GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Institut für Numerische und Angewandte Mathematik

Locating a median line with partial coverage distance

J. Brimberg, R. Schieweck, A. Schöbel

Nr. 32

Preprint-Serie des
Instituts für Numerische und Angewandte Mathematik
Lotzestr. 16-18
D - 37083 Göttingen

Locating a median line with partial coverage distance

Jack Brimberg[*†] Robert Schieweck[‡§¶] Anita Schöbel[‡¶]

December 2013

We generalize the classical median line location problem, where the sum of distances from a line to some given demand points is to be minimized, to a setting with *partial coverage distance*. In this setting, a demand point within a certain specified threshold distance r of the line is considered covered and its partial coverage distance is considered to be zero, while non-covered demand points are penalized an amount proportional to their distance to the covered region. The sum of partial coverage distances is to be minimized. We consider general norm distances as well as the vertical distance and extend classical properties of the median line location problem to the partial coverage case. We are finally able to derived a finite dominating set. While a simple enumeration of the finite dominating set takes $O(m^3)$ time, m being the number of demand points, we show that this can be reduced to $O(m^2 \log m)$ in the general case by plane sweeping techniques and even to $O(m)$ for the vertical distance and block norm distances by linear programming.

Keywords median line location, partial coverage, finite dominating set, plane sweeping, block norm

1 Introduction

Locating a straight line in continuous space in order to represent (or to fit) a given set of fixed points is a well-studied problem in facilities location theory. The problem was

[*]Royal Military College of Canada, 13 General Crerar Crescent Kingston, ON, K7K 7B4 Canada

[†]This author was supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (NSERC #205041-2008)

[‡]Institute for Numerical and Applied Mathematics, University of Göttingen, Lotzestr. 16–18, 37083 Göttingen, Germany

[§]This author was supported by the DFG through the Research Training Group 1023

[¶]This author was supported by the European Union Seventh Framework Programme (FP7-PEOPLE-2009-IRSES) under grant agreement no. 246647 within the OptALI project.

originally posed in terms of a minimization of the sum of weighted Euclidean distances from a set of fixed points (demand points or customers) to a line in the plane $\mathbb{R}^2$ [MT83, LC85]. The minmax criterion, i.e., minimizing the maximum distance to the line, has also been investigated [HT88]. In facility location, the linear facility, or line, can represent a new road or railway track through a given region, a dense series of communication towers, a main connector in an electrical circuit board, and so on. For an introduction to the linear facility location problem, see for example [LMW88]. Extensions of the basic model include the use of arbitrary norms in place of Euclidean distances [Sch99] and the location of a line in $\mathbb{R}^3$ [BJS02, BJS03, BCSS11]. A related problem to line location in the plane is the location of a hyperplane in $\mathbb{R}^n$, [MS98, MS01] and references therein. Locating a linear facility on a sphere translates to the location of a great circle on the sphere [BJS07].

In this paper we investigate a new extension of the median line problem in the plane where 'partial coverage distance' is used as the distance measure. The concept of partial coverage was recently introduced by [BJKS13] for the location of a point facility. The idea is that the distance function between a customer and the new facility is zero if the customer is within a threshold value r of the facility, and otherwise, becomes the closest distance from the customer to the boundary of a disc of diameter $2r$ centered at the facility. Thus, the partial coverage model attempts to combine the notion of coverage with the median objective. For further details and motivation see [BJKS13].

We extend the partial coverage concept to a linear facility in a straightforward manner. If a customer is within a specified threshold distance of r to the line, we consider that the customer is happy (i.e., covered) and associate a zero cost to that customer. Otherwise, a penalty is assessed which is proportional to the distance to the line in excess of r. The objective is then to minimize the sum of these penalty costs.

The location of a line with partial coverage is equivalent in effect to the location of a 'thick' linear facility which is a *strip* of width $2r$ (as measured by the given distance function). Thus, we obtain a new type of problem, which may also be referred to as *median strip location problem*, and which may have interesting applications in areas other than facility location such as regression analysis. Another interpretation of the problem is as follows: replace each fixed point by a disc of diameter $2r$ centered at the point; then locate a median line with respect to the discs. This problem has been studied for the Euclidean distance in [RT94].

It is well-known that an optimal solution of the median line problem always exists where two of the fixed points are coincident with the line when distance is measured by an arbitrary norm. Furthermore, if the norm is a smooth one, this incidence property even becomes a necessary condition [Sch99, MS99]. This important result reduces the problem to a search through a finite number of candidate solutions. One of the main findings of this paper is that the incidence property may be generalized to line location with partial coverage. Based on this, a finite dominating set can be developed also for the more general problem, although the candidate solutions are different than before.

The paper is organized as follows: Section 2 summarizes the notation and presents the mathematical model; Section 3 analyzes the model and establishes the incidence property and the finite dominating set. Section 4 presents different approaches derived from the properties obtained in the previous sections. These are an enumeration of the finite candidate set, an efficient sweep algorithm, and a linear programming formulation for the case of block norm distances. Finally, Section 5 gives our conclusions and some thoughts on further research.

2 The line location problem with partial coverage distance

The given information consists of a set of m fixed points (also demand points or customers) with known locations $p_i = (p_{i1}, p_{i2})^T \in \mathbb{R}^2$, $i = 1, \ldots, m$, and given weights $w_i > 0$, $i = 1, \ldots, m$, which may, for example, represent a demand for service or a flow in the location model. A *threshold value* r is also specified. If the distance to the linear facility exceeds r, a penalty cost proportional to the excess distance is charged; otherwise, the point is covered and the cost is zero. The distance to the line refers either to the closest distance measured by some given (arbitrary) norm, or to the *vertical distance* to the line. The latter is added because of its importance in regression analysis.

Let L denote the line (i.e., linear facility) to be located. We may define this line in terms of an unknown normal vector $n = (n_1, n_2)^T \in \mathbb{R}^2 \setminus \{0\}$ and some number $c \in \mathbb{R}$:

$$L = \{x \in \mathbb{R}^2 : n^T x = c\}. \tag{1}$$

If L is not a vertical line ($n_2 \neq 0$), we may set $n_2 = 1$ and obtain an equivalent representation:

$$L = \{x \in \mathbb{R}^2 : x_2 = sx_1 + b\} \tag{2}$$

where s and b are, respectively, the slope and intercept of the line. We write $L_{n,c}$ for a line in the normal vector parametrization or, if possible without ambiguity, $L_{s,b}$ for a line in the slope-intercept parametrization.

Let $k : \mathbb{R}^2 \to \mathbb{R}$ be a norm in the plane. Then the distance from any point $p \in \mathbb{R}^2$ to line L with respect to norm k is defined as

$$d(p, L) = \min\{k(p - x) : x \in L\}, \tag{3}$$

that is, the distance is the shortest one between p and L measured by the norm k. In [Man99] and [PC01] it is shown that for a line $L = L_{n,c}$ this distance can be computed as

$$d(p, L_{n,c}) = \frac{|n^T p - c|}{k^\circ(n)} \tag{4}$$

where $k^\circ : \mathbb{R}^2 \to \mathbb{R}^2$ denotes the dual norm of k. For the case of the vertical distance between a point p and a non-vertical line L (i.e., $n_2 \neq 0$) we obtain

$$d(p, L_{n,c}) = \frac{|n^T p - c|}{|n_2|} \tag{5}$$

or equivalently $d(p, L_{s,b}) = |p_2 - sp_1 - b|$ in the (s, b)-representation. Hence, the distance between a point and a (non-vertical) line with respect to a norm k is given as the vertical distance divided by the dual norm $k^\circ((s, -1)^T)$ of the normal vector. Notice that point-line distance, no matter if vertical or measured with a norm, can be written in the form (4) where k° is a convex function of n, namely the dual norm in the norm distance case and $k^\circ(n) := |n_2|$ in the vertical distance case. We will make use of this in the next sections to unify some of the proofs.

