Theoretical Computer Science 771 (2019) 9–22

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

A PTAS for the metric case of the optimum weighted source–destination communication spanning tree problem

Santiago Valdés Ravelo [a,*], Carlos Eduardo Ferreira [b]

[a] *Center of Computer Sciences, Federal University of Rio Grande, Brazil*
[b] *Institute of Mathematics and Statistics, University of Sao Paulo, Brazil*

ARTICLE INFO

Article history:
Received 15 December 2017
Received in revised form 12 November 2018
Accepted 12 November 2018
Available online 14 November 2018
Communicated by R. Klasing

Keywords:
Optimum communication spanning tree problem
Approximation algorithms
Polynomial-time approximation scheme
Metric problem

ABSTRACT

This work considers the optimum weighted source–destination communication spanning tree problem (WSDOCT), which is an NP-hard special case of the optimum communication spanning tree problem (OCT). Given an undirected graph $G = (V, E)$ with non-negative lengths $\omega(e)$ associated to the edges and non-negative routing weights $\sigma(u)$ and $\rho(u)$ respectively to sending and receiving communication of nodes $u \in V$, the objective is to find a spanning tree T of G, that minimizes:

$$\sum_{u,v\in V} (\sigma(u)\rho(v) + \rho(u)\sigma(v))\, d(T, u, v),$$

where $d(H, x, y)$ is the minimum distance between nodes x and y in a graph $H \subseteq G$. We present a polynomial time approximation scheme for the metric case of the WSDOCT. This result improves the until now best existing approximation algorithm for this problem. Also, we give a 2-approximation for the problem which improves the time complexity required by our PTAS to achieve this approximation ratio.

1. Introduction

Suppose we need to connect n servers in order to minimize the latency of the information delivery between them. If we know an estimate for the latency of delivering an information unit between each pair of servers and the amount of information units they share, then we are dealing with the optimum communication spanning tree problem (OCT), introduced by Hu in [6]. The OCT receives a graph $G = (V, E)$ with non-negative lengths associated to each edge and a non-negative communication requirement function $\psi(u, v)$ over all pair of nodes $u, v \in V$. The objective is to find a spanning tree T of G that minimizes the total communication cost. This cost is calculated by summing over all pair of nodes the communication requirement multiplied by the distance in the tree between them. Other interesting problems of the literature consider routing costs in graphs, such as, [2,3,8,13].

In this work we study the optimum weighted source–destination communication spanning tree problem (WSDOCT), a particular case of OCT introduced in [10]. The main applications of the WSDOCT lay in the area of network design problems: computer networks, telecommunication networks, transportation networks or supply networks. This problem considers the class of product-requirements incorporating directedness into the requirement model, in the sense that each

* Corresponding author.
E-mail addresses: ravelo@furg.br (S.V. Ravelo), cef@ime.usp.br (C.E. Ferreira).

https://doi.org/10.1016/j.tcs.2018.11.008
0304-3975/

node $u \in V$ has non-negative sending and receiving requirements $\sigma(u)$ and $\rho(u)$, and the requirement function is defined as $\psi(u,v) = \frac{1}{2}(\sigma(u)\rho(v) + \rho(u)\sigma(v))$.

The WSDOCT is NP-hard even when all the requirements are unitary (i.e. $\sigma(u) = \rho(u) = 1$ for all $u \in V$) [7]. This special case is known as the minimum routing cost spanning tree problem (MRCT), and in [17] the authors presented a reduction from the general to the metric case which implies that even with edge-length satisfying the triangular inequality the problem remains NP-hard. Also, in [17], were given a PTAS for the MRCT and using results of [1] a $O(\log^2 n)$-approximation for OCT. Later, results of [12] allowed to improve to $O(\log n)$ the approximation ratio for OCT which still is the best known ratio for the general and metric versions of OCT and WSDOCT.

The complexity of the WSDOCT justified the study of special cases of the problem, most of them NP-hard problems, as well. That are the cases of the minimum product-requirement communication spanning tree problem (PROCT) and minimum sum-requirement communication spanning tree problem (SROCT), introduced in [15]. PROCT is the case in which the requirements are equal for each node (i.e. $\sigma(u) = \rho(u)$ for all $u \in V$), and SROCT is the case in which all the receiving requirement are unitary (i.e. $\rho(u) = 1$ for all $u \in V$). In [15] a 1.577-approximation algorithm for PROCT and a 2-approximation for SROCT were presented. The approximation ratio for PROCT was improved in [16], where a PTAS was given.

Another particular case of WSDOCT is the weighted p-MRCT, which is also the particular case of SROCT where, given an integer p, only p nodes of the graph will have a positive sending requirement (i.e. the remaining nodes have zero sending requirement). When all the p nodes have sending requirement equals to 1, the problem is called p-MRCT. In [5] it was proved that 2-MRCT is NP-hard, this result was also proved in [14] where a PTAS for 2-MRCT and the metric case of weighted 2-MRCT were given.

Two other NP-hard particular cases of WSDOCT were introduced and studied in [10], the p-WSDOCT and the p-WDOCT. The p-WSDOCT is a generalization of p-MRCT where the receiving requirement for each node may be any non-negative value, while the p-WDOCT is a particular case of p-WSDOCT where all the non-zero sending requirement are 1. Also, in [10] and [11], PTASs were given for the metric versions of SROCT, p-MRCT, p-WDOCT and fixed parameter p-WSDOCT (where the fixed parameter was on the sending requirements).

In this work we give a 2-approximation algorithm for WSDOCT with time complexity $O(n^2 \log n + mn)$ and also we prove that there exists a PTAS for metric case of WSDOCT. Both the results improve the best known ratios for the problem.

This paper is organized as follows. In the next section we give some notation. In section 3 we present our PTAS for the metric WSDOCT. Subsections 3.1, 3.2 and 3.3, discuss over the proofs to achieve our result. In section 4 we show the 2-approximation for WSDOCT. Finally, in section 5 the conclusions and future work are given.

2. Definitions

Given a graph G we denote the set of its nodes as V_G and the set of its edges as E_G (when G is implicit by context we use V and E instead of V_G and 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 WSDOCT as:

Problem 1. WSDOCT*: Weighted Source Destination Communication Spanning Tree problem*

Input: *A graph G, a non-negative length function over the edges of G, $\omega : E \to \mathbb{Q}_+$, a non-negative sending requirement function over the nodes of G, $\sigma : V \to \mathbb{Q}_+$ and a non-negative receiving requirement function over the nodes of G, $\rho : V \to \mathbb{Q}_+$.*

Output: *A spanning tree T of G which minimizes the total routing cost:*
$c(T,\sigma,\rho) = \frac{1}{2}\sum_{u\in V}\sum_{v\in V}(\sigma(u)\rho(v) + \rho(u)\sigma(v))\,d(T,u,v)$.

Definition 2. Given a graph G and a non-negative sending and receiving requirement functions over the nodes of G, $\sigma : V \to \mathbb{Q}_+$ and $\rho : V \to \mathbb{Q}_+$, we denote $\sigma(G) = \sum_{u\in V_G}\sigma(u)$, $\rho(G) = \sum_{u\in V_G}\rho(u)$ and $n(G) = |V_G|$. When G is implicit by the context we use R_σ to denote $\sigma(G)$, R_ρ to denote $\rho(G)$ and n to denote $n(G)$.

3. PTAS for metric-WSDOCT

In this section we only consider the metric-WSDOCT(the metric case of WSDOCT), which is the particular case of WSDOCT 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 metric-WSDOCT we introduce the concept of k-star[1]:

[1] The definition of k-star used in this paper is the same used by [10,11,15–17], 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).

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. A k-star may be defined by (τ, S), being τ its core and $S = \{S_{u_1}, \cdots, S_{u_k}\}$, where S_{u_i} represents the set of leaves adjacent in the k-star to the node u_i of τ $(1 \le i \le k)$.

