New Mixed Integer Non Linear Programming ($MINLP$) Models for the Euclidean Steiner Tree Problem in R^n

Renan Vicente Pinto
Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro
renanvp@ufrrj.br

Hacène Ouzia
LIP6, Sorbonne Université, Paris
hacene.ouzia@sorbonne-universite.fr

Nelson Maculan
Universidade Federal do Rio de Janeiro, Rio de Janeiro
maculan@cos.ufrj.br

December 3, 2020

Abstract

Two new Mixed Integer Non Linear Programming (MINLP) models for the Euclidean Steiner Tree Problem in R^n will be presented in this work. The novelty of these models is the introduction of constraints that represent second-order cones, avoiding the problem of non-differentiability of continuous relaxation, which appears in other models. Computational results using the XPRESS software provided a more effective way to solve the Euclidean Steiner Tree Problem in R^n.

Keywords: Integer Programming, Euclidean Steiner tree problem, Steiner tree, Nonlinear optimization models, Mixed integer nonlinear optimization, Continuous relaxation, Second order cones.

1 Introduction

An important history of the Euclidean Steiner Tree Problem (ESTP) is presented in [2], but nothing is said on the optimization models to solve it. An interesting application of the ESTP using heuristics is in [12] and [11]. The first optimization model for this problem was presented in [10]. From this formulation, another formulation of the ESTP is found in [7], [8], [5], [13]. An overview of exact algorithms for the ESTP in n-dimensional space can be found in [3] and [6].

This paper is organized as follows. In the next section we use a graph defined in [10] on which all the mathematical optimization models presented in this work are based. It also shows how to change the coordinates of the given points so that they are all contained in a unitary hypercube.

Formulations derived from [10] and [13] will be presented in section 3, of which the last two are new and where second order cone constraints are introduced. These six mathematical optimization models are Mixed Integer Non Linear Programming (MINLP).

In section 4 known geometric cuts are presented. Considerations about tree isomorphism elimination, and constraints to eliminate isomorphism are also presented. Based on the enumeration scheme considered by Smith [14] symmetry elimination constraints are proposed. Computational results to compare the performance of the two new models are performed in the section 5.

2 Definitions

Given p different points in $\mathbb{R}^n$, the ESTP seeks to find a minimum tree that spans these points using or not extra points, which are called Steiner points. The length of each edge is the Euclidean distance between its ends.

We consider a special graph $G = (V, E)$ as follows (see [10]): let $P = \{1, 2, ..., p-1, p\}$ be the set of indices associated with the given points $x^1, x^2, \ldots, x^{p-1}, x^p \in \mathbb{R}^n$ and a set of indices $S = \{p+1, p+2, ..., 2p-3, 2p-2\}$ associated with the Steiner points $x^{p+1}, x^{p+2}, ..., x^{2p-3}, x^{2p-2} \in \mathbb{R}^n$. We take $V = P \cup S$ as the set of vertices of G. For $i, j \in V$, we denote an edge of G connecting vertices i and j as $[i, j]$. Thus we define $E_1 = \{[i, j] \mid i \in P,\ j \in S\}$, $E_2 = \{[i, j] \mid i < j,\ i, j \in S\}$ and $E = E_1 \cup E_2$. A tree which is an optimal solution for the ESTP is a sub-graph of $G = (V, E)$ (see [10]). It is very easy to verify that all Steiner points belong to the convex hull of the points $x^1, x^2, ..., x^{p-1}, x^p$. Let

$$||x^i - x^j|| = \sqrt{\sum_{k=1}^{n} (x_k^i - x_k^j)^2}$$

be the Euclidean distance between x^i and x^j. We compute

$$M = \max_{1 \le i < j \le p} ||x^i - x^j||,$$

which implies

$$||x^i - x^j|| \le M, \quad \forall\ [i, j] \in E. \tag{1}$$

For obtaining a better upper bound M, see Theorem 8.1 in [9]. Doing some algebra on (1), we have

$$\frac{1}{M}\sqrt{\sum_{k=1}^{n}(x_k^i - x_k^j)^2} \leq 1, \qquad \forall\, [i,j] \in E,$$

$$\sqrt{\sum_{k=1}^{n}\frac{(x_k^i - x_k^j)^2}{M^2}} \leq 1, \qquad \forall\, [i,j] \in E,$$

$$\sqrt{\sum_{k=1}^{n}\left(\frac{x_k^i}{M} - \frac{x_k^j}{M}\right)^2} \leq 1, \qquad \forall\, [i,j] \in E.$$

Thus, applying the change of variables $x_k^i := \dfrac{x_k^i}{M}$, $\forall\, i \in V$, $\forall\, k = 1,2,...,n$, and a translation so all coordinates will be non-negative, we can consider, without loss of generality

$$\max_{1\leq i<j\leq p} ||x^i - x^j|| = 1 \qquad \text{e} \qquad x_k^i \geq 0, \quad \forall\, k = 1,2,...,n,\ \forall\, i \in S.$$

In this case,

$$-1 \leq (x_k^i - x_k^j) \leq 1, \quad \forall\, k = 1,2,...,n,\ \forall\, [i,j] \in E. \tag{2}$$

