PTAS's for Some Metric p-source Communication Spanning Tree Problems*

Santiago V. Ravelo and Carlos E. Ferreira

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

Abstract. In this work we consider some NP-hard cases of the metric p-source communication spanning tree problem (metric p-OCT). Given an undirected complete graph $G = (V, E)$ with non-negative length $\omega(e)$ associated to each edge $e \in E$ satisfying the triangular inequality, a set $S \subseteq V$ of p vertices and non-negative routing requirements $\psi(u, v)$ between all pairs of nodes $u \in S$ and $v \in V$, the metric p-OCT's objective is to find a spanning tree T of G, that minimizes: $\sum_{u \in S} \sum_{v \in V} \psi(u, v) d(T, u, v)$, where $d(H, x, y)$ is the minimum distance between nodes x and y in a graph $H \subseteq G$. This problem is a particular case of the optimum communication spanning tree problem (OCT). We prove a general result which allows us to derive polynomial approximation schemes for some NP-hard cases of the metric p-OCT improving the existing ratios for these problems.

1 Introduction

In this work we consider NP-hard particular cases of the metric p-source optimum communication spanning tree problem (metric p-OCT) which is a particular case of the optimum communication spanning tree problem (OCT). In the OCT, introduced by Hu in 1974 ([Hu, 1974, Wu and Chao, 2004]), the input is an undirected graph $G = (V, E)$ with non-negative lengths $\omega(e)$ associated to each edge $e \in E$ and non-negative requirements $\psi(u, v)$ 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 given by $C(T) = \sum_{u \in V} \sum_{v \in V} \psi(u, v) d(T, u, v)$, where $d(H, x, y)$ denotes the minimum distance between the nodes x and y in the subgraph H of G. In the p-OCT it is additionally given a set of p nodes (sources) $S \subseteq V$, that are considered in the objective function: $C(T) = \sum_{u \in S} \sum_{v \in V} \psi(u, v)$ $d(T, u, v)$.

In [Johnson et al., 1978] it was proven, by a reduction from the 3-exact cover problem (3-EC), that the minimum routing cost spanning tree problem (MRCT) is NP-hard. MRCT is a particular case of OCT where for each pair of nodes the communication requirement between them is equal to one, i.e., $\psi(u, v) = 1$ for all $u, v \in V$. In [Wu et al., 2000c] a PTAS was given for the MRCT. Also in this

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

M.S. Rahman and E. Tomita (Eds.): WALCOM 2015, LNCS 8973, pp. 137–148, 2015.

work a reduction from the general case to the metric one is presented and it was proven that MRCT with edge-lengths that satisfy the triangular inequality is also NP-hard. Also, in [Wu et al., 2000c] a $O(log^2(n))$-approximation for OCT was presented applying a result from [Bartal, 1996] which was improved to a $O(log(n))$-approximation by [Talwar et al., 2003].

A particular case of p-OCT is weighted p-MRCT introduced in [Wu, 2002]. In this case a non-negative sending requirement $\sigma(u)$ is given for each source $u \in S$, and the requirement between a source $u \in S$ and a node $v \in V$ is $\psi(u, v) = \sigma(u)$. When the sending requirements σ for the p sources are equal to one, the problem is called p-MRCT. In [Wu, 2002] it was proven that 2-MRCT is NP-hard, moreover PTASs were shown for 2-MRCT and for the metric case of weighted 2-MRCT.

In [Wu et al., 2000a] was introduced the minimum sum-requirement communication spanning tree problem (SROCT), in which the sending requirement $\sigma(u)$ is given for each $u \in V$. Observe that weighted p-MRCT is a particular case of SROCT, where $\sigma(u) = u$ for all $u \in V - S$. Also, the authors presented a 2-approximation algorithm for this problem.

In [Wu, 2004] a 2-approximation algorithm was given for the metric p-OCT and a 3-approximation for the general 2-OCT.

In this work we present a general result that allows us to derive PTASs for natural special cases of p-OCT. We introduce three particular NP-hard cases of p-OCTand we derive PTASs for the metric cases of these problems. Also, we give new PTASs for the metric p-MRCT and for the fixed parameter weighted p-MRCT.