The idea of the PTAS is to select a fixed integer $k > 0$ depending on the desired approximation ratio and to iterate over all the possible sub-trees of the graph with k nodes. Then, for each sub-tree τ we approximate an optimal k-star with core τ. Finally, we select the k-star with minimum communication cost which will guarantee the approximation.

In order to approximate an optimal k-star with a given core τ we split the proof into three parts:

- First, we show how to obtain an optimal k-star with core τ if all the leaves (nodes out of τ) have binary receiving requirement (i.e. $r(u) \in \{0, 1\}$ for all $u \in V - V_\tau$).
- Then, we show that when the leaves have integer receiving requirement (i.e. $r(u) \in \mathbb{N}^+$) we can transform the problem to the case in which the receiving requirement of the leaves are binary.
- Finally, we show that for any receiving requirements on the leaves, there exists a reduction to integer receiving requirement on the leaves that guarantees some approximation.

This section is organized as follows. In subsection 3.1 we show how to obtain an optimal k-star with a given core τ when all the receiving requirements of the leaves are binary. In subsection 3.2 we show how to obtain an optimal k-star with a given core τ when the receiving requirements of the leaves are integers. In subsection 3.3 we show how to approximate an optimal k-star with core τ when the receiving requirements of the leaves are any rational. Finally, in subsection 3.4 we give the PTAS for metric-WSDOCT.

3.1. Optimal k-star with binary receiving requirement on leaves

Suppose we are given a sub-tree τ of G with k nodes and an instance of metric-WSDOCT where all the receiving requirements of the leaves are binary: $\rho(u) \in \{0, 1\}$ for each $u \in V - V_\tau$. In this subsection we will show how to obtain an optimal k-star with core τ for such instance, by modeling the problem as a minimum flow problem.

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

Definition 4. Given a k-star $T = (\tau, S)$ and a receiving requirement function $\rho : V \to \mathbb{Q}_+$ such that $\rho(u) \in \{0, 1\}$ for each $u \in V_{G-\tau}$, a **configuration** of T is $(\tau, L_\emptyset, L_+)$ where $L_\emptyset = \{l^\emptyset_{u_1}, ..., l^\emptyset_{u_k}\}$ and $L_+ = \{l^+_{u_1}, ..., l^+_{u_k}\}$ are two vectors of integers being $l^\emptyset_{u_i} = |S^0_{u_i}|$ and $l^+_{u_i} = |S^1_{u_i}|$, where $S^0_{u_i}$ is the set of nodes of S_{u_i} with receiving requirement equals to zero and $S^1_{u_i}$ is the set of nodes of S_{u_i} with receiving requirement equals to one $(1 \le i \le k)$.

Observe that the number of possible configurations with core τ is $O\left(n^k\right)$ which is polynomial in n. So, our proposal is to enumerate all possible configurations with core τ, finding an optimal k-star of each configuration, and finally select the best k-star among them.

We find an optimal k-star for metric-WSDOCT with a given configuration $(\tau, L_\emptyset, L_+)$ where $\rho : V_{G-\tau} \to \{0, 1\}$, by reducing the problem to an uncapacitated minimum cost flow problem (UMCF).

Problem 2. 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 guarantees 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 $e = \langle v, w \rangle$ ($\forall e \in E, v, w \in V$).

Given an instance $I = \langle G, \sigma, \rho, \omega, k \rangle$ of metric-WSDOCT with configuration $C = (\tau, L_\emptyset, L_+)$ where $\rho : V_{G-\tau} \to \{0, 1\}$, in order to reduce our problem to UMCF we define the set of nodes $U_\emptyset$ (U_+) obtained by creating the node $u_\emptyset$ (u_+) for each node u of τ. Then, we define the directed graph G' on the set of nodes $V_{G-\tau} \cup U_\emptyset \cup U_+$, which is bipartite and the bi-partition is given by the nodes in $V_{G-\tau}$ and outside this set. Also, we add an arc between each node of $U_\emptyset$ (U_+) to each node of $V_{G-\tau}$ in G' with receiving requirement zero (one). We associate, to each arc $\langle u_\emptyset, v \rangle$ $(\langle u_+, v \rangle)$, a cost equivalent to the value of assigning v as adjacent of u in a k-star with the given configuration. In order to calculate that cost, we have to consider the cost of routing the sending requirement of v to all nodes of τ assuming that each node $w \in V_\tau$ receives $(l^+_w + \rho(w))$ times the value $\sigma(v)$ (considering the transmission to the node w and the leaves adjacent to it that have receiving requirement one); also, when v has receiving requirement one, we must add the cost of routing the sending requirement of the entire graph $(R_\sigma - \sigma(v))$ to node v, which will pass by node u. Finally we set the demands to ensure

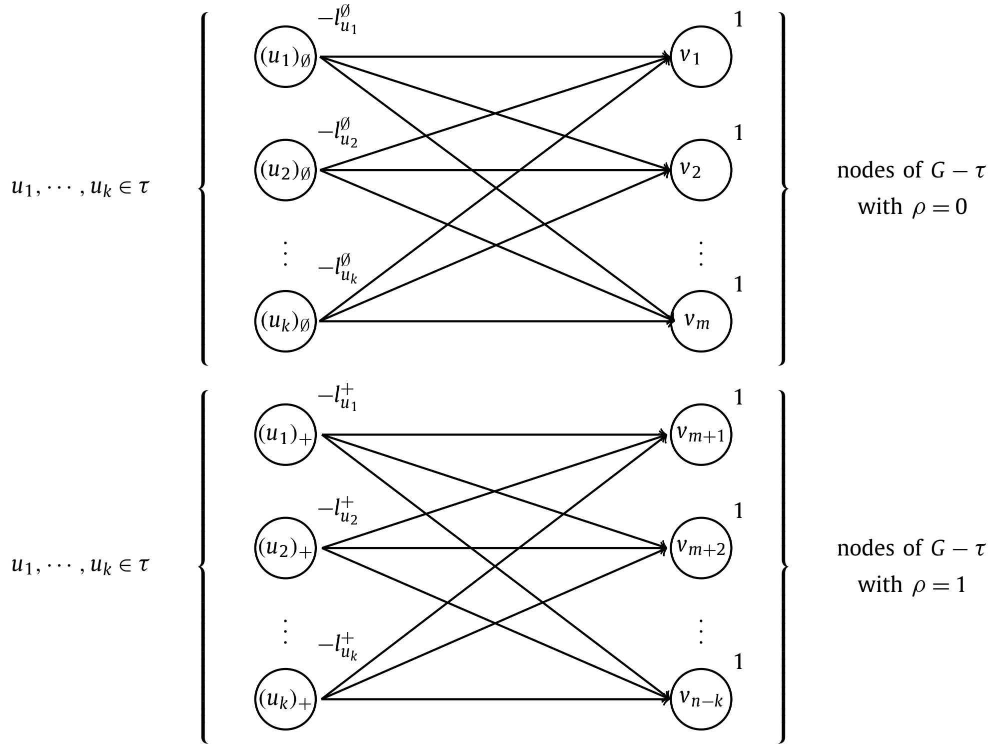

Fig. 1. Given an instance of WSDOCT with binary receiving requirement on the leaves and a configuration $(\tau, L_\emptyset, L_+)$. First, we create a partition with the nodes $(u_i)_\emptyset$ and $(u_i)_+$ where $u_i \in \tau$ and $1 \le i \le k$. We also associate to each node $(u_i)_\emptyset$ a demand with value $-l^{\emptyset}_{u_i}$ and to each node $(u_i)_+$ a demand with value $-l^{+}_{u_i}$. Then we create a second partition with the nodes $v_j \in G - \tau$ $(1 \le j \le n-k)$, associating to each one of those nodes a demand with value 1. Finally, we add an arc from each node $(u_i)_\emptyset$ to each node in the second partition with $\rho = 0$ and also from each node $(u_i)_+$ to each node in the second partition with $\rho = 1$. We set a cost to each arc $\langle u_\emptyset, v\rangle$ ($\langle u_+, v\rangle$) equivalent to the value of assigning v as adjacent of u in a k-star with the given configuration.

assignment between nodes out of $V_{G-\tau}$ and nodes in $V_{G-\tau}$. The following proposition gives us that reduction and Fig. 1 illustrates it.