The *partial coverage distance* between p and L is defined as follows for either case:

$$D(p, L) = \max\{d(p, L) - r, 0\} = [d(p, L) - r]^+, \tag{6}$$

where $[a]^+ = \max\{a, 0\}$ for all $a \in \mathbb{R}$. The new problem which we study here replaces the standard distance from point to line by the partial coverage distance:

$$\text{(MLPC)} \qquad \min_{\text{line } L} \sum_{i=1}^{m} w_i D(p_i, L). \tag{7}$$

We refer to the problem above as the *median line problem with partial coverage (MLPC)* and denote its objective by $f(L)$ or also

$$f(n, c) = \sum_{i=1}^{m} w_i \left[\frac{|n^T p_i - c|}{k^\circ(n)} - r \right]^+ \tag{8}$$

if L is parametrized by normal n and a real number c.

Since $(MLPC)$ is a generalization of a well-known line location problem, we will first of all explore connections to previously treated problems from the literature.

Lemma 1 (Connections to other line location problems)**.** *Let r_{max} be the optimal objective value of the* unweighted center line location problem *for points $p_1, \ldots, p_n$. Let a radius r be given and denote by z^* the optimal objective value of $(MLPC)$ for the same points, but possibly with weights. Then the following hold.*

1. *If $r = 0$, then $(MLPC)$ is the median line location problem.*

2. *$r < r_{max}$ if and only if $z^* > 0$ and $r \geq r_{max}$ if and only if $z^* = 0$.*

3. *If $r \geq r_{max}$ then any optimal solution to the unweighted center line location problem is optimal for $(MLPC)$. In case of $r = r_{max}$ a line L is optimal for $(MLPC)$ if and only if L is optimal for the unweighted center line problem.*

Proof. 1. This is obvious since $D(p, L) = d(p, L)$ for $r = 0$.

2.

$$\begin{aligned}
z^* = 0 \quad & \Leftrightarrow \quad \exists\, L\ : D(p_i, L) = 0 && \forall\, i = 1, \ldots, m \\
& \Leftrightarrow \quad \exists\, L\ : d(p_i, L) \leq r && \forall\, i = 1, \ldots, m \\
& \Leftrightarrow \quad \exists\, L\ : \max_{i=1,\ldots,m} d(p_i, L) \leq r \\
& \Leftrightarrow \quad r_{\max} = \min_{L} \max_{i=1,\ldots,m} d(p_i, L) \leq r
\end{aligned}$$

which proves both assertions because $f \geq 0$.

3. This can be seen easily from a calculation similar to that of the second case.

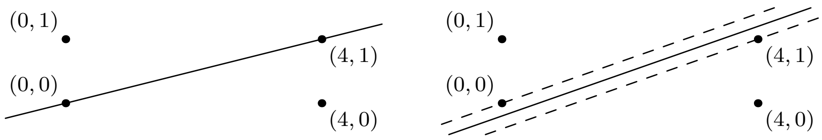
$\square$

This allows us to calculate an optimal solution to $(MLPC)$ in $O(m \log m)$ time as center line if $r \geq r_{\max}$ [Sch99]. If the threshold r is even larger, namely

$$r \geq \bar{r} := \max_{i,j=1,\ldots,n} d(p_i, p_j)$$

the problem becomes trivial and can be solved in constant time: every line L passing through one of the fixed points, say p_k, is optimal because $d(p_i, L) \leq d(p_i, p_k) \leq \bar{r}$ for all $i = 1, \ldots, m$. The interesting cases are those with some intermediate r, namely $r \in (0, r_{\max})$.

Before we turn to properties which help us to solve $(MLPC)$, it is worth noting, that there is no obvious analogon to Lemma 1.2 which allows us to solve $(MLPC)$ for small $r > 0$ by solving a *median* line location problem. More precisely, for any $r > 0$ – however small – the solution to $(MLPC)$ need not coincide with the solution to the median line location problem, i.e. $(MLPC)$ with $r = 0$, see Figure 1 for an example with $m = 4$ and $w_i \equiv 1$ using the Euclidean distance as a norm. This is a main difference to the location of a point with partial coverage where it is known that the solution of the partial coverage problem coincides with the solution of the median point problem for small coverage radii, see [BJKS13].



(a) An optimal median line.

(b) An optimal partial coverage line for small $r > 0$. Coverage zone is between dashed lines.

Figure 1: An optimal median line that does not stay optimal for $(MLPC)$ even for arbitrarily small $r > 0$.

Note that it is nevertheless possible under certain conditions that a median line remains optimal when allowing $r > 0$. It can be shown that this is the case for an odd number m of unweighted demand points in general position. There are also other (trivial) conditions, where optimality remains unchanged, e.g. if all p_i are collinear.

3 Properties and a finite dominating set

In order to be able to solve $(MLPC)$, we first derive some helpful properties. This section is inspired by related work that has been done on the median line problem without partial coverage and shows that many properties which hold for this problem can be generalized to our case. We start by showing an analogon to the well-known pseudo-halving property which states that every median line cuts the set of demand points approximately in half. To this end, denote for a fixed line L by

$$I_L^+ = \{i : d(p_i, L) > r \text{ and } p_i \text{ lies above } L\}$$

the index sets of demand points above L and having a strictly greater distance to L than the threshold value r. Furthermore, let analogously $\overline{I}_L^+$ be the index set of demand points above L with distance greater or equal than the threshold distance r, and let I_L^- and $\overline{I}_L^-$ be the corresponding index sets for demand points below L. Note that *above* should be replaced by *to the right* to deal with vertical lines. For each of these index sets the sum of corresponding weights is denoted by W_L^+, $\overline{W}_L^+$, W_L^-, and $\overline{W}_L^-$, respectively, e.g.

$$W_L^+ = \sum_{i \in I_L^+} w_i.$$

Recall that in the case $r = 0$, i.e. the median line location problem without partial coverage, every optimal line L satisfies $W_L^+ \leq \overline{W}_L^-$ and $W_L^- \leq \overline{W}_L^+$, see e.g. [Sch99]. The following theorem extends this to the case $r > 0$, i.e. the partial coverage setting.

Theorem 2 (Pseudo-halving property)**.** *Let L be optimal for $(MLPC)$. Then $W_L^+ \leq \overline{W}_L^-$ and $W_L^- \leq \overline{W}_L^+$.*

Proof. To show this, we prove the stronger statement, that a line L is optimal for $(MLPC)$ *restricted to some normal* n if and only if $W_L^+ \leq \overline{W}_L^-$ and $W_L^- \leq \overline{W}_L^+$. Then the assertion follows since an optimal line L (which has some normal, say n) for $(MLPC)$ must be optimal for $(MLPC)$ restricted to that normal n.

Clearly, for each partial coverage distance discussed here, induced by a norm or vertical, the objective $f(n, c)$ is convex in the right hand side c for each normal n, see (4) and (5). Hence every locally optimal solution is globally optimal for fixed n. We show now that the assertion of the theorem is equivalent to a local optimality condition.
Take some line $L = L_{n,c}$, w.l.o.g. n pointing upwards. Then there is an $\delta > 0$ such that for every upward translation $L^\varepsilon = L_{n,c+\varepsilon}$, $0 < \varepsilon < \delta$ of L the sets I_L^+ and $I_{L^\varepsilon}^+$ as well as $\overline{I}_L^-$ and $\overline{I}_{L^\varepsilon}^-$ coincide. Then it holds

$$\begin{aligned}
f(n, c+\varepsilon) &= \sum_{i \in I_{L^\varepsilon}^+} w_i \left(\frac{n^T p_i - c - \varepsilon}{k^\circ(n)} - r \right) - \sum_{i \in \overline{I}_{L^\varepsilon}^-} w_i \left(\frac{n^T p_i - c - \varepsilon}{k^\circ(n)} - r \right) \\
&= \sum_{i \in I_L^+} w_i \left(\frac{n^T p_i - c}{k^\circ(n)} - r \right) - \sum_{i \in \overline{I}_L^-} w_i \left(\frac{n^T p_i - c}{k^\circ(n)} - r \right) - \frac{\varepsilon(W_L^+ - \overline{W}_L^-)}{k^\circ(n)} \\
&= f(n, c) - \frac{\varepsilon(W_L^+ - \overline{W}_L^-)}{k^\circ(n)}.
\end{aligned}$$

We hence obtain that $f(n,c) \leq f(n,c+\varepsilon)$ if and only if $W_L^+ - \overline{W}_L^- \leq 0$. A similar calculation for downward translations yields $f(n,c-\varepsilon) = f(n,c) - \varepsilon(W_L^- - \overline{W}_L^+)/k^\circ(n)$ and hence $f(n,c) \leq f(n,c-\varepsilon)$ if and only if $W_L^- \leq \overline{W}_L^+$. Together, L is optimal for the normal-restricted $(MLPC)$ if and only if $W_L^+ \leq \overline{W}_L^-$ and $W_L^- \leq \overline{W}_L^+$. □

We will now come to one of the most important results from an algorithmically point of view, the *incidence property*. It is an extension of the incidence property for the median line location problem which states that there always is an optimal line which passes through two of the demand points [MN80, Sch99], thus giving rise to a finite dominating set and a simple enumeration algorithm. In our case with partial coverage we obtain the existence of an optimal line which has two of the given points at threshold distance r to the line.

