The Stop Location Problem with Realistic Traveling Time *

Emilio Carrizosa[1], Jonas Harbering[2], and Anita Schöbel[2]

1 **Universidad de Sevilla**
C/ Tarfia s/n, 41012 Sevilla, Spain
ecarrizosa@us.es

2 **Georg-August Universität Göttingen**
Lotzestraße 16–18, 37083 Göttingen, Germany
{jo.harbering,schoebel}@math.uni-goettingen.de

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 some given bus routes, or of railway stations along the tracks in a railway network. The positive effect of new stops is given by the better access of the customers to the public transport network, while the traveling time increases due to the additional stopping activities of the trains which is a negative effect for the customers.

Our goal is to locate new stops minimizing a *realistic* traveling time which takes acceleration and deceleration of the vehicles into account. We distinguish two variants: in the first (academic) version we locate p stops, in the second (real-world applicable) version the goal is to *cover* all demand points with a minimal amount of realistic traveling time. As in other works on stop location, covering may be defined with respect to an arbitrary norm. For the first version, we present a polynomial approach while the latter version is NP-hard. We derive a finite candidate set and an IP formulation. We discuss the differences to the model neglecting the realistic traveling time and provide a case study showing that our procedures are applicable in practice and do save in average more than 3% of traveling time for the passengers.

1998 ACM Subject Classification G.1.6 Optimization

Keywords and phrases Stop Location, Realistic Traveling Time, IP Formulation

Digital Object Identifier 10.4230/OASIcs.ATMOS.2013.80

1 Introduction

The acceptance of public transportation depends on various components such as convenience, punctuality, reliability, etc. In this paper, we address the question of convenience for the passengers. In particular, we investigate the problem of establishing additional stops (or stations) which on the one hand guarantee a good accessibility to the transportation network, but on the other hand do not increase the traveling time of passengers too much.

Due to their great potential for improving public transportation systems, several versions of the *stop location problem* (also called *station location problem*) have been considered by various authors in the last years, see [16] for a survey. In order to find "good" locations for

* The first author is partially supported by MTM2012-36163-C06-03 and P11-FQM-7603. The other authors are partially supported by the European Union Seventh Framework Programme (FP7-PEOPLE-2009-IRSES) under grant number 246647 and the New Zealand Government (project OptALI).

new stops, several objective functions are possible. One of the most frequently discussed goals is to minimize the number of stops such that each demand point is within a tolerable distance from at least one stop. The maximal distance that a customer is willing to tolerate is called *covering radius*, hence we call this type of stop location problem *SL-Cov* for short. For bus stops a covering radius of 400 m is common. In rail transportation, the covering radius is much larger (at least 2 km).

In the literature, stop location problems have been introduced in [1] and considered in [12, 10, 11, 6, 18], see also references therein. In these papers, the problem is treated in a discrete setting, i.e., a finite set is considered as potential new stops. [17] 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 this *continuous* version is given in [2], 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 traveling times due to additional stops is compared with the positive effect of shorter access times, the goal is to maximize the difference of the two effects.

Based on this application, variants of the continuous stop location problem have been treated in [5, 16, 17]. The problem has been solved for the case of two intersecting lines, see [7]. Algorithmic approaches for solving the underlying covering problem have been studied in [15, 9]. Complexity and approximation issues have been presented in [8].

Another objective function is to minimize the sum of distances from the customers to the public transportation system, i.e. the sum of the distances between the demand facilities and their closest stops, see [13, 14]. Recently, covering a set of OD-pairs with a given number of stops has been studied, see e.g., [4] and references therein.

Contribution. All the mentioned papers use a rough approximation of the traveling time by adding a penalty for each stop. Since trains have a long acceleration and deceleration phase this is unrealistic in practice. In this paper we consider stop location problems with *realistic* traveling time.

Structure of the paper. We develop and analyze the realistic traveling time function in Section 2. We then consider two variants of the stop location problem with realistic traveling time. In Section 3 we want to locate p stops minimizing the traveling time, while in Section 4 we want to cover all demand points with a set of stops, again with minimal realistic traveling time. While we present a polynomial algorithm for the former problem, the latter problem is NP hard. Nevertheless we are able to develop a finite dominating set which is the basis for an integer programming formulation. We compare our new model to the existing covering models (without realistic traveling time) and present a case study with numerical results. All proofs can be found in the appendix.

2 Stop location with realistic traveling time

In the stop location problems considered so far, the traveling time for passengers due to new stops is estimated by adding a penalty time_{pen} for every stop to be located. This is an exact estimate if the distance between two stops is larger than the distance needed for acceleration and deceleration, and if time_{pen} gives the loss of traveling time resulting from the additional stop. As an example, time_{pen} is estimated as two minutes for German regional trains. However, since trains accelerate slowly, this estimate is not realistic if the distance between two stops is rather short.

In this section we hence first introduce a function describing the realistic traveling time of a train between two consecutive stops. This function depends on the distance d between those two consecutive stops. Being able to compute realistic traveling times, we then define two variants of the stop location problem, both with realistic traveling time.

▶ **Lemma 1.** *[see also [3]] Let a maximum cruising speed $v_0 > 0$, an acceleration of $a_0 > 0$ and a deceleration of $b_0 > 0$ of a vehicle be given. Then the traveling time function depending on d, where d is the distance between two consecutive stops, is given as*

$$T(d) = \begin{cases} \sqrt{\frac{2(a_0+b_0)}{a_0 b_0} d} & \text{if } d \leq d^{max}_{v_0,a_0,b_0} \\ \frac{d}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0} & \text{if } d \geq d^{max}_{v_0,a_0,b_0} \end{cases} \quad \text{where} \quad d^{max}_{v_0,a_0,b_0} = \frac{v_0^2}{2a_0} + \frac{v_0^2}{2b_0}$$

The formula is a simple consequence from Newton's laws of motion. E.g., in [3] the traveling time function for the case $a_0 = b_0$ is introduced, and the practical relevance of this better estimate is analyzed for fire engines in New York City.

Note that $d^{max}_{v_0,a_0,b_0}$ is the point where the traveling time function turns from a square root behavior to a linear behavior.

