Annals of Operations Research 136, 211–227, 2005

Locating Stops Along Bus or Railway Lines—A Bicriteria Problem

ANITA SCHÖBEL schoebel@math.uni-goettingen.de
Institut für Numerische und Angewandte Mathematik, Georg-August Universität Göttingen, Lotzestr. 16-18, 37083 Göttingen

Abstract. In this paper we consider the location of stops along the edges of an already existing public transportation network. This can be the introduction of bus stops along given routes, or of railway stations along the tracks in a railway network. The goal is to achieve a maximal covering of given demand points with a minimal number of stops. This bicriteria problem is in general NP-hard. We present a finite dominating set yielding an IP-formulation as a bicriteria set covering problem. Using this formulation we discuss cases in which the bicriteria stop location problem can be solved in polynomial time. Extensions for tackling real-world instances are mentioned.

Keywords: set covering, bicriterial, location, dynamic programming, public transportation

Introduction

When designing or modifying a public transportation network, one has to decide about the number and the location of the stops (or stations). Unfortunately, the objective is not clear in this process, since even from a customer-oriented point of view, the following two conflicting effects of stops apply.

- On the one hand, many stops are advantageous, since they increase the accessibility for the customers. A customer is *covered* if the next stop is within a specified distance, called the *covering radius* (usually 400 m in bus transportation and 2 km in rail transportation).
- On the other hand, each additional stop increases the transportation time (e.g., by 2 minutes in rail transportation) for all trains or buses stopping there.

Consequently, it makes sense to establish as few stops as possible, in such a way that all customers are covered. For a given finite set of possible new locations, this has been done in the *discrete* stop location problem which turns out to be an unweighted set covering problem (as tackled in Toregas et al., 1971). In the context of stop location this set covering problem has been solved by Murray (2001b) using the Lagrangian-based set covering heuristic of Caprara, Fischetti, and Toth (1999), and applied in bus transportation in Brisbane, Australia, see Murray et al. (1998), Murray (2001a, 2001b). Recently, another discrete stop location model has been developed by Laporte, Mesa, and

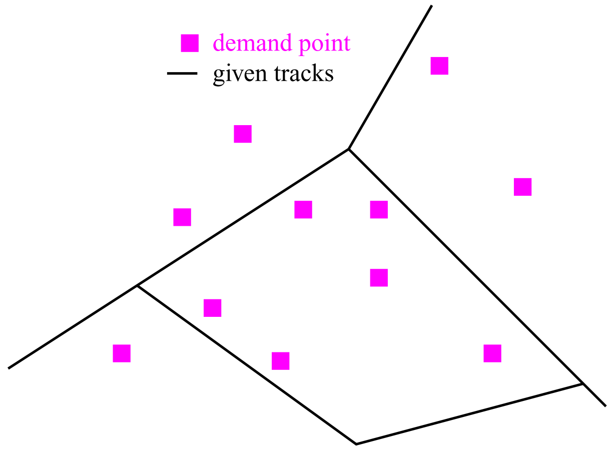

Figure 1. The set of tracks $\mathcal{T}$ and a set of demand points $\mathcal{D}$ in the plane.

Ortega (2002a). They investigate which candidate stops along one given line in Seville, Spain should be opened, taking into account constraints on the inter-station space. The problem is solved by a longest path algorithm in an acyclic graph.

On the other hand, in the *continuous* stop location problem, the whole track system (or the routes of the buses) may be used for locating stations. This problem was introduced in Hamacher et al. (2001) within a project with the largest German rail company (DB). In this paper, a genetic algorithm was used to minimize the average *door-to-door traveling time* of all customers. Minimizing the number or costs of the new stations while covering all demand points was discussed in Schöbel et al. (2002). A similar covering model has been considered in Kranakis (2002). An overview about continuous stop location is provided in Schöbel (2003). Planning not only the stops along the line but also the line itself occurs in network design problems, and has been investigated among others in Bruno, Ghiani, and Improta (1998) and Laporte, Mesa, and Ortega (2002b). A more general approach in this area is suggested in Current, ReVelle, and Cohon (1987).

In this paper we extend the continuous stop location problem as defined in Schöbel et al. (2002) to a bicriteria problem. We need the following notation (see figure 1).

Let $\mathcal{D} \subseteq \mathbb{R}^2$ be a given finite set of demand points, and PTN $= (V, E)$ be the current public transportation network, given as a set of already existing stations or breakpoints V and their direct connections E. Then the set $\mathcal{T}$ of all points of the linear embedding of the graph PTN represents the given track system (for railways) or the bus routes (for bus transportation). Given a distance measure γ_d (which may depend on the demand point d), a demand point d is *covered* by a stop $s \in \mathcal{T}$, if $\gamma_d(d, s) \leq r$. In the following we assume that γ_d is a norm-distance for each demand point d. To allow different distance functions for each demand point is due to the possibly different environments close to the demand points and allows the distance functions to be modeled more accurately. (Note that it is also possible to allow γ_d to be a distance derived from a gauge function. A gauge is defined similar to a norm, but without requiring symmetry, i.e., $\gamma_d(x, y) = \gamma_d(y, x)$ need not be satisfied, see, e.g., Minkowski (1967).

Let r be the specified covering radius. We define the *cover* of a set of stops $S \subseteq \mathcal{T}$ as

$$\text{cover}(S) = \{d \in D\colon \text{ there exists } s \in S \text{ such that } \gamma_d(d, s) \leq r\}.$$

The goal of the (unweighted) *continuous stop location problem (CSL)* as defined in Schöbel et al. (2002) is to find a set of (new) stops $S \subseteq \mathcal{T}$ with minimal cardinality, covering all demand points. This problem has been shown to be NP-complete.