This work is organized as follows. In section 2, we introduce some definitions. In section 3 a polynomial time algorithm to find an optimum k-star for a given k is presented. In section 4 we prove the main result of this work, which allows us to obtain approximation algorithms for different metric p-OCT problems. In section 5 we introduce particular cases of p-OCT giving a PTAS for the metric case of that problem. We finish the paper in section 6 with some conclusions.

2 Definitions

Unless specified all graphs in this work are undirected. 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 the context we use V or E instead of V_G or E_G and $n = |V|$). Also, when G has 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 a sub-graph H 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, for each positive integer p we can define the p-OCT as:

Problem 1. p-OCT - p-source Optimum Communication spanning Tree problem.

Input: A graph G, a non-negative length function over the edges of G, $\omega : E \to \mathbb{Q}_+$, a set of p sources $S \subseteq V$ and a non-negative routing requirement function from each source node to each node of V, $\psi : S \times V \to \mathbb{Q}_+$.

Output: A spanning tree T of G which minimizes the total requirement routing cost: $C(T) = \sum_{u \in S} \sum_{v \in V} \psi(u, v) d(T, u, v)$.

This paper considers the metric p-OCT, which is the particular case of p-OCT where the graph G is complete and the length function over the edges satisfies the triangular inequality. To find a feasible solution of the metric p-OCT we use a valid k-star[1]:

Definition 1. *Given a graph G, a set S of nodes of G and an integer $k \geq |S|$, 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 core τ is* **valid** *for the set S of nodes (or just valid) if τ contains all the nodes of S. Then, a* **valid** k**-star** *of G is a k-star of G with at least a valid core.*

The problem of finding an optimal valid k-star for the metric p-OCT can be defined as follows. Also, in section 3, we show how to solve it efficiently.

Problem 2. Optimum valid k-star for the metric p-OCT.

Input: A positive integer k and an instance of metric p-OCT: a complete graph G, a non-negative length function over the edges of G which satisfies the triangular inequality, $\omega : E \to \mathbb{Q}_+$, a set of p sources $S \subseteq V$ and a non-negative routing requirement function between each node of S and each node of V, $\psi : S \times V \to \mathbb{Q}_+$. (Notice that $k \geq p = |S|$)

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

3 Optimal Valid k-star for Metric p-OCT

First note that if k and p are constants, the number of possible valid cores of k-stars of G is polynomial. Indeed, since the core of a valid k-star must contain the p vertices in S and $k-p$ other vertices, one can enumerate the $\binom{n}{k-p} = O(n^{k-p})$ possibilities. For each different choice one has to enumerate all possible trees with k vertices and there are $O(k^k)$ possible trees. Then, the number of all possible valid cores is limited by $O(k^k n^{k-p})$. Our approach is to find an optimal valid k-star with core τ for each valid core τ, selecting the minimum k-star among them.

[1] The definition of k-star used in this paper is similar to the one used by [Wu et al., 2000c, Wu et al., 2000a, Wu et al., 2000b], 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).

Given a valid core τ, to obtain a valid k-star T with core τ each node of $V_G - V_\tau$ must be adjacent to some node of τ (i.e., these nodes will be leaves of T).

Let u_v be the node of τ adjacent to $v \in V_G - V_\tau$ in T. Since all nodes in S belong to τ then: $C(T) = C(\tau) + \sum_{v \in V_G - V_\tau} \sum_{w \in S} \psi(w,v)(d(\tau,w,u_v) + \omega(u_v,v))$.

Thus, in order to find the best vertex of the core τ to link each vertex $v \in V - V_\tau$ it suffices to consider the node $u_v^* \in V_\tau$ that minimizes:

$$\sum_{w \in S} \psi(w,v)(d(\tau,w,u_v) + \omega(u_v,v)).$$

To compute it efficiently, first we pre-calculate for each $w \in S$ and $u \in V_\tau$ all the distances $d(\tau,w,u)$ (it can be done in $O(|S|k) = O(pk)$). After that, we calculate for each pair of nodes $v \in V_G - V_\tau$ and $u \in V_\tau$ ($k(n-k)$ pairs) the cost of linking v to vertex u in τ which can be computed in $O(|S|) = O(p)$ using the pre-calculated distances. Therefore, we can obtain an optimal valid k-star with core τ in $O(k(n-k)p + pk) = O(npk)$ time.

