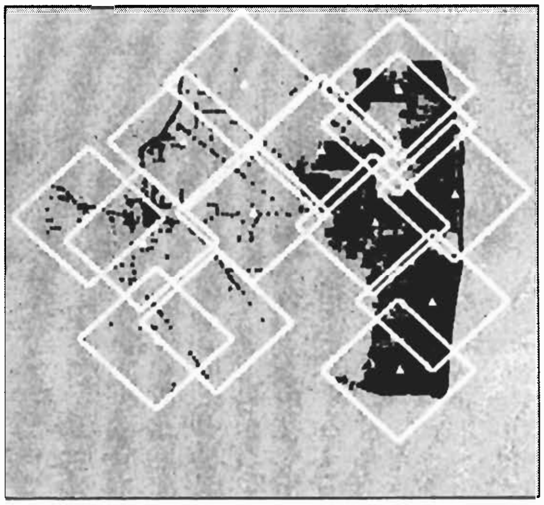

Fig. 3. Solution to aggregated problem providing UB, with radius 50K feet.

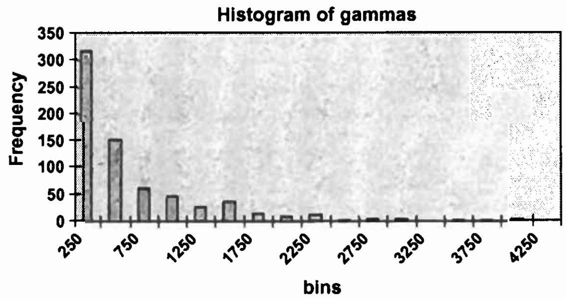

Fig. 4. Histogram of gammas for Fig. 2 problem; $\gamma_j = \min\{d(q_j, p_i) : p_i \in I_j\}$.

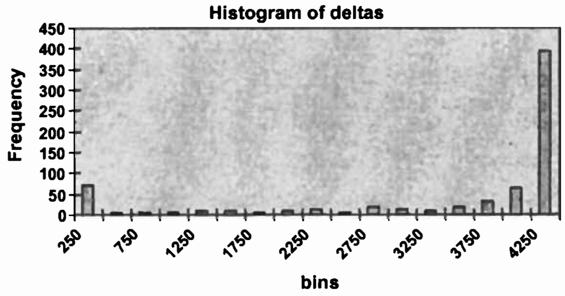

Fig. 5. Histogram of deltas for Fig. 2 problem; $\delta_j = \max\{d(q_j, p_i) : p_i \in I_j\}$.

a danger of obtaining an inaccurate solution by not using enough ADPs, or using an inferior aggregation method. Using enough ADPs can yield an optimal solution to the unaggregated original problem. If there are $|Q|$ ADPs then, prior to the use of reduction methods, the aggregated covering problem, posed as a 0–1 integer programming problem, will have $|Q|$ constraints (rows) and no more than $|Q|^2$ variables (columns). By contrast, under the same conditions the size of the original problem with m DPs is m rows and no more than m^2 columns. For the Palm Beach County problem, $m = 69{,}960$, while $|Q| \leqslant 900$.

6. Summary and conclusions

In this paper we have considered how DP, and solution space, aggregation can create error in a class of location models. We introduced two-sided error bound parameters α_1 and α_2 as measures of objective function error, and vectors β_1 and β_2 as measures of constraint and solution space aggregation error. An examination of the α and β bounds occurring in Section 3 and later shows that each such bound is nondecreasing in the terms $d(p_i, p_i')$ and in r. The conclusion is clear that the smaller these terms are the better the aggregation will be.

Via Theorem 1, we showed how constructing a relaxation and a restriction of the aggregated problem can provide optimality bounds on the original problem, and in Theorems 2 and 3, provided means of computing these bounds for SAND location models. We then considered covering and related center problems to illustrate the bounds, and for purposes of computational experimentation. These problems demonstrated that the bounds can be tight. Further, for covering and center problems, if each ADP is an original DP then the bounds of Theorems 2 and 3 can be improved. Computational experience for these problems seems encouraging. The ratio eb/t appears as if it may be a simple, useful measure for constraint aggregation error, and deserves further consideration. Because our results apply to a number of location models besides the covering and center models, there are considerable opportunities for further computational experimentation.