Theorem 1 [(Schöbel et al., 2002)]**.** (CSL) is NP-complete.

However, in a practical setting, one might not want to cover all demand points $\mathcal{D}$ but only a given percentage of the population. Hence let us assume that for each demand point d, we have given a weight w_d representing the number of customers who would like to use public transportation, if the next station was closer than r. Then the function

$$f_{\text{cover}}(S) = \sum_{d \in \text{cover}(S)} w_d$$

gives the number of (potential) customers which live closer than r to some stop in S.

Certainly, it is preferable to cover as many customers as possible, i.e, to maximize $f_{\text{cover}}(S)$. On the other hand, establishing many new stops is costly and increases the travel time for the customers in the trains (or buses), because each stop needs an additional time of, e.g., two minutes. Since this causes dissatisfaction for the customers we use

$$f_{\text{cost}}(S) = |S|$$

as a second objective function. The *bicriteria stop location problem* (*BSL*) can now be stated.

(BSL)

Given $G = (V, E)$ with its set of points of its planar embedding $\mathcal{T} = \bigcup_{e \in E} e \subseteq \mathbb{R}^2$, as well as a finite set of points $\mathcal{D} \subseteq \mathbb{R}^2$ with weights w_d and norms (or gauges) γ_d for all $d \in \mathcal{D}$, find a set $S \subseteq \mathcal{T}$ such that both

$$f_{\text{cost}}(S) = |S| \quad \text{and}$$
$$-f_{\text{cover}}(S) = -\sum_{d \in \text{cover}(S)} w_d$$

are minimized.

The remainder of the paper is organized as follows. In the next section we introduce the two e-constraint subproblems needed for a bicriteria analysis of (BSL) and transfer the finite candidate set of Schöbel et al. (2002) for the single objective problem to these problems in Section 2. In Section 3 we analyze the situation along a polygonal line and present a dynamic programming approach which finds all efficient solutions in this case in Section 4. Further research topics are mentioned in Section 5.

1. *e*-constraint problems

What we mean by "minimizing both" objective functions is to find Pareto solutions of the problem with respect to f_{cost} and f_{cover}. Recall (e.g., from textbooks as Steuer (1989) or Ehrgott (2000)) that if $S_1, S_2 \subseteq \mathcal{T}$ denote two feasible sets of stops, S_1 dominates S_2 if

$$f_{\text{cost}}(S_1) \leq f_{\text{cost}}(S_2) \quad \text{and}$$
$$f_{\text{cover}}(S_1) \geq f_{\text{cover}}(S_2),$$

where at least one of the inequalities is strict. Then a *Pareto solution* S^* is a feasible set of stops which is not dominated by any other feasible set of stops. If S^* is a Pareto solution, then the point

$$(f_{\text{cost}}(S^*), f_{\text{cover}}(S^*))$$

is called an *efficient* point.

To find Pareto solutions we can utilize the following two one-criteria problems.

(BSL-cost(Q)**)**: Given $\mathcal{D}$, $G = (V, E)$ with its set of points $\mathcal{T}$, weights w_d, and norms (or gauges) γ_d for all $d \in \mathcal{D}$, find a set $S^* \subseteq \mathcal{T}$ such that $f_{\text{cover}}(S^*) \geq Q$ and $f_{\text{cost}}(S^*)$ is minimal.

(BSL-cover(k)**)**: Given $\mathcal{D}$, $G = (V, E)$ with its set of points $\mathcal{T}$, weights w_d and norms (or gauges) γ_d for all $d \in \mathcal{D}$, find a set $S^* \subseteq \mathcal{T}$ such that $f_{\text{cost}}(S^*) \leq k$ and $f_{\text{cover}}(S^*)$ is maximal.

These problems are called the *e*-constraint problems resulting from (BSL). Note that (BSL-cost) resembles the *location set covering problem* introduced by Toregas (1971) and Toregas et al. (1971), while (BSL-cover) is related to the *Maximal Covering Location Problem*, see Church and ReVelle (1974) or White and Case (1974).

To utilize the *e*-constraint problems in our analysis we need the following result of Haimes and Chankong (1983).

Lemma 1. Let $Q, k \in \mathbb{N}$.

1. Let S be a unique optimal solution of (BSL-cost(Q)). Then S is a Pareto solution. If more than one optimal solution of (BSL-cost(Q)) exists, the solutions that additionally maximize f_{cover} are Pareto solutions.
2. Let S be a unique optimal solution of (BSL-cover(k)). Then S is a Pareto solution. If more than one optimal solution of (BSL-cover(k)) exists, the solutions that additionally minimize f_{cost} are Pareto solutions.

Using Lemma 1 to find Pareto solutions is known as the *e-constraint method*; see, e.g., Ehrgott (2000). Unfortunately, both e-constraint problems are hard to solve.

Corollary 1. (BSL) and the two e-constraint problems (BSL-cost) and (BSL-cover) are NP-hard, even if all weights w_d are equal to 1.

Proof. From Theorem 1 we know that finding a minimum cardinality set of stations covering all demand points is NP-hard. The *decision version* of both e-constraint problems (BSL-cost(Q)) and (BSL-cover(k)) is the following:

Given $\mathcal{D}$, $G = (V, E)$ with its planar embedding $\mathcal{T}$, weights w_d, norms (or gauges) γ_d, does there exist a set $S^* \subseteq \mathcal{T}$ such that $f_{\text{cost}}(S^*) \leq Q$ and $f_{\text{cover}}(S^*) \geq k$?