From the ideas above, we conclude that it is possible to find and optimal valid k-star in $O(k^{k+1}n^{k-p+1}p)$ time.

Lemma 1. *An optimum valid k-star for metric p-OCT with fixed $k \geq p$ can be found in $O(n^{k-p+1})$ time.*

4 Approximation Lemma (for Metric *p*-OCT Problems)

In this section we present the main result of the paper. First we introduce the notion of δ-balanced-path, $0 < \delta \leq \frac{1}{2}$. This definition is based on similar concepts for related problems introduced in [Wu et al., 2000c], [Wu et al., 2000a] and [Wu et al., 2000b]. Using it we derive a general lemma that applies for different special cases of metric p-OCT.

Definition 2. *Given a spanning tree T of G, we denote by S_T the minimal subtree of T which contains all the nodes in S. It is easy to see that every leaf of S_T must be a node of S.*

Definition 3. *We define $\psi(S',U) = \sum_{u \in S'} \sum_{v \in U} \psi(u,v)$ for every $S' \subseteq S$ and $U \subseteq V$.*

Definition 4. *Given a spanning tree T of G and a path $P = w_1, ..., w_h$ of T, we denote:*

- *$f_P = w_1$ and $l_P = w_h$ the endpoints of P;*
- *V_P^f: the set of nodes in T connected to P through vertex f_P (including f_P itself);*
- *V_P^m: the set of nodes in T connected to P through an internal node of P (including these nodes);*
- *V_P^l: the set of nodes in T connected to P through vertex l_P (including l_P itself).*

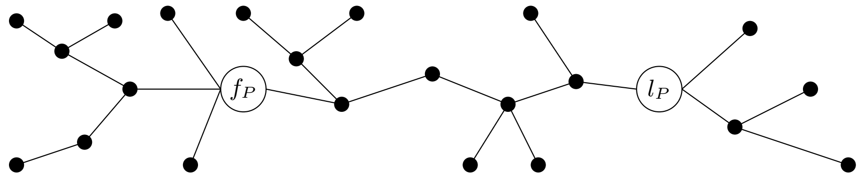

Fig. 1. Example of V_P^f, V_P^m and V_P^l for a path P of a tree T. Observe that V_P^f is the set of nodes to the left of f_P (including f_P), V_P^l is the set of nodes to the right of l_P (including l_P), V_P^m is the set containing the rest of the nodes and P is the path connecting f_P to l_P in T.

Notice that $V_P^f \cup V_P^m \cup V_P^l = V$ and these sets are disjoint. We also denote by S_P^i the set of vertices in $V_P^i \cap S$, where $i \in \{f, m, l\}$. We say that P is m**-source-free** *if $S_P^m = \emptyset$. If P is m-source-free, $S_P^f \neq \emptyset$ and $S_P^l \neq \emptyset$ we say that P is a* **connecting-source** *path.*

Now we introduce the definition for δ-balanced-path, that is an m-source-free path (P) for which the routing requirement delivered to its interior (V_P^m) is small, i.e., at most a portion (δ) of the routing requirement that passes through P. Formally:

Definition 5. *Given $0 < \delta \leq \frac{1}{2}$ and a spanning tree T of G, an m-source-free path P of T is a* δ**-balanced-path** *if $\psi(S_P^f \cup S_P^l, V_P^m) = \psi(S, V_P^m) \leq \delta\left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f)\right)$.*

The following proposition gives a basis for the main result in this work, since it provides a valid star whose cost is bounded by the cost of the given tree.

Proposition 1. *Consider $0 < \delta \leq \frac{1}{2}$, a spanning tree T of G and a set Y of internally disjoint δ-balanced-paths of T whose union results in S_T. Then there exists a valid $(|Y| + 1)$-star X such that: $C(X) \leq (1 + 2\delta)\, C(T)$. Where two paths are internally disjoint if the intersection of their sets of nodes is empty or contains nodes that are endpoints of both paths.*

