Stop Location Design in Public Transportation Networks: Covering and Accessibility Objectives[1]

Dwi Retnani Poetranto[a], Horst W. Hamacher[a2], Simone Horn[a], Anita Schöbel[b].

[a] Fachbereich Mathematik, Universität Kaiserslautern, Postfach 3049, Kaiserslautern, Germany

[b] Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen, Germany

Abstract

In *StopLoc* we consider the location of new stops along the edges of an existing public transportation network. Examples of StopLoc include the location of bus stops along some given bus routes or of railway stations along the tracks in a railway system. In order to measure the "convenience" of the location decision for potential customers in given demand facilities, two objectives are proposed. In the first one, we give an upper bound on reaching a closest station from any of the demand facilities and minimize the number of stations. In the second objective, we fix the number of new stations and minimize the sum of the distances between demand facilities and stations. The resulting two problems *CovStopLoc* and *AccessStopLoc* are solved by a reduction to a classical set covering and a restricted location problem, respectively. We implement the general ideas in two different environments - the plane, where demand facilities are represented by coordinates and in networks, where they are nodes of a graph.

[2]Corresponding author

1 Introduction

The acceptance of public transportation is depending on various components such as competitive fares, convenience, punctuality, reliability, etc. In this paper, we address the question of convenience. In particular, we investigate the problem of finding "good" locations for stops in order to allow potential customers to access the public transportation system within a reasonable amount of time (*StopLoc* problems).

Due to their great potential for improved public transportation systems, several versions of StopLoc have been considered by various authors in the last years. A survey on relevant theory and current solution methods can be found in the monograph of Schöbel [Sch06]. The public transportation system is defined by a *Public Transportation Network* (PTN), which is an undirected graph $PTN = (V, E)$ with node set V consisting of existing stops V_{stop} and breakpoints V_{break}, i.e., $V = V_{stop} \cup V_{break}$, and edge set E (containing direct connections between nodes in V). In the following, we assume that the PTN is embedded in the plane. The potential customers of the public transportation system are modeled as *demand facilities*, for which we assume that the distance to any point of the PTN is known. In order to find "good" locations, several objective functions are possible.

One of the most frequently discussed goals is to minimize the number of stops such that each demand facility is within a tolerable distance from at least one stop. The maximal distance that a customer is willing to tolerate is called *covering radius* and we therefore call this type of stop location problem *Covering Stop Location* problem, or *CovStopLoc* for short. For bus stops a covering radius of 400 m is common, meaning that customers are only willing to use public transportation, if the next stop is within a distance of at most 400 m. In rail transportation, the covering radius is larger and is usually assumed to be 2 km.

In the literature, CovStopLoc has been introduced in [Gle75] and is considered for example, in [MDSF98, Mur01a, Mur01b, LMO02], see also references therein. In these papers, StopLoc is treated in a discrete setting, i.e., a finite set is considered as potential new stops. [SHLW02] however, allow a continuous set of possible locations for the stops, for instance, all points on the current bus routes or railway tracks. An application of *continuous* CovStopLoc is given in [HLS$^+$01], where the authors report on a project with the largest German rail company (Deutsche Bahn) and consider the trade-off between the positive and negative effects of stops. The negative effect of longer travel times due to additional stops is compared with the positive effect of shorter access times, and the goal is to find locations for new stops that maximize the difference of the two effects.

Based on this application, variants of continuous CovStopLoc have been treated in [KPS$^+$03, Sch05]. The problem has been solved for the case of

two intersecting lines, see [MMW04]. Algorithmic approaches for solving the underlying covering problem have been studied in [RS04, MW04]. Recently, complexity and approximation issues have been presented in [MSW05].

In our paper, we will focus on two variants of StopLoc which will be considered in a planar and in a network environment. In the former case the demand is concentrated in points $p_m = (p_{m1}, p_{m2})$, $m = 1, \ldots, M$ of the Euclidean plane $\mathbb{R}^2$. In the latter case the demand originates in nodes of an extended network.

The first variant of StopLoc, CovStopLoc, was explained before. We review results for the planar case from the literature. The network case has to the best of our knowledge not been considered before. We present new results which were motivated by the thesis of Horn [Hor04].

In the Access Stop Location Problem (*AccessStopLoc*) we minimize the overall distance, i.e. the sum of the distances between the demand facilities and their closest stations. In spite of its obvious application potential, this problem has not been considered in the literature so far. AccessStopLoc was first introduced in the thesis of Poetranto [Poe04] where it is treated as a restricted median location problem.

The remainder of our paper is structured as follows. In the next section CovStopLoc and AccessStopLoc are formally introduced, both in the planar and network environment. In Section 3 we show that CovStopLoc can be reduced to a classical set covering problem, by identifying a finite dominating set (*FDS*). We will show that such an FDS can be found efficiently in both environments. Algorithms for solving AccessStopLoc are presented in Section 4. We give a reformulation of the problem as restricted location problem and show how to solve it for the planar and for the network case. We finish the paper by a conclusion, in which we sketch ideas on how to extend the theory presented in the previous sections. In particular, we consider an equity objective which can be tackled by either of the two approaches of Sections 3 or 4.

Throughout, we assume that the reader is familiar with basic concepts of location theory (see, e.g., [LMW88, FLMW92, Ham95, Das95, DH02]) .

3

2 Covering and Access Problems

In this section, we formally introduce the Covering and Access Stop Location Problem in planar and network environments.

The new stops are selected among the points of a given *public transportation network (PTN)*. A PTN consists of an undirected, connected graph $G = (V, E)$ where the edges $e = (v_i, v_j) \in E$ between nodes v_i and v_j are part of public transportation lines. We assume that the graph has a planar embedding such that the edges are represented as straight lines (this can always be enforced by adding nodes at crossing points of an originally given PTN). Each edge e has an associated *length* l_e which is a real number representing physical length or travel time. A *point* $s \in G$ of the graph $G = (V, E)$ is either a node or strictly contained in an edge $e = (v_i, v_j)$. In both cases, we often use the denotation $s = (e, t)$ where $0 \le t \le 1$ and where the distance between s and v_i is $t l_e$ and between s and v_j is $(1 - t) l_e$.

The potential users of the public transportation system are assumed to be represented by a finite set $\mathcal{P} = \{p_1, \ldots, p_M\}$ of *demand facilities*. For each demand facility $p \in \mathcal{P}$ and point $s \in G$ we assume that a *distance* $d(p, s) \in \mathbb{R}$ is known. If $\mathcal{S} \subseteq G$ is a subset of points, then

$$d(p, \mathcal{S}) := \min_{s \in \mathcal{S}} \ d(p, s). \tag{2.1}$$

In this paper, we consider two versions of stop location problems.