3 Optimization models

3.1 1^{st} Optimization Model, Maculan-Michelon-Xavier, 2000, [10]

$$(P1):\quad Minimize \sum_{[i,j]\in E} ||x^i - x^j||\, y_{ij},\ subject\ to: \tag{3}$$

$$\sum_{j\in S} y_{ij} = 1, \quad \forall\, i \in P, \tag{4}$$

$$\sum_{i\in P} y_{ij} + \sum_{k<j,k\in S} y_{kj} + \sum_{k>j,k\in S} y_{jk} = 3, \quad \forall\, j \in S, \tag{5}$$

$$\sum_{k<j,k\in S} y_{kj} = 1, \quad \forall\, j \in S - \{p+1\}, \tag{6}$$

$$\sum_{i\in P} y_{ij} \leq 2, \quad \forall\, j \in S, \tag{7}$$

$$x^i \in \mathbb{R}^n, \qquad \forall\, i \in S, \tag{8}$$

$$y_{ij} \in \{0,1\}, \qquad \forall\, [i,j] \in E. \tag{9}$$

If the binary variable $y_{ij} = 1$, then the edge (i, j) is in the solution. Note that constraints (7) are valid only for $p > 3$. Theoretically we do not need to consider constraints (5) and (7), but these constraints are valid for the model. If we consider $0 \leq x_k^i \leq 1$, for all $k = 1, \ldots, n$, for all $i \in P$, then (2) is also valid.

The continuous relaxation of $(P1)$ is not convex. About the use of this model, see [4].

3.2 2^{nd} Optimization Model, Fampa-Maculan, 2001, 2004, [7], [8]

$$(P2): \quad Minimize \sum_{[i,j] \in E} d_{ij}, \ subject\ to: \tag{10}$$

$$d_{ij} \geq ||x^i - x^j|| + y_{ij} - 1, \quad \forall\, [i,j] \in E, \tag{11}$$

$$(4), (5), (6), (7),$$

$$d_{ij} \geq 0, \qquad \forall\, [i,j] \in E, \tag{12}$$

$$x^i \in \mathbb{R}^n, \qquad \forall\, i \in S, \tag{13}$$

$$y_{ij} \in \{0, 1\}, \qquad \forall\, [i,j] \in E. \tag{14}$$

The continuous relaxation of $(P2)$ is convex.

3.3 3^{rd} Optimization Model, Ouzia-Maculan, 2018, [13]

As $y_{ij} \in \{0, 1\}$, we have $y_{ij} = y_{ij}^2, \ \forall\, [i,j] \in E$. Consider the objective function (3) in model (P1). We can rewrite the term in the summation as

$$||x^i - x^j||\ y_{ij} = y_{ij}^2 \sqrt{\sum_{k=1}^{n} (x_k^i - x_k^j)^2} \tag{15}$$

$$= \sqrt{y_{ij} \sum_{k=1}^{n} (x_k^i - x_k^j)^2} \tag{16}$$

$$= \sqrt{\sum_{k=1}^{n} y_{ij} (x_k^i - x_k^j)^2} \tag{17}$$

$$= \sqrt{\sum_{k=1}^{n} d_{ijk}^2}\ , \tag{18}$$

where, for all $[i,j] \in E$ and for all $k = 1, 2, ..., n$,

$$-y_{ij} \leq d_{ijk} \leq y_{ij}\ ,$$

$$-(1 - y_{ij}) + (x_k^i - x_k^j) \leq d_{ijk} \leq (x_k^i - x_k^j) + (1 - y_{ij}).$$

We reformulate (P1) as

$$(P3): \quad Minimize \sum_{[i,j]\in E} \sqrt{\sum_{k=1}^{n} d_{ijk}^2}, \; subject\ to: \tag{19}$$

$$-y_{ij} \le d_{ijk} \le y_{ij}, \;\; [i,j] \in E, \;\; k = 1,2,...,n. \tag{20}$$

$$-(1-y_{ij}) + (x_k^i - x_k^j) \le d_{ijk} \le (x_k^i - x_k^j) + (1-y_{ij}), \;\; [i,j] \in E, \;\; k = 1,2,...,n. \tag{21}$$

$$(4), (5), (6), (7),$$

$$x^i \in \mathbb{R}^n, \qquad \forall\, i \in S, \tag{22}$$

$$y_{ij} \in \{0,1\}, \qquad \forall\, [i,j] \in E. \tag{23}$$

The function $f(d_{ij}) = \sqrt{\sum_{k=1}^{n} d_{ijk}^2}$ is convex but it is not differentiable. We can use $\hat{f}(d_{ij}) = \sqrt{\lambda + \sum_{k=1}^{n} d_{ijk}^2}$ instead, where $\lambda > 0$. For example, $\lambda = 10^{-10}$.

The continuous relaxation of $(P3)$ is also convex.

3.4 4^{th} Optimization Model, Ouzia-Maculan, 2018, [13]

$$(P4): \quad Minimize \sum_{[i,j]\in E} \sqrt{d_{ij}}, \; subject\ to: \tag{24}$$

$$d_{ij} \ge \sum_{k=1}^{n} (x_k^i - x_k^j)^2 + (y_{ij} - 1)\, n, \quad \forall\, [i,j] \in E. \tag{25}$$

$$(4), (5), (6), (7),$$

$$d_{ij} \ge 0, \qquad \forall\, [i,j] \in E, \tag{26}$$

$$x^i \in \mathbb{R}^n, \qquad \forall\, i \in S, \tag{27}$$

$$y_{ij} \in \{0,1\}, \qquad \forall\, [i,j] \in E. \tag{28}$$

The objective function (24) is a concave function. Thus the continuous relaxation of $(P4)$ is not convex.

