A PTAS for the Metric Case of the Minimum Sum-Requirement Communication Spanning Tree Problem

Santiago V. Ravelo and Carlos E. Ferreira

Instituto de Matemática e Estatística, Universidade de São Paulo, Brasil
{ravelo,cef}@ime.usp.br

Abstract. This work considers the metric case of the minimum sum-requirement communication spanning tree problem (SROCT), which is an NP-hard particular case of the minimum communication spanning tree problem (OCT). Given an undirected graph $G = (V, E)$ with non-negative lengths $\omega(e)$ associated to the edges satisfying the triangular inequality and non-negative routing weights $r(u)$ associated to nodes $u \in V$, the objective is to find a spanning tree T of G, that minimizes: $\frac{1}{2} \sum_{u \in V} \sum_{v \in V} (r(u) + r(v))\, d(T, u, v)$, where $d(H, x, y)$ is the minimum distance between nodes x and y in a graph $H \subseteq G$. We present a polynomial approximation scheme for the metric case of the SROCT improving the until now best existing approximation algorithm for this problem.

1 Introduction

In this work we consider a particular case of the minimum communication spanning tree problem (OCT). The OCT was introduced by Hu in 1974. In the problem it is given an undirected graph $G = (V, E)$ with non-negative length $\omega(e)$ associated to each edge $e \in E$ and non-negative requirement $\psi(u, v)$ between each pair of nodes $u, v \in V$. The problem is to find a spanning tree T of G which minimizes the total communication cost: $C(T) = \sum_{u \in V} \sum_{v \in V} \psi(u, v) d(T, u, v)$, where $d(H, x, y)$ denotes the minimum distance between nodes x and y in the sub-graph H of G. ([1,2])

In [3] it was proved that the minimum routing cost spanning tree problem (MRCT) is NP-hard (by a reduction from the 3-exact cover problem (3-EC)). Observe that MRCT is a particular case of OCT where the requirement between all pair of nodes is equal to one ($\psi(u, v) = 1$ for all $u, v \in V$). In [4] a PTAS for the MRCT was given. The authors presented a reduction from the general to the metric case, which implies that MRCT with edge-lengths that satisfy the triangular inequality is also NP-hard. Also, in [4] an $O(\log^2(n))$-approximation was given for OCT applying a result from [5] which was later improved to a $O(\log(n))$-approximation by [6].

In [7], the minimum product-requirement communication spanning tree problem (PROCT) and the minimum sum-requirement communication spanning tree problem (SROCT) were introduced. In these problems each vertex $u \in V$

S. Ganguly and R. Krishnamurti (Eds.): CALDAM 2015, LNCS 8959, pp. 9–20, 2015.

has a non-negative routing weight $r(u)$. For PROCT the requirement is defined as $\psi(u,v) = \frac{1}{2}r(u)r(v)$, and for SROCT $\psi(u,v) = \frac{1}{2}(r(u)+r(v))$. Both problems are NP-hard. In [7] a 1.577-approximation algorithm for PROCT and a 2-approximation for SROCT are presented.

The approximation ratio for PROCT was improved in [8] where a PTAS was given. A particular case of SROCT is the weighted p-MRCT, were given an integer p, only p nodes of the graph will have a positive routing weight (i.e. the remaining nodes have zero weight). The particular case in which the p nodes have routing weight 1 is called p-MRCT. In [9] it was proved that 2-MRCT is NP-hard. It also was proved in [10] where PTASs for 2-MRCT and the metric case of weighted 2-MRCT were given.

To the best of our knowledge, there are no results improving the 2-approximation ratio for SROCT which is also the best known ratio for the metric case of SROCT (denoted by m-SROCT). Observe that this problem is also NP-hard, since MRCT is a particular case in which $r(u) = 1$ for all $u \in V$.

In this work we give a PTAS for m-SROCT improving the best previous known result for this problem. The idea of our algorithm was inspired in the previous PTASs for related problems such as MRCT and PROCT. This paper is organized as follows. In the next section we present some notation. In section 3 we show how to obtain an optimal k-star for SROCT in polynomial time for a fixed integer k. In section 4 we present a PTASfor the m-SROCT. Finally, in section 5 the conclusions and future work are given.

2 Definitions

Unless specified we consider all graphs as undirected graphs. Given a graph G we denote the set of its nodes by V_G and the set of its edges by E_G (when G is implicit by context we use V as V_G and E as E_G).

Definition 1. *Given a graph G with non-negative lengths associated to its edges, the* **length of a path** *in G is defined as the sum of the lengths of its edges (a path with no edges has length zero). The* **distance** *between node x and node y in H sub-graph of G is the length of a path with minimum length between x and y in H and is denoted by $d(H,x,y)$.*

Now we can define SROCT as:

Problem 1. SROCT - Sum-Requirement Communication Spanning Tree problem

Input: A graph G, a non-negative length function over the edges of G, $\omega : E \to \mathbb{Q}_+$ and a non-negative routing weight function over the nodes of G, $r : V \to \mathbb{Q}_+$.