Theorem 3 (Incidence property)**.** *If $r < r_{max}$ there is an optimal line L for $(MLPC)$ such that $d(p_j, L) = d(p_k, L) = r$ for at least two $j, k \in \{1, \ldots, m\}$ such that $p_j \neq p_k$. If d is induced by a smooth norm, then* every *optimal line satisfies this criterion.*

Proof. Without loss of generality, suppose that all p_i, $i = 1, \ldots, m$, are distinct. Otherwise duplicate points may be replaced by a single point with adjusted weight.

Similarly to the proof of the pseudo-halving property (Theorem 2), any line L may be shifted upwards or downwards without deteriorating the objective until there is one demand point, say p_j, at threshold distance from L, i.e. $d(p_j, L) = r$. We now perturb an optimal line $\tilde{L} = L_{\tilde{n},\tilde{c}}$ while keeping p_j at distance r from L – and still not increasing the objective – until we reach another point, say p_k, also at threshold distance. The argument is based on the observation that this perturbation is a locally quasi-concave process since the objective function is locally the ratio of a non-negative linear function and a positive convex function, see [ADSZ88]. This is now justified formally. Let $\tilde{L} = L_{\tilde{n},\tilde{c}}$ be fixed and consider any (n,c) in the region

$$R := \{(n',c') : I^+_{L_{n',c'}} = I^+_{\tilde{L}} \text{ and } I^-_{L_{n',c'}} = I^-_{\tilde{L}}\}.$$

For every $(n,c) \in R$ we obtain

$$\begin{aligned} f(n,c) &= \sum_{i \in I^+_{\tilde{L}}} w_i\Big(\frac{n^T p_i - c}{k^\circ(n)} - r\Big) + \sum_{i \in I^-_{\tilde{L}}} w_i\Big(\frac{c - n^T p_i}{k^\circ(n)} - r\Big) \\ &= \frac{1}{k^\circ(n)}\Big(\sum_{i \in I^+_{\tilde{L}}} w_i n^T p_i - \sum_{i \in I^-_{\tilde{L}}} w_i n^T p_i + c(W^-_{\tilde{L}} - W^+_{\tilde{L}}) - r k^\circ(n)(W^+_{\tilde{L}} + W^-_{\tilde{L}})\Big). \end{aligned} \tag{9}$$

We are interested in the behavior of $f(n,c)$ under the constraint that p_j is at threshold distance from $L_{n,c}$, i.e. under the constraint $n^T p_j - c = r k^\circ(n)$ if we assume w.l.o.g.

that p_j lies above $\tilde{L}$. Plugging this into (9) yields

$$\begin{aligned} f(n,c) &= \frac{1}{k^\circ(n)}\Big(\sum_{i\in I^+_{\tilde{L}}} w_i n^T p_i - \sum_{i\in I^-_{\tilde{L}}} w_i n^T p_i \\ &\quad + (n^T p_j - r k^\circ(n))(W^-_{\tilde{L}} - W^+_{\tilde{L}})\Big) - r(W^+_{\tilde{L}} + W^-_{\tilde{L}}) \\ &= \frac{1}{k^\circ(n)}\Big(\sum_{i\in I^+_{\tilde{L}}} w_i n^T(p_i - p_j) - \sum_{i\in I^-_{\tilde{L}}} w_i n^T(p_i - p_j) - 2rk^\circ(n)W^-_{\tilde{L}}\Big) \end{aligned} \tag{10}$$

which is quasi-concave on any region for n on which the index sets do not change. Thus a minimum is attained on the boundary of this region. Note that these regions are full-dimensional. Such a boundary corresponds to another point p_k being at distance r from the minimizing line.
To obtain the stronger result for smooth norms k, note that k° is strictly convex in this case and thus (10) is strictly quasi-concave in n, see [ADSZ88]. Hence a minimum is only attained on the boundary. □

A direct consequence of the incidence properties is an extension of the so-called *halving property* for smooth norms [Sch99]. This property states that every optimal line in the model without partial coverage satisfies $W^+_L < \overline{W}^-_L$ and $W^-_L < \overline{W}^+_L$. This holds also in the partial coverage setting.

Corollary 4 (Halving property)**.** *If d is induced by a smooth norm, then every optimal line L for $(MLPC)$ satisfies $W^+_L < \overline{W}^-_L$ and $W^-_L < \overline{W}^+_L$.*

Proof. Suppose there is an optimal line L with $W^+_L = \overline{W}^-_L$. Similar to the proof of Theorem 2, this line stays optimal if shifted by some small $\varepsilon > 0$ upwards where ε can be chosen so that there is no point at threshold distance r from L anymore. This is a contradiction to Theorem 3 and the case $W^-_L = \overline{W}^+_L$ is settled analogously. □

While Theorem 3 yields a finite dominating set in the median line location model without partial coverage, this is unfortunately not true in the partial coverage case, i.e. $r > 0$. More precisely, a line is not necessarily determined by two points which are at fixed distance r of the line.

For $d = d_{\text{ver}}$ this is the case if two demand points, $p = (x_p, y_p)^T$ and $q = (x_q, y_q)^T$, are vertically aligned, i. e. $x_p = x_q$ and $|y_p - y_q| = 2r$. Then any line L passing through $(x_p, \frac{y_p+y_q}{2})^T$ will satisfy $d_{\text{ver}}(p, L) = d_{\text{ver}}(q, L) = r$. Fortunately, this is the only case in which infinitely many lines have the same distance r from both p and q. If $|y_p - y_q| \neq 2r$ there is in fact no line with this property since if there was such a line L it would have to pass the midpoint of the segment joining p and q, otherwise $d(p, L) \neq d(q, L)$. But then $2r = d(p, L) + d(q, L) = |y_p - y_q| \neq 2r$, a contradiction. If on the other hand $x_p \neq x_q$, then a line with $d(p, L) = d(q, L) = r$ has to have p and q either above or below it, respectively. If e. g. p is below L then L has to pass through $(x_p, y_p + r)$. Thus there are 4 lines in total with p and q at distance r of it: each such line must pass through $(x_p, y_p + r)$ or $(x_p, y_p - r)$ and also through $(x_q, y_q + r)$ or $(x_q, y_q - r)$.

Theorem 6 settles the question for norm distances d. In order to proof it, we need a generalized version of the Cauchy-Schwarz inequality which can be found for example in [Mic93].

Lemma 5 (Generalized Cauchy-Schwarz)**.** *Let k be a norm on $\mathbb{R}^2$ and k° its dual. Then $v^T w \leq k(v)k^\circ(w)$ for any $v, w \in \mathbb{R}$ and if $v, w \neq 0$ equality holds if and only if $w = \lambda z$ for some $\lambda \in \mathbb{R}$ and a subgradient $z \in \partial k(v)$.*

Theorem 6 (Solution count)**.** *Let p and q be two distinct points in $\mathbb{R}^2$ and $r > 0$. Let k be a norm on $\mathbb{R}^2$.*

1. *If $2r > k(p - q)$, there are exactly two lines with p and q at distance r.*
2. *If $2r < k(p - q)$, there are exactly four lines with p and q at distance r.*
3. *If $2r = k(p - q)$ and k is smooth at $p - q$, there are exactly three lines with p and q at distance r.*
4. *If $2r = k(p - q)$ and k is non-smooth at $p - q$, there are infinitely many lines with p and q at distance r.*

For an illustration of Theorem 6, see Figure 2.