3.5 5^{th} Optimization Model, Maculan-Ouzia, 2019

Consider model (P2). Constraints (11) can be written as

$$d_{ij} + (1 - y_{ij}) \;\ge\; ||x^i - x^j||, \quad \forall\, [i,j] \in E. \tag{29}$$

We define new variables

$$z_{ij} = d_{ij} + (1 - y_{ij}), \quad \forall\ [i,j] \in E.$$

As $d_{ij} \geq 0$, from (12), we have $z_{ij} \geq 0,\ \forall\ [i,j] \in E$. Let us define

$$t_{ijk} = x^i_k - x^j_k, \quad \forall\ [i,j] \in E, \quad k = 1, 2, ..., n.$$

We can write

$$||x^i - x^j|| = \sqrt{\sum_{k=1}^{n} (x^i_k - x^j_k)^2} = \sqrt{\sum_{k=1}^{n} t^2_{ijk}}, \quad \forall\ [i,j] \in E.$$

We can write (29) as

$$z_{ij} \geq \sqrt{\sum_{k=1}^{n} t^2_{ijk}}, \quad \forall\ [i,j] \in E.$$

As both sides are positive, we can square them and replace constraints (11) by

$$z^2_{ij} \ \geq \ \sum_{k=1}^{n} t^2_{ijk}, \quad \forall\ [i,j] \in E, \tag{30}$$

$$z_{ij} = d_{ij} + (1 - y_{ij}), \quad \forall\ [i,j] \in E. \tag{31}$$

Notice that (30) and $z_{ij} \geq 0,\ \forall\ [i,j] \in E$, are second order cone constraints, see [17]. We obtain the following model:

$$(P5): \quad Minimize \sum_{[i,j] \in E} d_{ij},\ subject\ to: \tag{32}$$

$$z^2_{ij} \ \geq \ \sum_{k=1}^{n} t^2_{ijk}, \quad \forall\ [i,j] \in E, \tag{33}$$

$$z_{ij} = d_{ij} + (1 - y_{ij}), \quad \forall\ [i,j] \in E. \tag{34}$$

$$-y_{ij} \leq t_{ijk} \leq y_{ij},\ [i,j] \in E,\ k = 1, 2, ..., n. \tag{35}$$

$$-(1 - y_{ij}) + (x^i_k - x^j_k) \leq t_{ijk} \leq (x^i_k - x^j_k) + (1 - y_{ij}),\ [i,j] \in E,\ k = 1, 2, ..., n. \tag{36}$$

$$(4), (5), (6), (7),$$

$$d_{ij} \geq 0, \qquad \forall\ [i,j] \in E, \tag{37}$$

$$x^i \in \mathbb{R}^n, \qquad \forall\ i \in S, \tag{38}$$

$$y_{ij} \in \{0, 1\}, \qquad \forall\ [i,j] \in E. \tag{39}$$

3.6 6^{th} Optimization Model, Maculan-Ouzia-Pinto, 2020

In model (P5), for $[i,j] \in E$, as d_{ij} has a positive coefficient in the objective function (32), the model tends to minimize the value of this variable. As a consequence, from (34), the value of z_{ij} also tends to be minimized. We have

$$\begin{aligned} y_{ij} = 0 \quad & \Longrightarrow \quad t_{ijk} = 0, \quad \forall\, k = 1, \dots, n & \text{[from (35)]} \\ & \Longrightarrow \quad d_{ij} = 0 \quad \text{and} \quad z_{ij} = 1 & \text{[from (34) and (37)]} \end{aligned}$$

Also, from (34),

$$y_{ij} = 1 \quad \Longrightarrow \quad d_{ij} = z_{ij} = \sum_{k=1}^{n} t_{ijk}^2$$

From this results, the variables z_{ij} can be discarded. Constraints (33) and (34) will be replaced by

$$d_{ij}^2 \geq \sum_{k=1}^{n} t_{ijk}^2, \quad \forall\, [i,j] \in E.$$

Thus we will have:

$$(P6): \quad Minimize \sum_{[i,j] \in E} d_{ij}, \ subject\ to: \tag{40}$$

$$d_{ij}^2 \geq \sum_{k=1}^{n} t_{ijk}^2, \quad \forall\, [i,j] \in E. \tag{41}$$

$$-y_{ij} \leq t_{ijk} \leq y_{ij}, \ [i,j] \in E, \ k = 1, 2, ..., n. \tag{42}$$

$$-(1 - y_{ij}) + (x_k^i - x_k^j) \leq t_{ijk} \leq (x_k^i - x_k^j) + (1 - y_{ij}), \ [i,j] \in E, \ k = 1, 2, ..., n. \tag{43}$$

$$(4), (5), (6), (7),$$

$$d_{ij} \geq 0, \quad [i,j] \in E, \tag{44}$$

$$x^i \in \mathbb{R}^n, \ i \in S, \tag{45}$$

$$y_{ij} \in \{0, 1\}, \quad [i,j] \in E. \tag{46}$$

4 Improvements to all models

In this section we describe valid constraints that can be added to all proposed models.