Output: A spanning tree T of G which minimizes the total weighted routing cost:

$$C(T) = \sum_{u\in V}\sum_{v\in V}\frac{1}{2}(r(u)+r(v))d(T,u,v) = \sum_{u\in V}\sum_{v\in V} r(u)d(T,u,v).$$

Definition 2. *Given a graph G and a non-negative routing weight function over the nodes of G, $r : V \to \mathbb{Q}_+$, we denote $r(G) = \sum_{u \in V_G} r(u)$ and $n(G) = |V_G|$. When G is implicit by the context we use R to denote $r(G)$ and n to denote $n(G)$.*

This paper considers the m-SROCT, the metric case of SROCT, which is the particular case of SROCT where the graph G is complete and the length function over the edges satisfies the triangular inequality. In order to approximate an optimal solution of m-SROCT we introduce the concept of a k-star[1]:

Definition 3. *Given a graph G and a positive integer k, a* **k-star** *of G is a spanning tree of G with no more than k internal nodes (that is, at least $n - k$ leaves). A* **core** *of a k-star T of G is a tree resulting by eliminating $n - k$ leaves from T.*

Note that a k-star T can be represented by (τ, S), where τ is a core of T and $S = \{S_{u_1}, ..., S_{u_k}\}$ is a vector indexed by the nodes in τ where S_{u_i} is the set of leaves adjacent in T to $u_i \in V_\tau$ $(1 \le i \le k)$.

The problem of finding an optimal k-star for m-SROCT can be defined as:

Problem 2. Optimum k-star for m-SROCT

Input: A positive integer k and an instance of m-SROCT: a complete graph G, a non-negative length function over the edges of G which satisfies the triangular inequality, $\omega : E \to \mathbb{Q}_+$ and a non-negative routing weight function over the nodes of G, $r : V \to \mathbb{Q}_+$.

Output: A k-star T of G which minimizes the total weighted routing cost: $C(T) = \sum_{u \in V} \sum_{v \in V} r(u) d(T, u, v)$.

The next section shows an efficient algorithm to find an optimal k-star.

3 Optimal k-Star for m-SROCT

First we introduce the notion of configuration of a k-star:

Definition 4. *Given a k-star $T = (\tau, S)$ a* **configuration** *of T is (τ, L) where $L = \{l_{u_1}, ..., l_{u_k}\}$ is a vector of integers being $l_{u_i} = |S_{u_i}|$ $(1 \le i \le k)$. A configuration (τ, L)* **is over** *(k, G), where k is a positive integer and G is a graph, if τ is a tree of G with k nodes (that is, $\tau \subseteq G$ and $|V_\tau| = k$) and $\sum_{u \in V_\tau} l_u = n - k$.*

In [4] it was observed that given a complete graph G and a fixed positive integer k, the number of configurations over (k, G) is polynomial in n, resulting $O(k^k n^{2k-1})$. Then, given an instance $\langle G, \omega, r, k \rangle$ of the optimum k-star for m-SROCT, our proposal is to enumerate all possible configurations over (K, G),

[1] The definition of k-star used in this paper is the same used by [4,7,8], which is different from the usual definition of k-star in graph theory (a tree with k leaves linked to a single vertex of degree k).

finding an optimal k-star of each configuration, and finally select the best k-star among them.

We find an optimal k-star for an instance $\langle G, \omega, r, k\rangle$ of the optimum k-star for m-SROCT and a configuration (τ, L) over (k, G), reducing the problem to an uncapacitated minimum cost flow problem (UMCF).

Problem 3. UMCF - Uncapacitated Minimum Cost Flow problem

Input: A directed graph G, a cost function over the arcs $\omega : E \to \mathbb{Q}_+$ and a demand function over the nodes $r : V \to \mathbb{Z}$.

Output: An integer vector indexed by the arcs $X = (x_e)_{e\in E}$ which minimizes $C(X) = \sum_{e\in E} \omega(e)x_e$ and guaranties for each node $u \in V$:

$\sum_{e\in\delta^+(u)} x_e - \sum_{e\in\delta^-(u)} x_e = r(u)$,

where $e \in \delta^+(w)$ and $e \in \delta^-(v)$ iff $e = \langle v, w\rangle$ $(\forall e \in E, v, w \in V)$.

Proposition 1. *Given an instance $I = \langle G, \omega, r, k\rangle$ of the optimum k-star for m-*SROCT *and a configuration $c = (\tau, L)$ over (k, G), the problem of finding an optimal k-star with configuration c for I can be reduced in polynomial time to the* UMCF *with instance $I' = \langle G', \omega', r'\rangle$, where:*

- $V_{G'} = V_G$;
- $E_{G'} = \{(u, v)|u \in V_{G-\tau} \wedge v \in \tau\}$;
- $\omega'(u, v) = R\omega(u, v) + \sum_{w\in V_\tau} r(u)\left(d(\tau, v, w) + \omega(u, v)\right)(l_w+1) - 2r(u)\omega(u, v)$;
- *if* $u \in V_{G-\tau}$ *then* $r'(u) = -1$, *otherwise* $r'(u) = l_u$.