The shape and exact values of the function can be easily calculated; its main properties can be verified straightforwardly.

▶ **Lemma 2.** *$T(d)$ is continuous, differentiable, concave and monotonically increasing. Furthermore, for any d we have $\sqrt{\frac{2(a_0+b_0)}{a_0 b_0} d} \leq \frac{d}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0}$.*

The properties of T can be shown by easy calculations.

In the two variants of the stop location problem our objective is to minimize the (realistic) traveling time which is determined as follows.

Let $G = (V, E)$ be the given network in which the new stops should be located. Let $e = (i, j) \in E$ be an edge with length d_e. A point $s = (e, x) \in e$ is defined as the point on edge e with distance $d(i, s) = x$ and distance $d(s, j) = d_e - x$, $0 \leq x \leq d_e$. Note that $i = (e, 0)$ and $j = (e, d_e)$. The set of points of G is denoted as $\mathcal{S} = \bigcup_{e \in E} e$. The set of points between two points $s_1 = (e, x_1)$ and $s_2 = (e, x_2)$ on the same edge is denoted as $[s_1, s_2] = \{(e, x) : x_1 \leq x \leq x_2\}$.

A new stop s in the network may be any point $s = (e, x)$. We assume that all vertices V are existing stops.

Given a set $S \in \mathcal{S}$ of points of G, every set $S_e = S \cap e = \{s_1, \ldots, s_p\} \subseteq e$ of points on $e = (i, j)$ can be naturally ordered along the edge e such that $d(s_1, i) \leq \ldots \leq d(s_p, i)$. Let $\leq_e$ denote this ordering. Adding the points of S as new stops gives a subdivision of the network G, i.e. a new network $(V \cup S, E(S))$ (see Figure 2), where

$$E(S) = \{(s_i, s_j) : s_i = (e, x_i) \text{ and } s_j = (e, x_j) \text{ are consecutive on some } e \in E \text{ w.r.t. } \leq_e\}.$$

The length of an edge $e' = ((e, x_i), (e, x_j)) \in E(S)$ is given as $d_{e'} = |x_j - x_i|$.

Finally, given a set S of points on the graph G, we can define the (realistic) traveling time function as

$$g(S) := \sum_{e' \in E(S)} T(d_{e'}).$$

The stop location problem (SL) on $G = (V, E)$ is to locate a set of stops S which are points on G. Our objective is to minimize the (realistic) traveling time $g(S)$. This function can be seen as an intrinsical property of the network and an estimation for the traveling time of passengers without having information of their real paths and demands.

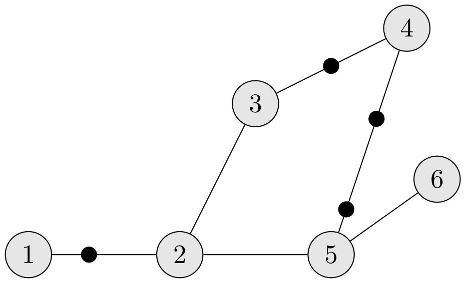

Figure 1 Locating $p = 4$ new stops on a Network $G = (V, E)$.

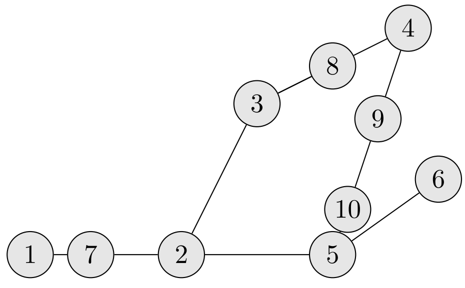

Figure 2 The new network $(V \cup S, E(S))$ with $p = 4$ new stops.

Without any constraints, $S = \emptyset$ would be the trivial optimal solution. We hence need to ensure that enough new stops are located. We consider the following two possibilities:

(SL-TT-p) Here, the goal is to locate p new stops on G. It is further required that the minimal distance between two stops is at least ϵ, i.e., $d_{e'} \geq \epsilon$ for all $e' \in E(S)$.

(SL-TT-Cov) This is an extension of the stop location problems considered in the literature, in which it is assumed that the network is embedded in the plane $\mathbb{R}^2$, and that a finite set of demand points $\mathcal{P} \subseteq \mathbb{R}^2$ is given. Furthermore, to measure the access times from the demand points to the railway network, a distance function $dist : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is given which has been derived from a norm, i.e. $dist(x, y) = \|y - x\|$ for a given norm $\| \cdot \|$. For a set $S \in \mathcal{S}$ we can now define the set of covered demand points as

▶ **Definition 3.** $cover_{\mathcal{P}}(S) = \{p \in \mathcal{P} : dist(p, s) \leq r \text{ for some } s \in \mathcal{S}\}$.

In (SL-TT-Cov) we look for a set S covering all demand points (i.e. $cover_{\mathcal{P}}(S) = \mathcal{P}$).

3 (SL-TT-p) Locating p stops

We start with locating a fixed number of p stops on a single line segment. We hence have $G = (V, E)$ where $V = \{i, j\}$ is the set of nodes and $E = \{e = (i, j)\}$ is one edge. Locating p new stops S on the edge e increases the traveling time for the passengers that want to traveling from i to j. Our goal is to minimize this traveling time using the realistic traveling time function $g(S)$:

(Line-SL-TT-p) Let $G = (\{i, j\}, \{e\})$ be one single edge, $v_0 > 0$, $a_0 > 0$, and $b_0 > 0$ and let a natural number $p > 0$, and $0 \leq \epsilon \leq \frac{d_e}{p+1}$ be given. Find a subset $S^* \subseteq \mathcal{S}$ with $|S^*| = p$ and $d_{e'} \geq \epsilon$ for all $e \in E(S^*)$ such that $g(S^*)$ is minimized.

Note that (Line-SL-TT-p) is not feasible if $d_e < (p + 1)\epsilon$.

We start by discussing the case of locating only $p = 1$ stop without further restriction (i.e. $\epsilon = 0$) since this instance contains the main idea for the general case.