Proof. Given two points p and q, there always exist two lines with p and q at distance r, regardless of $k(p - q)$: both are parallel to the line joining p and q, since

$$d(p, L) = \frac{n^T p - c}{k^\circ(n)} = r \quad \text{and} \quad d(q, L) = \frac{n^T q - c}{k^\circ(n)} = r \qquad \Rightarrow \qquad n^T(p - q) = 0,$$

and one is translated such that both p and q lie below it, the other is translated such that both p and q lie above it. The amount of translation must be chosen such that $d(p, L) = d(q, L) = r$.
It is more interesting to analyze if there exist lines L such that p and q lie on different sides of L. The number of such lines has to be added to each of the four cases stated in the theorem, namely 0, 2, 1, or infinitely many, as proved below. Together with the two lines with p and q on the same side of L this gives the stated number of lines.
If p and q are on different sides of L it holds

$$d(p, L) = \frac{n^T p - c}{k^\circ(n)} \qquad \text{and} \qquad d(q, L) = \frac{c - n^T q}{k^\circ(n)}$$

after appropriately selecting n. This means that $d(p, L) = d(q, L) = r$ if and only if $n^T(p - q)/k^\circ(n) = 2r$. We are hence interested in the solutions of

$$g(n) := \frac{n^T(p - q)}{k^\circ(n)} = 2r \tag{11}$$

and distinguish the following three cases.

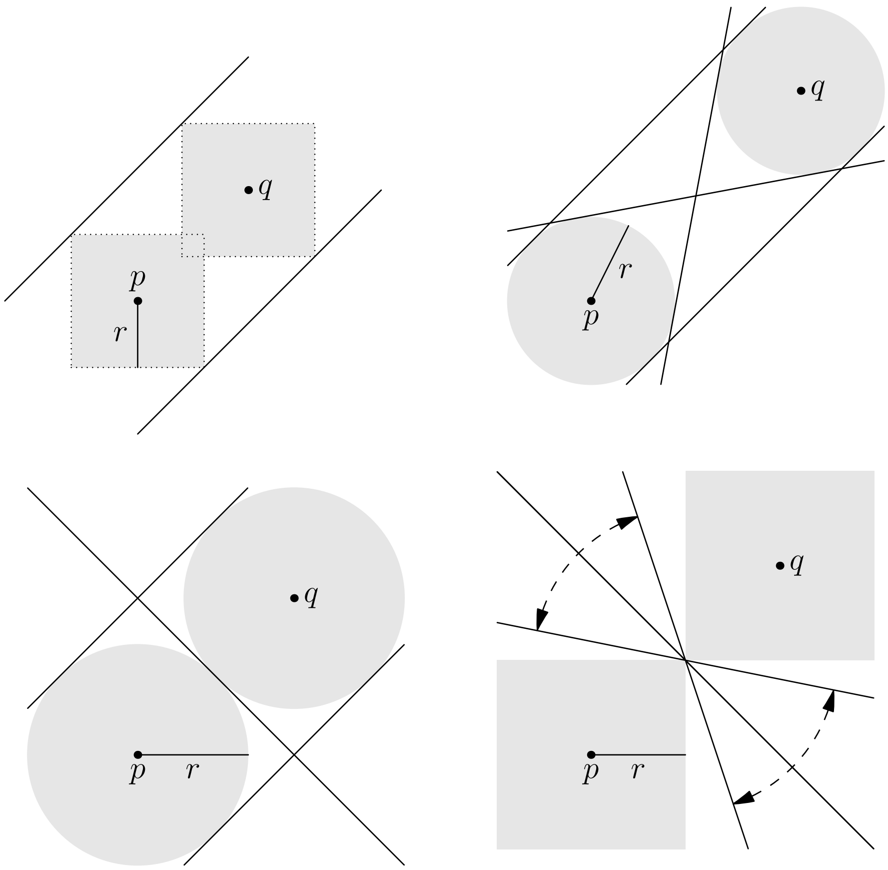


Figure 2: An illustration of the four cases of Theorem 6, cases 1 through 4 from top left to bottom right.

1. If $k(p-q) < 2r$ then (11) does not have a solution. Assume it had a solution. Then we get $n^T(p-q) = 2rk^\circ(n) > k(p-q)k^\circ(n)$ which is a contradiction to the Cauchy-Schwarz inequality, Lemma 5.

2. In the case $k(p-q) > 2r$ we construct two solutions each of which determines a different line. To this end, choose some subgradient $v \in \partial k(p-q)$ and choose $w \perp v$ as well as $u \perp (p-q)$ to define the cones

$$\begin{aligned} A :=&\{z \in \mathbb{R}^2 : z^T(p-q) \geq 0,\ z^T w \geq 0\} \setminus \{0\} \quad \text{and} \\ B :=&\{z \in \mathbb{R}^2 : z^T(p-q) \geq 0,\ z^T w \leq 0\} \setminus \{0\}. \end{aligned}$$

Now $u, v \in A$ and $-u, v \in B$ (if not, substitute v by $-v$ or w by $-w$) and it is $g(u) = g(-u) = 0$. Furthermore, $v \in \partial k(p-q)$ (or $-v \in \partial k(p-q)$), hence the generalized Cauchy-Schwarz inequality $v^T(p-q) \leq k^\circ(v)k(p-q)$ holds with equality. This yields $g(v) = \frac{v^T(p-q)}{k^\circ(v)} = k(p-q)$. Since A and B are both connected, for any r satisfying $0 < 2r < k(p-q)$ there are $n_A \in A$ and $n_B \in B$ with $g(n_A) = g(n_B) = 2r$ by the intermediate value theorem (see e.g. [Fle77, Thm. 2.8]). To show that n_A and n_B determine distinct lines, we have to show that there is no $\lambda \in \mathbb{R} \setminus \{0\}$, so that $n_A = \lambda n_B$. Assume on the contrary that such a λ exists. If $\lambda > 0$, then $n_A \in B$ and hence in $A \cap B$. Since $A \cap B$ contains only multiples of $v \in \partial k(p-q)$, this leads to $g(n_A) = k(p-q)$, again by the equality part of the Cauchy-Schwarz inequality, and thus to a contradiction to $g(n_A) = 2r < k(p-q)$. If $\lambda < 0$, then $n_A = \lambda n_B \notin A$ gives another contradiction and thus we have found two distinct lines with normals n_A and n_B, respectively.

Now we show that there are not more than two lines having p and q on different sides and distance r to both of them. Suppose there are three lines L_1, L_2 and L_3. Since they all have the same distance from p and q and also p and q lie on different sides of each L_i, they all intersect in the midpoint $x = \frac{1}{2}(p+q)$ of the line segment joining p and q. This holds since, w.l.o.g.,

$$n^T p - c = k^\circ(n)r \quad \text{and} \quad c - n^T = k^\circ(n)r \quad \Rightarrow \quad n^T(p-q) = 2rk^\circ(n).$$

It follows that

$$n^T\Big(\frac{p+q}{2}\Big) = n^T\Big(p + \frac{q-p}{2}\Big) = n^T p - \frac{1}{2}n^T(p-q) = n^T p - k^\circ(n)r = c.$$

Observe that x cannot be the closest point of any L_i to p since otherwise p would have distance $k(p-x) > r$ from each line.

Now we derive a contradiction to the statement that all three lines have distance r from p. Notice that any ray emanating from p which is not parallel to any of the three lines intersects each L_i in a different point y_i if it does not pass through x. Let the L_i be ordered in increasing order of the distances $k(y_i - p)$. Then the index $i = 2$ is the same for all such rays, i.e., L_2 is always between L_1 and L_3, see Figure 3. Consider the specific ray joining p and the norm-closest point y on L_2, i.e., $y = y_2$. Then $d(p, L_1) \leq k(y_1 - p) < k(y_2 - p) = d(p, L_2)$, a contradiction.

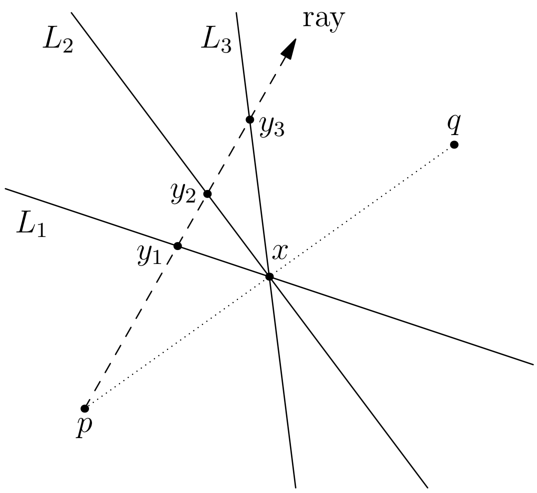


Figure 3: The line L_2 is always between lines L_1 and L_3.