The graph G' is a complete bipartite graph on the same node set V_G of G. The bi-partition is given by the nodes in τ and outside this set. The cost of arc $\langle u, v\rangle$ is equivalent to the value of assigning u as adjacent of v in a k-star with the given configuration. We have to consider the cost of sending the routing weight from u to all nodes of τ assuming that each node $w \in V_\tau$ receives $(l_w + 1)$ times the value $r(u)$ (considering the transmission to the node w and the leaves adjacent to it); also, we add the cost of sending the routing weight of the entire graph $(R - r(u))$ to node u, which must pass by node v. Finally the demands r' are set to ensure assignment between nodes out of τ and nodes in τ.

Proof. Since demands are integer we know that in any feasible solution the values x_e will be either zero or one. Moreover, exactly $n - k$ arcs of G' will have value 1. This guaranties that every feasible solution S' of the flow problem represents an assignment of leaves outside τ to be adjacent to nodes in τ for a k-star T of G with configuration (τ, L). Also, it is easy to see that any k-star T with configuration (τ, L) provides a feasible solution to the flow problem: connect node $u \in \tau$ to the l_u leaves adjacent to it in T.

Observe that[2]: $C(S') = C(T) - \sum_{u\in V_\tau}\sum_{v\in V_\tau} r(v)d(\tau, u, v)(l_u + 1)$, where $\sum_{u\in V_\tau}\sum_{v\in V_\tau} r(v)d(\tau, u, v)(l_u + 1)$ is the same for every solution with the same configuration. Then, an optimum of UMCF with instance I' is associated to an optimal k-star with configuration c of m-SROCT with instance I.

[2] A detailed proof of this fact can be found in the full version of the paper.

In order to obtain I' from I the cost of each arc in G' must be calculated. This can be done in $O((n-k)k^3)$. Defining the demands and the graph G' itself can be done in $O((n-k)k+n)$. Finally, obtaining the k-star T associated to a solution S' can be done in $O(n-k)$, while the complexity of calculating $C(T)$ would be $O(k^3)$. So, the reduction above can be done in $O(nk^3)$.

□

It is well known that UMCF can be solved in $O(n\log(n)(nk+n\log(n))) = O(n^2\log^2(n))$ (e.g. [11]). Then, finding an optimal k-star for m-SROCT with fixed k can be done efficiently.

Lemma 1. *The optimum k-star for m-SROCT with fixed k can be solve in $O(n^{2k+1}\log^2(n))$.*

4 PTAS for m-SROCT

In this section we prove that for $0 < \delta \leq \frac{1}{2}$ there exists a k-star, with k depending on δ, which is a $\frac{1}{1-\delta}$-approximation of m-SROCT. For that, from now on, we will consider an instance I of m-SROCT. Remember that $n = n(G)$ and $R = r(G)$.

The idea of the proof is similar to those presented in [4], [7] and [8]. Given $0 < \delta \leq \frac{1}{2}$ and a spanning tree T of G, we show the existence of a set Y of internally disjoint paths whose union results in a sub-tree S of T, such that the communication cost of each component $B \in T - S$ is at most a small fraction of the communication cost of T, which implies that most of the communication cost of T passes by S. Also, we prove that the size of Y is limited by a function of δ and we show how to construct a k-star from Y, where the value of k depends on the size of Y. The communication cost of the k-star approximates the communication cost of T by a factor of $\frac{1}{1-\delta}$.

4.1 Notation

First, in order to present the results of this section, we need some notation, which generalizes the notation given in [4], [7] and [8]:

Definition 5. *Given a spanning tree T of G, a set of edges H of T and a node u of T, $VB(T,H,u)$ is the set of nodes in the component of $T-H$ containing the vertex u.*

Definition 6. *Given a spanning tree T of G, a path $P = u_1,...,u_h$ of T, we denote by f_P (or f, when P is clear by the context) the first node of P and l_P (or l) the last node. We will use η to denote number of nodes and ρ to denote communication requirements. We define:*

For the number of nodes (figure 1 gives an example of these notation):

- $\eta_P^f = |VB(T,E_P,f)|$ *and* $\eta_P^l = |VB(T,E_P,l)|$, *are the number of nodes in the component of $T-E_P$ containing the first node of P and the component containing the last node of P,*

– $\eta_P^m = n - \eta_P^f - \eta_P^l$, *is the number of nodes of all the component of* $T - E_P$ *containing an internal node of* P,
– $N(P, v) = \sum_{i=2}^{h-1} |VB(T, E_P, u_i)|\, d(P, v, u_i)$, *is the sum over each internal node* u_i *of* P *of the number of nodes in the component of* $T - E_P$ *containing* u_i *times the distance in* P *from* u_i *to node* v,
– $N_f(P) = N(P, f)$, $N_l(P) = N(P, l)$, *represent the sums over each internal node* u_i *of* P *of the number of nodes in the component of* $T - E_P$ *containing* u_i *times the distance in* P *from* u_i *to the first node of* P *and to the last node of* P,
– $\eta_P^s = \max\{\eta_P^f, \eta_P^l\}$, $\eta_P^i = \min\{\eta_P^f, \eta_P^l\}$, *represents the number of nodes in the greater component of* $T - E_P$ *containing an extremal node of* P,
– *if* $\eta_P^f = \eta_P^s$ *then* $N(P) = N_f(P)$, *else* $N(P) = N_l(P)$.