In the *Covering Stop Location Problem (CovStopLoc)* a real number r is given as *radius*. It is assumed that customers will not use the public transportation system if the distance to a station is larger than r. In CovStopLoc we want to find a set of locations for stations, which is a subset $\mathcal{S} \subseteq G = (V, E)$ with minimal cardinality, such that every customer in one of the demand facilities can reach at least one of the new stations located in some $s \in \mathcal{S}$ with a distance of at most r. Hence we want to solve the following problem.

$$
\begin{aligned}
\text{minimize} \quad & |\mathcal{S}| \\
\text{subject to} \quad & d(p, \mathcal{S}) \le r \quad \forall p \in \mathcal{P} \\
& \mathcal{S} \subseteq G.
\end{aligned}
\qquad (CovStopLoc)
$$

In the *Accessibility Stop Location Problem (AccessStopLoc)* a positive integer k is given which is maximal number of stations which can be built. We want to find in AccessStopLoc a subset $\mathcal{S} \subseteq G = (V, E)$ with $|\mathcal{S}| \le k$ such that its overall distance to the demand facilities is minimized. If we measure the overall distance by the sum of the distances from $\mathcal{S}$ to all customers, the resulting problem is the following.

$$\text{minimize } f(\mathcal{S}) := \sum_{p \in \mathcal{P}} d(p, \mathcal{S})$$

$$\text{subject to } \quad \mathcal{S} \subseteq G \hspace{3cm} (AccessStopLoc)$$

$$|\mathcal{S}| \leq k$$

Both versions of the stop location problem include the distance $d(p, s)$ between demand facilities and points in the PTN in their formulations. We have two different environments, where distances are defined.

In the *planar environment* we use the given straight line planar embedding of the PTN $G = (V, E)$. Here, any point s of the graph G can be interpreted as a point $s = (s_1, s_2)$ in the Euclidean plane $\mathbb{R}^2$. If we additionally assume that for each $m = 1, \ldots, M$ the demand facility $p_m = (p_{m1}, p_{m2})$ is also a point in the Euclidean plane $\mathbb{R}^2$, then $d(p, s)$ can be any distance function in $\mathbb{R}^2$, for instance, the rectangular distance

$$d(p, s) := l_1(p, s) := |p_1 - s_1| + |p_2 - s_2|.$$

In the *network environment*, the PTN is part of a larger network, the *street and public transportation network (SPTN)* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = V \cup \mathcal{P}$ and $E \subseteq \mathcal{E}$. The edges in $\mathcal{E} \setminus E$ represent streets which can be used by potential passengers to reach one of the new stations. An example of a SPTN is given in Figure 1.
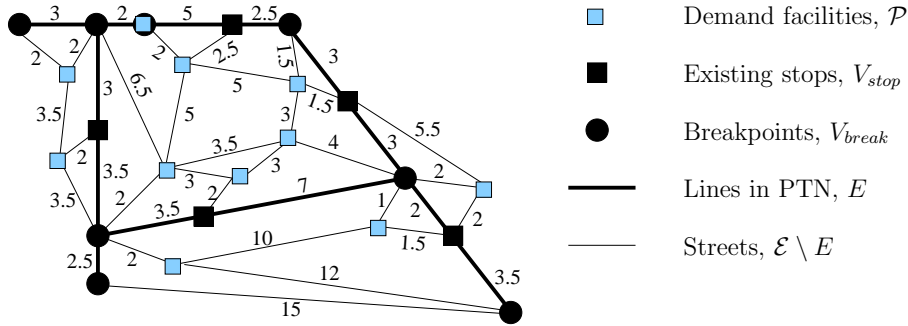

Figure 1: Street and Public Transportation Network.

The distance $d(p, s)$ is the shortest path distance in $\mathcal{G}$, if both p and s are nodes of $\mathcal{G}$, i.e. $p \in \mathcal{P}$ and $s \in V$. We denote this network distance in the following with $ND(p, s)$. If $s = (e, t)$ is a point of the PTN inside the edge $e = (v_i, v_j) \in E$ (i.e., with $0 < t < 1$) and $p \in \mathcal{P}$, then

$$ND(p, s) := \min \ \{ND(p, v_i) + tl_e, \ ND(p, v_j) + (1 - t)l_e\}. \hspace{1cm} (2.2)$$

5

The distance $ND(p, s)$ as a function of t if we fix $p \in \mathcal{P}$ and move $s = (e, t)$ along an edge e is illustrated in Figure 4.

Having defined two different versions of StopLoc and two environments, we can define four types of StopLoc. The key in tackling these four problem types is a discretization based on the following properties.

1. The *finite dominating set (FDS)* property reduces the continuous StopLoc to one, where only a finite candidate set $\mathcal{S}_{cand}$ has to be considered to choose locations for new stations.

2. The *polynomiality property* yields a set $\mathcal{S}_{cand}$ which is polynomial in the size of the problem.

If both properties are satisfied for one of our problems, the following algorithmic approach is straightforward.

Algorithm :	Generic algorithm for StopLoc
Input :	$PTN = (V, E)$, demand facilities $\mathcal{P} = \{p_1, ..., p_M\} \subset \mathbb{R}^2$.
Output :	$\mathcal{S}^*$ optimal stops.
Step 1 :	Determine the set of candidates $\mathcal{S}_{cand}$.
Step 2 :	Let $\mathcal{S}^* = \text{argmin } \{f(\mathcal{S}) : \mathcal{S} \in \mathcal{S}_{cand}\}$
Step 3 :	Output: $\mathcal{S}^*$

In the following sections we will consider this approach in more detail.

3 Covering

In this section we assume that a *radius* $r \in \mathbb{R}$ is given.

Definition 3.1. (Cover) A demand facility $p \in \mathcal{P}$ is *covered* by a point s in the PTN $G = (V, E)$ if $d(s, p) \leq r$. For a set $\mathcal{S}$ of points in G we define

$$cover(\mathcal{S}) = \{p \in \mathcal{P} : d(s, p) \leq r \text{ for some } s \in \mathcal{S}\}. \tag{3.1}$$

If $\mathcal{S} = \{s\}$ is a singleton we use the denotation *cover(s)* instead of $cover(\{s\})$. $\mathcal{S}$ is called a *cover* of $\mathcal{P}$ if all $p \in \mathcal{P}$ are covered by some $s \in \mathcal{S}$, i.e., for all $p \in \mathcal{P}$ there exists some $s \in \mathcal{S}$ such that $d(s, p) \leq r$.

The *Covering Stop Location Problem (CovStopLoc)* is thus the problem of finding a cover with smallest cardinality.

For any finite set $\mathcal{S}_{cand} \subseteq G$ of points in the PTN $G = (V, E)$ which contains candidates for the location of new stations we can define a binary matrix $A = (a_{ps})_{\substack{p \in \mathcal{P} \\ s \in \mathcal{S}_{cand}}}$ where

$$a_{ps} = \begin{cases} 1 & \text{if } d(p, s) \leq r \\ 0 & \text{else} \end{cases} \tag{3.2}$$

Since any subset $\mathcal{S} \subseteq \mathcal{S}_{cand}$ can be represented using its characteristic vector $x = (x_s)$ defined by

$$x_s = \begin{cases} 1 & \text{if } s \in \mathcal{S} \\ 0 & \text{else,} \end{cases} \tag{3.3}$$

the discrete CovStopLoc, where we allow only stops from the set $\mathcal{S}_{cand}$, can be written as the following integer program:

$$\min \quad \sum_{s \in \mathcal{S}_{cand}} x_s \tag{3.4}$$

$$s.t \quad \sum_{s \in \mathcal{S}_{cand}} a_{ps} x_s \geq 1, \quad \forall p \in \mathcal{P} \tag{3.5}$$

$$x_s \in \{0, 1\}, \quad \forall s \in \mathcal{S}_{cand} \tag{3.6}$$

The objective function (3.4) represents the number of selected candidates and the constraint (3.5) guarantees that each demand facility is covered by at least one selected candidate. The integer program (3.4 - 3.6) is obviously the well-known set-covering problem.