Proposition 1. *Given an instance $I = \langle G, \sigma, \rho, \omega, k\rangle$ of the optimum k-star for metric-*WSDOCT *and a configuration $C = (\tau, L_\emptyset, L_+)$ where $\rho : V_{G-\tau} \to \{0, 1\}$, the problem of finding an optimal k-star with configuration C for I can be reduced in polynomial time to the* UMCF *with instance $I' = \left\langle G', \omega', r'\right\rangle$, where:*

- $V_{G'} = V_{G-\tau} \cup \{u_\emptyset, u_+ : u \in V_\tau\}$;
- $E_{G'} = \{(u_\emptyset, v) | u \in V_\tau \wedge v \in V_{G-\tau} \wedge \rho(v) = 0\}$
  $\cup \{(u_+, v) | u \in V_\tau \wedge v \in V_{G-\tau} \wedge \rho(v) = 1\}$;
- $\omega'(u_\emptyset, v) = \sum_{w \in V_\tau} \sigma(v)\,(d(\tau, u, w) + \omega(u, v))\left(l^+_w + \rho(w)\right)$;
  $\omega'(u_+, v) = (R_\sigma - 2\sigma(v))\,\omega(u, v)$
  $+ \sum_{w \in V_\tau} \sigma(v)\,(d(\tau, u, w) + \omega(u, v))\left(l^+_w + \rho(w)\right)$;
- *if $u \in V_{G-\tau}$ then $r'(u) = 1$, otherwise $r'(u_\emptyset) = -l^{\emptyset}_u$ and $r'(u_+) = -l^+_u$.*

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 guarantees that every feasible solution S' of the flow problem represents an assignment of leaves of $V_{G-\tau}$ to be adjacent to nodes in τ for a k-star T of G with configuration $(\tau, L_\emptyset, L_+)$. Also, it is easy to see that any k-star T with configuration $(\tau, L_\emptyset, L_+)$ provides a feasible solution to the flow problem: for each $u \in \tau$ connect $u_\emptyset$ to the $l^{\emptyset}_u$ leaves with receiving requirement zero adjacent to u in T and connect u_+ to the l^+_u leaves with receiving requirement one adjacent to u in T.

Let S' be a solution for UMCF with instance I' obtained by the transformation described in Proposition 1 over an instance I of the optimum k-star for metric-WSDOCT. Let T be the associated k-star to S', $V_\emptyset$ the set of nodes of $V_{G-\tau}$ with receiving requirement zero and V_+ the set of nodes of $V_{G-\tau}$ with receiving requirement one. Also, for $v \in V_\emptyset$ ($v \in V_+$), $p(v)$ denotes the node in τ such that $p(v)_\emptyset$ ($p(v)_+$) is adjacent to v in S', then:

$$\begin{aligned}
c(S') &= \sum_{u \in V_{G-\tau}} \omega'(p(u), u) \\
&= \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(u)\,(d(\tau, p(u), v) + \omega(p(u), u))\left(l_v^+ + \rho(v)\right) \\
&\quad + \sum_{u \in \{v \mid v \in V_{G-\tau} \wedge \rho(v)=1\}} (R_\sigma - 2\sigma(u))\,\omega(p(u), u).
\end{aligned}$$

Observe that for every leaf $u \in V_{G-\tau}$ and node v of τ:

$$d(T, u, v) = d(\tau, p(u), v) + \omega(p(u), u).$$

Also, $R_\sigma = \sum_{v \in V_G} \sigma(v)$ and for every leaf $u \in V_{G-\tau}$, if $\rho(u) \neq 1$, then $\rho(u) = 0$. Then:

$$\begin{aligned}
c(S') &= \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(u) d(T, u, v)\left(l_v^+ + \rho(v)\right) \\
&\quad + \sum_{u \in V_{G-\tau}} \left(\sum_{v \in V_G} \sigma(v) - 2\sigma(u)\right) \rho(u)\omega(p(u), u) \\
&= \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(u)\rho(v) d(T, u, v) + \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(u) d(T, u, v) l_v^+ \\
&\quad + \sum_{u \in V_{G-\tau}} \sum_{v \in V_G} \sigma(v)\rho(u)\omega(p(u), u) + \sum_{u \in V_{G-\tau}} -2\sigma(u)\rho(u)\omega(p(u), u) \\
&= \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(u)\rho(v) d(T, u, v) \\
&\quad + \sum_{u \in V_{G-\tau}} \left(\sum_{v \in \tau} \sigma(u) d(T, u, v) l_v^+ + \sum_{v \in V_{G-\tau}} \sigma(u)\rho(v)\omega(p(v), v)\right) \\
&\quad + \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(v)\rho(u)\omega(p(u), u) + \sum_{u \in V_{G-\tau}} -2\sigma(u)\rho(u)\omega(p(u), u).
\end{aligned}$$

Notice that for every pair of nodes u, v of $V_{G-\tau}$:

$$d(T, u, v) = d(T, u, p(v)) + \omega(p(v), v).$$

Then:

$$\begin{aligned}
c(S') &= \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(u)\rho(v) d(T, u, v) \\
&\quad + \sum_{u \in V_{G-\tau}} \sum_{v \in V_{G-\tau}} \sigma(u)\rho(v) d(T, u, v) + \sum_{u \in V_{G-\tau}} 2\sigma(u)\rho(u)\omega(p(u), u) \\
&\quad + \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(v)\rho(u)\omega(p(u), u) + \sum_{u \in V_{G-\tau}} -2\sigma(u)\rho(u)\omega(p(u), u) \\
&= \sum_{u \in V_{G-\tau}} \sum_{v \in V_G} \sigma(u)\rho(v) d(T, u, v) + \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(v)\rho(u)\omega(p(u), u).
\end{aligned}$$

By considering the sending requirement from each node of τ:

$$\sum_{u \in V_\tau} \sum_{v \in V_\tau} \sigma(u) d(T, u, v)(l_v^+ + \rho(v)),$$

we get:

$$\begin{aligned}
c(S') &= \sum_{u \in V_{G-\tau}} \sum_{v \in V_G} \sigma(u)\rho(v) d(T, u, v) + \sum_{u \in V_{G-\tau}} \sum_{v \in V_\tau} \sigma(v)\rho(u)\omega(p(u), u) \\
&\quad + \sum_{u \in V_\tau} \sum_{v \in V_\tau} \sigma(u) d(T, u, v)(l_v^+ + \rho(v))
\end{aligned}$$

$$
\begin{aligned}
&- \sum_{u\in V_\tau}\sum_{v\in V_\tau}\sigma(u)d(T,u,v)(l_v^+ + \rho(v))\\
&= \sum_{u\in V_G}\sum_{v\in V_G}\sigma(u)\rho(v)d(T,u,v) - \sum_{u\in V_\tau}\sum_{v\in V_\tau}\sigma(u)d(T,u,v)(l_v^+ + \rho(v))\\
&= \frac{1}{2}\sum_{u\in V_G}\sum_{v\in V_G}\left(\sigma(u)\rho(v)+\rho(u)\sigma(v)\right)d(T,u,v)\\
&\quad - \sum_{u\in V_\tau}\sum_{v\in V_\tau}\sigma(u)d(T,u,v)(l_v^+ + \rho(v))\\
&= c(T,\sigma,\rho) - \sum_{u\in V_\tau}\sum_{v\in V_\tau}\sigma(u)d(T,u,v)(l_v^+ + \rho(v)).
\end{aligned}
$$

Observe that:

$$
\sum_{u\in V_\tau}\sum_{v\in V_\tau}\sigma(u)d(T,u,v)(l_v^+ + \rho(v)),
$$

is the same for every solution of metric-WSDOCT with instance I and configuration C. Then, an optimum of UMCF with instance I' is associated to an optimal k-star with configuration C of metric-WSDOCT with instance I.

