Georg-August-Universität Göttingen

Station Location
Complexity Issues

Steffen Mecke / Anita Schöbel / Dorothea Wagner

Preprint Nr. 2005-04

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

Station Location— Complexity Issues

Steffen Mecke[1], Anita Schöbel[2], and Dorothea Wagner[1] *

[1] University of Karlsruhe, {mecke, wagner}@ira.uka.de
[2] University of Göttingen, schoebel@math.uni-goettingen.de

Abstract. Summary of results on the complexity status of the STATION LOCATION problem.

1 Introduction

[KPS+02] [HLS+01] [MW04] [Sch02] [**?**]

2 $\mathcal{NP}$-Completeness

Problem 2.1 (STATION LOCATION). Given a geometric graph $G = (V, E)$, i.e. a set V of vertices in the plane (stations, switches, bends) and a set E of edges (*lines*) represented as straight line segments, a set $\mathcal{P}$ of points in the plane (*settlements*)and a fixed radius R. Find a minimum number of vertices $\mathcal{S}$ (*new stops*) on the edges such that $\mathcal{P} \subseteq \text{cov}(\mathcal{S})$, where $\text{cov}(\mathcal{S}) = \{x \in \mathbb{R}^2 : d(x, \mathcal{S}) \leq R\}$.

It was shown in [SHLW02] that there exists a finite set $\mathcal{C}$ of *candidate* locations for new stops which contains an optimum solution and which can be computed by an algorithm which is polynomial in the sizes of G and P.

Definition 2.2. *The matrix* $A^{\text{cov}} = (a_{pc})$ *with* $a_{pc} = \begin{cases} 1 & \textit{if } c \textit{ covers } p \\ 0 & \textit{otherwise} \end{cases}$ *(for all* $p \in \mathcal{P}, c \in \mathcal{C}$*) is called the* covering matrix *of an instance of* STATION LOCATION*.*

In the following we will assume that some fixed set of candidates is given and sometimes use the terms rows and stations (resp. columns and candidates) as synonyms.

Given this, STATION LOCATION can be seen as a specialization of the well-studied SET COVERING (aka. HITTING SET) problem. We use the following notation to describe it as a linear problem:

$$\begin{aligned} \min \quad & cx \\ \text{s.t. } A^{\text{cov}} & x \geq \underline{1}_M \\ & x \in \{0,1\}^N, \end{aligned} \tag{SCP}$$

---

*

where $\underline{1}_M \in \mathbb{R}^M$ denotes the vector consisting of M ones, $c \in \mathbb{R}^N$ contains the costs of the columns, and A^{cov} is an $M \times N$-matrix with elements $a_{mj} \in \{0, 1\}$, $m = 1, \ldots, M$, $j = 1, \ldots, N$. We may assume without loss of generality that A^{cov} neither has zero rows nor zero columns and that the costs c_j are positive.

The goal is to find an optimal solution x^*, or equivalently, an optimal set $\mathcal{N}^* \subseteq \mathcal{N} := \{1, \ldots, N\}$ of columns of A^{cov}, where $\mathcal{N}^* = \{n \in \mathcal{N} : x_n^* = 1\}$.

Theorem 2.3 ([HLS+01]). STATION LOCATION *is $\mathcal{NP}$-complete.*

Definition 2.4. *1. A matrix A^{cov} over $\{0, 1\}$ has the* **strong consecutive ones property (C1P)** *if all of its rows $m \in \{1, \ldots, M\}$ satisfy the following condition for all $j_1, j_2 \in \{1, \ldots, N\}$:*

$$a_{mj_1} = 1 \text{ and } a_{mj_2} = 1 \Longrightarrow a_{mj} = 1 \text{ for all } j_1 \le j \le j_2.$$

2. A matrix has the **consecutive ones property (C1P)** *if there exists a permutation of its columns such that the resulting matrix has the strong consecutive ones property*

3. Let l_m (r_m) be the index of the leftmost (rightmost) 1 in the m-th row of A^{cov}. A matrix A^{cov} with strong C1P is ***strictly monotone*** *if the sequence $(l_m)_{1 \le m \le M}$ and $(r_m)_{1 \le m \le M}$ are strictly increasing.*

Theorem 2.5 ([SHLW02]). STATION LOCATION *is polynomially solvable if the covering matrix has consecutive ones property. This is the case if no settlement can be covered from more than one line.*

Lemma 2.6. *Let $A = (A_1|A_2)$ where A_1 and A_2 both are strictly monotone matrices. Then the* SET COVERING *problem with coefficient matrix A can be solved in polynomial time.*

Proof.