Let d_e be the length of edge $e = (i, j)$. Let $s = (e, x)$ be a new stop on e. We want to determine x. Given any $x \in [0, d_e]$, the traveling time for passengers from i to j is

$$g(S) = g(x) = T(x) + T(d_e - x),$$

i.e., we have to solve $\min\{T(x) + T(d_e - x) : 0 \leq x \leq d\}$. Since T is a concave function, also $T(d_e - x)$ is concave, hence also $g(x)$. A concave function over a compact interval takes its minimum at one of the endpoints of the interval, we hence evaluate $g(0) = g(d_e) = T(0)+T(d_e)$ and obtain

▶ **Lemma 4.** *The only optima for (Line-SL-TT-p) for the case of locating $p = 1$ stop without a minimal distance (i.e. $\epsilon = 0$) are obtained at $s_1 = i$ and at $s_2 = j$.*

However, since i and j are already stops, this solution does not give any new stop and consequently is not what we want. We hence require a minimal distance of $\epsilon > 0$ between any pair of stops. This means to solve $\min\{T(x) + T(d - x) : \epsilon \leq x \leq d - \epsilon\}$ and again results in two optima on the boundary of the interval. We obtain:

▶ **Lemma 5.** *The only optima for (Line-SL-TT-p) for the case of locating $p = 1$ stop on the edge $e = (i, j)$ are obtained at $s_1 = (e, \epsilon)$ and $s_2 = (e, d_e - \epsilon)$.*

This result can be generalized for the case of locating $p \geq 2$ stops $s_1 = (e, x_1), \ldots, s_p = (e, x_p)$. We hence look for the values of $x_1, \ldots, x_p$.

Fixing $x_0 = 0$ and $x_{p+1} = d_e$, the respective optimization program is stated as

$$\begin{aligned}
&\min \sum_{l=1}^{p+1} T(x_l - x_{l-1}) \\
&\text{s.t. } x_l - x_{l-1} \geq \epsilon \quad \forall\, l = 1, \ldots, p+1 \\
&\qquad x_l \in \mathbb{R} \quad \forall l = 1 \ldots, p
\end{aligned}$$

First note, that for $\epsilon \geq d^{max}_{v_0,a_0,b_0}$ there is not much to worry about.

▶ **Lemma 6.** *If $\epsilon \geq d^{max}_{v_0,a_0,b_0}$, every feasible solution to (Line-SL-TT-p) has the same objective value.*

For $\epsilon < d^{max}_{v_0,a_0,b_0}$ we then find the following result.

▶ **Lemma 7.**
- *If $\frac{d_e}{p+1} < \epsilon < d^{max}_{v_0,a_0,b_0}$, any solution where all but two stops are at ϵ-distance to both of their neighbors, and the remaining two are at ϵ-distance to one of their neighbors, is optimal.*
- *If $\frac{d_e}{p+1} = \epsilon < d^{max}_{v_0,a_0,b_0}$, the unique solution where all stops are at ϵ-distance to both of their neighbors is optimal.*

We summarize that the p stops to be located are clustered together in an optimal solution along one edge. As can be seen easily, this behavior still holds if a complete network is given and p stops should be located there. Obviously such a solution is not realistic for practical purposes. Thus, in the following a different model will be considered which is more related to realistic needs.

4 (SL-TT-Cov) Covering all demand points

4.1 Feasibility and complexity of (SL-TT-Cov)

As seen in the previous section it does not make much sense just to add p new stops to an existing network. The main objective function used for locating new stops is usually a covering-type objective: With the new stops, one tries to cover as much demand as possible. Given a set of demand points $\mathcal{P}$ in the plane, we say that $p \in \mathcal{P}$ is covered by a set of stops $S \in \mathcal{S}$ if $dist(p, s) \leq r$ for some $s \in S$, where r is a fixed *covering radius*. The 'classic' stop location problem (SL-Cov) asks for a set of stops of minimal cardinality covering all demand points:

(SL-Cov) Let $G = (V, E)$ be a graph and a finite set of points $\mathcal{P} \subseteq \mathbb{R}^2$ be given. Find a subset $S^* \in \mathcal{S}$, such that $cover_{\mathcal{P}}(S^*) = \mathcal{P}$ and $|S^*|$ is minimized.

The goal of this section is is to cover all demand points with a set of stops S such that the realistic traveling time function $g(S)$ is minimal:

(SL-TT-Cov) Let $G = (V, E)$ be a graph, $\mathcal{P} \subseteq \mathbb{R}^2$ be a finite set of points, $v_0 > 0$, $a_0 > 0$ and $b_0 > 0$ be given. Find a subset $S^* \in \mathcal{S}$, such that $cover_{\mathcal{P}}(S^*) = \mathcal{P}$ and $g(S^*)$ is minimized.

(SL-TT-Cov) need not be feasible, but if it is it admits a finite solution whose objective value can be bounded.

▶ **Lemma 8.**
- *(SL-TT-Cov) has a solution if and only if $cover_{\mathcal{P}}(\mathcal{S}) = \mathcal{P}$.*
- *If (SL-TT-Cov) has a feasible solution, then it also has a finite solution and $g(S^*) \leq (|E| + |\mathcal{P}|) \cdot \left(\max_{e \in E} \frac{d_e}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right)$.*

While feasibility is easy to check, solving (SL-TT-Cov) is NP-hard.

▶ **Lemma 9.** *(SL-TT-Cov) is NP-hard.*

4.2 A finite dominating set for (SL-TT-Cov)

In the following we show that (SL-TT-Cov) can be reduced to a discrete problem by identifying a finite dominating set, i.e., a finite set of candidates $\mathcal{S}_{cand} \subseteq \mathcal{S}$, for which we know that it contains an optimal solution S^*, if the problem is feasible at all. Such a finite dominating set will enable us to derive an IP formulation in Section 4.3. It turns out that we can use nearly the same finite dominating set which has been used as candidate set for solving (SL-Cov) (see [17]). Throughout this section, let us assume that (SL-TT-Cov) is feasible, which can be tested (due to Lemma 8).

For an edge $e = (i, j) \in E$ we define

$$T^e(p) = \{s \in e : dist(p, s) \leq r\}$$

as the set of all points on the edge $e \subseteq \mathcal{S}$ that can be used to cover demand point p.