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 $c(T,\sigma,\rho)$ would be $O(k^3)$. So, the reduction above can be done in $O(nk^3)$. $\square$

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. [9]). Then, finding an optimal k-star for metric-WSDOCT with fixed k can be done efficiently.

Lemma 1. *Given τ, if $\rho(u)\in\{0,1\}$ for all $u\in V-V_\tau$, then the optimum k-star with core τ for metric-*WSDOCT *can be solved in $O\left(n^{k+2}\log^2(n)\right)$ time.*

3.2. Optimal k-star with integer receiving requirement on leaves

In this subsection we consider there are given a sub-tree τ of G with k nodes and an instance of metric-WSDOCT where $\rho(u)\in\mathbb{N}^+$ for each $u\in V-V_\tau$, and we show how to obtain an optimal k-star with core τ for metric-WSDOCT in $O\left(\left(n+R_\rho\right)^{k+2}\log^2\left(n+R_\rho\right)\right)$ time. The proposal is to transform the given instance of metric-WSDOCT into an instance where all the leaves have binary receiving requirement. In order to do that, we will split each leaf u with receiving requirement greater than 1 ($\rho(u)>1$) into $\rho(u)$ leaves with receiving requirement equals 1 and adjust the sending requirement and edge-lengths for those new leaves.

Let denote by $V_{>1}$ the set of leaves (nodes in $V-V_\tau$) with receiving requirement greater than 1. Also, define graph G', sending requirement function σ', receiving requirement function ρ' and length function over the edges ω' as follows:

- For each $u\in V_{>1}$ and $i\in\{1,\cdots,\rho(u)\}$, add node u_i to $V_{G'}$ where:
  - $\sigma'(u_i)=\frac{\sigma(u)}{\rho(u)}$, $\rho'(u_i)=1$ and
  - $\omega'(u_i,v)=\frac{\left(\sigma(u)\left(R_\rho-\rho(u)\right)+\rho(u)\left(R_\sigma-\sigma(u)\right)\right)}{\rho(u)\left(\frac{\sigma(u)}{\rho(u)}\left(R_\rho-1\right)+R_\sigma-\frac{\sigma(u)}{\rho(u)}\right)}\omega(u,v)\quad(v\in V_\tau)$.
- The rest of the nodes (those in $V-V_{>1}$) are added to $V_{G'}$ with the same sending and receiving requirement values. Also, the rest of the edges have the same length function.

Notice that the transformation above guarantees:

$$
\sum_{u\in V_{G'}}\sigma(u)=\sum_{u\in V_G}\sigma(u)=R_\sigma,\quad\text{and}\quad\sum_{u\in V_{G'}}\rho(u)=\sum_{u\in V_G}\rho(u)=R_\rho.
$$

Also, an optimal k-star X' with core τ for instance $\langle G',\sigma',\rho',\omega'\rangle$ satisfies that for every node $u\in V_{>1}$ all the nodes u_i ($i\in\{1,\cdots,\rho(u)\}$) are adjacent to the same node of τ in X'. In order to prove that, suppose there exist two nodes u_1 and u_2 obtained from some $u\in V_{>1}$ such that u_1 is adjacent to $v_1\in V_\tau$ in X' and u_2 is adjacent to $v_2\in V_\tau$, with $v_1\neq v_2$. Then the communication cost on X' results:

$$\begin{aligned}
c(X', \sigma', \rho') = &\sum_{u,v \in V_{G'}-\{u_1,u_2\}} d(X', u, v)\left(\sigma'(u)\rho'(v) + \rho'(u)\sigma'(v)\right) \\
&+ \sum_{u \in V_{G'}-\{u_1,u_2\}} d(X', u_1, u)\left(\sigma'(u_1)\rho'(u) + \sigma'(u)\right) \\
&+ \sum_{u \in V_{G'}-\{u_1,u_2\}} d(X', u_1, u)\left(\sigma'(u_2)\rho'(u) + \sigma'(u)\right) \\
&+ \left(\omega(u_1, v_1) + \omega(u_2, v_2) + d(X', u_1, u_2)\right)\left(\sigma'(u_1) + \sigma'(u_2)\right).
\end{aligned}$$

Without loss of generality suppose:

$$\begin{aligned}
&\sum_{u \in V_{G'}-\{u_1,u_2\}} d(X', u_1, u)\left(\sigma'(u_1)\rho'(u) + \sigma'(u)\right) + \omega(u_1, v_1)\left(\sigma'(u_1) + \sigma'(u_2)\right) \\
\le &\sum_{u \in V_{G'}-\{u_1,u_2\}} d(X', u_2, u)\left(\sigma'(u_2)\rho'(u) + \sigma'(u)\right) + \omega(u_2, v_2)\left(\sigma'(u_1) + \sigma'(u_2)\right).
\end{aligned}$$

Denote by X^* the k-star with core τ obtained by replacing the edge $\langle u_2, v_2\rangle$ by $\langle u_2, v_1\rangle$ in X', then:

$$\begin{aligned}
c(X^*, \sigma', \rho') = &\sum_{u,v \in V_{G'}-\{u_1,u_2\}} d(X', u, v)\left(\sigma'(u)\rho'(v) + \rho'(u)\sigma'(v)\right) \\
&+2 \sum_{u \in V_{G'}-\{u_1,u_2\}} d(X', u_1, u)\left(\sigma'(u_1)\rho'(u) + \sigma'(u)\right) \\
&+2\left(\omega(u_1, v_1)\right)\left(\sigma'(u_1) + \sigma'(u_2)\right) \\
\le &\, c(X', \sigma', \rho') + d(X', v_1, v_2)\left(\sigma'(u_1) + \sigma'(u_2)\right) \\
< &\, c(X', \sigma', \rho').
\end{aligned}$$

Arriving in a contradiction because X' is optimal. Since an optimal k-star X' with core τ for instance $\left\langle G', \sigma', \rho', \omega'\right\rangle$ satisfies that for every node $u \in V_{>1}$ all the nodes u_i $(i \in \{1, \cdots, \rho(u)\})$ are adjacent to the same node of $v \in V_\tau$ in X', then replacing in X' all the nodes u_i $(i \in \{1, \cdots, \rho(u)\})$ by u we obtain a k-star X with core τ which is a solution for metric-WSDOCT and instance $\langle G, \sigma, \rho, \omega\rangle$. Also, it is easy to see that for every k-star with core τ for $\langle G, \sigma, \rho, \omega\rangle$, there exists a k-star with core τ for $\left\langle G', \sigma', \rho', \omega'\right\rangle$ that is obtained by replacing each $u \in V_{>1}$ by the nodes u_i $(i \in \{1, \cdots, \rho(u)\})$.

Let X be a k-star with core τ for instance $\langle G, \sigma, \rho, \omega\rangle$ and X' the associated k-star with core τ and instance $\left\langle G', \sigma', \rho', \omega'\right\rangle$. Also, denote by $p_T(u)$ the node in τ adjacent to a leaf u in some k-star of core τ. Then:

$$\begin{aligned}
&c(X', \sigma', \rho') - c(X, \sigma, \rho) \\
= &\sum_{u \in V_{>1}} \sum_{i=1}^{\rho(u)} \omega'(u_i, p_{X'}(u_i))\left(\left(R_\sigma - \sigma'(u)\right)\rho'(u) + \sigma'(u)\left(R_\rho - \rho'(u)\right)\right) \\
&- \sum_{u \in V_{>1}} \omega(u, p_X(u))\left(\left(R_\sigma - \sigma(u)\right)\rho(u) + \sigma(u)\left(R_\rho - \rho(u)\right)\right) \\
= &\sum_{u \in V_{>1}} \rho(u)\omega'(u, p_X(u))\left(R_\sigma - \frac{\sigma(u)}{\rho(u)} + \frac{\sigma(u)}{\rho(u)}\left(R_\rho - 1\right)\right) \\
&- \sum_{u \in V_{>1}} \omega(u, p_X(u))\left(\left(R_\sigma - \sigma(u)\right)\rho(u) + \sigma(u)\left(R_\rho - \rho(u)\right)\right) \\
= &\,0.
\end{aligned}$$

Implying that X and X' have same communication cost. Then, if X' is optimal, the k-star X obtained from X' is an optimal k-star with core τ for metric-WSDOCT with instance $\langle G, \sigma, \rho, \omega\rangle$.