Analogously, for the communication requirements:

– $\rho_P^f = r\,(VB(T, E_P, f))$, $\rho_P^l = r\,(VB(T, E_P, l)$, $\rho_P^m = R - \rho_P^f - \rho_P^l$,
– $R(P, v) = \sum_{i=2}^{h-1} r\,(VB(T, E_P, u_i))\, d(P, v, u_i)$, $R_f(P) = R(P, f)$, $R_l(P) = R(P, l)$,
– $\rho_P^s = \max\{\rho_P^f, \rho_P^l\}$, $\rho_P^i = \min\{\rho_P^f, \rho_P^l\}$,
– *if* $\rho_P^f = \rho_P^s$ *then* $R(P) = R_f(P)$, *else* $R(P) = R_l(P)$.

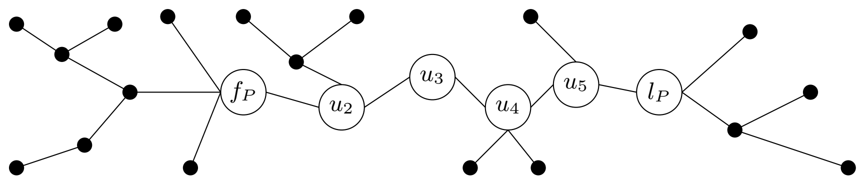

Fig. 1. Consider the above spanning tree T of a graph G, where all the edges have unitary weights and P is the path of T from node f_P to l_P. Observe that $VB(T, E_P, f_P)$ is the set of nodes to the left of f_P (including f_P), $VB(T, E_P, l_P)$ is the set of nodes to the right of l_P (including l_P), $VB(T, E_P, u_2)$ is the set of nodes containing u_2 and the three nodes above it, $VB(T, E_P, u_3) = \{u_3\}$, $VB(T, E_P, u_4)$ is the set of nodes containing u_4 and the two nodes below it, and $VB(T, E_P, u_5)$ is the set of nodes containing u_5 and the node above it. Then, $\eta_P^f = 9$, $\eta_P^l = 5$, $\eta_P^m = 10$, and thus $\eta_P^s = 9$ and $\eta_P^l = 5$. Also, $N_f(P) = |VB(T, E_P, u_2)| \times 1 + |VB(T, E_P, u_3)| \times 2 + |VB(T, E_P, u_4)| \times 3 + |VB(T, E_P, u_5)| \times 4 = 4 \times 1 + 1 \times 2 + 3 \times 3 + 2 \times 4 = 23$ and $N_l(P) = |VB(T, E_P, u_2)| \times 4 + |VB(T, E_P, u_3)| \times 3 + |VB(T, E_P, u_4)| \times 2 + |VB(T, E_P, u_5)| \times 1 = 4 \times 4 + 1 \times 3 + 3 \times 2 + 2 \times 1 = 27$, which yields $N(p) = 23$.

Now we introduce definitions for *separators*. A δ-separator is a sub-tree of a spanning tree T of G, whose deletion gives rise to components that are bounded (in the number of nodes, routing weight or both) by a factor δ of the total value (n or R). Formally:

Definition 7. *Given* $0 < \delta \le \frac{1}{2}$ *and a spanning tree* T *of* G*, a sub-tree* S *of* T *is a* **δ-η-separator** *of* T *if every component* B *of* $T - S$*, satisfies* $n(B) \le \delta n$*. If every component* B *of* $T - S$*, satisfies* $r(B) \le \delta R$*,* S *is a* δ-ρ*-separator of* T*. If both conditions apply,* S *is a* **δ-$\eta\rho$-separator** *of* T*.*

Also, we define δ-$\eta\rho$-path and δ-$\eta\rho$-spine:

Definition 8. *Given* $0 < \delta \le \frac{1}{2}$ *and a spanning tree* T *of* G*, a path* P *of* T *is a* **δ-$\eta\rho$-path** *of* T *if* $\eta_P^m \le \delta\frac{n}{6}$ *and* $\rho_P^m \le \delta\frac{R}{6}$*.*

Definition 9. *Given* $0 < \delta \le \frac{1}{2}$ *and a spanning tree* T *of* G*, a set* $Y = \{P_1, P_2, ..., P_l\}$ *of* δ-$\eta\rho$*-paths internally-disjoint of* T *is a* δ-$\eta\rho$*-spine, if* $S = \bigcup_{i=1}^{l} P_i$ *is a minimal* δ-$\eta\rho$*-separator of* T*.* $ext(Y)$ *denotes the endpoints set of all paths in* Y*.*

4.2 Approximation Lemma

Using those definitions we prove that for any $0 < \delta \le \frac{1}{2}$, any spanning tree T of G and any δ-$\eta\rho$-spine Y of T, there exists a $|ext(Y)|$-star with communication cost bounded by $\frac{1}{1-\delta}C(T)$. This lemma, together with lemma 4 are the basis of the main result of this work.