Since $T^e(p) = e \cap \{x \in \mathbb{R}^2 : dist(p, x) \leq r\}$ is the intersection of two convex sets, and contained in e, it turns out to be a line segment itself. This observation is due to [17].

▶ **Lemma 10** ([17])**.** *For each demand point $p \in \mathbb{R}^2$ the set $T^e(p)$ is either empty or an interval contained in edge e.*

Let f_p^e, l_p^e denote the endpoints of the interval $T^e(p)$ (which may coincide with the endpoints i, j of the edge e). We write $[f_p^e, l_p^e] = T^e(p)$. For each edge $e = (i, j)$ we define

$$\mathcal{S}_{cand}^e := \bigcup_{p \in \mathcal{P}} \{f_p^e, l_p^e\},$$

which can be ordered along the edge e with respect to $\leq_e$. Let the resulting set be given as $\mathcal{S}_{cand}^e = \{s_0, s_1, \ldots, s_{N_e}\}$. In the following we show that

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

is a finite dominating set for (SL-TT-Cov).

From [17] we know that moving a point $s \in \mathcal{S}$ until it reaches an element of $\mathcal{S}_{cand}$ does not change $cover_{\mathcal{P}}(\{s\})$.

▶ **Lemma 11** ([17]). *Let $s \in e$ for an edge e of E, and let $s_j, s_{j+1} \in \mathcal{S}_{cand}$ be two consecutive elements of the finite dominating set with $s_j <_e s <_e s_{j+1}$. Then*

$$cover_{\mathcal{P}}(\{s\}) \subseteq cover_{\mathcal{P}}(\{s_j\}) \cap cover_{\mathcal{P}}(\{s_{j+1}\}),$$

in particular, the cover of s does not decrease when moving s between s_j and s_{j+1}.

Now we are able to prove that $\mathcal{S}_{cand} = \bigcup_{e \in E} \mathcal{S}^e_{cand}$ is, indeed, a finite dominating set.

▶ **Theorem 12.** *Either (SL-TT-Cov) is infeasible, or there exists an optimal solution $S^* \subseteq \mathcal{S}_{cand}$.*

The number of candidates $|\mathcal{S}_{cand}| \leq 2|E||\mathcal{P}|$. Thus iterating leads to a number of $\mathcal{O}\left(2^{2|E||\mathcal{P}|}\right)$ different solutions to be tested. In the following, this is done by integer programming.

4.3 An integer programming formulation for (SL-TT-Cov)

Let $A^{cov} = (a_{p,s})_{p \in \mathcal{P}, s \in \mathcal{S}_{cand}}$ denote the covering matrix, given by

$$a_{ps} = \begin{cases} 1 & \text{if } p \in cover_{\mathcal{P}}(\{s\}) \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, let $E_{cand} = \{(s, s') : s, s' \in \mathcal{S}_{cand} \cup V \text{ and } s, s' \in e \text{ for some edge } e \in E\}$ be the set of all possible edges obtained by building any set of stops $S \subseteq \mathcal{S}_{cand}$. For those edges the distance $d_{e'}, e' \in E_{cand}$ can be precalculated. The IP formulation of the discrete version of (SL-TT-Cov) is then given by

$$\min \sum_{e \in E_{cand}} T(d_e) y_e \tag{1}$$

$$s.t. \sum_{s \in \mathcal{S}_{cand}} a_{ps} x_s \geq 1 \qquad \forall p \in \mathcal{P} \tag{2}$$

$$x_{s_i} + x_{s_j} - \sum_{\substack{s \in [s_i, s_j] \cap \mathcal{S}_{cand} \\ s \notin \{s_i, s_j\}}} x_s \leq y_{e'} + 1 \qquad \forall e' = (s_i, s_j) \in E_{cand} \tag{3}$$

$$x \in \{0,1\}^{|\mathcal{S}_{cand}|} \tag{4}$$

$$y \in \{0,1\}^{|E_{cand}|} \tag{5}$$

The variables x_i and y_e have the following interpretation.

$$x_{s_i} = \begin{cases} 1 & \text{if stop } s_i \in \mathcal{S}_{cand} \text{ is built.} \\ 0 & \text{otherwise.} \end{cases} \qquad y_e = \begin{cases} 1 & \text{if edge } e \in E_{cand} \text{ is built.} \\ 0 & \text{otherwise.} \end{cases}$$

Constraint (2) ensures that every demand point is covered by at least one stop. Constraints of type (3) ensure that an edge is considered in the objective function if and only if it is built, i.e. if and only if its two endpoints are stops and no candidate between the two endpoints is also a stop. Finally, the objective function (1) then gives the realistic traveling time:

▶ **Lemma 13.** *The above stated IP formulation is correct for (SL-TT-Cov).*

Since the proof is straight forward it is spared.

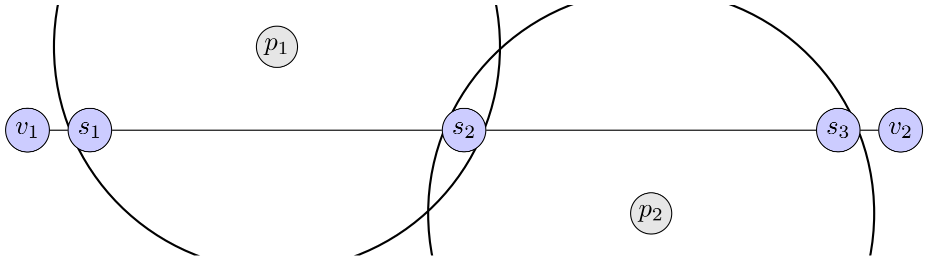

Figure 3 Example for SL-TT-Cov with more stops built than for SL-Cov.

▶ Remark. Consider the case, where $d_e \geq d^{max}_{v_0,a_0,b_0}$ for all $e \in E_{cand}$. Then the objective function can be rewritten as

$$\sum_{e' \in E_{cand}} T(d_{e'})y_e = \sum_{e \in E} T(d_e) + \left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right) \sum_{s \in S_{cand}} x_s,$$

thus the variables y_e can be eliminated, i.e., the objective function is equivalent to minimizing the number of new stops in this case. We conclude that (SL-Cov) and (SL-Cov-TT) are equivalent if $d_e \geq d^{max}_{v_0,a_0,b_0}$ for all $e \in E_{cand}$.