In order to show that every CovStopLoc can be written as set-covering problem, we need to establish the FDS and polynomiality properties introduced at the end of Section 2.

We begin with the definition of candidate set

$$\mathcal{S}_{cand} := \{s \in G : \exists\, p_m \in \mathcal{P} \text{ with } d(p_m, s) = r\} \tag{3.7}$$

where the distance function d is the rectangular distance for planar environment and the network distance ND given by (2.2).

In order to prove the dominance property of $\mathcal{S}_{cand}$ defined by (3.7), we call $\tilde{s} \in \mathcal{S}_{cand}$ a *neighbored candidate* of $s \in G$, if there exists a path $Path(s, \tilde{s})$ in the PTN with the following property. None of the points strictly between $\tilde{s}$ and s on the path is contained in $\mathcal{S}_{cand}$.

Lemma 3.2. *Let s be a point in G which is not a candidate and let $\tilde{s}$ be a neighbored candidate of s. Then*

$$cover(s) \subseteq cover(\tilde{s}) \tag{3.8}$$

Proof. Suppose $cover(s) \nsubseteq cover(\tilde{s})$. Then there exists some $p \in \mathcal{P}$ with

$$d(p, s) \le r \text{ and } d(p, \tilde{s}) > r$$

Since the distance function d is continuous and the path is connected, the intermediate value theorem yields some $\bar{s} \in Path(s, \tilde{s})$ such that $d(p, \bar{s}) = r$. Hence, $\bar{s} \in \mathcal{S}_{cand}$, contradicting the definition of s and $\tilde{s}$ as two points on G such that there exist no candidate along $Path(s, \tilde{s})$. $\qquad\square$

The dominance property for $\mathcal{S}_{cand}$ is now an immediate consequence of Lemma 3.2.

Theorem 3.3. *If CovStopLoc is feasible, then there exists an optimal solution $\mathcal{S} \subseteq \mathcal{S}_{cand}$.*

Proof. Let $\mathcal{S}$ be optimal and $\mathcal{S} \nsubseteq \mathcal{S}_{cand}$. Take any $s \in \mathcal{S} \setminus \mathcal{S}_{cand}$ and let $\tilde{s}$ be a candidate and $Path(s, \tilde{s})$ be a path not containing any other candidate point. We define

$$\bar{\mathcal{S}} = \mathcal{S} \setminus \{s\} \cup \{\tilde{s}\}\,.$$

Due to Lemma 3.2, $cover(s) \subseteq cover(\tilde{s})$ and consequently

$$\mathcal{P} = cover(\mathcal{S}) \subseteq cover(\bar{\mathcal{S}}),$$

i.e., $\bar{\mathcal{S}}$ is also feasible. Furthermore, $\left|\bar{\mathcal{S}}\right| \le |\mathcal{S}|$ and $\bar{\mathcal{S}}$ contains strictly less stops which are not in $\mathcal{S}_{cand}$. We can iterate this process until all stops in $\mathcal{S} \setminus \mathcal{S}_{cand}$ are replaced by stops in $\mathcal{S}_{cand}$. $\qquad\square$

To show the polynomiality property of $\mathcal{S}_{cand}$, we first consider CovStopLoc in the planar environment where the distance is measured with rectangular distance. The candidate set defined in (3.7) is given by the intersection points of G with the boundaries of the rectangular unit balls of radius r about the demand facilities p (see Figure 2).
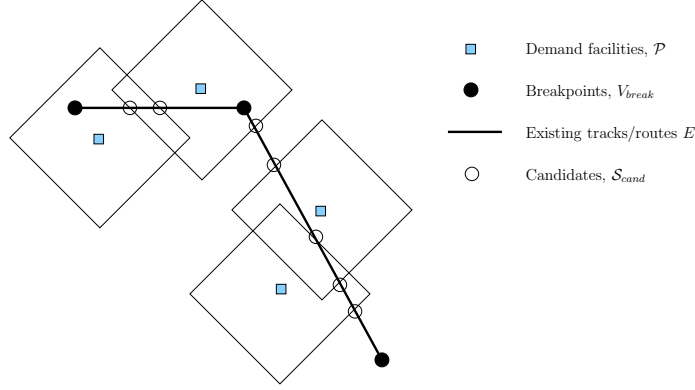
8

Figure 2: Candidates for CovStopLoc in planar environment.

In [SHLW02] it was shown that if $\mathcal{S}_{cand}$ is finite dominating set for the continuous StopLoc, then it consists of at most $\mathcal{O}(M|E|)$ candidates and hence satisfies the polynomiality property. Otherwise, if the boundary of some rectangular unit ball contains a linear piece which coincides with G in infinitely many points (see Figure 3), [SHLW02] proved that only the end points of the corresponding interval need to be considered. Hence, also in this case there are at most $\mathcal{O}(M|E|)$ candidates altogether.
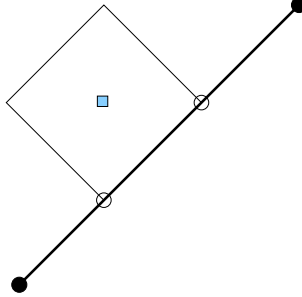

Figure 3: Candidates for CovStopLoc where some edge of the PTN coincides with a segment of some rectangular ball.

As a result, we get the following theorem.

Theorem 3.4. *In planar environment where the distance is measured with rectangular metric, the candidate set for CovStopLoc as defined in (3.7) is finite with polynomial cardinality.*

The resulting discrete optimization problem turns out to be a set covering problem. For some special cases (e.g. if the graph G consists of points along a straight line) the problem is solvable efficiently in polynomial time. Details can be found in [SHLW02, Sch06]. Moreover, it was shown that the dominance property of a finite candidate set of polynomial cardinality still

9

holds when covering is defined with respect to an arbitrary norm or even a gauge.

In the following, the distance between demand facilities $p \in \mathcal{P}$ and points $s \in \mathcal{S}$ is given by the network shortest path distance $ND(p,s)$ given by (2.2). The next theorem shows that in the network environment, $\mathcal{S}_{cand}$ is also a finite set of polynomial cardinality.