Lemma 2. *Given* $0 < \delta \le \frac{1}{2}$*, a spanning tree* T *of* G *and a* δ-$\eta\rho$*-spine* Y *of* T*, there exists a* $|ext(Y)|$*-star* X *of* G *satisfying* $C(X) \le \frac{1}{1-\delta}C(T)$*.*

In order to conclude that result, first we use the following three propositions[3].

Proposition 2. *Given* $0 < \delta \le \frac{1}{2}$ *a* δ-$\eta\rho$*-path* P *of a* δ-$\eta\rho$*-spine of a spanning tree* T *of* G*, then:*

$$
\begin{aligned}
&(R + n - \eta_P^m - \rho_P^m)\left(\eta_P^f \rho_P^l + \eta_P^l \rho_P^f\right)\omega(P) \\
&+ \left(\eta_P^l + \rho_P^l\right)\left(\omega(P)\left(\eta_P^m \rho_P^f + \eta_P^f \rho_P^m\right) + RN_l(P) + nR_l(P)\right) \\
&+ \left(\eta_P^f + \rho_P^f\right)\left(\omega(P)\left(\eta_P^m \rho_P^l + \eta_P^l \rho_P^m\right) + RN_f(P) + nR_f(P)\right) \\
\le\ & \frac{6 + 5\delta}{6}(R + n)\left(\eta_P^f \rho_P^l + \eta_P^l \rho_P^f + \eta_P^i \rho_P^m + \eta_P^m \rho_P^i\right)\omega(P) \\
&+ (R + n)\left(\left(\eta_P^s - \eta_P^i\right)R(P) + \left(\rho_P^s - \rho_P^i\right)N(P)\right).
\end{aligned}
$$

Proposition 3. *Given* $0 < \delta \le \frac{1}{2}$*, a spanning tree* T *of* G *and a* δ-$\eta\rho$*-spine* Y *of* T*, there exists a* $|ext(Y)|$*-star* X *of* G *which satisfies:*

$$
\begin{aligned}
C(X) \le & \sum_{P \in Y}\left(\eta_P^f \rho_P^l + \rho_P^f \eta_P^l\right)\omega(P) + \min\left\{\Delta_{fl}(P), \Delta_{lf}(P)\right\} \\
& + R\sum_{u \in V_G} d(T, u, S) + (n - 2)\sum_{u \in V_G} r(u) d(T, u, S).
\end{aligned}
$$

[3] The proofs of the propositions are intricate and do not bring any new insight into the problem, also they can be found in the full version of the paper.

Where $\Delta_{wz}(P) = \omega(P)\left(\eta_P^m \rho_P^w + \rho_P^m \eta_P^w\right) + RN_z(P) + (n-2)R_z(P)$, $w, z \in \{f, l\}$, *and* $S = \bigcup_{P \in Y} P$.

Observe that the result above shows the existence of a $|ext(Y)|$-star X of G (which comes from a δ-$\eta\rho$-spine Y of a spanning tree of G) and also gives us an upper bound for the associated communication cost of X. The next proposition gives us a lower bound for the communication cost of a spanning tree of G.

Proposition 4. *Given* $0 < \delta \leq \frac{1}{2}$, *a spanning tree* T *of* G *and a* δ-$\eta\rho$-*spine* Y *of* T, *then:*

$$\begin{aligned} C(T) \geq & \sum_{P \in Y} \left(\rho_P^l \eta_P^f + \rho_P^f \eta_P^l + \rho_P^i \eta_P^m + \eta_P^i \rho_P^m\right) \omega(P) \\ & + \sum_{P \in Y} \left(\eta_P^s - \eta_P^i\right) R(P) + \left(\rho_P^s - \rho_P^i\right) N(P) \\ & + (1-\delta)\left(n \sum_{u \in V_G} r(u) d(T, u, S) + R \sum_{u \in V_G} d(T, u, S)\right). \end{aligned}$$

Where $S = \bigcup_{P \in Y} P$.

Now we demonstrate the lemma 2, which states that if we are given a δ-$\eta\rho$-spine Y of a spanning tree T, then we can construct a star whose core lies on $ext(Y)$, such that its communication cost is bounded by $\frac{1}{1-\delta} C(T)$.