The number of constraints given by the candidate edges is of order $O(|\mathcal{S}_{cand}|^2)$:

▶ **Lemma 14.** *Suppose a network G=(V,E) and $\mathcal{S}_{cand}$ are given. Let $|E| = m$ and $|\mathcal{S}_{cand}| = n = \sum_{i=1}^{m} n_i$, where n_i for $i = 1, \ldots, m$ is the number of candidate stops on edge e_i. Then the number of candidate edges is given by* $|E_{cand}| = \sum_{i=1}^{m} \binom{n_i + 2}{2}$

Note that if there exists an $1 \leq i \leq m$ such that $n_i = 0$ then $T(d_{e_i})$ is a constant and thus does not have to be considered. In fact, the number of variables and constraints can then be reduced.

We close this section by a comparison between (SL-Cov) and (SL-Cov-TT). The following example shows a situation in which the realistic traveling time can be reduced by building two stops instead of only one.

▶ **Example 15.** In Figure 3 two demand points p_1 and p_2 have to be covered by stops on $e = (v_1, v_2)$. In order to minimize the number of stops it is sufficient to build only one stop, namely s_2, i.e., $s = \{s_2\}$ is an optimal solution. We compare S with the solution $\tilde{S} = \{s_1, s_2\}$, where s_1 and s_3 are close enough to v_1 and v_2 respectively. Assuming $d_{v_1,s_2}, d_{s_2,v_2} \geq d^{max}_{v_0,a_0,b_0}$, the traveling times can be computed as

$$\begin{aligned}
f(S) &= T(d_{v_1,s_2}) + T(d_{s_2,v_2}) = \frac{d_e}{v_0} + \frac{v_0}{a_0} + \frac{v_0}{b_0} \\
f(\tilde{S}) &= T(d_{v_1,s_1}) + T(d_{s_1,s_3}) + T(d_{s_3,v_2}) \\
&= \sqrt{\frac{2(a_0+b_0)}{a_0 b_0} d_{v_1,s_1}} + \frac{d_e - d_{v_1,s_1} - d_{s_3,v_2}}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0} + \sqrt{\frac{2(a_0+b_0)}{a_0 b_0} d_{s_3,v_2}}
\end{aligned}$$

and by letting d_{v_1,s_1} and d_{s_3,v_2} tend to 0, we see that $f(\tilde{S}) < f(S)$.

Other examples of the same pattern can be constructed which show that the number of stops in an optimal solution to (SL-TT-Cov) can differ by more than one from the number of stops in an optimal solution to (SL-Cov).

5 Experiments

Environment. All experiments were conducted on a PC with 24 six-core Intel Xenon X5650 Processor running at 2.67 GHz with 12 MB cache and a main memory of 94 GB. IPs were solved using Xpress Optimizer v23.01.05. The running time limit of the solver was set to 300 seconds.

Benchmark set. The southern part of the existing railway network of Lower Saxony, Germany, is used as the existing network $G = (V, E)$. From the same area the 34 largest cities are considered as demand points if they are not already close enough to an existing stop. This is the setting for the first benchmark set (LS=Lower Saxony). For our second benchmark set, stops which have only two adjacent tracks are removed and thus a set (LSR=Lower Saxony Reduced) with longer tracks and more uncovered demand points is obtained. This set has higher complexity.

The values for the traveling time are chosen according to the real properties, which are acceleration and deceleration of $0.7m/s^2$ and a cruising speed of $200km/h$. For a set of different radii ($r \in \{1750, 2100, 2450, \ldots, 12950\}$ (in meters)), we constructed instances containing all demand points which can be covered by r, i.e. $\mathcal{P}$ increases with the radius.

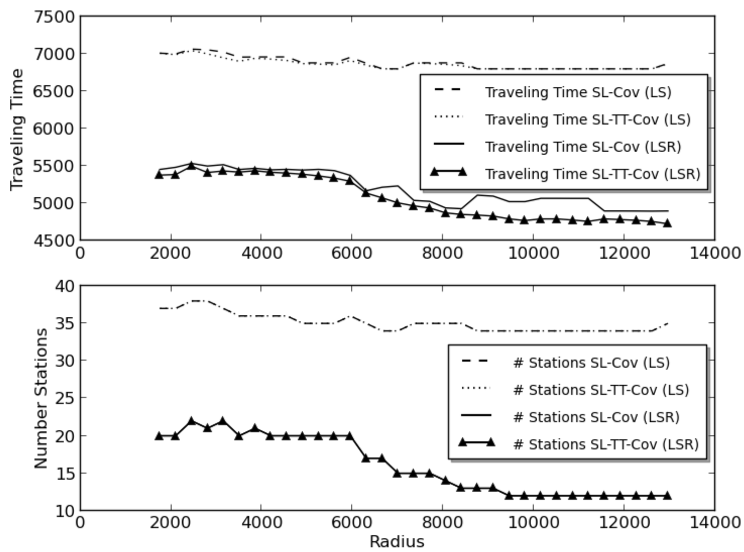

Figure 4 Traveling time with respect to the number of built stops.

Setup. The quality of the IP formulations for (SL-TT-Cov) and for (SL-Cov) (see [17] for the IP formulation of (SL-Cov)) are compared. To this end, for each of the benchmark sets and every radius $r \in \{1750, 2100, 2450, \ldots, 12950\}$ both models have been solved by Xpress Optimizer. Then for each run the quality of the solution is measured by evaluating the traveling time and the number of stops built.

Hypotheses. The evaluations are designed to approve or disprove the below stated hypotheses.

1. (SL-TT-Cov) performs better than (SL-Cov) in terms of traveling time.
2. (SL-TT-Cov) performs worse than (SL-Cov) in terms of number of stops.

Table 1 Average values of the objective functions for the solutions of (SL-Cov) and (SL-TT-Cov).