Proof. First we note that there exists a spanning tree $\overline{T}$ such that $S_{\overline{T}} = S_T$, all vertices of $\overline{T}$ that do not belong to S_T are leaves and $C(\overline{T}) \leq C(T)$. If all vertices of T that do not belong to S_T are leaves, then $\overline{T} = T$, otherwise there exists a leaf u in T that is not adjacent to any node of S_T. Let v be the nearest node to u in S_T. Since all the nodes of S are in S_T and the graph is metric, the spanning tree T' of G, resulting from removing the edge adjacent to u in T and adding the edge (v, u) satisfies $C(T') \leq C(T)$. It is easy to see that repeating this process to all the leaves that are not adjacent to a node in S_T we obtain tree $\overline{T}$. With this property we can suppose that in tree T all vertices in $V - S_T$ are leaves. Now we construct a valid $(|Y| + 1)$-star X of G as follows:

- The core τ of X has the set of nodes that are endpoints of the paths in Y. Two nodes $u, v \in \tau$ are adjacent in τ if in Y there exists a path with endpoints u and v. Since the paths in Y are internally disjoint and their union results in the tree S_T, we conclude that τ is a tree over the endpoints of the paths in Y.

- For every node $u \in \tau$ and for every leaf $v \in V - \tau$ adjacent to u in T we also include an edge (u, v) in X.
- Observe that each node $u \in T$ not included in X by the previous steps belongs to V_P^m for some path $P \in Y$. Then, we include edge (u, f_P) in X if $\omega(u, f_P) \leq \omega(u, l_P)$, otherwise we include edge (u, l_P) in X.

Our construction guarantees X to be a $(|Y| + 1)$-star of G with core τ. Then, we only need to analyze its associated communication cost, which can be calculated by adding over each edge e the communication amount passing over e times the length of e. Since every edge is a path with exactly two vertices we also use the notation given by definition 4 on the edges. Then:

$$\begin{aligned} C(X) &= \sum_{e \in E_X} \left(\psi(S_e^f, V_e^l) + \psi(S_e^l, V_e^f)\right) \omega(e) \\ &= \sum_{e \in E_\tau} \left(\psi(S_e^f, V_e^l) + \psi(S_e^l, V_e^f)\right) \omega(e) \\ &\quad + \sum_{e \in E_{X-\tau}} \left(\psi(S_e^f, V_e^l) + \psi(S_e^l, V_e^f)\right) \omega(e). \end{aligned}$$

Observe that, by construction, each edge $e \in E_\tau$ corresponds to a δ-balanced-path $P \in Y$, such that: $\psi(S_e^f, V_e^l) + \psi(S_e^l, V_e^f) \leq \psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)$. Also, by the triangular inequality: $\omega(e) \leq \omega(P)$, then:

$$\begin{aligned} C(X) \leq &\sum_{P \in Y} \Big(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\Big) \omega(P) \\ &+ \sum_{e \in E_{X-\tau}} \left(\psi(S_e^f, V_e^l) + \psi(S_e^l, V_e^f)\right) \omega(e). \end{aligned}$$

Notice that for every edge $e \in E_{X-\tau}$ one of its endpoints is a leaf outside of τ and the other one a node of τ. Let $p(u)$ be the node in τ adjacent in X to a leaf $u \in V - \tau$, then:

$$\begin{aligned} C(X) &\leq \sum_{P \in Y} \Big(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\Big) \omega(P) \\ &\quad + \sum_{u \in V_{X-\tau}} \psi(S, u) \omega(u, p(u)) \\ &\leq \sum_{P \in Y} \Big(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\Big) \omega(P) \\ &\quad + \sum_{P \in Y} \sum_{u \in V_P^m} \psi(S, u) \omega(u, p(u)) + \sum_{u \in V_I} \psi(S, u) \omega(u, p(u)), \end{aligned}$$

where V_I is the set of leaves in X that are not in τ and have the same adjacent node in both trees T and X. Notice that all leaves in T outside of τ adjacent (in

T) to some node of τ belong to V_I (that is, only nodes in V_P^m for some $P \in Y$ may be out of $V_I \cup V_\tau$).