3. Now let $k(p-q) = 2r$. In this case, the number of lines determined by the solutions of (11) equals the cardinality of the set $\partial k(p-q)$:
   - Let n be a solution to (11). Then $n^T(p-q) = 2rk^\circ(n) = k(p-q)k^\circ(n)$, hence the generalized Cauchy-Schwarz inequality holds with equality, and we conclude that $\lambda n \in \partial k(p-q)$.
   - On the other hand, if $\lambda n \in \partial k(p-q)$, then we know that $n^T(p-q) = k(p-q)k^\circ(n) = 2rk^\circ(n)$, i.e. (11) holds.

   If k is smooth at $p-q$ we have $|\partial k(p-q)| = 1$ and the number of lines are three, and if k is non-smooth at $p-q$, $|\partial k(p-q)| = \infty$ and hence infinitely many lines with p and q at distance r from them.

□

Now we are left with a finite candidate set except in the last case of Theorem 6 in which there is still one degree of freedom left. The following theorem exploits the local quasi-concavity of the point-line distance again and strengthens the assertions of Theorem 3 by fixing that degree of freedom.

Theorem 7 (Finite candidate set). *There is an optimal line L for $(MLPC)$ satisfying at least one of the following criteria:*

1. *Two demand points p_j and p_k are at threshold distance of L and lie on the same side of L.*
2. *Two demand points p_j and p_k are at threshold distance of L and lie on the different sides of L while $k(p_j - p_k) > 2r$.*

3. *Two demand points p_j and p_k are at threshold distance of L and lie on the different sides of L while $k(p_j - p_k) = 2r$ and the normal vector of L is an extremal direction of the cone $\{\lambda x : x \in \partial k(p_j - p_k), \lambda \geq 0\}$.*

This determines a finite candidate set for $(MLPC)$.

Proof. We know from Theorem 3 that there is an optimal solution L that has at least two points, say p_j and p_k, both at distance r from the line L. If p_j and p_k are on the same side of L then the first criterion is fulfilled. If not, then we know $k(p_j - p_k) \geq 2r$, compare the proof of Theorem 6. If in fact $k(p_j - p_k) > 2r$ holds, the second criterion is fulfilled. Consider now the case $k(p_j - p_k) = 2r$.
The idea is to construct a solution L' with an objective not worse than that of L which fulfills one of the three criteria. To this end perturb L while keeping p_j and p_k at distance r and on different sides of L as long as the subdifferential $\partial k(p_j - p_k)$ permits. By an analogous argument as in the proof of Theorem 3 this is a quasi-concave process (for the details see below) as long as no point enters or leaves the coverage zone. This means that there are two possibilities to finally fix the line while not deteriorating the objective:

1. a line L' with at least three demand points at threshold distance, two of them being p_j and p_k
2. a line L' with p_j and p_k at threshold distance and an extremal direction of $\{\lambda x : x \in \partial k(p_j - p_k), \lambda \geq 0\}$ as a normal vector.

In the second case obviously L' satisfies the third criterion. Moreover, there are only finitely many possibilities for L' since a cone in $\mathbb{R}^2$ cannot have more than two extremal directions.In the first case, it is clear that fixing three points at threshold distance of L' forces two of them, say p_j and p_k, to lie on the same side of L'. Since both of them have *the same* distance from L', the first criterion of the theorem is satisfied.
We finally give the details of the quasi-concave process: Consider the objective function of $(MLPC)$ according to (9), but restricted to a feasible region that guarantees $d(p_j, L_{n,c}) = d(p_k, L_{n,c}) = r$ as well as $\overline{I}_L^+ = \overline{I}_{L_{n,c}}^+$ and $\overline{I}_L^- = \overline{I}_{L_{n,c}}^-$

$$
\begin{aligned}
\min_{n,c} \quad & \frac{1}{k^\circ(n)}\Big(\sum_{i \in I_{\bar{L}}^+} w_i n^T p_i - \sum_{i \in I_{\bar{L}}^-} w_i n^T p_i + c(W_{\bar{L}}^- - W_{\bar{L}}^+) - rk^\circ(n)(W_{\bar{L}}^+ + W_{\bar{L}}^-)\Big) && (12)\\
\text{s. t.} \quad & n^T p_j = n^T p_k + 2rk^\circ(n) \\
& |n^T p_i - c| \geq rk^\circ(n) && i \in \overline{I}_L^+ \cup \overline{I}_L^- \\
& |n^T p_i - c| \leq rk^\circ(n) && i \in I \setminus (\overline{I}_L^+ \cup \overline{I}_L^-).
\end{aligned}
$$

Observe that the second constraint is equivalent to two linear constraints, since we have by the generalized Cauchy-Schwarz inequality [Mic93]

$$n^T p_j = n^T p_k + 2rk^\circ(n) \qquad \Leftrightarrow \qquad n^T(p_j - p_k) = k(p_j - p_k)k^\circ(n)$$

which is fulfilled if and only if $n \in \{\lambda x : x \in \partial k(p_j - p_k), \lambda \geq 0\}$. Since this is a cone in $\mathbb{R}^2$ this condition is again equivalent to $n \in \{x : x^T v_1 \geq 0 \text{ and } x^T v_2 \geq 0\}$ for some $v_1, v_2 \in \mathbb{R}^2$ which are orthogonal to the two extremal directions of that cone. Now substitute these two linear inequalities for the first constraint in (12) and assume that we have an optimal solution where none of the constraints is fulfilled with equality. Then, by continuity, there is a small environment around that solution in the feasible region and, by quasi-concavity, a minimum is also attained at the boundary, where one constraint is satisfied with equality. This corresponds to another point at threshold distance (if one of the relative position preserving constraints is fulfilled with equality in the minimum) – and thus to an optimal L' fulfilling the first criterion – or a minimizing n which is an extremal direction of the subdifferential cone (if $n^T v_1 = 0$ or $n^T v_2 = 0$) – and thus the third criterion.

Summarizing, we have that there is an optimal solution to $(MLPC)$ which satisfies at least one of the three stated criteria. From Theorem 6 and this proof we also see that each criterion is fulfilled by only finitely many lines and thus we have a finite dominating set. □

4 Solution approaches

4.1 Enumeration

Theorem 7 enables us to solve $(MLPC)$ in a straightforward way: for each pair (p_j, p_k) of demand points, determine the finitely many candidate lines as stated in the three cases of Theorem 7. Their number is two (if $k(p_j - p_k) < 2r$), three (if $k(p_j - p_k) = 2r$ and k is smooth at $p_j - p_k$) or four (if $k(p_j - p_k) > 2r$ or $k(p_j - p_k) = 2r$ and k is non-smooth at $p_j - p_k$). Then calculate the objective for each candidate line and keep the one with the smallest value.

Lemma 8. *The enumeration approach takes $O(n^3)$ time.*

Proof. For $O(m^2)$ pairs of fixed points, we determine the finitely many candidate lines as stated in the three cases of Theorem 7. This can be done by solving equation (11) (at least numerically) in constant time, i.e., it does not depend on the number of fixed points. We then evaluate each candidate line in $O(m)$ time, giving us a total of $O(m^3)$ time for the approach. □

Since the Euclidean and the Manhattan norm are among the most common norms in location problems we exemplarily show in the following how to calculate all candidate partial coverage lines for two points p and q at distance r from the line. Note that the calculation of all candidates determined by two points p and q in the case of $d = d_{\text{ver}}$ can be easily derived from the argumentation on page 3 preceding Theorem 6.

The Euclidean case We start with the Euclidean norm k_2 and p and q on the same side of the line $L = L_{n,c}$ which is to be determined. Since L has distance r from both, p and q, it is clear that the normal n of L is perpendicular to $p - q$. Then the first line with p and q on the same side is determined by $c = n^T p + rk_2^\circ(n)$ and the second one by $c = n^T p - rk_2^\circ(n)$. Then $d(p, L) = d(q, L) = r$. Note that $k_2^\circ = k_2$.