Theorem 3.5. *Given an instance of the network CovStopLoc with edge set E in the underlying PTN and demand facilities $\mathcal{P}$. Let*

$$\mathcal{S}_{cand} := \{s \in G : \exists\, p_m \in \mathcal{P} \text{ with } ND(p_m, s) = r\}$$

be the candidate set defined in (3.7). Then

$$|\mathcal{S}_{cand}| \leq 2\,|E|\,|\mathcal{P}|\,. \tag{3.9}$$

Proof. Consider a demand facility $p \in \mathcal{P}$ and a point $s = (e, t)$ on edge $e \in E$. From the definition of network distance given by (2.2), it follows that if we move s along e, then $ND(p,s) = ND(p,(e,t))$ is continous, concave, and piecewise linear in t with at most two linear pieces with slopes l_e and $-l_e$ respectively (see Figure 4).
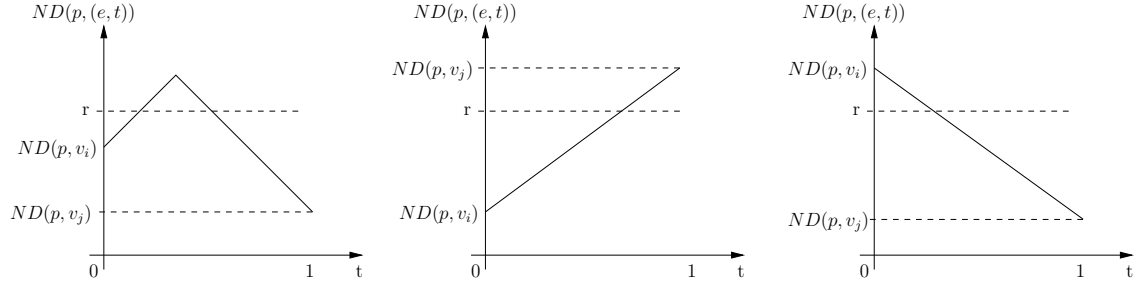

Figure 4: Network Distance.

Thus, given $r \in \mathbb{R}$, there exist at most two points on edge e that have distance exactly r from p, and the result follows. $\square$

10

4   Accessibility

The *Accessibility Stop Location Problem (AccessStopLoc)* as introduced in Section 2 is

$$\text{minimize } f(\mathcal{S}) := \sum_{p \in \mathcal{P}} d(p, \mathcal{S})$$
$$\text{subject to } \quad \mathcal{S} \subseteq G \qquad\qquad (AccessStopLoc)$$
$$|\mathcal{S}| \le k$$

Recall that $G = (V, E)$ is the public transportation network (PTN) and $\mathcal{S} \subseteq G$ indicates that $\mathcal{S}$ is a subset of the points of the PTN. Thus AccessStopLoc is a special case of the *restricted location problem* as discussed, for instance, in [Ham95, NH95, HS97]. More specifically, AccessStopLoc is a *restricted k-median problem*, where the (at most) k new locations are required to be points of the PTN.

Two problems are considered, completely redesigning all stops and opening additional stops. In the first problem, we assume that the customers will be served only by the new stops, thus we look for $\mathcal{S}^* \subseteq G$ that minimizes the objective function:

$$\min \ f(\mathcal{S}) = \sum_{p \in \mathcal{P}} d(p, \mathcal{S}) \tag{4.1}$$

In the second model, we assume that some stations do already exist in a subset $V_{stop} \subseteq V$ and the customers can go to both the existing and the new stops, whichever is closer. For this model, the objective function is:

$$\min \ f'(\mathcal{S}) = \sum_{p \in \mathcal{P}} d(p, V_{stop} \cup \mathcal{S}) \tag{4.2}$$

where $d(p, V_{stop} \cup \mathcal{S}) := \min_{s \in V_{stop} \cup \mathcal{S}} d(p, s)$ is the distance between a demand facility p and the set $V_{stop} \cup \mathcal{S}$ of existing and new stops.

As for the covering objective function, we will show in the following subsections that AccessStopLoc can be solved using the FDS and polynomiality property in the planar as well as in the network environment.

4.1   Accessibility Problem in Planar Environment

In this section, we will explain our ideas for solving the planar AccessStopLoc using the rectangular distance

$$d(p, s) := |p_1 - s_1| + |p_2 - s_2|.$$

As will be shown in Section 5, these ideas can be carried over to other distance functions, for instance, polyhedral gauges.

4.1.1 Single Location AccessStopLoc

We begin with the special case of establishing only a single new stop. Moreover, we first assume that all demand facilities are served by the new stop, i.e., there is so far no stop in the PTN. (Later in this subsection we will relax this condition.)

Using the definition of the rectangular distance and the denotation $f(s)$ instead of $f(\{s\})$ and denote the index set $\{1, \ldots, M\}$ by $\mathcal{M}$ we can rewrite Equation (4.1).

$$
\begin{aligned}
\min \ f(s) &= \sum_{m \in \mathcal{M}} l_1(p_m, s) \\
&= \sum_{m \in \mathcal{M}} |p_{m1} - s_1| + |p_{m2} - s_2| .
\end{aligned}
\tag{4.3}
$$

We sort the first and second coordinates of p_m in non-decreasing order,

$$
\begin{aligned}
p_{m_1 1} \le p_{m_2 1} \le \quad \ldots \quad \le p_{m_M 1} \\
p_{n_1 2} \le p_{n_2 2} \le \quad \ldots \quad \le p_{n_M 2} .
\end{aligned}
$$

In the case where $p_{m_i 1} = p_{m_{i+1} 1} = \ldots = p_{m_j 1}$ we omit $p_{m_{i+1} 1}, \ldots, p_{m_j 1}$. The list of first coordinates then becomes $\tilde{p}_{11}, \ldots, \tilde{p}_{P1}$ $(P \le M)$, where $\tilde{p}_{11} < \ldots < \tilde{p}_{P1}$. Moreover, we define weights $v_{i1} := |\{p_{m1} : p_{m1} = \tilde{p}_{i1}\}|$ for $\tilde{p}_{i1}$, $i = 1, \ldots, P$. The list $\tilde{p}_{12} < \ldots < \tilde{p}_{Q2}$ and weights $v_{12}, \ldots, v_{Q2}$, $(Q \le M)$ with respect to the second coordinates of p_m are defined analogously. Additionally we define $\tilde{p}_{01} = \tilde{p}_{02} = -\infty$ and $\tilde{p}_{P+1,1} = \tilde{p}_{Q+1,2} = \infty$.

Now we get a decomposition of $\mathbb{R}^2$ into rectangles

$$
\begin{aligned}
< i, j > := \{ (x_1, x_2) \quad : \quad &\tilde{p}_{i1} \le x_1 \le \tilde{p}_{i+1,1}, \\
&\tilde{p}_{j2} \le x_2 \le \tilde{p}_{j+1,2} \}
\end{aligned}
$$

for $i \in \mathcal{P}_0 := \{0, 1, \ldots, P\}$ and $j \in \mathcal{Q}_0 := \{0, 1, \ldots, Q\}$.

Definition 4.1. (Construction lines) The lines determining the rectangles $< i, j >$,

$$
\mathcal{L} := \{ x_1 = \tilde{p}_{m1} : m \in \{1, \ldots, P\} \} \cup \{ x_2 = \tilde{p}_{n2} : n \in \{1, \ldots, Q\} \}
$$

are called *construction lines*.