4.1 Geometric cuts

For each given point (terminal) x^i, $i \in P$, calculate its minimum distance to any other given point, i.e., let

$$\eta_i = \min_{j \in P,\, j \neq i} \|x^i - x^j\|, \quad \forall\, i \in P.$$

As proved in [15], two terminal points x^i and x^j may be connected to the same Steiner point only if

$$\|x^i - x^j\| \ \leq\ \eta_i + \eta_j. \tag{47}$$

It is also shown that this property can be strengthened to

$$\|x^i - x^j\| \ \leq\ \sqrt{\eta_i^2 + \eta_j^2 + \eta_i\,\eta_j}. \tag{48}$$

Using (47), we can add to all proposed models the following constraints:

$$y_{is} + y_{js} \leq 1,\ \forall\, i < j \in P,\ \forall\, s \in S \text{ such that } \|x^i - x^j\| > \eta_i + \eta_j. \tag{49}$$

Or, using (48),

$$y_{is} + y_{js} \leq 1,\ \forall\, i < j \in P,\ \forall\, s \in S \text{ such that } \|x^i - x^j\| > \sqrt{\eta_i^2 + \eta_j^2 + \eta_i\,\eta_j}. \tag{50}$$

4.2 Tree isomorphism elimination

Consider the two topologies for solutions depicted in Figure 1 for an instance of the ESTP with $p = 5$ terminal points. The only difference between them is the exchange of labels in nodes 6 and 7.

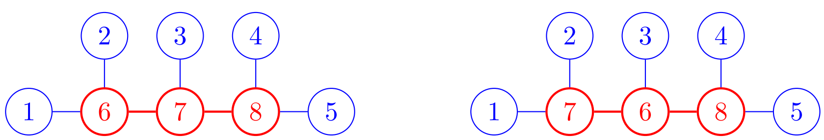


Figure 1: Considering the subtree spanned by the Steiner points (in red), we have two isomorphic trees.

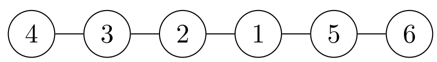
The two solutions in Figure 1 are the same from the application's point of view. In both of them, terminals 1 and 2 will be connected to the same Steiner point; terminal 3 will be connected to the Steiner point in the middle; and terminals 4 and 5 will be connected to the same Steiner point.

When two trees have the same topology and the only difference between them is the labeling, we call them *isomorphic*. We can improve the model adding constraints to cut isomorphic solutions (trees). In this section, we focus on the subtree spanned by the Steiner points.

According to [1] and [16], isomorphic trees having $r \geq 3$ vertices can be eliminated if we consider that the trees must satisfy the following properties:

- The root node of the tree (labeled 1) is in the middle of the longest path in the graph. Therefore, the root node must have at least two children.
- Consider the disjoint subtrees S_i rooted in each child of the root node . Let h_i be the height of S_i. The heights must be in not-ascending order $(h_1 \geq h_2 \geq \ldots)$.
- $h_1 - h_2 \leq 1$, which implies that or $h_1 = h_2$ or $h_1 = h_2 + 1$.
- If $h_1 = h_2$, then S_1 has more vertices than S_2.
- If $h_1 = h_2 + 1$, then $m - 2 \leq r - m + 2$, where m is the label of the root vertex of S_2.

The authors represent a tree by a list of the levels of the vertices, visited in depth-first search. For example, the canonical representation of the graph



is [1 2 3 4 2 3]. A representation of the graph is said to be *canonical* if it satisfies the properties described above. We will write constraints that allow only trees in the canonical form.

Level sequence

The representation of the tree is encoded in the integer variables $l_j,\ j \in S$. We have

$$l_{p+1} = 1 \tag{51}$$
$$l_j \leq L_{\max}, \qquad \forall j \in S,\ j > p+1 \tag{52}$$
$$2 \leq l_j \leq l_{j-1} + 1, \qquad \forall j \in S,\ j > p+1 \tag{53}$$

The highest level of a vertex in a canonical representation is $L_{max} = \left\lfloor \frac{p-2}{2} \right\rfloor + 1$, which can be demonstrated using the tree with the shape of a line.

Identifying the children of the root node

Consider the binary variables $seg_j,\ j = p+3, \ldots, 2p-2$, such that $seg_j = 1$ if node j is the root of the subtree S_2. It also means that the position of the second ocurrence of the number 2 in the representation of the tree is j, which will be the value of the variable m. We have

$$\sum_{j \in S,\, j > p+2} seg_j = 1 \tag{54}$$
$$y_{p+1,j} \geq seg_j \qquad \forall j \in S,\ j > p+2 \tag{55}$$
$$m = \sum_{j \in S,\, j > p+2} j\, seg_j \tag{56}$$

Consider the binary variables $terc_j,\ j = p+4,\ldots,2p-2$, such that $terc_j = 1$ if node j is the root of the subtree S_3, in case it exists. It also means that the position of the third ocurrence of the number 2 in the representation of the tree is j, which will be the value of the variable mm. We have

$$\sum_{j\in S,\, j>p+3} terc_j = \sum_{j\in S,\, j>p+1} y_{p+1,j} - 2 \tag{57}$$

$$y_{p+1,j} \geq terc_j, \quad \forall j \in S,\ j > p+3 \tag{58}$$

$$mm = \sum_{j\in S,\, j>p+3} j\, terc_j + (2p-1)\left(1 - \sum_{j\in S,\, j>p+3} terc_j\right) \tag{59}$$

$$m \leq mm - 1 \tag{60}$$

If S_3 doesn't exist, (59) forces $mm = 2p - 1$.