Proof. Let $S = \bigcup_{P \in Y} P$ be the minimal δ-$\eta\rho$-separator associated with Y and X the $|ext(Y)|$-star of G given by proposition 3, then:

$$\begin{aligned} C(X) \leq & \sum_{P \in Y} \left(\eta_P^f \rho_P^l + \rho_P^f \eta_P^l\right) \omega(P) + \min\left\{\Delta_{fl}(P), \Delta_{lf}(P)\right\} \\ & + R \sum_{u \in V_G} d(T, u, S) + (n-2) \sum_{u \in V_G} r(u) d(T, u, S) \\ \leq & \sum_{P \in Y} \frac{R + n - \rho_P^m - \eta_P^m}{R + n - \rho_P^m - \eta_P^m} \left(\eta_P^f \rho_P^l + \rho_P^f \eta_P^l\right) \omega(P) + \min\left\{\Delta_{fl}(P), \Delta_{lf}(P)\right\} \\ & + R \sum_{u \in V_G} d(T, u, S) + n \sum_{u \in V_G} r(u) d(T, u, S). \end{aligned}$$

Since the minimum between two numbers is less than or equal to their weighted median, we have:

$$\begin{aligned} C(X) \leq & \sum_{P \in Y} \frac{R + n - \rho_P^m - \eta_P^m}{R + n - \rho_P^m - \eta_P^m} \left(\eta_P^f \rho_P^l + \rho_P^f \eta_P^l\right) \omega(P) \\ & + \sum_{P \in Y} \frac{\eta_P^l + \rho_P^l}{\eta_P^l + \rho_P^l + \eta_P^f + \rho_P^f} \Delta_{fl}(P) + \frac{\eta_P^f + \rho_P^f}{\eta_P^l + \rho_P^l + \eta_P^f + \rho_P^f} \Delta_{lf}(P) \\ & + R \sum_{u \in V_G} d(T, u, S) + n \sum_{u \in V_G} r(u) d(T, u, S). \end{aligned}$$

Since every path P satisfies $R+n-\eta_P^m-\rho_P^m=\eta_P^f+\eta_P^l+\rho_P^f+\rho_P^l$, and every $P\in Y$ is a δ-$\eta\rho$-path, then by applying the result of Proposition 2 we conclude:

$$\begin{aligned}C(X)\leq&\sum_{P\in Y}\frac{\left(\frac{6+5\delta}{6}\right)(R+n)}{R+n-\rho_P^m-\eta_P^m}\left(\eta_P^f\rho_P^l+\rho_P^f\eta_P^l+\eta_P^i\rho_P^m+\rho_P^i\eta_P^m\right)\omega(P)\\&+\sum_{P\in Y}\frac{n+R}{R+n-\rho_P^m-\eta_P^m}\left(\left(\eta_P^s-\eta_P^i\right)R(P)+\left(\rho_P^s-\rho_P^i\right)N(P)\right)\\&+R\sum_{u\in V_G}d(T,u,S)+n\sum_{u\in V_G}r(u)d(T,u,S).\end{aligned}$$

Notice that any δ-$\eta\rho$-path satisfies: $\rho_P^m+\eta_P^m\leq\frac{\delta}{6}R+\frac{\delta}{6}n=\frac{\delta}{6}(n+R)$, then:

$$\begin{aligned}C(X)\leq&\sum_{P\in Y}\frac{\left(\frac{6+5\delta}{6}\right)(R+n)}{R+n-\frac{\delta}{6}(R+n)}\left(\eta_P^f\rho_P^l+\rho_P^f\eta_P^l+\eta_P^i\rho_P^m+\rho_P^i\eta_P^m\right)\omega(P)\\&+\sum_{P\in Y}\frac{n+R}{R+n-\frac{\delta}{6}(R+n)}\left(\left(\eta_P^s-\eta_P^i\right)R(P)+\left(\rho_P^s-\rho_P^i\right)N(P)\right)\\&+R\sum_{u\in V_G}d(T,u,S)+n\sum_{u\in V_G}r(u)d(T,u,S)\\=&\frac{6+5\delta}{6-\delta}\sum_{P\in Y}\left(\eta_P^f\rho_P^l+\rho_P^f\eta_P^l+\eta_P^i\rho_P^m+\rho_P^i\eta_P^m\right)\omega(P)\\&+\frac{6}{6-\delta}\sum_{P\in Y}\left(\left(\eta_P^s-\eta_P^i\right)R(P)+\left(\rho_P^s-\rho_P^i\right)N(P)\right)\\&+\frac{1}{1-\delta}(1-\delta)\left(R\sum_{u\in V_G}d(T,u,S)+n\sum_{u\in V_G}r(u)d(T,u,S)\right).\end{aligned}$$

Since $\frac{6+5\delta}{6-\delta}=\frac{(6+5\delta)(1-\delta)}{(6-\delta)(1-\delta)}=\frac{6-\delta-5\delta^2}{(6-\delta)(1-\delta)}<\frac{6-\delta}{(6-\delta)(1-\delta)}=\frac{1}{1-\delta}$ and $\frac{6}{6-\delta}<\frac{6+5\delta}{6-\delta}<\frac{1}{1-\delta}$, by applying Proposition 4 we obtain: $C(X)\leq\frac{1}{1-\delta}C(T)$.