We have assumed that DP aggregations are given. While this may be true, often they must be determined. The following literature deals with determining such aggregations for specific models having SAND structure: Francis and Rayco (1996), Francis et al. (1999), Rayco et al. (1997, 1999), Zhao and Batta (1999, 2000). Our work now opens the question of how to determine good aggregations for the class of SAND models we have studied.

Acknowledgements

This research is supported in part by the US National Science Foundation, Grant DMI-9908124.

References

Bowerman, R.L., Calamai, P.H., Hall, B., 1999. The demand partitioning method for reducing aggregation errors in p-median problems. Computers and Operations Research 26, 1097–1111.

Current, J.R., Schilling, D.A., 1987. Elimination of source A and B errors in p-median problems. Geographical Analysis 19, 95–110.

Current, J.R., Schilling, D.A., 1990. Analysis of errors due to demand point aggregation in the set covering and maximal covering location problems. Geographical Analysis 22, 116–126.

Daskin, M.S., Haghani, A.E., Khanal, M., Malandraki, C., 1989. Aggregation effects in maximum covering models. Annals of Operations Research 18, 115–139.

Dyer, M., Frieze, A.M., 1985. A simple heuristic for the p-center problem. Operations Research Letters 3, 285–288.

Emir-Farinas, H., Francis, R.L., 2004. Demand point aggregation for some covering location problems. Annals of Operations Research, forthcoming.

Erkut, E., Bozkaya, B., 1999. Analysis of aggregation errors for the *p*-median problem. Computers and Operations Research 26, 1075–1096.

Francis, R.L., Lowe, T.J., 1992. On worst-case aggregation analysis for network location problems. Annals of Operations Research 40, 229–246.

Francis, R.L., Rayco, M.B., 1996. Asymptotically optimal aggregation for some unweighted *p*-center problems with rectilinear distances. Studies in Locational Analysis 10, 25–36.

Francis, R.L., Lowe, T.J., Rushton, G., Rayco, M.B., 1999. Aggregation methods for multi-facility location problems: Strategies for containing error. Geographical Analysis 31, 67–87.

Francis, R.L., Lowe, T.J., Tamir, A., 2000. Aggregation error bounds for a class of location models. Operations Research 48, 294–307.

Francis, R.L., Lowe, T.J., Tamir, A., 2002. Demand point aggregation of location models. In: Drezner, Z., Hamacher, H. (Eds.), Facility Location: Applications and Theory. Springer-Verlag.

Francis, R.L., McGinnis, L.F., White, J.A., 1992. Facility Layout and Location: An Analytical Approach, second ed. Prentice Hall, New Jersey.

Garfinkel, R.S., Nemhauser, G.L., 1972. Optimal set covering: A survey. In: Geoffrion, A.M. (Ed.), Perspectives on Optimization: A Collection of Expository Articles. Addison Wesley, Reading, MA.

Hillsman, E.L., Rhoda, R., 1978. Errors in measuring distances from populations to service centers. Annals of Regional Science 12, 74–88.

Hodgson, M.J., Neuman, S., 1993. A GIS approach to eliminating source C aggregation error in *p*-median models. Location Science 1, 155–170.

Hooker, J.N., Garfinkel, R.S., Chen, C.K., 1991. Finite dominating sets for network location problems. Operations Research 39, 100–118.

Kolen, A., Tamir, A., 1990. Covering problems. In: Mirchandani, P.B., Francis, R.L. (Eds.), Discrete Location Theory. John Wiley & Sons, New York.

Plastria, F., 2001. On the choice of aggregation points for continuous *p*-median problems: A case for the gravity center. TOP 9, 217–242.

Rayco, M.B., Francis, R.L., Lowe, T.J., 1997. Error-bound driven demand point aggregation for the rectilinear distance *p*-center model. Location Science 4, 213–235.

Rayco, M.B., Francis, R.L., Tamir, A., 1999. A *p*-center grid-positioning aggregation procedure. Computers and Operations Research 26, 1113–1124 (Special Issue on Aggregation & Disaggregation for Operations Research).

Zemel, E., 1985. Probabilistic analysis of geometric location problems. SIAM Journal on Algebraic and Discrete Methods 6, 189–200.

Zhao, P., Batta, R., 1999. Analysis of centroid aggregation for the Euclidean distance *p*-median problem. European Journal of Operational Research 113, 147–168.

Zhao, P., Batta, R., 2000. An aggregation approach to solving the network *p*-median problem with link demands. Networks 36, 233–241.