Identifying the nodes in the respective S_i

Consider the binary variables b_{jk}, $j = p+2,\ldots,2p-2$, $k = 1,\ldots,3$. If node j belongs to S_k, then $b_{jk} = 1$. And $b_{jk} = 0$, otherwise. We have

$$b_{j1} + b_{j2} + b_{j3} = 1 \qquad \forall j \in S,\ j > p+1 \tag{61}$$

$$j - m + 1 \leq (j - p - 2)(1 - b_{j1}) \qquad \forall j \in S,\ j > p+1 \tag{62}$$

$$m - j \leq (2p - 2 - j)(1 - b_{j2}) \qquad \forall j \in S,\ j > p+1 \tag{63}$$

$$j - mm + 1 \leq (j - p - 3)(1 - b_{j2}) \qquad \forall j \in S,\ j > p+1 \tag{64}$$

$$mm - j \leq (2p - 1 - j)(1 - b_{j3}) \qquad \forall j \in S,\ j > p+1 \tag{65}$$

Calculating the heights

Consider the binary variables q_{jk}, $j = p+2,\ldots,2p-2$, $k = 1,\ldots,3$. If node j has the highest level in S_k, then $q_{jk} = 1$. And $q_{jk} = 0$, otherwise. We have

$$\sum_{j\in S, j>p+1} q_{jk} = 1 \qquad \forall k = 1,\ldots,2 \tag{66}$$

$$\sum_{j\in S, j>p+1} q_{j3} = \sum_{j\in S,\, j>p+1} y_{p+1,j} - 2 \tag{67}$$

$$q_{jk} \leq b_{jk} \qquad \forall j \in S,\ j > p+1, \forall k = 1,\ldots,3 \tag{68}$$

Consider the variable h_k representing the height of $S_k,\ k = 1,\ldots,3$. We know

$$h_k = \max_{j\in S,\, j>p+1} b_{jk}\, l_j.$$

We use the McCormick inequalities, where $w_{jk} = b_{jk}\, l_j$:

$$w_{jk} \le h_k \le w_{jk} + L_{\max}(1 - q_{jk}) \qquad \forall j \in S,\ j > p+1,\ \forall k = 1,\dots,3 \tag{69}$$

$$b_{jk} \le w_{jk} \le L_{\max}\, b_{jk} \qquad \forall j \in S,\ j > p+1,\ \forall k = 1,\dots,3 \tag{70}$$

$$L_{\max}\, b_{jk} - L_{\max} + l_j \le w_{jk} \le l_j + b_{jk} - 1 \qquad \forall j \in S,\ j > p+1,\ \forall k = 1,\dots,3 \tag{71}$$

Properties from the article

$$h_1 \ge h_2 \ge h_3 \tag{72}$$

$$h_1 - h_2 \le 1 \tag{73}$$

$$2m - mm \ge p + 2 - (p-5)(h_1 - h_2) \tag{74}$$

$$m \le M + (2p - 2 - M)(1 - h_1 + h_2) \tag{75}$$

Where $M = \left\lfloor \dfrac{3p+2}{2} \right\rfloor$.

Linking variables l and y

If $y_{ij} = 1$, then we must have $l_i + 1 = l_j$:

$$(2 - L_{\max})(1 - y_{ij}) + l_j - 1 \ \le\ l_i \ \le\ l_j - 1 + L_{\max}(1 - y_{ij}). \tag{76}$$

Also, nodes i and j must belong to S_k, for some k:

$$-(1 - y_{ij}) + b_{jk} \ \le\ b_{ik} \ \le\ b_{jk} + (1 - y_{ij}). \tag{77}$$

In this paragraph, consider $i, j \in S$, $i < j$. We will use a binary variable ρ_{ij} that takes value 1 if $l_i + 1 = l_j$ (node i could be father of node j). We define the variable $w_{ij} = |l_i + 1 - l_j|$. We want $\rho_{ij} = 1 \Leftrightarrow w_{ij} = 0$. For the modulus linearization, we will use the binary variable s_{ij}. For all $i, j \in S$, $i < j$,

$$l_i + 1 - l_j \le w_{ij} \le l_i + 1 - l_j + 2L_{\max}\, s_{ij} \tag{78}$$

$$-(l_i + 1 - l_j) \le w_{ij} \le -(l_i + 1 - l_j) + 2L_{\max}(1 - s_{ij}) \tag{79}$$

$$\rho_{ij} \ge 1 - w_{ij} \tag{80}$$

$$L_{\max}\, \rho_{ij} + w_{ij} \le L_{\max}. \tag{81}$$

Finally, the father of node j is $f_j = \max\{\, i\, \rho_{ij} \mid p + 1 < i < j \le 2p - 2\}$.

$$i\, \rho_{ij} \ \le\ f_j \ \le\ i\, \rho_{ij} + (j-1)(1 - y_{ij}) \qquad \forall i, j \in S,\ i < j \tag{82}$$

4.3 Symmetry elimination constraints

In the previous section, we considered the isomorphism only in the subtree spanned by the Steiner points. We now focus on the entire tree, including the terminal points.

We will use the idea presented by Smith in his paper [14]. The author presents a bijection between every possible full Steiner topology on p terminal points and vectors $a \in \mathbb{N}^{p-3}$, where

$$1 \leq a_i \leq 2i+1, \qquad i = 1, \ldots, p-3. \tag{83}$$

If there are $p = 3$ terminal points, there is only one full Steiner topology, as in Figure 2, which corresponds to the null vector. The edges are also labeled. The construction of the tree is by induction on p, by adding one Steiner point and one terminal point at a time. To generate the full Steiner topology corresponding to a given vector $(a_1, \ldots, a_{p-3})$ from the topology corresponding to $(a_1, \ldots, a_{p-4})$, place the new Steiner point in the middle of edge a_{p-3} and connect the new terminal point to it. The label of the new edges will be $2p-4$ and $2p-3$.