For every $P \in Y$ and every node $u \in V_P^m$, $p(u)$ is one of the endpoints of the path P. Thus $\omega(u, p(u)) = \min\{\omega(u, f_P), \omega(u, l_P)\}$:

$$C(X) \leq \sum_{P \in Y} \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\right) \omega(P) + \sum_{P \in Y} \sum_{u \in V_P^m} \psi(S, u) \min\{\omega(u, f_P), \omega(u, l_P)\} + \sum_{u \in V_I} \psi(S, u)\omega(u, p(u)).$$

Let $q(u)$ be the node in S_T adjacent to $u \in T - S_T$ and for $u \in S_T$ consider $q(u) = u$, then:

$$C(X) \leq \sum_{P \in Y} \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\right) \omega(P) + \sum_{P \in Y} \sum_{u \in V_P^m} \psi(S, u) \min\{\omega(u, f_P), \omega(u, l_P)\} + \sum_{u \in V_I} \psi(S, u)\omega(u, q(u)).$$

By the triangular inequality $\omega(u, f_P) \leq \omega(q(u), f_P) + \omega(u, q(u))$ for $P \in Y$ and $u \in V_P^m$ (the same applies for l_P). Since $q(u) \in P$, $\omega(q(u), f_P) \leq \omega(P)$ and $\omega(q(u), l_P) \leq \omega(P)$. So, $\min\{\omega(u, f_P), \omega(u, l_P)\} \leq \omega(P) + \omega(u, q(u))$. Then:

$$\begin{aligned}
C(X) &\leq \sum_{P \in Y} \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\right) \omega(P) \\
&\quad + \sum_{P \in Y} \sum_{u \in V_P^m} \psi(S, u) \left(\omega(P) + \omega(u, q(u))\right) + \sum_{u \in V_I} \psi(S, u)\omega(u, q(u)) \\
&= \sum_{P \in Y} \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + \psi(S, V_P^m)\right) \omega(P) \\
&\quad + \sum_{P \in Y} \omega(P)\psi(S, V_P^m) + \sum_{u \in V_{T-\tau}} \psi(S, u)\omega(u, q(u)) \\
&\leq \sum_{P \in Y} \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f) + 2\psi(S, V_P^m)\right) \omega(P) \\
&\quad + \sum_{u \in V_{X-\tau}} \psi(S, u)\omega(u, q(u)).
\end{aligned}$$

Since each $P \in Y$ is δ-balanced, $\psi(S, V_P^m) \leq \delta \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f)\right)$. Also, $\omega(u, q(u)) = 0$ for each node u in S_T, then:

$$\begin{aligned}
C(X) &\leq \sum_{P \in Y} (1 + 2\delta) \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f)\right) \omega(P) \\
&\quad + \sum_{u \in V_{T-S_T}} \psi(S, u)\omega(u, q(u)) \\
&\leq (1 + 2\delta)\, C(T).
\end{aligned}$$

□

The above lemma gives us a general result: given $0 < \delta \leq \frac{1}{2}$, if a metric p-OCT problem Π satisfies that for any spanning tree T of G there exists a set of k internally disjoint δ-balanced-paths whose union results in S_T, then for each T there exists a $(k+1)$-star X_T such that $C(X_T) \leq (1+2\delta)C(T)$. Thus, for an optimal spanning tree T^* there exists a valid $(k+1)$-star of G which is a $(1+2\delta)$-approximation for Π. Then, an optimal valid $(k+1)$-star of G is a $(1+2\delta)$-approximation for Π. Remember that, by lemma 1, we can find an optimal valid $(k+1)$-star in $O(n^{k-p+2})$ (and therefore an approximation for the optimum value).

Lemma 2. *Consider a metric p-OCT problem for which we can guarantee for every spanning tree T the existence of k internally disjoint δ-balanced-paths whose union results in S_T. Thus, there exists a $(1+2\delta)$-approximation algorithm with time complexity $O(n^{k-p+2})$ where $0 < \delta \leq \frac{1}{2}$.*

The following proposition provides a sufficient condition for the existence of an internally disjoint δ-balanced-path set whose union results in S_T for any spanning tree T of a metric p-OCT problem.

Proposition 2. *Consider a metric p-OCT problem. Given $0 < \delta \leq \frac{1}{2}$, if it satisfies that every connecting-source path of any spanning tree T of G can be divided in, at most k internally disjoint δ-balanced-paths, then there exists a set of internally disjoint δ-balanced-paths of T with at most $2k(p-1)$ elements whose union results in S_T.*

Proof. Let T be a spanning tree of G, and consider the corresponding sub-tree S_T, whose leaves are nodes of S (S_T has at most p leaves), and S_T is also a tree. Then, the number of nodes with degree greater than two in S_T is at most $p-1$.