□

4.3 Existence of Bounded δ-$\eta\rho$-Spine

In the next lemma we show that there exists a δ-$\eta\rho$-spine Y of T whose $|ext(Y)|$ is bounded by a function of δ.

Lemma 3. *Given $0<\delta\leq\frac{1}{2}$ and a spanning tree T of G, there exists a δ-$\eta\rho$-spine Y of T satisfying $|ext(Y)|\leq 3\left(\left\lceil\frac{6}{\delta}\right\rceil^2-11\left\lceil\frac{6}{\delta}\right\rceil+1\right)$.*

Proof. Consider a minimal δ-ρ-separator S_ρ of T and a minimal δ-η-separator S_η of T. If S_ρ and S_η have at least one node in common, then define $S'=S_\rho\cup S_\eta$ and obviously S' is a δ-$\eta\rho$-separator. If S_ρ and S_η have no nodes in common, then since both are trees, S_η must be included in a component of $T-S_\rho$. But, S_ρ

is a δ-ρ-separator of T, then every component of $T - S_\rho$ has weight bounded by δR, so the path P in T connecting S_ρ to S_η satisfies $\rho_P^m < \delta R$. Analogously, P also satisfies $\eta_P^m < \delta n$. Then, P can be divided into 6 paths each one with weight bounded by $\frac{\delta R}{6}$ and another 6 paths each one with number of nodes bounded by $\frac{\delta n}{6}$. Since each division uses 5 internal nodes, in the worst case, using 10 internal nodes we obtain a division of P in δ-$\eta\rho$-paths, and $S' = S_\rho \cup S_\eta \cup P$ is a δ-$\eta\rho$-separator.

By modifying[4] a proof of [4], [7] and [8], we prove that there exists an Y'_ρ and an Y'_η sets of internally-disjoint δ-$\eta\rho$-paths, satisfying $\cup_{P\in Y'_\rho} P = S_\rho$, $\cup_{P\in Y'_\eta} P = S_\eta$ and $\left|ext(Y'_\rho)\right|, \left|ext(Y'_\eta)\right| \le \left\lceil \frac{6}{\delta} \right\rceil^2 - 11 \left\lceil \frac{6}{\delta} \right\rceil + 1$.

For each path $P \in Y'_\rho$ if P contains internal nodes in $ext(Y'_\eta)$, divide P on those nodes to create new internally-disjoint δ-$\eta\rho$-paths and put those new paths in Y_ρ, otherwise add P to Y_ρ. Analogously, define Y_η from Y'_η and $ext(Y'_\rho)$. Observe that no path of $Y_\eta \cup Y_\rho$ has an internal node in $ext(Y_\eta) \cup ext(Y_\rho)$, also Y_η and Y_ρ are sets of internally-disjoint δ-$\eta\rho$-paths such that $\cup_{P\in Y_\rho} P = S_\rho$, $\cup_{P\in Y_\eta} P = S_\eta$ and:

$$|ext(Y_\rho \cup Y_\eta)| \le 2\left(\left\lceil \frac{6}{\delta} \right\rceil^2 - 11 \left\lceil \frac{6}{\delta} \right\rceil + 1\right).$$

Notice that, since $S_\eta \cup S_\rho$ is acyclic, each path of Y_ρ internally-intersects at most one path in Y_η and vice-verse.

If there are two paths $P_\eta \in Y_\eta$ and $P_\rho \in Y_\rho$ whose internal-intersection is not empty and their end-points do not belong to their intersection, then no other path of Y_η intersects any path of Y_ρ, and by removing from P_η the internal nodes of the intersection we add at most two new extremal points (the end-points of the intersection). Then $Y' = (Y_\eta - P_\eta) \cup Y_\rho \cup (P_\eta - (P_\eta \cap P_\rho))$ is a set of internally-disjoint δ-$\eta\rho$-paths which satisfies $\cup_{P\in Y'} P = S'$ and:

$$|ext(Y')| \le 2\left(\left\lceil \frac{6}{\delta} \right\rceil^2 - 11 \left\lceil \frac{6}{\delta} \right\rceil + 1\right) + 2.$$

Otherwise, if no path of Y_η intersects any path of Y_ρ then, as seen before, there exists a path P connecting S_η to S_ρ that can be divided in at most 11 δ-$\eta\rho$-paths, and the union of those paths with Y_η and Y_ρ results in a set Y' of internally-disjoint δ-$\eta\rho$-paths such that $\cup_{P\in Y'} P = S'$ and:

$$|ext(Y')| \le 2\left(\left\lceil \frac{6}{\delta} \right\rceil^2 - 11 \left\lceil \frac{6}{\delta} \right\rceil + 1\right) + 10.$$

The last possibility is that at least one path of Y_η internally-intersects a path of Y_ρ and each not-empty intersection between a path of Y_η and a path of Y_ρ contains at least one endpoint. Then, remove from each path in Y_η the internal nodes of the intersection with each path in Y_ρ (notice that a path of Y_η at most internally-intersects one path in Y_ρ). In this case the number of new extremal