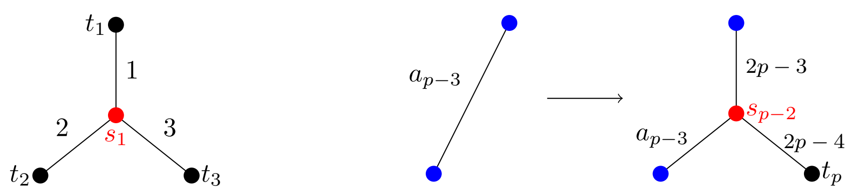


Figure 2: Tree construction from Smith's vector [14].

If there are $p = 4$ terminal points, only 3 full Steiner topologies will be considered, which correspond to the $\mathbb{N}^1$-vectors 1, 2 and 3, as in Figure 3. The tree is constructed from the 3-terminal full topology depicted in Figure 2. For instance, when $a_1 = 1$, a new Steiner point will be placed in the middle of edge 1 and connected to a new terminal. It is analogous for $a_1 = 2$ and $a_1 = 3$.

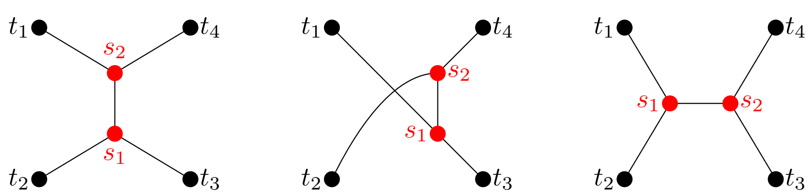


Figure 3: Smith's representation for full Steiner topologies when there are 4 terminal points.

To work only with full Steiner topologies that correspond to the vectors described in (83), we define the binary variables

$$v_{ij} \in \{0, 1\}, \qquad i = 1, \ldots, p-3, \quad j = 1, \ldots, 2i+1, \tag{84}$$

that assume value 1 if $a_i = j$. We must have

$$\sum_{j=1}^{2i+1} v_{ij} = 1, \quad \forall\, i = 1, \ldots, p-3. \tag{85}$$

We assume all points are labeled as follows: terminal points are labeled $1, \ldots, p$ and Steiner points are labeled $p+1, \ldots, 2p-2$. To control the edges labeling, we define, for all $i = 0, \ldots, p-3,\ j = 1, \ldots, 2i+3,\ k \in \{1, 2\},\ l = 1, \ldots, 2p-2$, the binary variables

$$e_{ijkl} \in \{0, 1\}. \tag{86}$$

The vertices to which an edge labeled j is incident may change after a step of the tree reconstruction method (see edge a_{p-3} in Figure 2). In (86), the index j is the label of the edge and the index $k = 1, 2$ refers to each of the two vertices the edge is incident to. The index i is the iteration of the tree construction method and l is the vertex label. We will then have $e_{ijkl} = 1$ if, in stage i of the method, vertex k of edge j is the one labeled l. We must have

$$\sum_{l=1}^{2p-2} e_{ijkl} = 1, \quad \forall i, j, k. \tag{87}$$

along with some initial conditions (corresponding to the null-vector or the 3-terminal topology in Figure 2):

$$e_{0j1j} = 1, \qquad j = 1, 2, 3, \tag{88}$$

$$e_{0j2(p+1)} = 1, \qquad j = 1, 2, 3. \tag{89}$$

In iteration $i > 0$, the new Steiner point will be placed in the middle of edge a_i, or, in another notation, in the middle of edge j, such that $v_{ij} = 1$. We write, for all $i = 1, \ldots, p-3,\ j = 1, \ldots, 2i+3,\ l = 1, \ldots, 2p-2$,

$$e_{ij1l} = e_{(i-1)j1l}, \tag{90}$$

$$-v_{ij} + e_{(i-1)j2l} \leq e_{ij2l} \leq e_{(i-1)j2l} + v_{ij}, \tag{91}$$

$$v_{ij} \leq e_{ij2(i+1+p)} \leq 2 - v_{ij}. \tag{92}$$

For the two new edges added at each iteration, we have, for all $i = 1, \ldots, p-3,\ l = 1, \ldots, 2p-2$,

$$e_{i(2i+2)1(i+3)} = 1, \tag{93}$$

$$e_{i(2i+2)2(i+1+p)} = 1, \tag{94}$$

$$e_{i(2i+3)1l} = \sum_{j=1}^{2i+1} v_{ij} \cdot e_{(i-1)j2l}, \tag{95}$$

$$e_{i(2i+3)2(i+1+p)} = 1. \tag{96}$$

We can linearize the binary product in equation (95) using the McCormick inequalities. Finally, we relate to the variables y_{ij}. For all $i = 1, \ldots, 2p-2,\ j = p+1, \ldots, 2p-2,\ i < j,\ k = 1, \ldots, 2p-3$,

$$y_{ij} \geq e_{(p-3)k1i} + e_{(p-3)k2j} - 1. \tag{97}$$

5 Computational results

Consider the instances of a regular octagon (8 points in $\mathbb{R}^2$) inscribed in a circle of radius one and a unit cube (8 points in $\mathbb{R}^3$), both described in Table 1 and depicted in Figure 4.