We now show the finite dominating set (FDS) property for the intersection points between edges e of the PTN and the construction lines as candidate set.
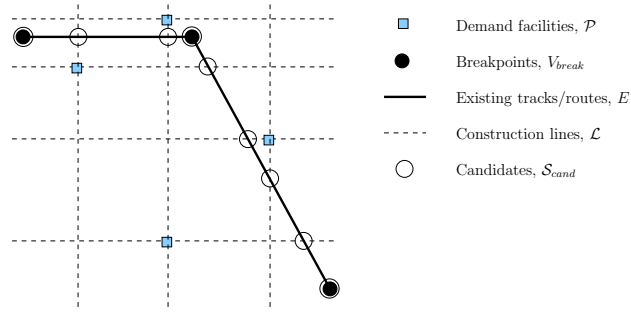
Figure 5: Candidates for AccessStopLoc in planar environment.

Definition 4.2. (Candidate Set) $\mathcal{S}_{cand} := V_{break} \cup \mathcal{C}$ where

$$\mathcal{C} := \{e \cap l : e \in E, l \in \mathcal{L}, \text{ and } |e \cap l| = 1\}.$$

The candidate set $\mathcal{S}_{cand}$ is illustrated in Figure 5.

Theorem 4.3. $\mathcal{S}_{cand}$ *is an FDS for AccessStopLoc in the case of a single stop, i.e., there exists an optimal solution* $s^* \in \mathcal{S}_{cand}$.

Proof. From (4.3) we know that the objective function f is piecewise linear in the plane, in particular, f is a linear function on each rectangle $< i, j >$, $i \in \mathcal{P}_0$, $j \in \mathcal{Q}_0$. Recall from Section 2 that the PTN has a planar representation, where each edge $e \in E$ is a line segment in $\mathbb{R}^2$. Thus, for each linear part $e \subseteq \mathcal{S}$, the intersection $e \cap < i, j >$ is either empty or again a line segment, and f is linear on $e \cap < i, j >$. Hence, the linear programming problem

$$\min\{f(s) : s \in e \cap < i, j >\}$$

is either infeasible or has an optimal solution at an endpoint of the segment $e \cap < i, j >$. An endpoint of $e \cap < i, j >$ is either a breakpoint, hence in V_{break}, or an intersection point $e \cap l$ for some construction line l not containing e, hence in $\mathcal{C}$. ☐

Note that $\mathcal{C}$ is contained in the set of intersection points $\mathcal{S} \cap \mathcal{L}$ between the tracks and the construction lines, but neglecting cases where a construction line l coincides with a linear segment e of the tracks. Such a situation is illustrated in Figure 6. As the proof of Theorem 4.3 shows, the intersection $l \cap e$ need not be considered in this case.

After establishing that $\mathcal{S}_{cand}$ is a finite dominating set for AccessStopLoc, we can use the algorithm described at the end of Section 2 to find $\{s^*\}$, the optimal solution of single AccessStopLoc.

Next we establish the polynomiality property. The candidates $s_1, ..., s_L$ can be determined in $\mathcal{O}(|V| + M|E|)$. The effort to compute the objective
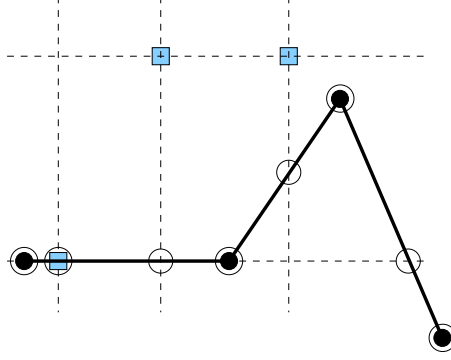
Figure 6: Candidates for AccessStopLoc where some edge of the PTN coincides with a segment of some construction line.

value for each candidates is $\mathcal{O}(M)$, thus the overall complexity of the algorithm is $\mathcal{O}(M^2|E|)$.

Next, we remove the condition that all demand facilities have to be served by the new stop, i.e., we consider a subset $V_{stop} \neq \emptyset$ of the node set $V = V_{stop} \cup V_{break}$ as well as a single new one. To distinguish this problem from the former one, we denote it by AccessStopLoc'. We will see that several results of the previous model, with some modifications, can be carried over to this model.
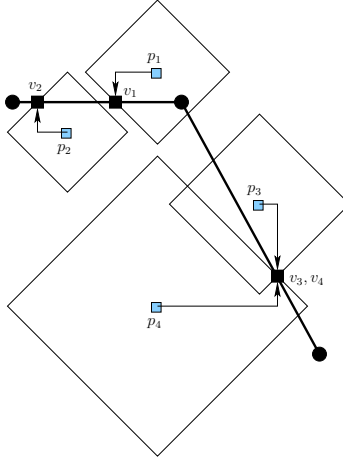

Figure 7: Graphical solution idea.

For each demand facility $p_m, m \in \mathcal{M}$, we determine an existing stop

14

$v_m \in V_{stop}$ that is closest to p_m.

$$v_m \quad \in \quad \underset{v \in V_{stop}}{\operatorname{argmin}} \, l_1(p_m, v)$$

$$d_m \quad := \quad l_1(p_m, v_m)$$

For $s \in \mathbb{R}^2$, we have to decide whether p_m is closer to s than to the current closest stop v_m. This is easily done using the rectangular circle

$$D_m \quad := \quad \left\{ s \in \mathbb{R}^2 : l_1(p_m, s) \leq d_m \right\}, m \in \mathcal{M}. \tag{4.4}$$

with center p_m and radius d_m. Customers in demand facility p_m will prefer a new stop in s to the previous stops if and only if s lies inside this circle. Therefore, we only need to consider points that lie in the interior of the set

$$D \quad := \quad \bigcup_{m \in \mathcal{M}} D_m, \tag{4.5}$$

i.e., in the interior of at least one circle (see Figure 7). Clearly,

$$f'(s) \leq f'(v) = f'(\emptyset) = \sum_{m \in \mathcal{M}} d_m =: d \quad \forall \, v \in V_{stop}, s \in \mathbb{R}^2. \tag{4.6}$$

and

$$f'(s) < d \quad \Leftrightarrow \quad s \in int(D). \tag{4.7}$$

Definition 4.4. (Construction lines) The construction lines for AccessStopLoc' are

$$\mathcal{L}' = \{ l \cap int(D) : l \in \mathcal{L} \}$$

The construction lines $l \in \mathcal{L}'$ decompose D_m into four quadrants (see Figure 8 for an illustration):

$$D_{m1} \quad := \quad \{(x_1, x_2) \in D_m : x_1 \geq p_{m1}, x_2 \geq p_{m2}\}$$
$$D_{m2} \quad := \quad \{(x_1, x_2) \in D_m : x_1 < p_{m1}, x_2 \geq p_{m2}\}$$
$$D_{m3} \quad := \quad \{(x_1, x_2) \in D_m : x_1 < p_{m1}, x_2 < p_{m2}\}$$
$$D_{m4} \quad := \quad \{(x_1, x_2) \in D_m : x_1 \geq p_{m1}, x_2 < p_{m2}\}$$