Consider now the collection of paths Y constructed from S_T such that each path has as endpoints nodes with degree different of two in S_T or nodes of S, and all the internal nodes of these paths are not in S and has degree two in S_T. It is easy to see that the number of paths in Y is at most $2p-2$ (that is, the number of nodes in S which contains the leaves of S_T plus the number of nodes with degree greater than two in S_T minus one). Also, observe that each path in Y is a connecting-source path, so it can be divided in at most k internally disjoint δ-balanced-paths. Then, by applying that division we obtain a set of δ-balanced-paths with at most $2k(p-1)$ elements whose union results in S_T. □

Finally, from the results given by lemma 2 and proposition 2 we conclude:

Lemma 3. *Consider a metric p-OCT problem. If, for this problem it holds that for every connecting-source path P of any spanning tree of G there exists at most k internally disjoint δ-balanced-paths whose union results in P, there exists a $(1+2\delta)$-approximation algorithm with time complexity $O(n^{(2k-1)(p-1)+1})$ where $0 < \delta \leq \frac{1}{2}$.*

In next section we use this result to prove PTASs for three different metric p-OCT problems, for which the condition given above holds.

5 PTAS for Three p-OCT Metric Problems

In the previous section we provide a general result to obtain good approximations for p-OCT metric problems. In this section we show some special cases for which lemma 3 applies and consequently we are able to provide PTASs.

The first problem we introduce is the p-source Weighted Source Destination Optimum Communication spanning Tree problem (p-WSDOCT) which is a particular case of the p-OCT, where there is a sending requirement for each source node and also all the nodes have a receiving requirement associated:

Problem 3. p-WSDOCT - p-source Weighted Source Destination Optimum Communication spanning Tree problem.

Input: A graph G, a non-negative length function over the edges of G, $\omega : E \to \mathbb{Q}_+$, a set of p sources $S \subseteq V$, a positive sending requirement $\sigma : S \to \mathbb{Q}_+$ and a non-negative receiving requirement $\lambda : V \to \mathbb{Q}_+$. The requirement function between $u \in S$ and $v \in V$ is given by $\psi(u, v) = \sigma(u)\lambda(v)$.

Output: A spanning tree T of G which minimizes the total requirement routing cost: $C(T) = \sum_{u\in S}\sum_{v\in V} \sigma(u)\lambda(v)d(T, u, v)$.

Now, using the result of lemma 3 we prove the following theorem:

Theorem 1. *There exists a* PTAS *for metric* p-WSDOCT *with fixed parameter. Given* $0 < \delta \leq \frac{1}{2}$, *the algorithm can find a* $(1 + 2\xi\delta)$-*approximation in* $O\left(n^{\left(2\frac{1}{\delta-\delta^2}-1\right)(p-1)+1}\right)$ *time complexity where* $\xi = \frac{\sum_{u\in S}\sigma(u)}{\min\{\sigma(u)\}_{u\in S}}$.

Proof. Consider a connecting-source path P for any spanning tree T of G, where:

$$\sum_{u\in V_P^m} \lambda(u) \leq \delta\,(1-\delta) \sum_{u\in V} \lambda(u).$$

Observe that:

$$\begin{aligned}
\psi(S, V_P^m) &= \left(\sum_{u\in S}\sigma(u)\right)\left(\sum_{u\in V_P^m}\lambda(u)\right) \leq \left(\sum_{u\in S}\sigma(u)\right)\delta\,(1-\delta)\sum_{u\in V}\lambda(u)\\
&\leq \left(\sum_{u\in S}\sigma(u)\right)\delta\left(1-(\delta-\delta^2)\right)\sum_{u\in V}\lambda(u)\\
&= \left(\sum_{u\in S}\sigma(u)\right)\delta\left(\sum_{u\in V}\lambda(u) - (\delta-\delta^2)\sum_{u\in V}\lambda(u)\right)\\
&\leq \left(\sum_{u\in S}\sigma(u)\right)\delta\left(\sum_{u\in V}\lambda(u) - \sum_{u\in V_P^m}\lambda(u)\right).
\end{aligned}$$