Set (LS)	SL-Cov	SL-TT-Cov	Set (LSR)	SL-Cov	SL-TT-Cov
Traveling Time $g(S)$	6878.62	6863.34	Traveling Time $g(S)$	5201.62	5069.24
# Built stops $\|S\|$	35.09	35.09	# Built stops $\|S\|$	16.06	16.06

3. With increasing radius and a fixed set of demand points, the traveling time decreases.
4. The difference in performance between (SL-TT-Cov) and (SL-Cov) is more evident on (LSR) than on (LS).
5. The running time of (SL-TT-Cov) is exponential in the number of candidates.
6. As the acceleration tends to infinity, the traveling time of the solutions of (SL-TT-Cov) and (SL-Cov) tend to the traveling time with constant speed.

Results. Table 1 summarizes our results calculating the average values of the two objective functions for all instances.

Hypotheses 1 and 2. Considering the benchmark sets (LS) and (LSR) the solutions of (SL-Cov) and (SL-TT-Cov) in terms of the number of built stops do not vary at all. In terms of the resulting traveling time on (LS) only small differences are recognizable. However, on the benchmark set (LSR) the results show bigger differences. The average traveling time of (SL-Cov) can be reduced by more than 3% by using (SL-TT-Cov). These messages can clearly be confirmed by Figure 4. Hence, from the experiments we can approve hypotheses 1 and disprove 2.

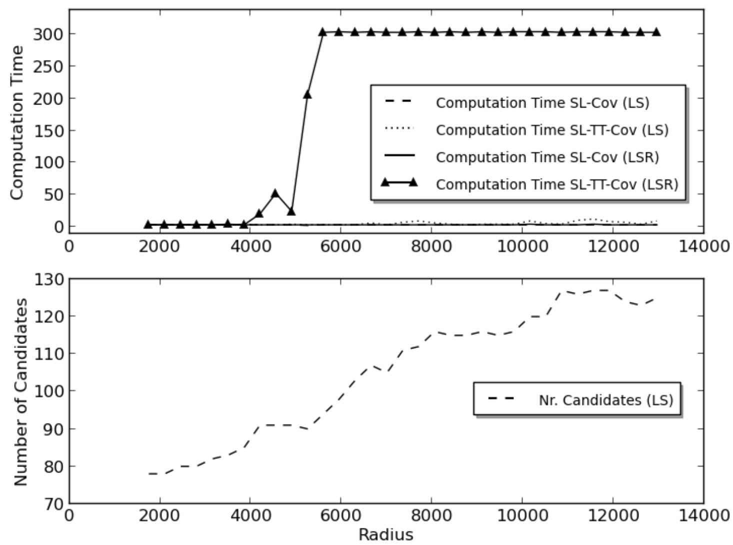

Figure 5 Computing time with respect to the number of candidates.

Hypothesis 3. Figure 4 clearly shows that in both sets (LS) and (LSR) the traveling time decreases with increasing radius r. This makes sense as for the same demand points but increasing radius possibly more demand points can be covered by the same stop. In some instances the number of stops or the traveling time increases with the radius, which is due to the increased number of demand points which require more stops to be built. In general though, we can detect that the traveling time decreases with increasing radius. Thus, the hypothesis 3 can be approved.

Hypothesis 4. Also hypothesis 4 can be approved by the results depicted in Figure 4. Note that the vertices in the underlying network G are always stops. Thus, for (LS) already 35 stops are fixed to be built, which explains why there is no big difference between (SL-Cov) and (SL-TT-Cov). For (LSR) though, the number of previously fixed stops is only 10, i.e., the models (SL-Cov) and (SL-TT-Cov) have more freedom to find a solution and hence the difference in terms of the traveling time is bigger. Hypothesis 4 can hence be approved.

Hypothesis 5. On the other hand the more freedom is granted to the models, the higher is the complexity and subsequently the higher is the running time. Figure 5 depicts the running time of (SL-Cov) and (SL-TT-Cov) for (LS) and (LSR). The maximal running time for (SL-TT-Cov) is set to 300 seconds and the solution obtained if the algorithm exceeds this limit is usually not optimal. In the experiments, all solution obtained were at least feasible, and although not necessarily optimal, the solutions for (SL-TT-Cov) have lower traveling times than the solutions for (SL-Cov). Figure 5 hence approves the hypothesis 5.

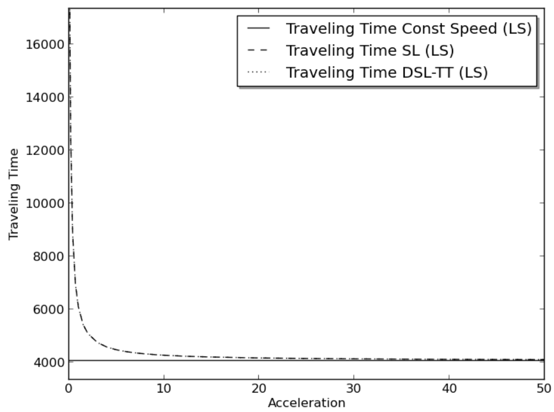

Figure 6 Relation between traveling time and acceleration

Hypothesis 6. Finally, Figure 6 depicts the behavior of the traveling time for increasing acceleration and deceleration. To this end, we solved (SL-Cov) and (SL-TT-Cov) on (LS) with a fixed radius of 3500 meters for different acceleration and deceleration values. It is assumed that acceleration and deceleration are always equal. The traveling time is compared to the function summing up all edge lengths and dividing by the cruising speed. This is the traveling time function assuming a constant speed. For increasing acceleration we can detect that the traveling time T tends to the value of the traveling time assuming constant speed. Taking into account the shape of $T(d)$ it means that for increasing a_0 and b_0 the acceleration and deceleration phases become shorter. Thus, hypothesis 6 can be approved.

6 Conclusion and further research

In this paper we included a realistic traveling time function in stop location problems. We derived a finite dominating set and an IP formulation and showed the applicability of the model on two different benchmark sets. It turns out that the solutions of (SL-TT-Cov) usually outperform the solutions of (SL-Cov) with a trade-off of having higher running times.

Further research on this topic goes into two directions. First, we assumed that all vertices of the existing network are built as stops. However, it may be better to close or move some of these. In order to model this appropriately, an integration with line planning is necessary. Secondly, the traveling time for the passengers could be even more realistic if OD-pairs are considered. Minimizing their traveling time leads to a different model and thus analysis.