Defining $Q = \sum_{d \in \mathcal{D}} w_d$ shows that the decision version of (CSL) is a special case of the decision version of both (BSL-cost(Q)) and (BSL-cover(k)) and thus both e-constraint problems are NP-hard. □

We now discuss the two lexicographic optimal solutions, which we know are Pareto solutions.

- Maximizing f_{cover} as first objective means that we have to cover **all** demand points that can be covered, i.e., all demand points $d \in \text{cover}(\mathcal{T})$. This yields (CSL), if we define

$$\mathcal{D}' = \mathcal{D} \cap \text{cover}(\mathcal{T})$$

  as the set of demand points to be covered, and hence this problem is NP-hard (see Theorem 1).
- On the other hand, minimizing f_{cost} leads to a trivial problem since it can be solved easily by not installing any stop at all.

Note that (BSL-cover(k)) was investigated in Kranakis et al. (2002) for the case of one single straight-line track and a special case with two parallel straight-line tracks. For both cases, polynomial time algorithms using dynamic programming were developed with a time complexity of $O(k|\mathcal{D}|^2)$ for the single track case. Moreover, it is shown that along one straight line track, (BSL-cover(k)) is equivalent to the one-dimensional (uncapacitated) k-facility location problem. Due to Hassin and Tamir (1991), (BSL-cover(k)) can hence be solved in $O(|\mathcal{D}|^2)$ time.

2. Integer programming formulations

To derive integer programming formulations we proceed as follows. For an edge $e \in E$ with endpoints v_1^e, v_2^e we define

$$\mathcal{T}^e(d) = \{s \in e : \gamma_d(d, s) \leq r\}$$

as the set of all points on the edge $e \subseteq \mathcal{T}$ that can be used to cover demand point d, and

$$\mathcal{T}(d) = \{s \in \mathcal{T} : \gamma_d(d, s) \leq r\}.$$

Note that $s \in \mathcal{T}(d)$ if and only if $d \in \text{cover}(s)$. The following simple observation will become important later.

Lemma 2. For each demand point $d \in \mathbb{R}^2$ the set $\mathcal{T}^e(d)$ is either empty or an interval contained in edge e.

Proof. Note that $\mathcal{T}^e(d) = e \cap \{x \in \mathbb{R}^2 : \gamma_d(d, x) \leq r\}$ is the intersection of two convex sets, namely, of the line segment e and the unit ball of the norm (or gauge) γ_d about d. Consequently, $\mathcal{T}^e(d)$ itself is a convex set contained in a line segment and hence either empty or a line segment itself. □

Let f_d^e, l_d^e denote the endpoints of the interval $\mathcal{T}^e(d)$ (which may coincide with the endpoints v_1^e, v_2^e of the edge e). We write

$$\left[f_d^e, l_d^e\right] = \mathcal{T}^e(d).$$

Along the lines of Schöbel et al. (2002) we can now derive a finite dominating set $\mathcal{S} \subseteq \mathcal{T}$ as follows. For each edge let

$$\mathcal{S}^e = \bigcup_{d \in \mathcal{D}} \left\{f_d^e, l_d^e\right\} \cup \left\{v_1^e, v_2^e\right\}$$

be the set of all endpoints of intervals $\mathcal{T}^e(d)$. This set can be ordered along the edge e (e.g., by starting in v_1^e and moving to v_2^e), resulting in a set

$$\mathcal{S}^e = \left\{s_0, s_1, \ldots, s_{N_e}\right\},$$

and we write $v_1^e = s_0 < s_1 < \cdots < s_{N_e} = v_2^e$ to indicate the order of the points with respect to $v_1^e < v_2^e$. In the following we show that

$$\mathcal{S} = \bigcup_{e \in E} \mathcal{S}^e$$

is a finite dominating set for the bicriteria stop location problem. For an illustration of $\mathcal{S}$ we refer to figure 2.

The next lemma states that we can always improve the cover of a stop by moving the stop to an appropriate point in $\mathcal{S}$.

Lemma 3 [(Schöbel et al., 2002)]. Let e be an edge of E, and let $s \in]s_j, s_{j+1}[_e$ for some $j \in \{0, 1, \ldots, N_e - 1\}$. Then

$$\text{cover}(s) \subseteq \text{cover}(s_j) \cap \text{cover}(s_{j+1}).$$

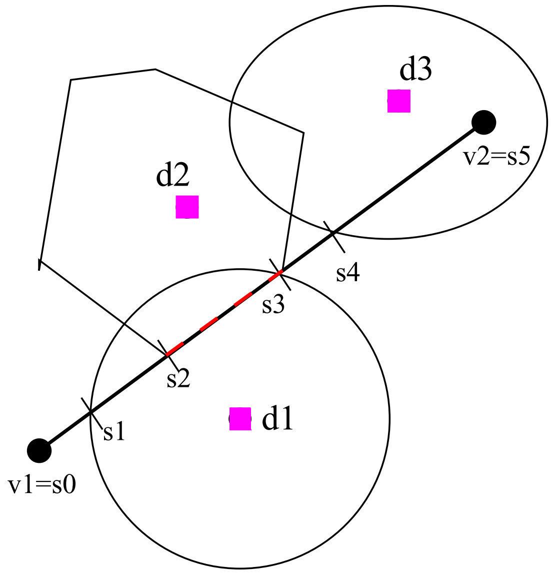

Figure 2. The set of candidates along one linear piece $e \in E$.