The case of p and q on different sides of $L = L_{n,c}$ is slightly more complicated. As in the proof of Theorem 6, n is determined by the equation $|n^T(p-q)| = 2rk_2^\circ(n)$ or equivalently $n^T(p-q) = 2rk_2^\circ(n)$ or $n^T(p-q) = -2rk_2^\circ(n)$. Letting $n = (n_1, n_2)^T$, assuming w.l.o.g. $n_1^2 + n_2^2 = 1$ and substituting $n_2 = \sqrt{1 - n_1^2}$ this becomes, for $\sigma \in \{-1, 1\}$,

$$\begin{aligned} & n_1(p_1 - q_1) + \sqrt{1 - n_1^2}(p_2 - q_2) = 2r\sigma \\ \implies \quad & n_1^2 \cdot [(p_1 - q_1)^2 + (p_2 - q_2)^2] - n_1 \cdot 4r\sigma(p_1 - q_1) + 4r^2 - (p_2 - q_2)^2 = 0 \end{aligned}$$

and the solutions of this quadratic equation are given by

$$n_1 = \frac{2r\sigma(p_1 - q_1) \pm (p_2 - q_2)\sqrt{k_2(p-q)^2 - 4r^2}}{k_2(p-q)^2}. \tag{13}$$

Note that this has no real root if $k_2(p-q) < 2r$ (which corresponds to the first case of Theorem 6), a double root if $k_2(p-q) = 2r$ (third case), and two distinct real roots if $k_2(p-q) > 2r$ (second case). Note that the Euclidean norm is a smooth norm, hence all possible cases are covered. Among the candidate solutions for (13) we pick the ones that actually satisfy $|n^T(p-q)| = 2r$. Finally, c is determined by the fact that the point $\frac{p+q}{2}$ is contained in $L = L_{n,c}$ with $d(p, L) = d(q, L) = r$ and p and q on different sides of L.

The Manhattan case For the Manhattan norm k_1, the case of p and q on the same side of $L = L_{n,c}$ is completely analogous to the Euclidean norm. n is again perpendicular to $p - q$ and $c = n^T p \pm rk_1^\circ(n)$. Note that k_1° is the Chebyshev norm.

This means we have to solve

$$|n^T(p-q)| = 2rk_1^\circ(n) = 2r \max\{|n_1|, |n_2|\} \tag{14}$$

in the case of p and q on different sides of $L = L_{n,c}$. From (14) we get by assuming $k_1^\circ(n) = 1$ and $p_2 - q_2 \neq 0$ w.l.o.g. and with the triangle inequality

$$\begin{aligned} |n_1(p_1 - q_1) + n_2(p_2 - q_2)| = 2r \quad & \implies \quad |p_1 - q_1| + |p_2 - q_2| \geq 2r \\ & \implies \quad |p_2 - q_2| \geq 2r - |p_1 - q_1|. \end{aligned}$$

Thus one can choose $\tilde{n}_1 \in [-1, 1]$ and $\tilde{n}_2 \in \{-1, 1\}$ such that

$$\tilde{n}_2(p_2 - q_2) = 2r - |p_1 - q_1| \qquad \text{and} \qquad \tilde{n}_2(p_2 - q_2) = 2r - \tilde{n}_1(p_1 - q_1)$$

and we have that

$$|\tilde{n}^T(p-q)| = \tilde{n}_1(p_1 - q_1) + \tilde{n}_2(p_2 - q_2) = 2r \max\{|\tilde{n}_1|, |\tilde{n}_2|\}$$

i.e., $(\tilde{n}_1, \tilde{n}_2)^T$ is a normal vector of one candidate line. If $p_1 = q_1$, the second line is determined by inverting the sign of n_1, if $p_1 \neq q_1$ the second line is obtained by reversing the roles of $p_1 - q_1$ and $p_2 - q_2$. Notice that, in the case $p_1 = q_1$ and $|p_2 - q_2| = 2r$, which corresponds to the third case in Theorem 6, it holds in fact $n_2 \in \{-1, 1\}$ and n_1 can be chosen arbitrarily in $[-1, 1]$. As in the Euclidean case, c is obtained readily since $\frac{p+q}{2} \in L$.

4.2 Sweeping

In order to explore the possibilities of solving $(MLPC)$ faster for arbitrary norms k we propose an alternative approach based on a plane-sweeping technique for arrangements of lines by [EW86]. This allows us to efficiently update the objective function, namely in constant time per candidate line, instead of linear time as in the brute force approach used for enumeration in Section4.1.

For the sweeping algorithm we restrict ourselves to non-vertical lines. This is not a severe restriction, since one can, for *any* norm k, treat the vertical line problem separately as special case. This results in a one-dimensional problem. Using that also this problem satisfies an incidence property, namely, that there is one point at distance r to the vertical line, it can be solved by sorting the fixed points $p_1, \ldots, p_n$ by their x-coordinate and then evaluate $O(m)$ candidate lines which have at least one of the fixed points at distance r. This can be done in a straight-forward way in $O(m^2)$ time or by linear programming in $O(m)$ time.

Neglecting the case of a vertical line we can apply a well-known duality transform *, mapping non-vertical lines to points in $\mathbb{R}^2$ and vice versa, which can be found e. g. in [Mat02]:

$$L = L_{s,b} \mapsto L^* = (s, -b), \qquad p = (x, y) \mapsto p^* = L_{x,-y}.$$

This transform preserves the vertical distance $d_{\text{ver}}(p, L) = d_{\text{ver}}(L^*, p^*)$ and above-below relationships, i. e. p lies above L if and only if L^* lies above p^*. Hence also the partial coverage distance $D(p, L) = D(L^*, p^*)$ is preserved in the vertical distance case. We call the original space the primal space and the transformed space the dual space. The horizontal coordinate of a dual point hence is the slope of the corresponding line in primal space.

By the above considerations of the duality transform, $(MLPC)$ is now to find a dual point L^* which has at least two dual lines p_i^* at distance r, a consequence of Theorem 3. We can enumerate all these dual points L^* satisfying this property by sweeping the arrangement of dual lines $p_1^*, \ldots, p_n^*$ starting at an arbitrary horizontal coordinate and sweeping first to the right from that point and then to the left from the same point. During the sweep we calculate the objective values of all candidates and return the best one in the end.

Say we start the sweeping at horizontal coordinate a and sweep to the right. The sweeping to the left is completely analogous. We determine the best dual point L_a^* for this particular horizontal coordinate a. This means, finding the best intercept for a line with given slope. Since we know that the objective function w.r.t the intercept is piecewise linear and convex (see the proof of Theorem 2) this can be done, e.g. by linear programming in $O(m)$ time. If the solution is not unique, we choose it such, that it has at least one of the dual lines p_i^* at distance r. Moreover, we determine the top-down order $(1)_a, \ldots, (m)_a$ of the p_i^* at the current horizontal coordinate in the dual space and the indices $i_a^{+,u}$ and $i_a^{+,l}$ as well as $i_a^{-,u}$ and $i_a^{-,l}$ such that $p^*_{(i_a^{+,u})_a}$ is the first dual line above L^* with $d(L^*, p^*_{(i_a^{+,u})_a}) \geq r$ and $p^*_{(i_a^{+,l})_a}$ the last dual line above L^* with $d(L^*, p^*_{(i_a^{+,l})_a}) \leq r$ The other two indices store the position of the corresponding dual lines below L^*. These will be needed during the sweep to allow for efficient updates of the objective function. Now we have to find the next candidate along the horizontal axis. There are only the following possibilities for an event that asks for an action to be performed, when increasing the horizontal coordinate from a to $A > a$. We calculate the horizontal coordinate A of each possibly next event, but in fact only the one with least A actually occurs.

1. Two lines $p^*_{(i)_A}$ and $p^*_{(i+1)_A}$ intersect at A.
2. L_a^* has only one of the p_i^*, say $p_{i_0}^*$, at distance r. W.l.o.g. let $p_{i_0}^*$ be below L_a^*. Then p_{i_0} is also below L_A^* and either
   a) L_A^* has at least two of the p_i^* at distance r, both above L_A^* or
   b) L_A^* has at least two of the p_i^* at distance r, one above and one L_A^*
3. L_a^* has more than one of the p_i^* at distance r and either (a) or (b) as above occur at A.