Octagon		Cube		
1	0.5	0	0	0
0.8535	0.8535	1	0	0
0.5	1	0	1	0
0.1464	0.8535	0	0	1
0	0.5	1	1	0
0.1464	0.1464	1	0	1
0.5	0	0	1	1
0.8535	0.1464	1	1	1

Table 1: Instances of an octagon and a cube.

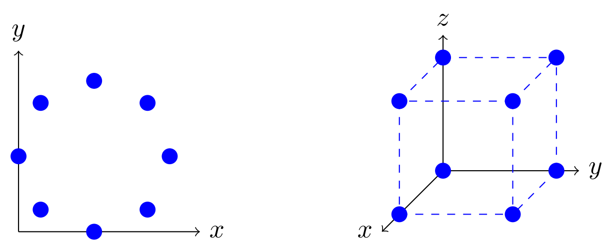


Figure 4: Image of the instances of an octagon and a cube.

To cope with the non-differentiability of the norm function in models (P1), (P2), (P3) and (P4), we have introduced a parameter $\lambda > 0$ as follows:

$$\|x\| \quad \approx \quad \sqrt{\lambda + \sum_{k=1}^{n} (x_k)^2} \tag{98}$$

In Table 2 we report results for the instance of the unit cube using model (P1), where the norm function is approximated by the function in (98). The second column f^* represents the best objective function value returned by the solver. There is a high sensitivity to the value of λ. In our subsequent tests, we have fixed λ =1e-10.

All experiments were performed on an Intel i5-4200U, CPU 1.6GHz and 6 GB of RAM, using AMPL Version 20200501 (Linux x86_64 (gcc 7.4.1)). As solver Xpress 8.8.0(35.01.01) can deal with second order cone constraints, it was used to solve models (P5) and (P6). But Xpress can't handle non-quadratic nonlinear constraints. For this reason, for models (P1)-(P4), we used solver Knitro 12.1.0 instead.

λ	f^*
1e-12	6.508687025
1e-11	6.477778467
1e-10	**6.228036322**
1e-9	6.243924003
1e-8	6.367537078
1e-7	6.477946866
1e-6	6.310132578
1e-5	6.670187305
1e-4	6.484672826
1e-3	6.312216035

Table 2: Different results depending on the value of λ.

In Table 3, we report the best value of the objective function returned by the respective solver and the total running time for the instance of the octagon for all models. Each model is tested alone and also with the addition of one of the three sets of valid constraints: geometric cuts (50); or constraints (51)-(82) to eliminate isomorphism in Steiner-Steiner subtree; or constraints (84)-(97) to eliminate isomorphism in Steiner tree. Models (P5) and (P6) obtain the optimal solution in all tested cases and all three sets of valid constraints indeed improved the model, reducing considerably the running time. Models (P1), (P2), (P3) and (P4) present an unpredictable behavior, due to nonlinearity in the four models and nonconvexity in models (P1) and (P4).

		(P1)	(P2)	(P3)	(P4)	(P5)	(P6)
	f^*	3.0913	3.1968	3.5339	2.68	2.6788	2.6788
	$t(s)$	0	0	10304	7723	761	596
using	f^*	2.6788	3.4516	3.5339	3.4484	2.6788	2.6788
(50)	$t(s)$	0	0	597	0	88	72
using	f^*	3.0653	2.6788	3.5714	3.1973	2.6788	2.6788
(51)-(82)	$t(s)$	38	2559	2458	140	48	43
using	f^*	3.5577	2.9407	3.5339	3.11175	2.6788	2.6788
(84)-(97)	$t(s)$	1006	5161	3403	1705	28	26

Table 3: Results for the instance of the octagon.

Constraints (50) can be used together with constraints (51)-(82) or together with constraints (84)-(97). We tested these combinations in models (P5) and (P6). The results are in Table 4.

We repeated the same tests with the instance of the unit cube. The results are reported in Table 5. No geometric cuts are added in this instance, as no pair of points in the cube satisfy the condition in (50). Once more, models (P5) and (P6) obtained the optimal solution and the two other sets of valid constraints improved the model. The behavior of models (P1), (P2), (P3) and (P4) is again unsatisfactory.

		(P5)	(P6)
using (50)	f^*	2.6788	2.6788
and (51)-(82)	$t(s)$	5	3
using (50)	f^*	2.6788	2.6788
and (84)-(97)	$t(s)$	4	4

Table 4: More results for the instance of the octagon.

		(P1)	(P2)	(P3)	(P4)	(P5)	(P6)
	f^*	6.228	6.5088	6.9891	6.4891	6.1961	6.1961
	$t(s)$	0	4668	5672	0	3388	3073
using	f^*	6.1961	6.5088	7.0009	250834774.3	6.1961	6.1961
(51)-(82)	$t(s)$	2	4019	5246	585	149	110
using	f^*	6.3102	6.5088	7.0009	6.9551	6.1961	6.1961
(84)-(97)	$t(s)$	724	6262	2666	8	85	73

Table 5: Results for the instance of the cube.