Theorem 2. $\mathcal{S}$ is a finite dominating set for (BSL-cost(Q)), (BSL-cover(k)), and (BSL). More precisely,

- If (BSL-cost(Q)) is feasible there exists an optimal solution $S^* \subseteq \mathcal{S}$.
- If (BSL-cover(k)) is feasible there exists an optimal solution $S^* \subseteq \mathcal{S}$.
- Let (k, Q) be an efficient solution of (BSL). Then there exists a Pareto solution $S \in \mathcal{S}$ with $f_{\text{cost}}(S) = k$ and $f_{\text{cover}}(S) = Q$.

Proof. Given some Pareto set S^*, we iteratively construct a set $S' \subseteq \mathcal{S}$ by moving stops of the given set S^* into points of $\mathcal{S}$ without changing the objective function values as follows. Let $s \in S^* \setminus \mathcal{S}$ be a point in the optimal solution and let e be the edge of s. Then determine two consecutive points $s_j, s_{j+1} \in \mathcal{S}^e$ such that s lies between s_j and s_{j+1}. According to Lemma 3 we know that $\text{cover}(s) \subseteq \text{cover}(s_j)$, hence

$$S' = S^* \setminus \{s\} \cup \{s_j\}$$

satisfies

$$\begin{aligned} f_{\text{cover}}(S^*) &\leq f_{\text{cover}}(S') \quad \text{and} \\ f_{\text{cost}}(S^*) &\geq f_{\text{cost}}(S'), \end{aligned}$$

i.e., S' is at least as good as S^* with respect to both criteria. Proceeding like this for all points in $S^* \setminus \mathcal{S}$ proves the result. $\square$

Using Theorem 2, (BSL) and its two e-constraint problems can be formulated as integer programs. As decision variable we define

$$x_s = \begin{cases} 1 & \text{if candidate } s \text{ is chosen as a new stop} \\ 0 & \text{otherwise} \end{cases}.$$

To keep track of the population covered by the new stops, we also have to know which demand points are covered and which are not. We therefore define another set of binary variables

$$y_d = \begin{cases} 1 & \text{if demand point } d \text{ is covered} \\ 0 & \text{otherwise} \end{cases},$$

and let $w = (w_{d_1}, w_{d_2}, \ldots, w_{d_{|\mathcal{D}|}})$ and $\underline{1} \in \mathbb{R}^{|\mathcal{S}|}$ be the vector with a 1 in each component.

Furthermore, we can store the covering information in the following *covering matrix* $A^{\text{cov}} = (a_{ds})$ with

$$a_{ds} = \begin{cases} 1 & \text{if } d \in \text{cover}(s) \text{ (or, equivalently, if } s \in \mathcal{T}(d)) \\ 0 & \text{otherwise} \end{cases},$$

The IP model of (BSL) can now be formulated as

$$\begin{aligned} \min \quad & \begin{pmatrix} \underline{1}x \\ -wy \end{pmatrix} \\ \text{s.t.} \quad & A^{\text{cov}}x \geq y \\ & x \in \{0,1\}^{|\mathcal{S}|} \\ & y \in \{0,1\}^{|\mathcal{D}|}. \end{aligned}$$

The IP model for (BSL-cost(Q)) is

$$\begin{aligned} \min \quad & \underline{1}x \\ \text{s.t.} \quad & A^{\text{cov}}x - y \geq 0 \\ & wy \geq Q \\ & x \in \{0,1\}^{|\mathcal{S}|} \\ & y \in \{0,1\}^{|\mathcal{D}|}, \end{aligned}$$

and (BSL-cover(k)) is given by

$$\begin{aligned} \max \quad & wy \\ \text{s.t.} \quad & A^{\text{cov}}x - y \geq 0 \\ & \underline{1}x \leq k \\ & x \in \{0,1\}^{|\mathcal{S}|} \\ & y \in \{0,1\}^{|\mathcal{D}|}. \end{aligned}$$

3. Bicriteria stop location along a polygonal line

We now analyze the situation along a polygonal line $\mathcal{T}$.

Lemma 4. If $\mathcal{T}$ is a polygonal line and $\mathcal{T}(d)$ is connected for each demand point d, then A^{cov} has the consecutive ones property, i.e., in each row of A^{cov} the ones appear consecutively.

Proof. Let $a_{ds_1} = a_{ds_2} = 1$ for $s_1 < s_2$. We then have to show that $a_{ds} = 1$ for all s with $s_1 < s < s_2$. Take a candidate s on the polygonal line between s_1 and s_2. From $a_{ds_1} = a_{ds_2} = 1$ we know that $s_1, s_2 \in \mathcal{T}(d)$. Hence, since $\mathcal{T}(d)$ is connected, also $s \in \mathcal{T}(d)$ and hence $a_{ds} = 1$. □

Note that the assumption of Lemma 4 is always satisfied if $\mathcal{T}$ consists of one single edge only, an observation which was first noted in Schöbel et al. (2002). Generalizations and decomposition results that can be used to apply this fact to more complex networks are given in Schöbel (2003).

To illustrate the condition of Lemma 4 we consider figures 3 and 4. In figure 3 an example of a polygonal line not satisfying the condition of Lemma 4 and with a coefficient matrix without consecutive ones property is given.