Using the denotation

$$\mathcal{M}_s := \{m \in \mathcal{M} : s \in D_m\} \subseteq \mathcal{M}$$

for all $s \in \mathbb{R}^2$, we can rewrite the objective function (4.2).

$$
\begin{aligned}
f'(s) \;=\; & \sum_{m \in \mathcal{M}} l_1(p_m, V_{stop} \cup \{s\}) && (4.8) \\[2mm]
=\; & \sum_{m \in \mathcal{M}} \min\{d_m, l_1(p_m, s)\} && (4.9) \\[2mm]
=\; & \sum_{m \in \mathcal{M} \setminus \mathcal{M}_s} d_m + \sum_{m \in \mathcal{M}_s} l_1(p_m, s) && (4.10) \\[2mm]
=\; & \sum_{m \in \mathcal{M} \setminus \mathcal{M}_s} d_m + \\
& \sum_{m: s \in D_{m1}} ((s_1 - p_{m1}) + (s_2 - p_{m2})) + \\
& \sum_{m: s \in D_{m2}} ((p_{m1} - s_1) + (s_2 - p_{m2})) + \\
& \sum_{m: s \in D_{m3}} ((p_{m1} - s_1) + (p_{m2} - s_2)) + \\
& \sum_{m: s \in D_{m4}} ((s_1 - p_{m1}) + (p_{m2} - s_2)) && (4.11)
\end{aligned}
$$

From (4.11), we conclude that f' is linear in each D_{mi}, $m \in \mathcal{M}$, $i = 1, \ldots, 4$. This fact can be used to establish a finite candidate set $\mathcal{S}'_{cand}$ that contains a global optimal solution for AccessStopLoc' (see Figure 8).

Definition 4.5. (Candidate Set) $\mathcal{S}'_{cand} := V'_{break} \cup \mathcal{C}'$ where

$$
\begin{aligned}
V'_{break} \;:=\;& V_{break} \cap int(D) \\
\mathcal{C}' \;:=\;& \{e \cap l : e \in E, l \in \mathcal{L}', \text{ and } |e \cap l| = 1\}.
\end{aligned}
$$

Recall that $V_{break} \subseteq V$ is the set of breakpoints of PTN $G = (V, E)$.

Theorem 4.6. *The optimal solution for AccessStopLoc' is contained in $\mathcal{S}'_{cand}$.*

Proof. We use a similar argument as in the proof of Theorem 4.3. Due to (4.11), f' is linear in each D_{mi}, $m \in \mathcal{M}$, $i = 1, \ldots, 4$. Since the tracks are assumed to be piecewise linear in the plane, f' is again linear on $e \cap D_{mi}$ for each linear part e. If $e \cap D_{mi}$ is not an empty set, then f' attains its minimal at the endpoint of $e \cap D_{mi}$, that is either at a breakpoint, hence in V'_{break}, or an intersection point $e \cap l$ for some construction line l not containing e, hence in $\mathcal{C}'$, or at an intersection $e \cap \partial D_m$ where ∂D_m is the boundary of D_m.

It will be shown in the following lemma that $f'(s)$ will not achieve a minimum at a point $\tilde{s} \in \partial(D_m), m \in \mathcal{M}$, even though it may be an endpoint of $e \cap D_{mi}$.

Thus, we only need to consider $s \in \mathcal{S}'_{cand} = V'_{break} \cup \mathcal{C}'$. $\qquad \square$
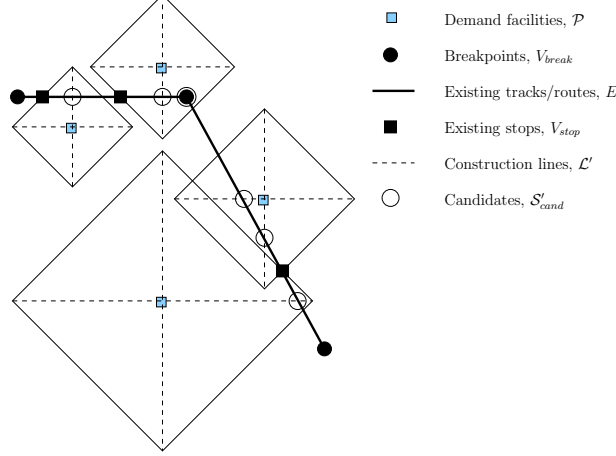
Figure 8: Candidates for AccessStopLoc' in planar environment.

Lemma 4.7. $f'(s)$ *can not achieve a minimum at points* $s \in \partial(D_m), m \in \mathcal{M}$.

Proof. Let $\tilde{s} \in G \cap \partial D_{\tilde{m}}$ for some $\tilde{m} \in \mathcal{M}$. To show that $\tilde{s}$ cannot be a minimum, we consider the piecewise linear functions

$$g_m(s) := \min\{d_m, l_1(p_m, s)\}$$

giving the distance from p_m to the closest facility in $V_{stop} \cup \{s\}$. Let $U := U(\tilde{s}) \cap G$ be a (sufficiently small) neighborhood about $\tilde{s}$ On U we have:

- g_m has a local maximum in s if $s \in G \cap \partial D_m$,

- g_m is linear if $s \notin G \cap \partial D_m$.

Summing up, we obtain that $f(s) = \sum_{m \in \mathcal{M}} g_m(s)$ cannot have a minimum at $\tilde{s}$, hence $\tilde{s}$ is not optimal. $\qquad\square$

Figure 9 illustrates the function $f'(s)$ along a track $e \in \mathcal{S}$.

Similar as for AccessStopLoc, the proof of Theorem 4.6 shows that we do not need to consider the intersection $l \cap e$ where the construction line $l \in \mathcal{L}'$ coincides with a linear segment e of the tracks. It follows that the candidate set $\mathcal{S}'_{cand}$ is finite, consisting of at most $\mathcal{O}(M|E|)$ candidates.

The computation of v_m, d_m, and D_m needs $\mathcal{O}(M|V|)$ time and the effort to compute objective value for each candidates is $\mathcal{O}(M)$. Hence, the overall complexity of the algorithm as described at the end of Section 2 for single AccessStopLoc' is $\mathcal{O}(M^2|E|)$.

It can happen that $\mathcal{S}'_{cand} = \emptyset$, namely if there does not exist any point on the tracks that improves the objective value $f'(v), v \in V_{stop}$. Thus, in such a situation it is not reasonable to add any station to our PTN.
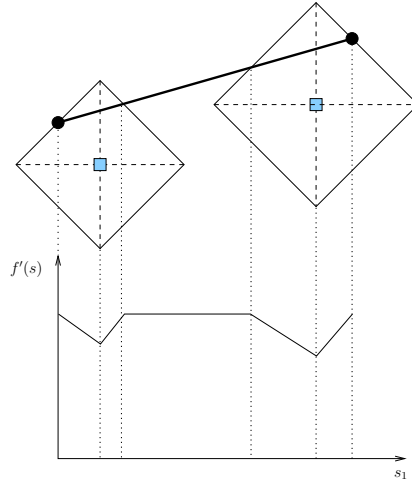
17

Figure 9: $f'(s)$ along one track.