Theorem 2.7. STATION LOCATION *is $\mathcal{NP}$-complete in the strong sense, even for the case that no settlement can be covered from more than two tracks.*

Proof. By reduction from PLANAR VERTEX COVER. In [GJS74] it has been shown that this problem remains $\mathcal{NP}$-complete even for planar graphs with maximum degree 6. We can further constrain this to maximum degree 3: PLANAR DEG-6 VERTEX COVER $\propto$ PLANAR DEG-3 VERTEX COVER (PD3VC). Therefor we replace every node of degree six by the gadget of eleven nodes shown in Fig. 1. A very similar procedure works for nodes of degree four and five. So that the resulting graph G' has $|V| + 10v_6 + 8v_5 + 6v_4$ nodes, maximum degree 3, and is still planar (v_6, v_5, and v_4 are the numbers of nodes of degree six, five and four, resp., in the original graph G). A node cover of size K in G exists if and only if a node cover of size $K' := K + 5v_6 + 4v_5 + 3v_4$ exists in G'.

The next step is to reduce PD3VC to STATION LOCATION. There exists a planar orthogonal unit grid drawing of G' with $O(n^2)$ area and at most $2n + 4$

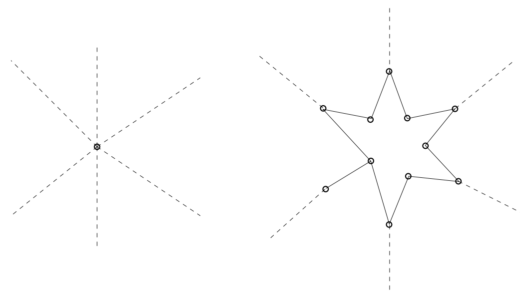


Fig. 1. PLANAR DEG-6 VERTEX COVER $\propto$ PLANAR DEG-3 VERTEX COVER

bends which is constructible in polynomial time (cf. [dBETT99], Theorem 4.16). We construct an instance of STATION LOCATION as follows: Let $R = 1/4$ and construct the settlements and candidates as follows. The unit grid cuts every edge into segments of unit length. Let S be the set of all these segments. Each grid segment of an edge has either two, one, or zero nodes of G' at its ends. First, replace every node in V by a candidate. Then, replace the segments by settlements and candidates according to the gadgets shown in Fig. 2. The result of this construction is sketched in Fig. 3.

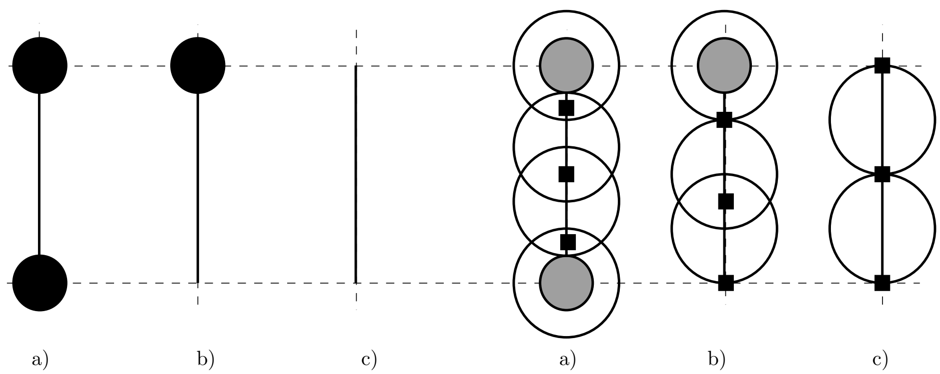


Fig. 2. Three gadgets (right) for the three different types of segments (left). Settlements are depicted by squares, candidates for stations as small discs. The big circles indicate the covering radius. The grid is dashed.