In this example, $\mathcal{T}$ is a polygonal line consisting of three nodes. Numbering the candidates from left to right, $A^{\text{cov}}_{\text{Fig3}}$ is given by

$$A^{\text{cov}}_{\text{Fig3}} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix},$$

which cannot be reordered to satisfy the consecutive ones property.

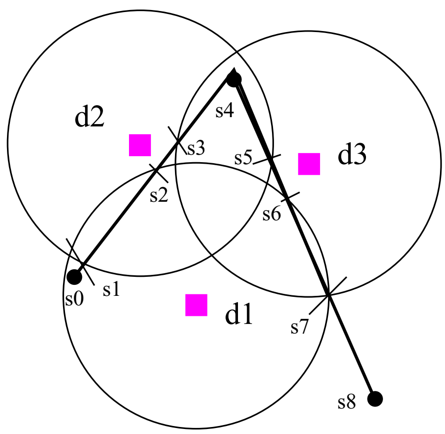

Figure 3. An instance of (BSL) on a polygonal line where $\mathcal{T}(d_1)$ is not connected, and without consecutive ones property.

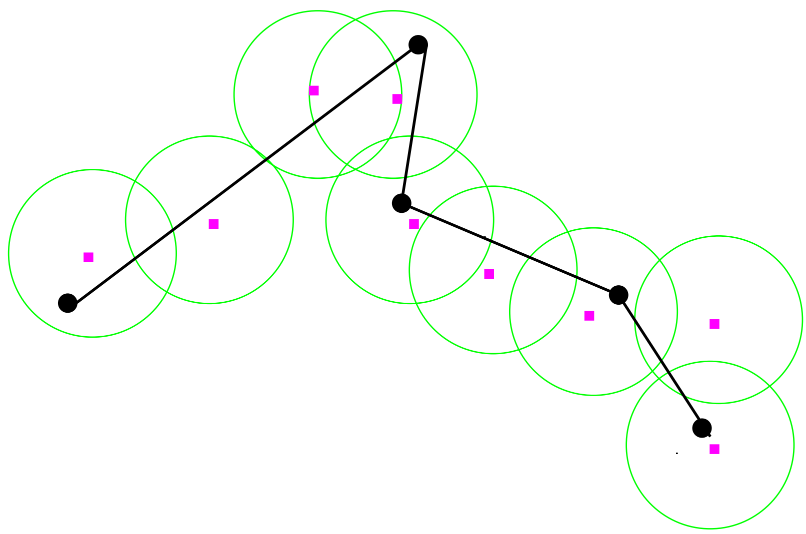

Figure 4. An instance of (BSL) on a polygonal line satisfying that all sets $\mathcal{T}(d)$ are connected, and hence having the consecutive ones property.

On the other hand, figure 4 shows an example for a polygonal line together with a set of demand points $\mathcal{D}$, in which all sets $\mathcal{T}(d)$ are connected. Hence, the covering matrix $A^{\text{cov}}_{\text{Fig4}}$ of this example satisfies the consecutive ones property.

The importance of Lemma 4 is due to the fact that matrices having the consecutive ones property are totally unimodular such that in this case the stop location problem (CSL) can be solved efficiently by linear programming methods. Unfortunately, even if A^{cov} has the consecutive ones property and $w_d = 1$ for all $d \in \mathcal{D}$, this property need not hold for the constraint versions of our problem (BSL) as the following example demonstrates.

Consider figure 5 and note that the coefficient matrix in this small example is

$$A^{\text{cov}}_{\text{Fig5}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

which has the consecutive ones property.

(**BSL**-cost(Q)): Although $A^{\text{cov}}_{\text{Fig5}}$ has the consecutive ones property this does not yield a totally unimodular coefficient matrix for (BSL-cost(Q)). Namely, the coefficient matrix of (BSL-cost) in the example shown in figure 5 is given as

$$\begin{pmatrix} 1 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix},$$

which is not totally unimodular.

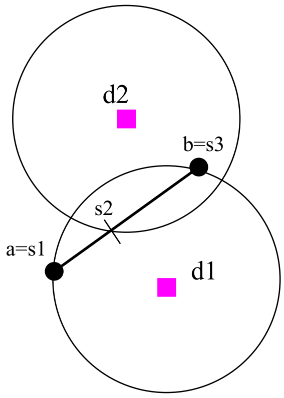

Figure 5. The coefficient matrix of (BSL-cost) is not totally unimodular.

(**BSL**-cover(k)): On the other hand, using again the example depicted in figure 5, the coefficient matrix of (BSL-cover(k)) is given by

$$\begin{pmatrix} 1 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 0 & 0 \end{pmatrix},$$

which does not have the consecutive ones property, but still is a totally unimodular matrix.

This observation holds in general.

Lemma 5. Let A^{cov} have the consecutive ones property and assume that $w_d = 1$ for all $d \in \mathcal{D}$. Then (BSL-cover(k)) can be solved by linear programming.

Proof. Note that $\begin{pmatrix} A^{\text{cov}} \\ 1\ 1\ \dots\ 1 \end{pmatrix}$ has the consecutive ones property and hence is totally unimodular. Thus, also $\begin{pmatrix} A^{\text{cov}} \\ -1\ -1\ \dots\ -1 \end{pmatrix}$ is totally unimodular and hence the coefficient matrix

$$\begin{pmatrix} A^{\text{cov}} & -I \\ -1\ -1\ \dots\ -1 & 0\ 0\ \dots\ 0 \end{pmatrix}$$

of the IP-formulation of (BSL-cover(k)) also satisfies this property. Consequently, the result follows from integer programming theory, see, e.g., Nemhauser, and Wolsey (1988). □

Based on this observation (although not true for arbitrary weights w_d) we suggest to solve a family of e-constraint problems of type (BSL-cover) to find all efficient solutions of (BSL) in the case that A^{cov} has the consecutive ones property. In the next section we show how this can be done efficiently by dynamic programming.