The first case is simple. In order to keep the top-down order updated, we exchange $(i+1)_A := (i)_a$ and $(i)_A := (i+1)_a$ and also adjust $i_A^{+,u}$, $i_A^{+,l}$, $i_A^{-,u}$, and $i_A^{-,l}$ if necessary. Consider now the second case. Since L_a^* with one dual line p_i^* at distance r is optimal for horizontal coordinate a, L_A^*, $A > a$ with the same p_i^* at distance r remains optimal as long as none of the above events occur due to Theorem 2, since there is no change in the weights above and below when passing from L_a^* to L_A^*. Thus we keep increasing a until an event of type 2(a) or 2(b) occurs. A type 2(a) event implies that two of the dual lines p_i^* intersect and can be found as a type 1 event. In addition to the actions performed, when the event is only of type 1, we also evaluate the objective since these events yield candidate solutions for optimality according to Theorem 3. The horizontal coordinates of possible type 2(b) events can be found in constant time as solutions of

$$|n^T(p_j - p_{i_0})| = 2rk^\circ(n), \tag{15}$$

for each $j \in \{i_A^{+,u}, i_A^{+,l}, i_A^{-,u}, i_A^{-,l}\}$ where $n = (A, -1)$.
The third case can be reduced to the second one: if L_a^* has more than one dual line at distance r, one can always tell by the weights of dual lines above and below, respectively, which one of those lines should be kept at distance r when moving on to horizontal coordinate A and is going to play the role of $p_{i_0}^*$ in the distinction of the possible events when continuing the sweep, compare also Theorem 2. The other dual lines at distance r can be ignored and we are in the second case.

Lemma 9. *($MLPC$) can be solved in $O(m^2 \log m)$ time with the sweeping approach.*

Proof. We first calculate the $O(m^2)$ intersection points of the dual lines $p_1^*, \dots, p_n^*$ and sort them by their horizontal coordinates in $O(m^2 \log m)$ time. We then sweep to the right and to the left from an arbitralily chosen horizontal coordinate in the dual space. Each direction of the sweep ends if no further events in the direction of sweeping are found. It takes only constant time to find the next event and the objective can also be evaluated in constant time, if we keep track of the quantities

$$X_a^+ := \sum_{j=1}^{i_a^{+,u}} w_{(j)_a}(d(p_{(j)_a}, L_a^*) - r) \qquad \text{and} \qquad X_a^- := \sum_{j=i_a^{-,l}}^{n} w_{(j)_a}(d(p_{(j)_a}, L_a^*) - r)$$

which can be viewed as aggregate distances. The objective is then $X_a^+ + X_a^-$. If we can now assure an amount of $O(m^2)$ events during the sweep to the left and to the right, this approach has a running time of $O(m^2 \log n)$, the sorting of the intersection points being the bottleneck. Clearly each pair of points p and q with $k(p-q) \neq 2r$ gives rise to a constant number of events by the first two assertions of Theorem 6. If $k(p-q) = 2r$, it suffices to examine a constant number of solutions in (15) (if there is more than one at all) by Theorem 7, namely the two solutions $(A_{\min}, -1)$ and $(A_{\max}, -1)$ with minimal and maximal A, respectively. Thus we have $O(m^2)$ events in total for both directions of the sweep. The correctness of the sweeping algorithm follows directly from Theorem 7. □

4.3 A linear programming formulation

We start by considereing the vertical distance. As used before (see page 3) we may parametrize L by its slope s and intercept b in this case, since a vertical line can only be optimal if all demand points lie on that line.
We obtain $d(p_i, L_{s,b}) = |p_{i2} - sp_{i1} - b|$ for a point $p_i = (p_{i1}, p_{i2}) \in \mathbb{R}^2$ and ($MLPC$) can be written as the following linear program

$$\begin{aligned}
\min \quad & \sum_{i=1}^{m} w_i D_i & & \\
\text{s. t.} \quad & D_i \geq p_{i2} - sp_{i1} - b - r & i = 1, \dots, m \qquad & (16) \\
& D_i \geq -p_{i2} + sp_{i1} + b - r & i = 1, \dots, m \qquad & (17) \\
& s, b \in \mathbb{R},\ d_1, \dots, d_m \geq 0. & & (18)
\end{aligned}$$

In an optimal solution to this program, the variables D_i contain the partial coverage distance of the line $L_{s,b}$ to point p_i. This holds, since (16) and (17) together are equivalent to $D_i + r \geq |p_{i2} - sp_{i1} - b|$, and the minimization of the $D_i \geq 0$ forces them to become

$$D_i = \max\{0, |p_{i2} - sp_{i1} - b| - r\}.$$

This linear program is of the form considered in [Zem84] and hence a linear time algorithm is available.

The case of a block norm distances can be reduced to vertical distance case in the following way. For the point-line distance for a block norm k with fundamental direction $e_1, \ldots, e_F$ it holds that

$$d(p, L) = \min_{k=1,\ldots,F} \underbrace{\min \left\{ |\lambda| : p + \lambda e_k \in L, \ \lambda \in \mathbb{R} \right\}}_{=:d_k(p,L)},$$

where the optimal index k^* in the outer minimization depends only on the slope of L, see [Sch99]. Thus, the objective function becomes

$$f(s, b) = \min_{k=1,\ldots,F} \sum_{i=1}^{m} [d_k(p_i, L) - r]^+. \tag{19}$$

Moreover, it holds

$$d_k(p, L) = \frac{1}{l_k} d_{\mathrm{ver}}(T_{\alpha_k}(p), L) \qquad \forall\, k = 1, \ldots, F$$

where l_k is the Euclidean length of e_k and T_{α_k} the rotation about the origin by α_k, the angle subtended by the positive x-axis and e_k. Hence we can solve $(MLPC)$ in the block norm case by solving F linear programs of the form

$$\begin{array}{lll} \min & \displaystyle\sum_{i=1}^{m} w_i D_i & \\ \text{s. t.} & D_i \geq \dfrac{1}{l_k}(p_{i2}^{(\alpha_k)} - p_{i1}^{(\alpha_k)} s - b) - r & i = 1, \ldots, m \\ & D_i \geq -\dfrac{1}{l_k}(p_{i2}^{(\alpha_k)} - p_{i1}^{(\alpha_k)} s - b) - r & i = 1, \ldots, m \\ & s, b \in \mathbb{R}, \ d_1, \ldots, d_m \geq 0 & \end{array} \tag{20}$$

for $k = 1, \ldots, F$, where $(p_{i1}^{(\alpha_k)}, p_{i2}^{(\alpha_k)}) = T_{\alpha_k}(p_i)$. According to (19), the linear program with the minimal objective value among them determines the optimal solution to $(MLPC)$ with block norm k. These considerations hence yield the following result.

Lemma 10. *$(MLPC)$ with a block norm k having F fundamental directions can be solved in $O(mF)$ time by solving F linear programs of the form* (20).

The linear porgramming approach is therefore substantially faster then the sweeping in $O(m^2 \log m)$ time if the number F of fundamental directions is small, e.g. in the case of Manhattan or Chebyshev distances.

5 Conclusion

In this paper we have generalize the classical median line model to a setting with partial coverage where a demand point sufficiently close to a line, i.e. closer than some fixed threshold distance r, is considered covered. A covered point contributes no cost to the objective function. Demand points which are not within the threshold distance incur a penalty cost proportional to the distance to the zone covered by the line. The $(MLPC)$ model represents a compromise between the classical median objective and the center objective, see Lemma 1: for $r = 0$ the median model is reproduced and for a certain $r_{\max}$ our problem is equivalent to the center line problem.

To solve the $(MLPC)$ for arbitrary distances induced by a norm we generalized classical results for the median model, in particular, we could establish an incidence property in Theorem 3 which led to a finite candidate set (Theorem 7) and allowing to solve the problem by pure enumeration in $O(m^3)$ time. We were able to reduce the enumeration time by applying plane sweeping techniques in Section 4.2 to $O(m^2 \log m)$. For the special case of block norms and vertical distances, a linear programming formulation which has the structure required for the linear time algorithm proposed in [Zem84] can be found in Section 4.3.

Since another interpretation of the median line problem with partial coverage is the location of a median line to approximate norm disks of radius r, an interesting generalization would be the approximation of norms disks of different radii (or even arbitrary sets). First results can be found in [Sch13].
Another obvious extension is the location of a line or hyperplane with partial coverage in higher dimensions. In the case of block norms an experimental comparison of the sweeping approach in Section 4.2 and of the linear programming formulation in Section 4.3 would be interesting.