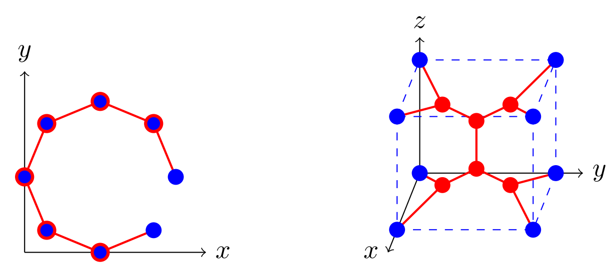


Figure 5: In the left, solution for the instance of the octagon with approximated value 2.6788. In the right, solution for the instance of the cube with approximated value 6.1961.

For other instances, we decided to report results only for model (P6) with constraints (50) and (84)-(97). We generated three instances, each one with 11 random terminals in the unit cube, and three more instances, each one with 12 random terminals in the unit cube. These instances were solved using AMPL Version 20200501 (Linux x86_64 (gcc 7.4.1)) and solver XPRESS 8.8.0(35.01.01). These experiments were performed on a SGI H2106 equipped with four AMD 6376 CPU 2.30GHz, with 16-cores each, and a total of 256 GB of RAM. The results are shown in Tables 6 and 7.

6 Conclusion

Models (P5) and (P6) brought more optimism to the ESTP solution for $n \geq 3$.

Acknowledgements
This work was partially supported by the National Council for Scientific and

	Instance 1	Instance 2	Instance 3
time(s)	639	1595	920
nodes	954405	2239626	1338733

Table 6: Instances where terminals are 11 random points in the unit cube.

	Instance 1	Instance 2	Instance 3
time(s)	2864	25390	32043
nodes	4059902	27410846	30527560

Table 7: Instances where terminals are 12 random points in the unit cube.

Technological Development - CNPq, under grant 302435/2019-0, COPPETEC Foundation, and by CAPES-MEC.

References

[1] T. Beyer and S. M. Hedetniemi. Constant time generation of rooted trees. *SIAM Journal on Computing*, 9(4):706–712, 1980.

[2] M. Brazil, R.L. Graham, D. A. Thomas, and M. Zachariasen. On the history of the Euclidean Steiner tree problem. *Archive for History of Exact Sciences*, 68:327–354, 2014.

[3] V. Costa, M.H.C. Fampa, and N. Maculan. Um modelo matemático para o problema euclidiano de Steiner em R^n. In *A Investigação Operacional em Portugal: Novos Desafios, Novas Ideias*, pages 145–158. Editora IST PRESS, Lisboa, Portugal, 2016.

[4] C. D'Ambrosio, M.H.C. Fampa, J. Lee, and S. Vigerske. Formulation of the Euclidean Steiner tree problem in n-space: missing proofs. *Optimization Letters*, 14:409–415, 2018.

[5] M.H.C. Fampa and K.M. Anstreicher. An improved algorithm for computing Steiner minimal trees in euclidean d-space. *Discrete Optimization*, 5(2):530–540, 2008.

[6] M.H.C. Fampa, J. Lee, and N. Maculan. An overview of exact algorithms for the Euclidean Steiner tree problem in n-space. *International Transactions in Operational Research (ITOR)*, 23(4):861–874, 2016.

[7] M.H.C. Fampa and N. Maculan. A new relaxation in conic form for the Euclidean Steiner problem in R^n. *RAIRO - Operations Research*, 35(4):283–394, 2001.

[8] M.H.C. Fampa and N. Maculan. Using a conic formulation for finding Steiner minimal trees. *Numerical Algorithms*, 35(4):315–330, 2004.

[9] M.H.C. Fampa, N. Maculan, V. Costa, R. Pinto, and B. Sartini. Numerical solutions of the Euclidean Steiner tree problem in n-space. *Report-Systems Engineering and Computer Sciences, COPPE, Federal University of Rio de Janeiro, Brazil*, 2015.

[10] N. Maculan, Ph. Michelon, and A.E. Xavier. The Euclidean Steiner tree problem in R^n : a mathematical programming formulation. *Annals of Operations Research*, 96:209–220, 2000.

[11] F. Montenegro, N. Maculan, G. Plateau, and P. Boucher. *Essays and Surveys in Metaheuristics*, chapter New Heuristics for the Euclidean Steiner Problem in $\mathbb{R}^n$, pages 509–524. Kluwer Academic Publishers, 2001.

[12] F. Montenegro, J.R.A. Torreão, and N. Maculan. Microcanonical optimization for the Euclidean Steiner problem in R^n with application to phylogenetic inference. *Physical Review E*, 68:056702–1– 056702–5, 2003.

[13] H Ouzia and N. Maculan. Mixed integer nonlinear optimization models for the Euclidean Steiner tree problem in R^d. Technical Report https://hal.archives-ouvertes.fr/hal-02293105, HAL archives-ouvertes, 2019.

[14] W. D. Smith. How to find steiner minimal trees in euclidean d-space. *Algorithmica*, 7:137–177, 1992.

[15] J.W. Van Laarhoven and K.M. Anstreicher. Geometric conditions for Euclidean Steiner trees in $\mathbb{R}^d$. *Computational Geometry*, 46:520–531, 2013.

[16] R. Wright, B. Richmond, A. Odlyzko, and B. McKay. Constant time generation of free trees. *SIAM Journal on Computing*, 15(2):540–548, 1986.

[17] G. Xue and Y. Ye. An efficient algorithm for minimizing a sum of euclidean norms with applications. *SIAM Journal on Optimization*, 4(7):1017–1036, 1997.