4. Dynamic programming approach

To develop a dynamic programming approach for (BSL) we first investigate cover(S) in more detail. Again, consider a polygonal line $\mathcal{T}$ satisfying the assumption of Lemma 4 and let $\mathcal{S}$ be the set of candidates. We assume that the candidates have been ordered along $\mathcal{T}$, e.g., from left to right.

Lemma 6. Let $\mathcal{T}$ be a polygonal line satisfying that $\mathcal{T}(d)$ is connected for all $d \in \mathcal{D}$. Let $S = \{s_1, \ldots, s_p\} \subseteq \mathcal{T}$ be any set of points with $s_1 < \cdots < s_p$. Then for all $i = 1, \ldots, p-1$ we have

$$\text{cover}(s_{i+1}) \setminus \text{cover}\{s_1, \ldots, s_i\} = \text{cover}(s_{i+1}) \setminus \text{cover}(s_i).$$

Proof. Since "$\subseteq$" is trivial, we only need to verify "$\supseteq$".

To this end, let $d \in \text{cover}(s_{i+1}) \setminus \text{cover}(s_i)$. We show that $d \notin \text{cover}(s_j)$ for all $j \leq i$. Assume to the contrary that $d \in \text{cover}(s_j)$ for some $j < i$ but that $d \in \text{cover}(s_{i+1})$. This means that $s_j \in \mathcal{T}(d)$ and $s_{i+1} \in \mathcal{T}(d)$, and, since $\mathcal{T}(d)$ is connected due to our assumption also $s_i \in \mathcal{T}(d)$, a contradiction to $d \notin \text{cover}(s_i)$. □

As an example, consider figure 2 and note that, e.g.,

$$\text{cover}(s_6) \setminus \text{cover}(s_4) = \text{cover}(s_6) - \text{cover}(\{s_2, s_3, s_4\}) = \{d_3\},$$

while in figure 3, $\mathcal{T}(d_1)$ is not connected, and

$$\text{cover}(s_7) \setminus \text{cover}(s_6) \neq \text{cover}(s_7) \setminus \text{cover}(\{s_1, \ldots, s_7\}.$$

Lemma 6 suggests a standard dynamic programming approach to solve (BSL-cover(k)). This approach has been derived by transferring the Bellman-Ford algorithm (see Bellman, 1958; Ford, and Fulkerson, 1962) for k-cardinality constrained shortest paths to set covering problems along the lines of Section 4.3 in Schöbel (2003). Note that the resulting approach for solving single-criteria problems of type (BSL-cover) along a straight line has also been developed directly in the context of stop location by Kranakis (2002) and an improved version of this approach has been suggested in Hassin and Tamir (1991) for the k-facility location problem along a line.

In this paper we embed such an approach for solving BSL-cover(k) in our algorithm for finding all efficient solutions. In step 4 of our approach we are looking for a set of stops, all of them smaller than (i.e. on the left hand side of) a given stop s_j which itself should be contained in S, i.e., we solve subproblems of type

$$(\mathbf{P}(k, s_j)) \quad \max\{f_{\text{cover}}(S) : S \subseteq \{s_1, \dots, s_j\}, s_j \in S, \text{ and } |S| \leq k\}.$$

Note that we start with the optimal solution $\{s_j\}$ of $\text{P}(1, s_j)$ in step 3. To obtain the solution of $\text{P}(k, s_j)$ within step 4 we can use the previously calculated solutions for $k-1$. To this end, we set (in step 2)

$$w_{ij} = \sum_{d \in \text{cover}(s_j) \setminus \text{cover}(s_i)} w_d \quad \text{for } i < j \quad \text{and}$$
$$W = f_{\text{cover}}(\mathcal{S}).$$

W denotes the maximum weight that we can cover, if we choose all candidates as new stops, while w_{ij} gives the gain if we add s_j to a set of stops containing s_i as its rightmost stop. Denoting the optimal solution of $\text{P}(k, s_j)$ by $h^k(s_j)$ we iteratively calculate

$$h^k(s_j) = \max_{i: s_i < s_j} w_{ij} + h^{k-1}(s_i).$$

in each sub iteration of step 4. For the sake of simplicity we use the standard dynamic programming approach in step 4 but remark that since the weights w_{ij} satisfy the concavity property, the implementation of Galil, and Park (1990) and Hassin, and Tamir (1991) leads to a better time complexity of our algorithm.

Finally, the optimal solution of (BSL-cover(k)) can be obtained as the best solution over all optimal solutions of $\text{P}(k, s_j)$ over $j = 1, \dots, N$ (step 5). In step 6, we use Lemma 1 to obtain all efficient solutions. We first state the algorithm and then show the correctness of the above statements.

Algorithm : Finding all efficient solutions of (BSL)

Input: $\mathcal{D}$, a polygonal line $\mathcal{T}$ with connected sets $\mathcal{T}(d)$, weights w_d for all $d \in \mathcal{D}$.

Output: All efficient solutions of (BSL), and a Pareto solution for each of them.

Step 1. Derive the set of candidates $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$ as in Theorem 2 and order them along $\mathcal{T}$.