References

[ADSZ88] Mordecai Avriel, Walter E. Diewert, Siegfried Schaible, and Israel Zang. *Generalized Concavity*, volume 36 of *Mathematical Concepts and Methods in Science and Engineering.* Plenum Press, New York, 1988.

[BCSS11] R. Blanquero, E. Carrizosa, A. Schöbel, and D. Scholz. Location of a line in the three-dimensional space. *EJOR*, pages 14–20, 2011.

[BJKS13] J. Brimberg, H. Juel, M.-C. Körner, and A. Schöbel. On models for continuous facility location with partial coverage. *JORS*, 2013. to appear.

[BJS02] J. Brimberg, H. Juel, and A. Schöbel. Linear facility location in three dimensions - models and solution methods. *Operations Research*, 50(6):1050–1057, 2002.

[BJS03] J. Brimberg, H. Juel, and A. Schöbel. Properties of 3-dimensional line location models. *Annals of Operations Research*, 122:71–85, 2003.

[BJS07] J. Brimberg, H. Juel, and A. Schöbel. Locating a circle on a sphere. *Operations Reserach*, 55(4):782–791, 2007.

[EW86] H. Edelsbrunner and E. Welzl. Constructing belts in two-dimensional arrangements with applications. *SIAM J. Comput.*, 15(1):271–284, 1986.

[Fle77] Wendell Fleming. *Functions of Several Variables.* Undergraduate Texts in Mathematics. Springer, New York, 1977.

[HT88] M.E. Houle and G.T. Toussaint. Computing the width of a set. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(5):761–765, 1988.

[LC85] D. T. Lee and Y. T. Ching. The power of geometric duality revisited. *Information Processing Letters*, 21(3):117 – 122, 1985.

[LMW88] Robert F. Love, James G. Morris, and George O. Wesolowsky. *Facilities Location: Models and Methods.* North-Holland, 1988.

[Man99] Olvi L. Mangasarian. Arbitrary-norm separating plane. *Operations Research Letters*, 24(1-2):15 – 23, 1999.

[Mat02] Jiří Matoušek. *Lectures on Discrete Geometry*, volume 212 of *Graduate Texts in Mathematics.* Springer, New York, 2002.

[Mic93] Christian Michelot. The mathematics of continuous location. *Studies in Locational Analysis*, 5:58–83, 1993.

[MN80] James G. Morris and John P. Norback. A simple approach to linear facility location. *Transportation Science*, 14(1):1–8, 1980.

[MS98] H. Martini and A. Schöbel. Median hyperplanes in normed spaces — a survey. *Discrete Applied Mathematics*, 89:181–195, 1998.

[MS99] H. Martini and A. Schöbel. A characterization of smooth norms. *Geometriae Dedicata*, 77:173–183, 1999.

[MS01] H. Martini and A. Schöbel. Median and center hyperplanes in Minkowski spaces — a unifying approach. *Discrete Mathematics*, 241:407–426, 2001.

[MT83] Nimrod Megiddo and Arie Tamir. Finding least-distances lines. *SIAM J. Algebraic Discrete Methods*, 4(2):207–211, 1983.

[PC01] F. Plastria and E. Carrizosa. Gauge distances and median hyperplanes. *J. Optim. Theory Appl.*, 110(1):173–182, 2001.

[RT94] Jean-Marc Robert and Godfried T. Toussaint. Linear approximation of simple objects. *Computational Geometry*, 4(1):27–52, 1994.

[Sch99] A. Schöbel. *Locating Lines and Hyperplanes — Theory and Algorithms.* Number 25 in Applied Optimization Series. Kluwer, 1999.

[Sch13] R. Schieweck. Lower bounds for line location problems via demand regions. Preprint 28, Institute for Numerical and Applied Mathematics, University of Göttingen, 2013.

[Zem84] Eitan Zemel. An $O(n)$ algorithm for the linear multiple choice knapsack problem and related problems. *Inform. Process. Lett.*, 18(3):123–128, 1984.

Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
D - 37083 Göttingen

Telefon: 0551/394512
Telefax: 0551/393944

Email: trapp@math.uni-goettingen.de URL: http://www.num.math.uni-goettingen.de

Verzeichnis der erschienenen Preprints 2013

Number	Authors	Title
2013 - 1	M. Ehrgott, J. Ide, A. Schöbel	Minmax Robustness for Multi-objective Optimization Problems
2013 - 2	M. Goerigk, M. Gupta, J. Ide, A. Schöbel, S. Sen	The Uncertain Knapsack Problem with Queries
2013 - 3	M. Siebert, M. Goerigk	An Experimental Comparison of Periodic Timetabling Models
2013 - 4	F. Dunker, S. Hoderlein, H. Kaido	Random Coefficients in Static Games of Complete Information
2013 - 5	F. Dunker, J.-P. Florens, T. Hohage, E. Mammen	Iterative Estimation of Solutions to Noisy Nonlinear Operator Equations in Nonparametric Instrumental Regression
2013 - 6	F. Dunker, J.-P. Florens, T. Hohage, E. Mammen	Iterative Estimation of Solutions to Noisy Nonlinear Operator Equations in Nonparametric Instrumental Regression
2013 - 7	C. Buchheim, R. Hübner, A. Schöbel	Ellipsoid bounds for convex quadratic integer programming
2013 - 8	T. Hohage, F. Werner	Convergence Rates for Inverse Problems with Impulsive Noise
2013 - 9	K. Kuhn, A. Raith, M. Schmidt, A. Schöbel	Bicriteria robust optimisation
2013 - 10	R. Schaback	Direct Discretizations with Applications to Meshless Methods for PDEs
2013 - 11	R. Schaback	Greedy Sparse Linear Approximations of Functionals from Nodal Data

Number	Authors	Title
2013 - 12	M. Bozzini, M. Rossini, R. Schaback, E. Volonte	Kernels via Scale Derivatives
2013 - 13	O. Davydov, R. Schaback	Error Bounds for Kernel-Based Numerical Differentiation
2013 - 14	R. Schaback	A computational tool for comparing all linear PDE solvers (- Error-optimal methods are meshless -)
2013 - 15	M. Mohammadi, R. Mokhtari, R. Schaback	Simulating the 2D Brusselator system in reproducing kernel Hilbert space
2013 - 16	Y.C. Hon, R. Schaback	Direct Meshless Kernel Techniques for Time–Dependent Equations
2013 - 17	Y.C. Hon, M. Zhong, R. Schaback	The Meshless Kernel-Based Method of Lines for the Heat Equation
2013 - 18	M. Bozzini. L. Lenarduzzi. M. Rossini, R. Schaback	Interpolation with variably scaled kernels
2013 - 19	T. Hohage, F. Le Louer	A spectrally accurate method for the dielectric obstacle scattering problem and applications to the inverse problem
2013 - 20	T. Hohage, F. Le Louer	A spectrally accurate method for the dielectric obstacle scattering problem and applications to the inverse problem
2013 - 23	M. Schmidt, A. Schöbel	Timetabling with Passenger Routing
2013 - 24	J. Ide, E. Köbis	Concepts of Robustness for Multi-Objective Optimization Problems based on Set Order Relations

Number	Authors	Title
2013 - 25	J. Ide, E. Köbis, D. Kuroiwa, A. Schöbel, C. Tammer	The relation between multicriteria robustness concepts and set valued optimization
2013 - 26	J. Ide, M. Tiedemann, S. Westphal, F. Haiduk	An Application of Deterministic and Robust Optimization in the Wood Cutting Industry
2013 - 27	J. Ide, A. Schöbel	Robustness for uncertain multi-objective optimization
2013 - 28	R. Schieweck	Lower bounds for line location problems via demand regions
2013 - 29	A. Schöbel, S. Schwarze	Finding Delay-Resistant Line Concepts using a Game-Theoretic Approach
2013 - 30	M. Goerigk, A. Schöbel	Recovery-to-Optimality: A New Two-Stage Approach to Robustness with an Application to Aperiodic Timetabling
2013 - 31	M. Goerigk, A. Schöbel	Algorithm Engineering in Robust Optimization
2013 - 32	J. Brimberg, R. Schieweck, A. Schöbel	Locating a median line with partial coverage distance