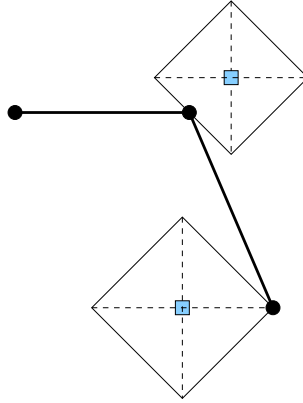

Figure 10: An example where the existing PTN can not be improved.

Notice that $\mathcal{S}'_{cand} \subseteq \mathcal{S}_{cand}$. Moreover, if $\mathcal{S}'_{cand} = \mathcal{S}_{cand}$, then AccessStopLoc and AccessStopLoc' are equivalent. Such a situation happens, for example, if all the existing stops are at infinity. In practice, this can be interpreted as follows: if all the existing stops are far away from the set of demand facilities, we do not need to consider them in our planning to open some additional stop. All the demand facilities will be served only by the new stop, and the problem is reduced to the problem of completely redesigning all stops.

4.1.2 Multiple Location AccessStopLoc

Now we are ready to analyze the general stop location problem where we want to establish at most k stops. In what follows we will only consider AccessStopLoc, but all the arguments also hold for AccessStopLoc'.

We know that the optimal solution for the single stop problem is contained in the candidate set $\mathcal{S}_{cand}$. The next theorem shows that this holds also for the general problem.

Theorem 4.8. *For AccessStopLoc, there exists an optimal solution $\mathcal{S}^* \subseteq \mathcal{S}_{cand}$, $|\mathcal{S}^*| \leq k$*

Proof. Suppose we have already a set of solutions $\tilde{\mathcal{S}}$ for the set of demand facilities $\mathcal{P}$ with $|\tilde{\mathcal{S}}| \leq k$, $\tilde{\mathcal{S}} \nsubseteq \mathcal{S}_{cand}$. Take $\tilde{s} \in \tilde{\mathcal{S}} \setminus \mathcal{S}_{cand}$ and consider only demand facilities that are served by $\tilde{s}$. Use the algorithm given in Section 2 to find the single stop that minimizes the total distance from these demand facilities. Due to Theorem 4.3 we obtain an optimal solution $\tilde{s}' \in \mathcal{S}_{cand}$. By reassigning the covering of the demand facilities, we will get at least the same objective value. Iterate this process until all stops in $\tilde{\mathcal{S}} \setminus \mathcal{S}_{cand}$ are replaced by stops in $\mathcal{S}_{cand}$. $\qquad\qquad\square$

However, AccessStopLoc is $\mathcal{NP}$-hard even if we only consider the finite dominating set $\mathcal{S}_{cand}$.

Theorem 4.9. *AccessStopLoc is $\mathcal{NP}$-hard*

Proof. The proof is a reduction of the rectangular k-median problem to the decision version of AccessStopLoc.
(Rectangular k-median problem) Given a set $\mathcal{P} = \{p_1, \ldots, p_M\} \subset \mathbb{R}^2$ of points in the plane and positive rational number k. Is there a set $S \subset \mathbb{R}^2$ of k points such that if $l_1(p_m, S)$ is the length of shortest l_1 travel path from p_m to the closest point in S, then $\sum_{m=1}^{M} l_1(p_m, S) \leq k$?
The rectangular k-median problem is known to be $\mathcal{NP}$-hard (see [MS84]).
It can easily be shown (see [Poe05]) that rectangular k-median is polynomially transformable to an instance of AccessStopLoc as follows:

- Leave $\mathcal{P}$ and k as it is.

- Define G such that it contains sufficiently large parts of the construction lines, where V consists of the end points of the segments and of the intersection points between pairs of the segments. G can now be defined as a grid graph with V as nodes and edges corresponding to the linear pieces between the nodes.

$\square$

Heuristic methods based on a node partitioning scheme and some numerical results are presented in [Poe05].

4.2 Accessibility Problem in Street Network Environment

To complete our analysis on stop location models, we discuss next the accessibility problem on a street network as defined in Section 2.

In locational analysis, the problem of locating k facilities on a graph $G = (V, E)$ to serve the customers such that the sum of the weighted shortest distances between customers and facilities is minimized is commonly referred as k-median problem. If the customers are only located on the vertices of G, k-median has the node optimality property (see [Hak65]), meaning that an optimal solution exists on the vertices. Due to this property, one must only examine all subsets of V containing k vertices to find the optimal solution. However, even if we restrict the facilities to be a subset of vertices, solving this problem is non-trivial. The number of possible solutions is exponentially increasing in the number n of vertices and the number k of facilities. Kariv and Hakimi [HK79] showed that the k-median problem is $\mathcal{NP}$-hard, even when the network has a simple structure, e.g. if it is a planar graph with maximum vertex degree three.

In the street network environment, our stop location model can be considered as a variant of the restricted k-median problem. Given the SPTN $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where vertex set $\mathcal{V}$ contains the demand facilities $\mathcal{P}$, breakpoints V_{break}, and existing stops V_{stop} (which can be empty), we define the weights of vertices $v \in \mathcal{P}$ as equal to one, whereas the weights of all other vertices is set to zero. Since the set of demand facilities $\mathcal{P}$ is a subset of $\mathcal{V}$, the node optimality property also holds for AccessStopLoc. That means, our candidate set contains only the set of breakpoints, which is obviously finite.

Definition 4.10. (Candidate Set) $\mathcal{S}_{cand} := V_{break}$

We remark that the set of demand facilities is not included in the candidate set, even though $\mathcal{P} \subseteq \mathcal{V}$. This is due to the fact that the potential location is restricted only on the existing PTN, hence we only need to consider the set of vertices on the PTN, which by definition consists of the breakpoints and of the existing stops. For AccessStopLoc we assume the set of existing stops is empty, thus the above definition of S_{cand} follows.

20

For AccessStopLoc', the arguments used in the planar case can be carried over to the network environment. For each demand facility $p_m \in \mathcal{P}$, we determine a closest stop $v_m \in V_{stop}$, that is

$$v_m \in \operatorname*{argmin}_{v \in V_{stop}} ND(p_m, v)$$

The objective function is improved if we locate a new stop in the interior of D, where D is defined as in (4.5) but the radius is of course measured by the network distance. That means, the set $\mathcal{S} \setminus int(D)$ can be considered as forbidden region. The node optimality property is however still valid for AccessStopLoc', leading to the result that we only need to consider the set of breakpoints that lie on $int(D)$. Following the notation from the planar case, we denote this set by V'_{break}.

Definition 4.11. (Candidate Set) $\mathcal{S}'_{cand} := V'_{break}$

With this result, we have shown that in the network environment we can also discretize the continuous problem. Solving AccessStopLoc and AccessStopLoc' on SPTN can be done by any method for the restricted k-median problem.