Step 2. Let $W = f_{\text{cover}}(\mathcal{S})$ and $w_{ij} = \sum_{d \in \text{cover}(s_j) \setminus \text{cover}(s_i)} w_d$ for all $i < j$ with $i, j \in \{1, \dots, N\}$

Step 3. Let for all $j = 1, \dots, N$: $h^1(s_j) = \text{cover}(s_j)$, $S^1(s_j) = \{s_j\}$, $k = 2$.

Step 4. `For all` $j = 1, \ldots, N$:

$$h^k(s_j) = \max_{i:s_i<s_j} w_{ij} + h^{k-1}(s_i)$$

`If` $h^k(s_j) = w_{i^*j} + h^{k-1}(s_{i^*})$ `let` $S^k(s_j) = S^{k-1}(s_{i^*}) \cup \{s_j\}$.

Step 5. `Let` $h^k = \max_{j=1,\ldots,N} h^k(s_j) =: h^k(s^*)$ `and let` $S^k = S^k(s^*)$.

- `If` $h^k = W$ `then set` $K^* = k$ `and stop.`
- `Otherwise` $k = k + 1$ `and goto step 4.`

Step 6. `Output:` $Eff = \{(h^k, k) : k = 1, \ldots, K^*\}$ `with corresponding Pareto solutions` S^k, $k = 1, \ldots, K^*$.

To show the correctness of the algorithm we need the following lemmas.

Lemma 7. $S^k(s_j)$ is an optimal solution of $(P(k, s_j))$ with objective value $h^k(s_j)$.

Proof. We use induction over k. For $k = 1$ the optimal solution of (P$(1, s_j)$) is $S^1(s_j) = \{s_j\}$. Now assume that $S^{k-1}(s_j)$ is the optimal solution of (P$(k-1, s_j)$) for any fixed s_j. For the induction step we first note that (P(k, s_j)) is equivalent to

$$\max\{f_{\text{cover}}(S' \cup \{s_j\}) : S' \subseteq \{s_1, \ldots, s_{j-1}\}, \text{ and } |S'| \le k - 1\}.$$

Now calculate for any $S' = \{s_{i_1}, s_{i_2}, \ldots, s_{i_p}\}$ with $s_{i_1} < s_{i_2} < \ldots s_{i_p} < s_j$

$$\begin{aligned} f_{\text{cover}}(S' \cup \{s_j\}) &= \sum_{d \in \text{cover}(S' \cup \{s_j\})} w_d \\ &= f_{\text{cover}}(S') + \sum_{d \in \text{cover}(s_j) \setminus \text{cover}(S')} w_d \\ &= f_{\text{cover}}(S') + \sum_{d \in \text{cover}(s_j) \setminus \text{cover}(s_{i_p})} w_d \quad \text{due to Lemma 6} \\ &= f_{\text{cover}}(S') + w_{i_p j} \end{aligned}$$

Hence, (P(k, s_j)) can further be rewritten as

$$\max\{f_{\text{cover}}(S') + w_{i_p j} : i_p \in \{s_1, \ldots, s_{j-1}\}, S' \subseteq \{s_1, \ldots, s_{i_p}\}, \\ i_p \in S', \text{ and } |S'| \le k - 1\}$$

and it becomes clear that the set S' in this formulation is an optimal solution of (P$(k-1, s_{i_p})$). Using the induction hypothesis we finally obtain that (P(k, s_j)) is equivalent to

$$\max\{f_{\text{cover}}(S^{k-1}(i_p)) + w_{i_p j} : i_p \in \{s_1, \ldots, s_{j-1}\}\}$$

which shows the result. □

Corollary 2. S^k is an optimal solution of (BSL-cover(k)) and its objective value is h^k.

Proof. This consequence follows from Lemma 7 and the definition of S^k in step 5 of the algorithm. □

Finally, to apply Lemma 1 we need the following result.

Lemma 8. For $k \leq K^*$ any optimal solution S^* of $(BSL - \text{cover}(k))$ satisfies $|S^*| = k$.

Proof. Let S be an optimal solution of (BSL-cover(k)) for some $k < K^*$. This means that $f_{\text{cover}}(S) = f_{\text{cover}}(S^k) < W$ due to Corollary 2 and step 6 of the algorithm. Hence there exists $s \notin S$ such that

$$\sum_{d \in \text{cover}(s) \setminus \text{cover}(S)} w_d > 0$$

and hence $f_{\text{cover}}(S \cup \{s\}) > f_{\text{cover}}(S)$. If $|S| \leq k-1$ this yields that $S \cup \{s\}$ does not contain more than k stops and hence is feasible for (BSL-cover(k)), which is a contradiction to the optimality of S. □

Theorem 3. The algorithm finds all efficient solutions of (BSL).

Proof. For each $k \leq K^*$ we know from Corollary 2 that S^k is an optimal solution of (BSL-cover(k)) . Furthermore, Lemma 8 shows that all optimal solutions of (BSL-cover(k)) consist of the same number k of stops. Hence (h^k, k) is an efficient solution according to Lemma 1.

On the other hand, no solution S with $|S| > K^*$ is Pareto, since such a solution S is always dominated by S^{K^*}, because

$$|S| > K^* = |S^{K^*}| \text{ and}$$
$$f_{\text{cover}}(S) \leq W = f_{\text{cover}}(S^{K^*}).$$

□

Since the number of candidates $|S|$ is at most twice the number of demand points for a polygonal line satisfying the assumptions of the algorithm, the worst-case complexity of the algorithm (finding all efficient solutions of the bicriteria stop location problem) is given by $O(K^*|\mathcal{D}|^2)$ where K^* is the minimum number of stops needed to cover all demand points in cover($\mathcal{T}$). Using the concavity property of the weights and the resulting better implementation of step 4 due to Galil and Park (1990) and Hassin and Tamir (1991), one can reduce the overall time complexity to $O(|\mathcal{D}|^2)$ to find *all* efficient solutions of (BDM).