In the previous subsection we show how to find in $O\left(n^{k+2}\log^2(n)\right)$ computational time, an optimal k-star with core τ for metric-WSDOCT when $\rho(u) \in \{0, 1\}$ for all $u \in V - V_\tau$ which is the case of ρ', where $n(G') \le n + R_\rho$. Leading us to the following result.

Lemma 2. *Given an instance $I = \langle G, \sigma, \rho, \omega\rangle$ for metric-*WSDOCT *and a sub-tree τ of G, such that for every $u \in V - V_\tau$ $\rho(u) \in \mathbb{N}$, an optimal k-star ($k = n(\tau)$) with core τ for metric-*WSDOCT *can be found in $O\left(\left(n + R_\rho\right)^{k+2}\log^2\left(n + R_\rho\right)\right)$ time complexity.*

3.3. Approximation ratio of optimal k-star for metric-WSDOCT

In this subsection we consider an instance I of metric-WSDOCT, $0 < \delta \leq \frac{1}{2}$ and a sub-tree τ of G with k nodes. We will transform the given instance I into an instance I' where all receiving requirement of the leaves are integers, and we will show that an optimal k-star with core τ of I' is a $(1+8\delta)$-approximation for some optimal special k-star X with core τ of I (if such X exists).

First we need to define the special k-star mentioned above, which is the **δ-balanced k-star**. This definition depends on the concept of **δ-$\sigma\rho$-separator**.

Definition 5. Given $0 < \delta \leq \frac{1}{2}$, a spanning tree T of a graph G and two non-negative functions σ and ρ of sending and receiving requirements over the set of nodes V, a sub-tree S of T is a **δ-$\sigma\rho$-separator** of T if every component B of $T - S$, satisfies $\sigma(B) \leq \delta R_\rho$ and $\rho(B) \leq \delta R_\sigma$.

Definition 6. Given $0 < \delta \leq \frac{1}{2}$, a k-star T of a graph G and two non-negative functions σ and ρ of sending and receiving requirements over the set of nodes V. We say, that T is a **δ-balanced k-star** if its core τ is a minimal δ-$\sigma\rho$-separator of T.

Our goal is to approximate an optimal δ-balanced k-star with core τ. In order to achieve that, we show how to transform any instance of metric-WSDOCT to an instance in which every node out of τ have integer receiving requirement. The idea is to identify a threshold such that every leaf u with receiving requirement lower than the threshold gets $\rho(u) = 0$. Then, we modify the new receiving requirement to guarantee that each leaf gets an integer receiving requirement and the sum of them is a value that depends on n and δ.

Denote by $V_L = V - V_\tau = \{u_1, \cdots, u_{n-k}\}$, the set of leaves to be added to τ in order to build a solution, and consider a sorting over the elements of V_L that if $i < j$ then $\rho(u_i) \leq \rho(u_j)$ for all $i, j \in \{1, \cdots, n-k\}$. Also, let m be the maximum value between 1 and $n-k$, such that $\sum_{i=1}^{m} \rho(u) \leq \delta^2 R_\rho$, and consider $V_L^0 = \{u_1, \cdots, u_m\}$. Then, define a new receiving requirement function ρ', such that $\rho'(u) = 0$ if $u \in V_L^0$, otherwise $\rho'(u) = \rho(u)$.

Consider a value μ such that if $V_L - V_L^0 = \emptyset$ then $\mu = 1$, otherwise let μ be the minimal value of ρ' on the set $V_L - V_L^0$, and define ρ'' a new receiving requirement function where $\rho''(u) = \left\lfloor \frac{\rho'(u)}{\delta\mu} \right\rfloor$ if $u \in V_L - V_L^0$, otherwise $\rho''(u) = \frac{\rho'(u)}{\delta\mu}$. Observe that for every node u, $\frac{1}{\delta\mu}\rho'(u) \geq \rho''(u)$ and $(1+2\delta)\,\rho''(u) \geq \frac{1}{\delta\mu}\rho'(u)$. Then, for any k-star T with core τ:

$$\frac{1}{\delta\mu\,(1+2\delta)} c(T, \sigma, \rho') \leq c(T, \sigma, \rho'') \leq \frac{1}{\delta\mu} c(T, \sigma, \rho').$$

Let X'' be an optimal solution of metric-WSDOCT with requirement function ρ'', then, for any k-star T with core τ:

$$\frac{1}{\delta\mu\,(1+2\delta)} c(X'', \sigma, \rho') \leq c(X'', \sigma, \rho'') \leq c(T, \sigma, \rho'') \leq \frac{1}{\delta\mu} c(T, \sigma, \rho').$$

So, $c(X'', \sigma, \rho') \leq (1+2\delta) c(T, \sigma, \rho')$, for any k-star T with core τ.

Now we find an upper bound for the communication cost of a k-star. For any k-star T with core τ:

$$c(T, \sigma, \rho) = c(T, \sigma, \rho') + \sum_{i=1}^{m} \rho(u_i) \sum_{v \in V} d(T, u_i, v)\sigma(v).$$

Also, for every $i \in \{1, \cdots, m\}$, u_i is a leaf of T and:

$$d(T, u_i, v) = d(T, u_i, p_T(u_i)) + d(T, p_T(u_i), v),$$

then:

$$\begin{aligned} c(T, \sigma, \rho) = \; & c(T, \sigma, \rho') \\ & + \sum_{i=1}^{m} \rho(u_i) \sum_{v \in V_L^0} \left(d(T, u_i, p_T(u_i)) + d(T, p_T(v), v)\right)\sigma(v) \\ & + \sum_{i=1}^{m} \rho(u_i) \sum_{v \in V_L^0} d(T, p_T(u_i), p_T(v))\sigma(v) \\ & + \sum_{i=1}^{m} \rho(u_i) \sum_{v \in V - V_L^0} \left(d(T, u_i, p_T(u_i)) + d(T, p_T(u_i), v)\right)\sigma(v). \end{aligned}$$

Suppose X is an optimal δ-balanced k-star with core τ and requirement function ρ. Denote by $\omega(\tau)$ the sum of all the edge lengths in τ, since the graph is metric, for every $u \in V_L^0$,

$$d(T, u, p_T(u)) \le d(X, u, p_X(u)) + d(\tau, p_T(u), p_X(u)) \le d(X, u, p_X(u)) + \omega(\tau).$$

Then:

$$\begin{aligned} c(T, \sigma, \rho) \le c(T, \sigma, \rho') &+ \sum_{i=1}^{m} 3\rho(u_i) R_\sigma \omega(\tau) \\ &+ \sum_{i=1}^{m} \rho(u_i) \sum_{v \in V_L^0} \left(d(X, u_i, p_X(u_i)) + d(X, p_X(v), v)\right) \sigma(v) \\ &+ \sum_{i=1}^{m} \rho(u_i) \sum_{v \in V - V_L^0} \left(d(X, u_i, p_X(u_i)) + d(X, p_X(u_i), v)\right) \sigma(v). \end{aligned}$$

Since $c(X'', \sigma, \rho') \le (1 + 2\delta)\, c(X, \sigma, \rho')$, by using in the above equation X'' instead of T, we get:

$$\begin{aligned} c(X'', \sigma, \rho) &\le (1 + 2\delta)\, c(X, \sigma, \rho) + 3R_\sigma \omega(\tau) \sum_{i=1}^{m} \rho(u_i) \\ &\le (1 + 2\delta)\, c(X, \sigma, \rho) + 3\delta^2 R_\sigma R_\rho \omega(\tau). \end{aligned}$$

Now we analyze the value of $3\delta^2 R_\sigma R_\rho \omega(\tau)$. In order to do that, the following proposition gives us a lower bound for the communication requirement passing over any edge of τ in X.