References

1 J. Gleason. A set covering approach to bus stop allocation. *Omega*, 3:605–608, 1975.
2 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), 2001.
3 P. Kolesar, W. Walker, and J. Hausner. Determining the relation between fire engine travel times and travel distances in new york city. *Operations Research*, 23:614–627, 1975.
4 M.-C. Körner, J.A. Mesa, F. Perea, A. Schöbel, and D. Scholz. A maximum trip covering location problem with an alternative mode of transportation on tree networks and segments. *TOP*, 2012. published online, DOI 10.1007/s11750-012-0251-y.
5 E. Kranakis, P. Penna, K. Schlude, D.S. Taylor, and P. Widmayer. Improving customer proximity to railway stations. In *CIAC*, pages 264–276, 2003.
6 G. Laporte, J.A. Mesa, and F.A. Ortega. Locating stations on rapid transit lines. *Computers and Operations Research*, 29:741–759, 2002.
7 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.
8 S. Mecke, A. Schöbel, and D. Wagner. Stop location - complexity and approximation issues. In *5th workshop on algorithmic methods and models for optimization of railways*, number 06901 in Dagstuhl Seminar proceedings, 2006.
9 S. Mecke and D. Wagner. Solving geometric covering problems by data reduction. In *Proceedings of European Symposium on Algorithms (ESA)*, pages 760–771, 2004.
10 A. Murray. Strategic analysis of public transport coverage. *Socio-Economic Planning Sciences*, 35:175–188, 2001.
11 A. Murray. A coverage models for improving public transit system accessibility and expanding access. *Annals of Operations Research*, 123:143–156, 2003.
12 A. Murray, R. Davis, R.J. Stimson, and L. Ferreira. Public transportation access. *Transportation Research D*, 3(5):319–328, 1998.
13 A. Murray and X. Wu. Accessibility tradeoffs in public transit planning. *J. Geographical Syst.*, 5:93–107, 2003.
14 D. Poetranto, H.W. Hamacher, S. Horn, and A. Schöbel. Stop location design in public transportation networks: Covering and accessibility objectives. *TOP*, 17(2):335–346, 2009.
15 N. Ruf and A. Schöbel. Set covering problems with almost consecutive ones property. *Discrete Optimization*, 1(2):215–228, 2004.
16 A. Schöbel. *Optimization in public transportation. Stop location, delay management and tariff planning from a customer-oriented point of view.* Optimization and Its Applications. Springer, New York, 2006.
17 A. Schöbel, H.W. Hamacher, A. Liebers, and D. Wagner. The continuous stop location problem in public transportation. *Asia-Pacific Journal of Operational Research*, 26(1):13–30, 2009.
18 C. Wu and A. Murray. Optimizing public transtit quality and system access: the multiple-route, maximal covering/shortest path problem. *Environment and Planning B: Planning and Design*, 32:163–178, 2005.

A Appendix

Proof. (Lemma 6) Let $S = \{(e, x_1), \ldots, (e, x_p)\}$ with $x_1 < \ldots < x_p$ be a feasible solution, and let $x_0 = 0$ and $x_{p+1} = d_e$. Then

$$g(S) = \sum_{l=1}^{p+1} T(x_l - x_{l-1}) = \sum_{l=1}^{p+1} \frac{(x_l - x_{l-1})}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0} = \frac{d_e}{v_0} + \frac{v_0(p+1)}{2}\left(\frac{1}{a_0} + \frac{1}{b_0}\right),$$

which is independent of S. ◀

Proof. (Lemma 7) Note that $T(x - y)$ is a concave function in (x, y) on $\{(x, y) : x \geq y\}$, hence $g(x_1, \ldots, x_p) = \sum_{l=1}^{p+1} T(x_l - x_{l-1})$ is also concave. The minimum of the above program is hence taken at an extreme point of the feasible polyhedral set $P = \{(x_1, \ldots, x_p) : x_l + \epsilon \leq x_{l+1}, l = 0, \ldots, p\}$. P has exactly $p + 1$ extreme points given by

$$x^h = (x_0 + \epsilon, x_0 + 2\epsilon, \ldots, x_0 + (h-1)\epsilon, x_{p+1} - (p - (h-1))\,\epsilon, \ldots, x_{p+1} - 2\epsilon, x_{p+1} - \epsilon)$$

for $h = 1, \ldots, p + 1$. Evaluating the objective function at an extreme point x^h yields

$$\begin{aligned} g(x^h) &= \sum_{i=0}^{p} T(x^h_{i+1} - x^h_i) \\ &= pT(\epsilon) + T\left(x_{p+1} - p\epsilon + (h-1)\epsilon - x_0 - (h-1)\epsilon\right) = pT(\epsilon) + T(d_e - p\epsilon) \end{aligned}$$

which is independent of h. Hence, any of the extreme points is optimal. ◀

Proof. (Lemma 8) The first part of the lemma is obvious. For the second part, let (SL-TT-Cov) be feasible. Then there exists some point $s_p \in \mathcal{S}$ such that $dist(p, s) \leq r$ for every demand point $p \in \mathcal{P}$. Choose $S := \{s_p : p \in \mathcal{P}\}$ as a feasible solution. Each stop $s \in S$ adds a new edge to $E(S)$, hence $|E(S)| = |E| + |\mathcal{P}|$. Let $e' = (i, j) \in E(S)$ be a new edge with $i = (\bar{e}, x_i), j = (\bar{e}, x_j)$ for some $\bar{e} \in E$. Then we estimate $d_{e'} \leq d_{\bar{e}} \leq \max_{e \in E} d_e$, and since T is monotone we obtain

$$T(d_{e'}) \leq \max_{e \in E} T(d_e) \overset{(I)}{\leq} \max_{e \in E} \frac{d_e}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0}, \text{ where (I) is a result from Lemma 2.}$$

$$\text{Hence, } g(S^*) \leq g(S) \leq |E(S)| \max_{e \in E} T(d_e) \leq (|E| + |\mathcal{P}|) \max_{e \in E} \frac{d_e}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0}.$$