5. Conclusion

In this paper we developed a model for the bicriteria stop location problem and proposed an efficient solution approach for determining a Pareto solution for each efficient solution in the special case that the set of tracks is given by a polygonal line with connected intervals $\mathcal{T}(d)$ for each demand point d. Investigating the real-world data of German rail (DB) all over Germany, it turns out that this assumption is *almost* satisfied in practice. Dealing with a few demand points not satisfying the connectedness of $\mathcal{T}(d)$ (or, more specific, with a few rows of the covering matrix not satisfying the consecutive ones property) has been treated in Ruf, and Schöbel (2004).

Moreover, the extension of the results to demand regions instead of demand points is investigated. For some first results in this area we refer to Schöbel and Schröder (2003).

References

Bellman, R. (1958). "On a Routing Problem." *Quarterly Applied Mathematics* 16, 87–90.

Bruno, G., G. Ghiani, and G. Improta. (1998). "A multi-Modal Approach to the Location of a Rapid Transit Line." *European Journal of Operational Research* 104(2), 321–332.

Caprara, A., M. Fischetti, and P. Toth (1999). "A Heuristic Method for the Set Covering Problem." *Operations Research* 47(5), 730–743.

Church, R. and C. ReVelle (1974). "The Maximal Covering Location Problem." *Papers of the Regional Science Association* 32, 101–118.

Current, J., C. ReVelle, and J. Cohon. (1987). "The Median Shortest Path Problem: A Multiobjective Approach to Analyze Cost vs. Accessibility in the Design of a Transportation Network." *Transportation Science* 21(3), 490–503.

Ehrgott, M. (2000). *Multiple Criteria Optimization.* Vol. 491 of *Lecture Notes in Economics and Mathematical Systems* Springer, Berlin.

Ford, L.R. and D.R. Fulkerson (1962). *Flows in Networks*. Princeton University Press.

Galil, Z. and K. Park (1990). "A Linear-Time Algorithm for Concave One-Dimensional Dynamic Programming." *Information Processing Letters* 33, 309–311.

Haimes, Y.Y. and V. Chankong (1983). *Multiobjective Decision Making—Theory and Methodology*. North Holland, New York.

Hamacher, H.W., A. Liebers, A. Schöbel, D. Wagner, and F. Wagner (2001). "Locating New Stops in a Railway Network." *Electronic Notes in Theoretical Computer Science* 50(1).

Hassin, R. and A. Tamir (1991). "Improved Complexity Bounds for Location Problems on the Real Line." *Operations Research Letters* 10, 395–402.

Kranakis, E., P. Penna, K. Schlude, D.S. Taylor, and P. Widmayer (2002). "Improving Customer Proximity to Railway Stations." Technical report, ETH Zürich.

Laporte, G., J.A. Mesa, and F.A. Ortega (2002a). "Locating Stations on Rapid Transit lines." *Computers and Operations Research* 29, 741–759.

Laporte, G., J.A. Mesa, and F.A. Ortega (2002b). "Maximizing Trip Coverage in the Location of a Single Rapid Transit Alignment." ISOLDE IX, Fredericton & St. Andrews, New Brunswick, Canada.

Minkowski, H. (1967). *Gesammelte Abhandlungen, Band 2*. Chelsea Publishing Company, New York.

Murray, A. (2001a). "Coverage Models for Improving Public Transit System Accessibility and Expanding Access." Technical report, Department of Geography, Ohio State University.

Murray, A. (2001b). "Strategic Analysis of Public Transport Coverage." *Socio-Economic Planning Sciences* 35, 175–188.

Murray, A., R. Davis, R.J. Stimdon, and L. Ferreira (1998). "Public Transportation Access." *Transportation Research D* 3(5), 319–328.

Nemhauser, G.L. and L.A. Wolsey. (1988). *Integer and Combinatorial Optimization*. Wiley.

Ruf, N. and A. Schöbel. (2004). "Set Covering Problems with almost Consecutive ones Property." *Discrete Optimization* 1(2), 215–228.

Schöbel, A. (2003). *Customer-Oriented Optimization in Public Transportation.* Habilitationsschrift, Universität Kaiserslautern.

Schöbel, A., H.W. Hamacher, A. Liebers, and D. Wagner. (2002). "The Continuous Stop Location Problem in Public Transportation." Technical report, Universität Kaiserslautern. Report in Wirtschaftsmathematik 81.

Schöbel A., and M. Schröder. (2003). "Covering Population Areas by Railway Stops." In *Proceedings of OR 2002, Klagenfurt*, Springer.

Steuer, R.E. (1989). *Multiple Criteria Optimization: Theory, Computation, and Application*. Krieger Publishing Company, Malabar, Florida.

Toregas, C. (1971). " The Location of Emergency Service Facilities." *Operations Research* 19, 1363–1373.

Toregas, C., R. Swain, C. ReVelle, and L. Bergman. (1971). "The Location of Emergency Facilities." *Operations Research* 19, 1363–1373.

White, J. and K. Case. (1974). "On Covering Problems and the Central Facility Location Problem." *Geographical Analysis* 6, 281–293.