Proposition 2. *Given an optimal δ-balanced k-star X with core τ and an edge $e \in E_\tau$, the communication requirement passing over e in X is at least $\frac{\delta}{2} R_\sigma R_\rho$.*

Proof. Let X_u and X_v be the sub-trees of X after removing edge e, then the communication requirement passing over e in X is:

$$\sigma(X_u)\rho(X_v) + \rho(X_u)\sigma(X_v) = \sigma(X_u)\left(R_\rho - \rho(X_u)\right) + \rho(X_u)\left(R_\sigma - \sigma(X_u)\right).$$

Since X is a δ-balanced k-star and τ is minimal, then:

$$\begin{aligned} \sigma(X_u) > \delta R_\sigma \quad &\text{or} \quad \rho(X_u) > \delta R_\rho \qquad \text{and} \\ R_\sigma - \sigma(X_u) > \delta R_\sigma \quad &\text{or} \quad R_\rho - \rho(X_u) > \delta R_\rho. \end{aligned}$$

Without loss of generality suppose $\sigma(X_u) > \delta R_\sigma$. Then, we analyze two possibilities:

- If $R_\sigma - \sigma(X_u) \le \delta R_\sigma$, then $R_\rho - \rho(X_u) > \delta R_\rho,\quad \sigma(X_u) \ge R_\sigma - \delta R_\sigma \ge \frac{R_\sigma}{2}$, and the communication requirement passing over e in X is at least:
$$\frac{R_\sigma}{2}\delta R_\rho + \rho(X_u)\left(R_\sigma - \sigma(X_u)\right) \ge \frac{\delta}{2} R_\sigma R_\rho.$$
- Otherwise, the communication requirement passing over e in X is at least:
$$\delta R_\sigma \left(R_\rho - \rho(X_u)\right) + \rho(X_u)\left(\delta R_\sigma\right) = \delta R_\sigma R_\rho \ge \frac{\delta}{2} R_\sigma R_\rho. \quad \square$$

Proposition 2 allows us to obtain:

$$c(X, \sigma, \rho) \ge \sum_{e \in E_\tau} \frac{\delta}{2} R_\sigma R_\rho \omega(e) \ge \frac{\delta}{2} R_\sigma R_\rho \omega(\tau).$$

Then:

$$c(X'', \sigma, \rho) \le (1 + 2\delta)\, c(X, \sigma, \rho) + 3\delta\left(2c(X, \sigma, \rho)\right) \le (1 + 8\delta)\, c(X, \sigma, \rho).$$

The previous subsections showed us how to find X'' in computational time $O\left((n + R_{r''})^{k+2} \log^2 (n + R_{r''})\right)$. Since $\delta^2 R_\rho \le (m+1)\,\mu \le n\mu$ and $R_{r''} \le \frac{R_\rho}{\mu}$, then $R_{r''} \le \frac{nR_\rho}{\delta^2 R_\rho} = \frac{n}{\delta^2}$.

Lemma 3. *Given $0 < \delta \leq \frac{1}{2}$ and τ a sub-tree of G with k nodes, if there exists a δ-balanced k-star of G with core τ, then there exists an $(1+8\delta)$-approximation for the optimum δ-balanced k-star with core τ for metric-*WSDOCT*, that can be found in $O\left(\left(n+\frac{n}{\delta^2}\right)^{k+2}\log^2\left(n+\frac{n}{\delta^2}\right)\right)$ time complexity.*

3.4. PTAS

In order to obtain our PTAS, we use a result of [11], in which it was proved that for any constant $0 < \delta \leq \frac{1}{2}$ and any instance of metric-WSDOCT, an optimal δ-balanced k-star with $k = \left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$ is a $(\frac{1}{1-\delta})$-approximation of the optimum value, but no algorithm to find such approximation was given.

Then, given an instance of metric-WSDOCT and $0 < \delta \leq \frac{1}{2}$. Our algorithm proposal is to enumerate over all possible sub-trees τ of G that can be core of a k-star, with $k = \left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$. There are $O(k^k)$ possible trees over k nodes and $\binom{n}{k} = O(n^k)$ possible ways of selecting k nodes among the n nodes of G, so the total of possibilities to enumerate is $O\left(k^k n^k\right) = O\left(n^k\right)$.

As shown in the previous subsections, for each enumerated τ we find in $O\left(\left(n+\frac{n}{\delta^2}\right)^{k+2}\log^2\left(n+\frac{n}{\delta^2}\right)\right)$ time, an $(1+8\delta)$-approximation to an optimal δ-balanced k-star with core τ. Then, by selecting the minimum communication cost k-star among all of them, we obtain a $(\frac{1+8\delta}{1-\delta})$-approximation for WSDOCT.

Theorem 1. *There exists a* PTAS *for metric-*WSDOCT *such that, for any constant* $0 < \delta \leq \frac{1}{2}$ *a* $\left(1+\frac{9\delta}{1-\delta}\right)$*-approximation can be found with time complexity* $O\left(n^k\left(n+\frac{n}{\delta^2}\right)^{k+2}\log^2\left(n+\frac{n}{\delta^2}\right)\right)$, *where* $k = \left(3\left(\left\lceil\frac{6}{\delta}\right\rceil^2 - 11\left\lceil\frac{6}{\delta}\right\rceil + 1\right)\right)$.

4. 2-Approximation for WSDOCT

Theoretically the PTAS shown in the previous section represents a good step forward on the study of the metric-WSDOCT, but in practice the time complexity is not feasible. Also, the above discussion does not bring any new light on the complexity of the general WSDOCT. For these reasons, in this section we generalize the result in [15] for SROCT and give a 2-approximation algorithm not only for the metric case but also for the general case of WSDOCT, with time complexity $O(n^2\log n + mn)$, where m is the number of edges of G ($m = |E_G|$).

The algorithm is simple. For each node u of G calculate a shortest-path spanning tree rooted at u and select the minimum communication tree among them. A shortest-path spanning tree rooted at a node u, is a spanning tree T of G, where the distance in T from u to each node v of G is equal to the minimum distance from u to v in G (i.e. $d(T,u,v) = d(G,u,v)$ for all $v \in V_G$). For example, Dijkstra's algorithm for minimum distances from a source find such a tree [4]. Since each tree can be calculated in $O(n\log n + m)$, and we calculate $O(n)$ trees, the time complexity of the algorithm is $(n^2\log n + mn)$.

Now we prove that the above algorithm guarantees a 2-approximation of the optimum value. Before that we need to introduce the concepts of σ**-centroid** and ρ**-centroid**.

Definition 7. Given an instance of WSDOCT and a spanning tree T of G, a node u is an σ-centroid (ρ-centroid) of T if for each component C of T resulting after remove u from T, satisfies: $\sigma(C) \leq \frac{R_\sigma}{2}$ ($\rho(C) \leq \frac{R_\rho}{2}$).

The idea of the proof is first to establish a lower bound on the communication cost of any spanning tree T of G using the path P between the σ-centroid and the ρ-centroid of T. Then, we demonstrate that the communication cost of any spanning tree T of G is at least a half of the minimum between a shortest-path spanning tree rooted at the σ-centroid of T and shortest-path spanning tree rooted at the ρ-centroid of T. Finally, we conclude the 2-approximation algorithm.

In order to simplify the notation, from now on we will use $c(T)$ to represent the communication cost $c(T,\sigma,\rho)$ of any spanning tree T of G.