[4] Such modification can be found in the full version of the paper.

points will be at most $|Y'_\eta|$ and the set Y' defined by the union of Y_ρ with the modified Y_η is a set of internally-disjoint δ-$\eta\rho$-paths that satisfies $\cup_{P\in Y'} P = S'$ and:

$$|ext(Y')| \le 3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right).$$

Since, for $0 < \delta \le \frac{1}{2}$: $\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right) \ge 12^2 - 11(12) + 1 = 13 > 10$, then we always can obtain a set Y' of internally-disjoint δ-$\eta\rho$-paths which satisfies $\cup_{P\in Y'} P = S'$ and:

$$|ext(Y')| \le 3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right).$$

If S' is a minimal δ-$\eta\rho$-separator, then $Y = Y'$ is a δ-$\eta\rho$-spine. Otherwise, exists a minimal δ-$\eta\rho$-separator $S \subset S'$ and by deleting from each path in Y' the elements that are not contained in S we obtain a δ-$\eta\rho$-spine Y of T satisfying:

$$|ext(Y)| \le 3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right).$$

□

4.4 PTAS

Using lemmata 2 and 3 we can state the following proposition:

Proposition 5. *Given $0 < \delta \le \frac{1}{2}$ and a spanning tree T of G, there exists a $\left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$-star X of G, such that $C(X) \le \frac{1}{1-\delta}C(T)$.*

Let T^* be an optimal spanning tree for m-SROCT over G, by proposition 5 for any $0 < \delta \le \frac{1}{2}$, there exists a $\left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$-star X of G such that $C(X) \le \frac{1}{1-\delta}C(T^*)$. Since an optimal $\left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$-star X^* of G guarantees $C(X^*) \le C(X)$, then $C(X^*) \le \frac{1}{1-\delta}C(T^*)$.

Lemma 4. *Given $0 < \delta \le \frac{1}{2}$ an optimal $\left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$-star of G is a $\frac{1}{1-\delta}$-approximation for m-SROCT.*

The results of lemmata 1 and 4 complete the necessary tools for providing the PTAS:

Theorem 1. *There exists a PTAS for m-SROCT, such that a $\left(1 + \frac{\delta}{1-\delta}\right)$-approximation can be found in $O\left(n^{6\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)+1}\log^2(n)\right)$ time complexity where $0 < \delta \le \frac{1}{2}$.*

5 Conclusions

In this work we present a PTAS for m-SROCT, a NP-hard particular case of OCT. The best previously known result for this problem was a 2-approximation algorithm due to [7]. Many questions remain open regarding OCT and related problems. One could improve the approximation ratio for SROCT or other particular case of OCT. In future works we will attempt to answer this question for some of these problems.

Acknowledgments. This research is supported by the following projects: FAPESP 2013/03447-6, CNPq 477203/2012-4 and CNPq 302736/2010-7.

References

1. Hu, T.C.: Optimum communication spanning trees. SIAM J. Comput. 3(3), 188–195 (1974)
2. Wu, B.Y., Chao, K.M.: Spanning Trees and Optimization Problems. Chapman & Hall / CRC (2004) ISBN: 1584884363
3. Johnson, D.S., Lenstra, J.K., Rinnooy Kan, A.H.G.: The complexity of the network design problem. Networks 8, 279–285 (1978)
4. Wu, B.Y., Lancia, G., Bafna, V., Chao, K.M., Ravi, R., Tang, C.Y.: A polynomial time approximation scheme for minimum routing cost spanning trees. SIAM J. on Computing 29(3), 761–778 (2000)
5. Bartal, Y.: Probabilistic approximation of metric spaces and its algorithmic applications. In: Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science, pp. 184–1963 (1996)
6. Talwar, K., Fakcharoenphol, J., Rao, S.: A tight bound on approximating arbitrary metrics by tree metrics. In: Proceedings of the 35th Annual ACM Symposium on Theory of Computing, pp. 448–455 (2003)
7. Wu, B.Y., Chao, K.M., Tang, C.Y.: Approximation algorithms for some optimum communication spanning tree problems. Discrete and Applied Mathematics 102, 245–266 (2000)
8. Wu, B.Y., Chao, K.M., Tang, C.Y.: A polynomial time approximation scheme for optimal product-requirement communication spanning trees. J. Algorithms 36, 182–204 (2000)
9. Farley, A.M., Fragopoulou, P., Krumme, D., Proskurowski, A., Richards, D.: Multi-source spanning tree problems. Journal of Interconnection Networks 1(1), 61–71 (2000)
10. Wu, B.Y.: A polynomial time approximation scheme for the two-source minimum routing cost spanning trees. J. Algorithms 44, 359–378 (2002)
11. Orlin, B.J.: A faster strongly polynomial minimum cost flow algorithm. Operations Research 41(2), 338–350 (1993)