Note that, after all segments have been replaced, there are exactly $|V| + 2|S|$ candidates and $3|S|$ settlements. Further note, that settlements are covered from candidates from different segments if and only if the corresponding segments are adjacent and no settlement is covered by more than two candidates. Finally, let

$$K'' := K' + |S| \ .$$

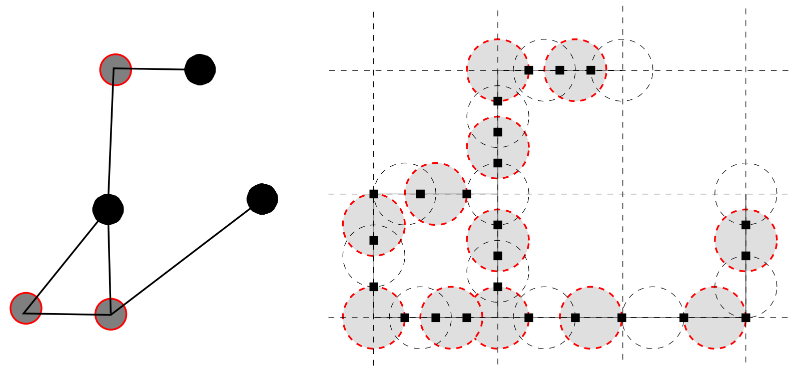


Fig. 3. PD3VC $\propto$ Station Location (with vertex cover resp. station cover in grey)

A vertex cover of cardinality K' in G' exists if and only if there is a solution with cardinality K'' for the constructed instance of Station Location. As the construction works in polynomial time, this implies the $\mathcal{NP}$-completeness of this variant of Station Location.

Corollary 2.8 *The* Station Location *problem remains $\mathcal{NP}$-complete even for the case that the covering matrix can be written as $(A|B)$, where A and B have the consecutive ones property (even if $(A|B)$ has exactly 2 ones per row, A has no more than one one per row and B has no more than two ones per row).*

Proof. Consider the instance of Station Location and the graph G' constructed in the above proof. There are two classes of candidates: Candidates corresponding to nodes of G' and candidates on edges of G'. Assign columns corresponding to candidates of the first class to A and columns corresponding to candidates of the second class to B. Order the columns of B in the natural way, namely corresponding to their order on the edges between two nodes of G' (cf. Fig. 3). Then the following properties hold:

1. No row of A has more than one entry, because every station can only be reached by one class-A node. It follows that A has C1P.
2. An ordering of the columns of B following the above rule exists. For every row covered by columns of B the (up to two) columns covering it are consecutive.

It follows that A and B have C1P and no row of $(A|B)$ has more than two non-zero entries.

3 Approximation

In [LY94] it is shown that Set Covering cannot be approximated within a factor of $\log(n)$ unless some likely assumption on complexity classes holds. Using the so called *shifting technique* of [HM85] it was shown in [FCB01] that a PTAS exists for COVERING BY DISCS, which is similar to STATION LOCATION except for the fact that the locations for new stations are not restricted. [GGRV01] found a PTAS for a problem which is almost identical to STATION LOCATION, however "cheating" a bit by assuming that stations cannot be arbitrarily close to each other. However, these techniques cannot be easily adapted to our more general problem, although the could probably work well in practice.

3.1 A block-based reformulation

Definition 3.1. *If A^{cov}_m is a row of A^{cov} let bl_m be its number of blocks of consecutive ones.*

If a matrix has the consecutive ones property, i.e., $bl_m = 1$ for all rows m, the permutation of the columns making the ones appear consecutively can be found by using the algorithm of [BL76,MPT98]. This algorithm can be performed in O(MN). Without loss of generality we can therefore assume that a matrix with consecutive ones property is already ordered, i.e. we assume that its ones already appear consecutively in all of its rows. We say that a SET COVERING problem has C1P if its covering matrix A^{cov} has C1P.