Lemma 4. *Given an instance of* WSDOCT *and a spanning tree T of G. If P is the path in T between the σ-centroid and the ρ-centroid of T, then:*

$$c(T) \geq \frac{1}{2}\sum_{u\in V_G}\left(R_\rho\sigma(u) + R_\sigma\rho(u)\right)d(T,u,P) + \frac{1}{2}R_\sigma R_\rho\omega(P).$$

Where, $d(T,u,P)$ denotes the distance in T from node u to the nearest node of P, and $\omega(P)$ is the sum over the lengths of the edges in P.

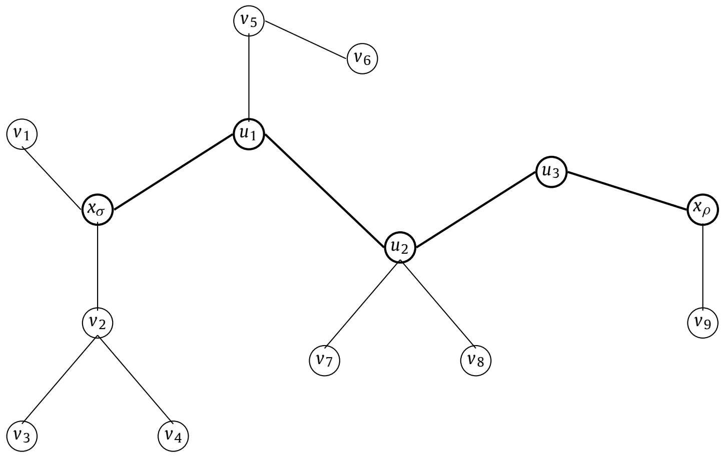

Fig. 2. Consider a tree T and the path (in T) between the σ-centroid x_σ and the ρ-centroid x_ρ. Notice that $CP(v_2) = CP(v_3) = CP(v_4) = \{v_2, v_3, v_4\}$ and $CP(v_5) = CP(v_6) = \{v_5, v_6\}$, while for the rest of the nodes in T, $CP(u) = \{u\}$. Also, observe that $p(v_1) = p(v_2) = p(v_3) = p(v_4) = p(x_\sigma) = x_\sigma$, $p(v_5) = p(v_6) = p(u_1) = u_1$, $p(v_7) = p(v_8) = p(u_2) = u_2$, $p(u_3) = u_3$ and $p(v_9) = p(x_\rho) = x_\rho$.

Proof. Consider a spanning tree T of G:

$$\begin{aligned} c(T) &= \frac{1}{2} \sum_{u \in V_G} \sum_{v \in V_G} (\sigma(u)\rho(v) + \rho(u)\sigma(v))\, d(T, u, v) \\ &= \sum_{u \in V_G} \sum_{v \in V_G} \sigma(u)\rho(v) d(T, u, v). \end{aligned}$$

Denote by $T - P$ the graph resulting from removing the nodes and edges of P from T. Then, for each node $u \notin P$ define $CP(u)$ as the set of nodes in the same component of u in $T - P$ and for each node $u \in P$, $CP(u) = \{u\}$. Also, we define as $p(u)$ the node of P with minimum distance to v in T, when $u \in P$, $p(u) = u$. Fig. 2 exemplifies these notations. Then:

$$\begin{aligned} c(T) \geq &\sum_{u \in V_G} \sum_{v \notin CP(u)} \sigma(u)\rho(v)\, (d(T, u, P) + d(T, v, P)) \\ &+ \sum_{u \in V_G} \sum_{v \in V_G} \sigma(u)\rho(v) d(P, p(u), p(v)) \\ = &\sum_{u \in V_G} \sum_{v \notin CP(u)} \sigma(u)\rho(v) d(T, u, P) + \sum_{u \in V_G} \sum_{v \notin CP(u)} \sigma(u)\rho(v) d(T, v, P) \\ &+ \sum_{u \in V_G} \sum_{v \in V_G} \sigma(u)\rho(v) d(P, p(u), p(v)) \\ = &\sum_{u \in V_G} \sigma(u) d(T, u, P) \sum_{v \notin CP(u)} \rho(v) + \sum_{u \in V_G} \rho(u) d(T, u, P) \sum_{v \notin CP(u)} \sigma(v) \\ &+ \sum_{u \in V_G} \sum_{v \in V_G} \sigma(u)\rho(v) d(P, p(u), p(v)). \end{aligned}$$

Notice that for each node $u \notin P$ we have:

$$\sigma(CP(u)) \leq \frac{R_\sigma}{2} \quad \text{and} \quad \rho(CP(u)) \leq \frac{R_\rho}{2},$$

and for each node $u \in P$ $d(T, u, P) = 0$. Then:

$$\begin{aligned} c(T) \geq &\sum_{u \in V_G} \sigma(u) d(T, u, P) \frac{R_\rho}{2} + \sum_{u \in V_G} \rho(u) d(T, u, P) \frac{R_\sigma}{2} \\ &+ \sum_{u \in V_G} \sum_{v \in V_G} \sigma(u)\rho(v) d(P, p(u), p(v)). \end{aligned}$$

If we denote by $P(T,u,v)$ the path in T between nodes u and v, then:

$$\begin{aligned}
c(T) \geq & \sum_{u\in V_G} \sigma(u) d(T,u,P)\frac{R_\rho}{2} + \sum_{u\in V_G} \rho(u) d(T,u,P)\frac{R_\sigma}{2} \\
& + \sum_{u\in V_G}\sum_{v\in V_G} \sigma(u)\rho(v) \sum_{e\in P(T,u,v)\cap P} \omega(e) \\
= & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) \\
& + \sum_{e\in P}\omega(e)\sum_{u\in V_G}\sum_{v\in\{w|w\in V_G\wedge e\in P(T,u,w)\}} \sigma(u)\rho(v).
\end{aligned}$$

For each edge e of T, denote T_{e1} and T_{e2}, the sub-trees obtained after remove e from T. Then:

$$\begin{aligned}
c(T) \geq & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) \\
& + \sum_{e\in P}\omega(e)\left(\sigma(T_{e1})\rho(T_{e2}) + \rho(T_{e1})\sigma(T_{e2})\right).
\end{aligned}$$

For each edge e of P we have the σ-centroid in one of the sub-trees T_{e1} or T_{e2}, and the ρ-centroid in the other sub-tree. Without loss of generality suppose the σ-centroid of T is in T_{e1}, then the ρ-centroid is in T_{e2}, and $\sigma(T_{e1}) \geq \frac{R_\sigma}{2}$ and $\rho(T_{e2}) \geq \frac{R_\rho}{2}$. Then:

$$\begin{aligned}
c(T) \geq & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) \\
& + \sum_{e\in P}\omega(e)\Big(\sigma(T_{e1})\rho(T_{e2}) + \big(R_\rho - \rho(T_{e2})\big)\left(R_\sigma - \sigma(T_{e1})\right)\Big) \\
c(T) \geq & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) \\
& + \sum_{e\in P}\omega(e)\left(\frac{R_\sigma R_\rho}{2} + \left(2\sigma(T_{e1}) - R_\sigma\right)\rho(T_{e2}) + \left(\frac{R_\sigma}{2} - \sigma(T_{e1})\right)R_\rho\right) \\
\geq & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) \\
& + \sum_{e\in P}\omega(e)\left(\frac{R_\sigma R_\rho}{2} + \left(2\sigma(T_{e1}) - R_\sigma\right)\frac{R_\rho}{2} + \left(\frac{R_\sigma}{2} - \sigma(T_{e1})\right)R_\rho\right) \\
= & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) + \sum_{e\in P}\omega(e)\frac{R_\sigma R_\rho}{2} \\
= & \frac{1}{2}\sum_{u\in V_G} \big(\sigma(u)R_\rho + \rho(u)R_\sigma\big) d(T,u,P) + \frac{1}{2}R_\sigma R_\rho \omega(P). \quad \square
\end{aligned}$$

The following lemma guarantees that for every spanning tree T there exists a node u such that the shortest-path spanning tree rooted at u has communication cost at must twice the communication cost of T.