◀

Proof. (Lemma 9) To see that (SL-TT-Cov) is NP-hard, we reduce it to the discrete stop location in a network: Given a network embedded in the plane, a set of demand points, and a finite candidate set $\mathcal{S}_{cand}$, find a set $S^* \subseteq \mathcal{S}_{cand}$ with minimal cardinality covering all demand points. This problem is NP-hard, also if $V \subseteq \mathcal{S}_{cand}$, see [17]. Let an instance of the discrete stop location problem be given. Determine $\underline{m} := \min\{d(s, s') : s, s' \in \mathcal{S}_{cand}$, and $s, s' \in e$ for some $e \in E\}$ as the closest distance between two candidate locations on the same edge. Note that for $a_0, b_0 \to \infty$ we obtain that $d^{max}_{a_0,b_0,v_0} \to 0$. Hence choose a_0, b_0 large enough such that $d^{max}_{a_0,b_0,v_0} \leq \underline{m}$. We claim that a solution to the discrete stop location problem with $|S| \leq K$ exists if and only if a solution S^* to (SL-TT-Cov) exists with $g(S^*) \leq \sum_{e \in E)} \frac{d_e}{v_0} + (|E| + K)\left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right)$.

To see this, note that for $d_{e'} \geq d^{max}_{a_0,b_0,v_0}$ for all $e' \in E(S)$ the objective function of (SL-TT-Cov) reduces to

$$g(S) = \sum_{e' \in E(S)} \left(\frac{d_{e'}}{v_0} + \frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right) = \sum_{e \in E} \frac{d_e}{v_0} + |E(S)|\left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right). \tag{6}$$

"$\Rightarrow$" Let S be a solution to (SL) with $|S| \leq K$. Then there exists another optimal solution $S^* \subseteq \mathcal{S}_{cand}$. Then S^* is feasible for (SL-TT-Cov) and $d_{e'} \geq \underline{m} \geq d^{max}_{a_0,b_0,v_0}$ for all $e' \in E(S^*)$. Hence $g(S) \leq \sum_{e\in E} \frac{d_e}{v_0} + (|E|+K)\left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right)$.

"$\Leftarrow$" Let S^* be a solution to (SL-TT-Cov) with $g(S^*) \leq \sum_{e\in E} \frac{d_e}{v_0} + (|E|+K)\left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right)$. Again, there exists $S \subseteq \mathcal{S}_{cand}$ with $g(S^*) = g(S)$. Since $d_{e'} \geq \underline{m} \geq d^{max}_{a_0,b_0,v_0}$ for all $e' \in E(S^*)$ we have

$$\sum_{e\in E)} \frac{d_e}{v_0} + (|E|+K)\left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right) \geq g(S) = g(S^*) = \sum_{e\in E} \frac{d_e}{v_0} + |E(S)|\left(\frac{v_0}{2a_0} + \frac{v_0}{2b_0}\right),$$

from which we conclude $|E(S)| \leq |E| + K \Leftrightarrow |S| \leq K$.

◀

Proof. (Theorem 12) Let $S^* \subseteq \bigcup_{e\in E, p\in\mathcal{P}} T^e(p)$ be optimal, but $S^* \not\subseteq \mathcal{S}_{cand}$. The goal is to replace each $\tilde{s} \in S^* \setminus \mathcal{S}_{cand}$ by a point in $\mathcal{S}_{cand}$ without loosing feasibility or optimality. To this end, take some $\tilde{s} \in S^* \setminus \mathcal{S}_{cand}$. If $\tilde{s} \in V$, then $\tilde{s}$ can be removed, since the vertices are existing stops. Thus, we can assume that $\tilde{s} \notin V$, i.e. $\tilde{s} = (e,x) \in E$. Now find the following points on edge e:

- $s_j = (e,x_j), s_{j+1} = (e,x_{j+1}) \in \mathcal{S}_{cand}$ with $s_j <_e \tilde{s} <_e s_{j+1}$ for two consecutive elements of $\mathcal{S}^e_{cand}$ (if they exist on e), and
- $s_{left}, s_{right} \in (S^* \cup V) \cap e$ with $s_{left} <_e \tilde{s} <_e s_{right}$ for the two direct neighbors of $\tilde{s}$ on e (which always exist)

We now investigate the objective function if we move $\tilde{s}$. For all s with $s_{left} = (e,x_{left}) <_e s = (e,x) <_e s_{right} = (e,x_{right})$ the objective function $h(x) := g(S \setminus \tilde{s} \cup \{(e,x)\})$ is given as

$$h(x) = \sum_{e'\in E(S\setminus\tilde{s}\cup\{(e,x)\}} T(d_{e'}) = const + T(x - x_{left}) + T(x_{right} - x)$$

where the constant part is independent of the choice of $s = (e,x)$. As in Lemma 4, $h(x)$ is concave in x on the segment between s_{left} and s_{right}. Furthermore, from Lemma 11 we know that $cover_{\mathcal{P}}(\{s\}) \supseteq cover\{\tilde{s}\}$ for all $s = (e,x)$ between s_j and s_{j+1}. Now consider the minimization problem

$$\min\{h(x) = T(x - x_{left}) + T(x_{right} - x) : \max\{x_{left}, x_j\} \leq x \leq \min\{x_{right}, x_{j+1}\}\}.$$

Due to the concavity of $h(x)$ we know that an optimal solution $x^* \in \{x_{left}, x_{right}, x_j, x_{j+1}\}$ exists.

1. In case that $x^* = x_j$ or $x^* = x_{j+1}$ we may replace $\tilde{s}$ by $s = (e,x^*) \in \mathcal{S}_{cand}$ and hence obtain a feasible solution with the same objective value.
2. In case that $x^* = x_{left}$ or $x^* = x_{right}$ we may delete $\tilde{s}$ since the new solution is still feasible and has the same objective value.

In both cases, we have reduced the number of points in $S^* \setminus \mathcal{S}_{cand}$. Proceeding with remaining points of S^* which do not belong to $\mathcal{S}_{cand}$ finishes the proof. ◀

Proof. (Lemma 14) The sum is obtained since the number of candidate edges on each edge e of the original graph G can be calculated independently by $\binom{n+2}{2}$. ◀