Since SET COVERING problems in which A^{cov} has the consecutive ones property can be solved efficiently the idea is to split each row m with $bl_m > 1$ into bl_m rows, each of them satisfying the consecutive ones property, and to require that at least one of these rows needs to be covered. We remark that the condition of the above definition will turn out to be necessary to ensure an efficient behavior of our solution approach, but still there remain instances that cannot be solved in reasonable time by our approach although satisfying the *almost consecutive ones property*. `Another criterion to classify well-solvable matrices will be made precise at the end of this paper.`

Now consider a zero-one matrix A^{cov} with M rows, such that $bl_m = 1$ for $m = 1, \ldots, p)$, and $bl_m > 1$ in the remaining rows $p + 1, \ldots, M$.

For the ith block of consecutive ones in row m let

- $f_{m,i}$ be the column of the first 1 of block i and
- $l_{m,i}$ be the column of its last 1.

This means, that

$$a_{mj} = \begin{cases} 1 & \text{if there exists } i \in \{1, \ldots, bl_m\} \text{ such that } f_{m,i} \le j \le l_{m,i} \\ 0 & \text{otherwise.} \end{cases}$$

Consider a row A^{cov}_m of A^{cov} with $bl_m > 1$. According to the transformation introduced in [?] we replace A^{cov}_m by bl_m rows,

$$B_{m,1}, B_{m,2}, \ldots, B_{m,bl_m}$$

each of them containing only one single block of row A_m, i.e., we define the jth element of row $B_{m,i}$ as

$$(B_{m,i})_j = \begin{cases} 1 & \text{if } f_{m,i} \le j \le l_{m,i} \\ 0 & \text{otherwise.} \end{cases}$$

Due to [?] the SET COVERING problem (SCP) is hence equivalent to

$$\begin{array}{ll} \min & cx \\ s.t. & A_m^{\text{cov}} x \geq 1 \quad \text{for } m = 1, \ldots, p \\ & B_{m,i} x \geq y_{m,i} \quad \text{for } m = p+1, \ldots, M, i = 1, \ldots, bl_m \\ & \sum_{i=1}^{bl_m} y_{m,i} \geq 1 \quad \text{for } m = p+1, \ldots, M \\ & y_{m,i} \in \{0,1\} \quad \text{for } m = p+1, \ldots, M, i = 1, \ldots, bl_m \\ & x \in \{0,1\}^N. \end{array} \tag{SCP'}$$

It is more convenient to rewrite (SCP') in matrix form. To this end, we define

- the matrix A as the first p rows of A^{cov},
- $bl = \sum_{m=p+1}^{M} bl_m$ as the total number of blocks in the "bad" rows of A^{cov}, i.e., in rows of A^{cov} without consecutive ones property,
- I as the $bl \times bl$ identity matrix,
- B as the matrix containing the bl rows $B_{m,i}$, $m = p+1, \ldots, M, i = 1, \ldots, bl_m$ and
- C as a matrix with $M - p$ rows and bl columns, with elements

$$c_{ij} = \begin{cases} 1 & \text{if } \sum_{m=p+1}^{p+i-1} bl_m < j \leq \sum_{m=p+1}^{p+i} bl_m \\ 0 & \text{otherwise.} \end{cases}$$

(SCP') can hence be reformulated as

$$\begin{array}{lll} \min & cx & \\ \text{s.t.} & Ax & \geq \underline{1}_p \\ & Bx - Iy & \geq \underline{0}_{bl} \\ & Cy & \geq \underline{1}_{M-p} \\ & x & \in \{0,1\}^N, \\ & y & \in \{0,1\}^{bl}. \end{array} \tag{SCP'}$$

The constraint $Cy \geq \underline{1}_{M-p}$ makes sure that at least one block of each row A_m^{cov} with $m \geq p+1$ is covered.

Note that all three matrices A,B, and C have the consecutive ones property. Unfortunately, the coefficient matrix of (SCP') does not have the consecutive ones property, and is in general even not totally unimodular.

3.2 Approximation

We first clarify the complexity status of SET COVERING problems with at most k blocks of consecutive ones per row.