Lemma 5. *Given an instance of* WSDOCT *and a spanning tree T of G, there exists a node u such that any shortest-path spanning tree X rooted at u satisfies:*

$$c(X) \leq 2c(T).$$

Proof. Notice that for any spanning tree T and every node x:

$$c(T) = \frac{1}{2}\sum_{u\in V_G}\sum_{v\in V_G}\left(\sigma(u)\rho(v) + \rho(u)\sigma(v)\right) d(T,u,v)$$

$$\leq \frac{1}{2}\sum_{u\in V_G}\sum_{v\in V_G}\left(\sigma(u)\rho(v)+\rho(u)\sigma(v)\right)\left(d(T,u,x)+d(T,v,x)\right)$$

$$\begin{aligned}
c(T) &\leq \frac{1}{2}\sum_{u\in V_G}\sum_{v\in V_G}\left(\sigma(u)\rho(v)+\rho(u)\sigma(v)\right)d(T,u,x)\\
&\quad+\frac{1}{2}\sum_{u\in V_G}\sum_{v\in V_G}\left(\sigma(u)\rho(v)+\rho(u)\sigma(v)\right)d(T,v,x)\\
&=\sum_{u\in V_G}\sum_{v\in V_G}\left(\sigma(u)\rho(v)+\rho(u)\sigma(v)\right)d(T,u,x)\\
&=\sum_{u\in V_G}d(T,u,x)\left(\sigma(u)\sum_{v\in V_G}\rho(v)+\rho(u)\sum_{v\in V_G}\sigma(v)\right)\\
&=\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)d(T,u,x).
\end{aligned}$$

Then, if we denote by X_σ a shortest-path spanning tree rooted at the σ-centroid x_σ of T, by X_ρ a shortest-path spanning tree rooted at the ρ-centroid x_ρ of T, and by P the path in T between this centroids x_σ and x_ρ:

$$\begin{aligned}
c(X_\sigma) &\leq \sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)d(X,u,x_\sigma)\\
&\leq \sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)d(T,u,x_\sigma)\\
&= \sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)\left(d(T,u,p(u))+d(P,p(u),x_\sigma)\right).
\end{aligned}$$

Analogously:

$$c(X_\rho) \leq \sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)\left(d(T,u,p(u))+d(P,p(u),x_\rho)\right).$$

The minimum between two values is not greater than their average value, then:

$$\begin{aligned}
&\min\left\{c(X_\sigma),c(X_\rho)\right\}\\
&\leq \frac{1}{2}\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)\left(d(T,u,p(u))+d(P,p(u),x_\sigma)\right)\\
&\quad+\frac{1}{2}\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)\left(d(T,u,p(u))+d(P,p(u),x_\rho)\right)
\end{aligned}$$

$$\begin{aligned}
\min\left\{c(X_\sigma),c(X_\rho)\right\} &\leq \frac{1}{2}\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)\left(2d(T,u,p(u))+\omega(P)\right)\\
&=\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)d(T,u,p(u))\\
&\quad+\frac{\omega(P)}{2}\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)\\
&=\sum_{u\in V_G}\left(\sigma(u)R_\rho+\rho(u)R_\sigma\right)d(T,u,p(u))+R_\sigma R_\rho\omega(P).
\end{aligned}$$

By using the result of the Lemma 4:

$$\min\left\{c(X_\sigma),c(X_\rho)\right\}\leq 2c(T).$$

Then, for one of the nodes x_σ or x_ρ in T, any shortest-path spanning tree rooted at it guarantees the result of the lemma. □

Suppose T^* is an optimal spanning tree for an instance of WSDOCT, by Lemma 5 there exists a node u such that any shortest-path spanning tree rooted at u is a 2-approximation for T^*. Then, obtaining a shortest-path spanning tree for each node of G and selecting a minimum among them, guarantees to find a 2-approximation for WSDOCT. Since we analyzed before, such algorithm has $O(n^2 \log n + mn)$ time complexity.

Theorem 2. *There exists a* 2*-approximation algorithm for* WSDOCT *with time complexity* $O(n^2 \log n + mn)$.

5. Conclusions

In this work we present a PTAS for metric-WSDOCT and also a 2-approximation for WSDOCT, both problems are NP-hard particular cases of OCT. The best previously known result for these problems was a $O(\log(n))$-approximation algorithm due to [12,17]. This result generalizes previous results for particular cases of the studied problems, these are the cases of MRCT, PROCT and SROCT ([11,16,17]). Many questions remain open regarding OCT and related problems. One could improve the approximation ratio for WSDOCT or other particular case of OCT. One could obtain an inapproximability result for OCT and some of its particular cases. In future works we will attempt to answer these questions for some of these problems.

References

[1] Y. Bartal, Probabilistic approximation of metric spaces and its algorithmic applications, in: Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science, 1996, pp. 184–1963.
[2] D. Bilò, L. Gualà, G. Proietti, Finding best swap edges minimizing the routing cost of a spanning tree, Algorithmica 68 (2014) 337–357.
[3] L. Chen, S. Hsieh, L. Hung, R. Klasing, Approximation algorithms for the p-hub center routing problem in parameterized metric graphs, in: Combinatorial Algorithms. IWOCA 2018, in: Lecture Notes in Computer Science, vol. 10979, 2018, pp. 115–127.
[4] E.W. Dijkstra, A note on two problems in connexion with graphs, Numer. Math. 1 (1959) 269–271.
[5] A.M. Farley, P. Fragopoulou, D. Krumme, A. Proskurowski, D. Richards, Multi-source spanning tree problems, J. Interconnect. Netw. 1 (2000) 61–71.
[6] T.C. Hu, Optimum communication spanning trees, SIAM J. Comput. 3 (1974) 188–195.
[7] D.S. Johnson, J.K. Lenstra, A.H.G. Rinnooy Kan, The complexity of the network design problem, Networks 8 (1978) 279–285.
[8] T. Kim, S. Chung Seo, D. Kim, Distributed formation of degree constrained minimum routing cost tree in wireless ad-hoc networks, J. Parallel Distrib. Comput. 83 (2015) 143–158.
[9] B.J. Orlin, A faster strongly polynomial minimum cost flow algorithm, Oper. Res. 41 (1993) 338–350.
[10] S.V. Ravelo, C.E. Ferreira, Ptas's for Some Metric p-Source Communication Spanning Tree Problems, WALCOM 2015, 2015.
[11] S.V. Ravelo, C.E. Ferreira, A ptas for the metric case of the minimum sum-requirement communication spanning tree problem, Discrete Appl. Math. 228 (2017) 158–175.
[12] K. Talwar, J. Fakcharoenphol, S. Rao, A tight bound on approximating arbitrary metrics by tree metrics, in: Proceedings of the 35th Annual ACM Symposium on Theory of Computing, 2003, pp. 448–455.
[13] B. Wu, C.Y. Hsiao, K.M. Chao, The swap edges of a multiple-sources routing tree, Algorithmica 50 (2008) 299–311.
[14] B.Y. Wu, A polynomial time approximation scheme for the two-source minimum routing cost spanning trees, J. Algorithms 44 (2002) 359–378.
[15] B.Y. Wu, K.M. Chao, C.Y. Tang, Approximation algorithms for some optimum communication spanning tree problems, Discrete Appl. Math. 102 (2000) 245–266.
[16] B.Y. Wu, K.M. Chao, C.Y. Tang, A polynomial time approximation scheme for optimal product-requirement communication spanning trees, J. Algorithms 36 (2000) 182–204.
[17] B.Y. Wu, G. Lancia, V. Bafna, K.M. Chao, R. Ravi, C.Y. Tang, A polynomial time approximation scheme for minimum routing cost spanning trees, SIAM J. Comput. 29 (2000) 761–778.