5   Conclusion

In this paper we have discussed solution approaches to four classes of stop location problems, the covering and the accessibility problem in planar and network environments. Our solution algorithms are based on a discretization, in which a candidate set $\mathcal{S}_{cand}$ is defined, which has the finite dominating set and the polynomiality property. In the AccessStopLoc we distinguish the case of existing stops and of planning from scratch. It should be noted that these two problems are equivalent for the CovStopLoc (see [Sch06]).

Several other variations of the problems are possible, three of them should be mentioned below.

First, the rectangular distance function used for planar StopLoc problems can be generalized to polyhedral gauge distances.

A *polyhedral gauge* is defined by its unit ball B, which is a convex polyhedron in the plane $\mathbb{R}^2$ containing the origin $(0,0)$ in its interior. Let $exp(B) = \{v_1, \ldots, v_R\}$ be the extreme points of B. The fundamental directions of B are the half-lines starting at the origin with direction $v_r, r \in \{1, \ldots, R\}$. Rooting the fundamental directions as construction lines at each demand facility $p_m, \forall m \in \mathcal{M}$, the candidate set is defined as the union of the breakpoints and the intersection points between edges $e \in E$ and construction lines. With similar arguments as for the rectangular distance, we can show that this candidate set is indeed a finite dominating set.

Another interesting variant of the StopLoc is the *Center Stop Location Problem* with respect to a given positive integer k defined by

$$
\begin{aligned}
\text{minimize } g(\mathcal{S}) &:= \max_{p \in \mathcal{P}} \ d(p, \mathcal{S}) \\
\text{subject to } \quad \mathcal{S} &\subseteq G \\
|\mathcal{S}| &= k
\end{aligned}
\qquad (CovStopLoc)
$$

This problem can be tackled by using the results either of CovStopLoc (by a binary search strategy) or of AccessStopLoc (by looking at constraint center location problems instead of median location problems).

Another straightforward extension is to take the number of potential passengers in each demand facility into account by using weights for each $p \in \mathcal{P}$. This leads to weighted objective functions in the set-covering problem and in the restricted location problem in Sections 3 and 4.

References

[Das95] M. S. Daskin. *Network and Discrete Location: Models, Algorithms, and Applications.* Wiley, New York, 1995.

[DH02] Z. Drezner and H. W. Hamacher, editors. *Facility Location: Application and Theory.* Springer, Berlin, 2002.

[FLMW92] R. L. Francis, F. Leon, J. McGinnis, and J. A. White. *Facility Layout and Location: An Analytical Approach.* Prentice Hall, New York, 2nd edition, 1992.

[Gle75] J. Gleason. A set covering approach to bus stop allocation. *Omega*, 3:605–608, 1975.

[Hak65] S. L. Hakimi. Optimum distribution of switching centers in a communication network and some related graph theoritic problems. *Operation Research*, 13:462–475, 1965.

[Ham95] H. W. Hamacher. *Mathematische Lösungsverfahren für planare Standortprobleme.* Vieweg Verlag, 1995.

[HK79] S. L. Hakimi and O. Kariv. An algorithmic approach to network location problems. part 2 : The p-medians. *SIAM Journal on Applied Mathematics*, 37:539–560, 1979.

[HLS$^+$01] H. W. Hamacher, A. Liebers, A. Schöbel, D. Wagner, and F. Wagner. Locating new stops in a railway network. *Electronic Notes in Theoretical Computer Science*, 50(1):1–11, 2001.

[Hor04] S. Horn. Verkehrsplanung und schulmathematik. Master's thesis, Fachbereich Mathematik, Technische Universität Kaiserslautern, Germany, 2004.

[HS97] H.W. Hamacher and A. Schöbel. A note on center problems with forbidden polyhedra. *Letters of Operations Research*, 20:165–169, 1997.

[KPS$^+$03] E. Kranakis, P. Penna, K. Schlude, D.S. Taylor, and P. Widmayer. Improving customer proximity to railway stations. In *Proceedings of the 5th conference on algorithms and complexity*, number 2653 in Lecture Notes in Computer Science, 2003.

[LMO02] G. Laporte, J. A. Mesa, and F. A. Ortega. Locating stations on rapid transit lines. *Computers and Operation Research*, 39:741–759, 2002.

[LMW88] R. F. Love, J. G. Morris, and G. O. Wesolowsky. *Facility Location: Models and Methods.* Elsevier Science Publishing Co.,Inc., New York, 1988.

[MDSF98] A. Murray, R. Davis, R. J. Stimson, and L. Ferreira. Public transportation access. *Transportation Research D*, 3(5):319–328, 1998.

[MMW04] M.F. Mammana, S. Mecke, and D. Wagner. The station location problem on two intersecting lines. *Electronic Notes in Theoretical Computer Science*, 92:52–64, 2004.

[MS84] N. Megiddo and K. J. Supowit. On the complexity of some common geometric location problems. *SIAM Journal on Computing*, 13:182–196, 1984.

[MSW05] S. Mecke, A. Schöbel, and D. Wagner. Stop location - complexity and approximation issues. In *Proceedings of ATMOS 2005*, 2005.

[Mur01a] A. Murray. Coverage models for improving public transit system accessibility and expanding access. Technical report, Department of Geography, Ohio State University, 2001.

[Mur01b] A. Murray. Strategic analysis of public transport coverage. *Socio-Economic Planning Sciences*, 35:175–188, 2001.

[MW04] S. Mecke and D. Wagner. Solving geometric covering problems by data reduction. In *Proceedings of European Symposium on Algorithms (ESA)*, pages 760–771, 2004.

[NH95] S. Nickel and H. W. Hamacher. Restricted planar location problem and applications. *Naval Research Logistics*, 42:967–992, 1995.

[Poe04] D. R. Poetranto. Stop location problem in public transportation networks. Master's thesis, Fachbereich Mathematik, Technische Universität Kaiserslautern, Germany, 2004.

[Poe05] D. R. Poetranto. A restricted median location model for stop location design in public transportation networks. In *Operational Research Paripatetic Postgraduate Programme (ORP3)*, pages 297–310, Valencia, Spain, 2005. Universidad Politécnica de Valencia.

[RS04] N. Ruf and A. Schöbel. Set covering problems with almost consecutive ones property. *Discrete Optimization*, 1(2):215–228, 2004.

[Sch05] A. Schöbel. Locating stops along bus or railway lines – a bicriteria problem. *Annals of Operations Research*, 2005. to appear.

[Sch06] A. Schöbel. *Optimization in Public Transportation.* Springer, 2006. to appear.

[SHLW02] A. Schöbel, H. W. Hamacher, A. Liebers, and D. Wagner. The continous stop location problem in public transportation networks. Report in Wirtschaftsmathematik 81/2002, Fachbereich Mathematik, Technische Universität Kaiserslautern, 2002. http://kluedo.ub.uni-kl.de/volltexte/2002/1338/ps/gelb_81.ps.