Let k be an upper bound on the number of blocks in each row of A, i.e., such that

$$bl_m \leq k \text{ for all } m = 1, \ldots, M.$$

Corollary 3.2 *For $k = 1$ the* SET COVERING *problem is polynomially solvable, for all fixed $k \geq 2$ the problem is NP-hard.*

Algorithm 1: Linear Programming Heuristisc

Input: $M \times N$ matrix A

Output: approximate solution $\tilde{x}$

1 Solve LP-Relaxation of the reformulation (**) to obtain a solution (x^r, y^r).

2 **for** $m := 1, \ldots, M$ **do**

3 Find an index $i(m)$ with $y^r_{m,i(m)} \geq y^r_{m,i}$ for all $k = 1, \ldots, bl_i$.

4 Define

$$\tilde{y}_{m,i} = \begin{cases} 1 & \text{if } i = i(m) \\ 0 & \text{otherwise.} \end{cases}$$

5 Solve $\min\{cx : Bx \geq \tilde{y}, x \in \{0,1\}^n\}$ to obtain $\tilde{x}$.

6 **return** $\tilde{x}$.

We suggest the following algorithm, for which we will show that it provides a k approximation, if k is an upper bound on the number of blocks of consecutive ones per row.

Note that Algorithm 1 can be solved by linear programming, since in Step 3, the coefficient matrix has the consecutive ones property. More efficent approaches for solving SET COVERING problems with consecutive ones property can be found in [NW88,**?**,**?**].

Theorem 3.3. *Algorithm 1 is a k approximation algorithm, where*

$$k = \max_{m=1,\ldots,M} bl_m.$$

Proof. Let (x^*, y^*) be an optimal solution and (x^r, y^r) be an optimal solution of the linear programming relaxation of (SCP'). This means that

$$cx^r \leq cx^*. \tag{1}$$

Now note that

$$ky^r_{m,i} \geq \tilde{y}_{m,i}. \tag{2}$$

This trivially holds for $\tilde{y}_{m,i} = 0$, while for $\tilde{y}_{m,i} = 1$ we know that

$$\begin{aligned} y^r_{m,i} &= \max_{k=1,\ldots,bl_m} y^r_{m,k} \\ &\geq \frac{1}{bl_m} \sum_{k=1,\ldots,bl_m} y^r_{m,k} \\ &\geq \frac{1}{bl_m} \quad \text{since } Cy \geq \underline{1}_{M-p} \\ &\geq \frac{1}{k} \end{aligned}$$

Moreover,

$$\min\{cx : Bx \geq \tilde{y}\} = \min\{cx : Bx \geq \tilde{y}, x \in \{0,1\}^N\},$$

since in any optimal solution of the latter, $x \leq 1$, and the the integrality constraint $x \in \mathbb{N}^N$ can be deleted since B has the consecutive ones property and hence is totally unimodular.

Now estimate $B(kx^r)$ as

$$B(kx^r) = kBx^r \geq ky^r \geq \tilde{y},$$

where the last inequality is due to (2). In other words, kx^r is feasible for $\{x : Bx \geq \tilde{y}\}$, and hence we get

$$\begin{aligned} kcx^r &\geq \min\{cx : Bx \geq \tilde{y}\} \\ &= \min\{cx : Bx \geq \tilde{y}, x \in \{0,1\}^N\} \\ &= c\tilde{x} \end{aligned}$$

Combining the latter with (1) we finally obtain

$$c\tilde{x} \leq kcx^r \leq kcx^*.$$

Theorem 3.4. *Is there a factor c algorithm for for* Station Location*? So far not...*

Theorem 3.5. *Is there a PTAS for* Station Location*? So far not...*

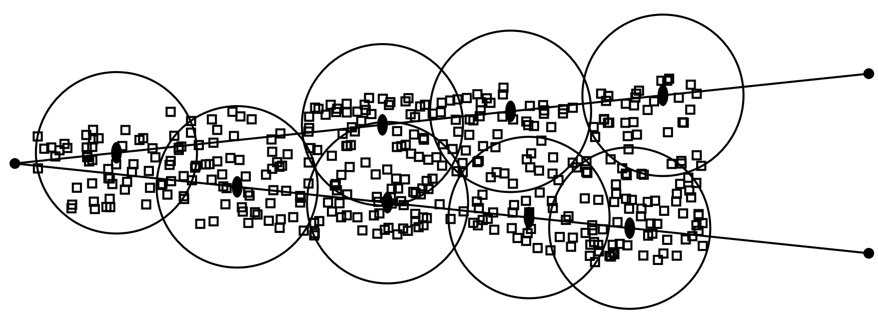
Theorem 3.6. *There is no FPTAS for* Station Location *since it is strongly $\mathcal{NP}$-complete.*

4 Parameterized Complexity

Theorem 4.1. Station Location *is solvable in $O(\text{poly}(m, n) \cdot 2^k)$ if the distance between the first and last 1 is not greater than k for every column of A.*

5 Conclusion

Remark 5.1. C1P for columns is more important than for rows



Niedermayer, Rossmanith 2002 JDA, Hitting Set kernel
Nemhauser Trotte Math. Prog, 1975 Vertex Cover

References

[BL76] K. S. Booth and G. S. Lueker. Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms. *Journal of Computer and System Science*, 13(3):335–379, 1976.

[dBETT99] Giuseppe di Battista, Peter Eades, Roberto Tamassia, and Ioannis G. Tollis. *Graph Drawing*. Prentice Hall, Inc., 1999.

[FCB01] Massimo Franceschetti, Matthew Cook, and Jehoshua Bruck. A geometric theorem for approximate disk covering algorithms. Technical report, California Institute of Technology, 2001.

[GGRV01] Matthias Galota, Christian Glaßer, Steffen Reith, and Heribert Vollmer. A polynomial-time approximation scheme for base station positioning in umts networks. In *DIALM '01: Proceedings of the 5th international workshop on Discrete algorithms and methods for mobile computing and communications*, pages 52–59. ACM Press, 2001.

[GJS74] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified np-complete problems. In *STOC '74: Proceedings of the sixth annual ACM symposium on Theory of computing*, pages 47–63. ACM Press, 1974.

[HLS+01] Horst W. Hamacher, Annegret Liebers, Anita Schöbel, Dorothea Wagner, and Frank Wagner. Locating new stops in a railway network. 50(1), 2001.

[HM85] Dorit S. Hochbaum and Wolfgang Maass. Approximation schemes for covering and packing problems in image processing and vlsi. *J. ACM*, 32(1):130–136, 1985.

[KPS+02] Evangelos Kranakis, Paolo Penna, Konrad Schlude, David Scot Taylor, and Peter Widmayer. Improving customer proximity to railway stations. In *ATMOS 2002*. ATMOS, 2002.

[LY94] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, September 1994.

[MPT98] João Meidanisa, Oscar Portob, and Guilherme P. Telles. On the consecutive ones property. *Discrete Applied Mathematics*, 88(1-3):325–354, 1998.

[MW04] Steffen Mecke and Dorothea Wagner. Solving Geometric Covering Problems by Data Reduction. In *Proceedings of the 12th Annual European Symposium on Algorithms, Bergen, Norway*, volume 3221 of *Lecture Notes in Computer Science*. Springer, 2004.

[NW88] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization*. Wiley, 1988.

[Sch02] Anita Schöbel. *Customer-Oriented Optimization in Public Transportation*. PhD thesis, University of Kaiserslautern, 2002. Habilitation.

[SHLW02] Anita Schöbel, Horst W. Hamacher, Annegret Liebers, and Dorothea Wagner. The continuous stop location problem in public transportation networks. *Report in Wirtschaftsmathematik*, 81, 2002.

Institut fr Numerische und Angewandte Mathematik
Universitt Gttingen
Lotzestr. 16-18
D - 37083 Gttingen

Telefon: 0551/394512
Telefax: 0551/393944

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

Verzeichnis der erschienenen Preprints:

2005-01	A. Schöbel, S. Scholl:	Line Planning with Minimal Traveling Time
2005-02	A. Schöbel	Integer programming approaches for solving the delay management problem
2005-03	A. Schöbel	Set covering problems with consecutive ones property
2005-04	S. Mecke, A. Schöbel, D. Wagner	Station location - Complex issues