Since $V = V_P^l \cup V_P^f \cup V_P^m$:

$$\psi(S, V_P^m) \leq \left(\sum_{u\in S}\sigma(u)\right)\delta\left(\sum_{u\in V_P^l}\lambda(u) + \sum_{u\in V_P^f}\lambda(u)\right)$$

$$\leq \left(\sum_{u \in S} \sigma(u)\right) \delta \frac{\sum_{u \in S_P^f} \sigma(u)}{\min\{\sigma(u)\}_{u \in S}} \sum_{u \in V_P^l} \lambda(u)$$

$$+ \left(\sum_{u \in S} \sigma(u)\right) \delta \frac{\sum_{u \in S_P^l} \sigma(u)}{\min\{\sigma(u)\}_{u \in S}} \sum_{u \in V_P^f} \lambda(u)$$

$$= \frac{\sum_{u \in S} \sigma(u)}{\min\{\sigma(u)\}_{u \in S}} \delta \left(\psi(S_P^f, V_P^l) + \psi(S_P^l, V_P^f)\right).$$

So, P is a $\frac{\sum_{u \in S} \sigma(u)}{\min\{\sigma(u)\}_{u \in S}} \delta$-balanced-path and every connecting-source path of T can be divided in at most $\frac{1}{\delta - \delta^2}$ paths like P, which is a connecting-source path that satisfies $\sum_{u \in V_P^m} \lambda(u) \leq \delta (1 - \delta) \sum_{u \in V} \lambda(u)$.

Then, by applying lemma 3 we conclude that for metric p-WSDOCT there exists a $\left(1 + 2\frac{\sum_{u \in S} \sigma(u)}{\min\{\sigma(u)\}_{u \in S}} \delta\right)$-approximation algorithm with time complexity $O\left(n^{\left(2\frac{1}{\delta - \delta^2} - 1\right)(p-1)+1}\right)$. □

Another particular case of p-WSDOCT is the p-source Weighted Destination Optimum Communication spanning Tree problem (p-WDOCT), where all vertices of S have unitary sending requirement ($\sigma(S) = 1$).

Problem 4. p-WDOCT - p-source Weighted Destination Optimum Communication spanning Tree problem.

Input: A graph G, a non-negative length function over the edges of G, $\omega : E \to \mathbb{Q}_+$, a set of p sources $S \subseteq V$ and a non-negative receiving requirement $\lambda : V \to \mathbb{Q}_+$. The requirement function between $u \in S$ and $v \in V$ is given by $\psi(u, v) = \lambda(v)$.

Output: A spanning tree T of G which minimizes the total requirement routing cost: $C(T) = \sum_{u \in S} \sum_{v \in V} \lambda(v) d(T, u, v)$.

Observe that p-WDOCT is a particular case of p-WSDOCT in which:

$$\xi = \frac{\sum_{u \in S} \sigma(u)}{\min\{\sigma(u)\}_{u \in S}} = \frac{p}{1} = p.$$

Then, using theorem 1 we conclude the following result for p-WDOCT:

Corollary 1. *There exists a* PTAS *for metric* p-WDOCT*, which can find a* $(1 + 2p\delta)$*-approximation in* $O\left(n^{\left(2\frac{1}{\delta - \delta^2} - 1\right)(p-1)+1}\right)$ *time complexity where* $0 < \delta \leq \frac{1}{2}$.

Notice that the weighted p-MRCT is also a particular case of p-WSDOCT and the p-MRCT is a particular case of p-WDOCT so the results above imply also new PTASs for the fixed parameter metric weighted p-MRCT and for the metric p-MRCT.

Table 1. Comparison of our approach with current state of art in the literature. As it can be seen, for almost all the problems our approach improve the previous approximation ratios, only for the case of 2-MRCT it does not give a better solution (we obtain the same approximation scheme but with a greater complexity time). In the case of the PTASs it was considered the approximation $(1+\epsilon)$ being $\epsilon \in (0,1]$.

Problem	Best Approach In Literature			Our Approach	
	Ratio	Complexity	References	Ratio	Complexity
2-MRCT	PTAS	$\mathbf{O}\left(\mathbf{n}^{\frac{1}{\epsilon}+1}\right)$	[Wu, 2002]	PTAS	$O\left(n^{\frac{8}{\epsilon(2-\epsilon)}}\right)$
p-MRCT	2	$O\left(n^3\right)$	[Wu et al., 2000a]	**PTAS**	$O\left(n^{\left(\frac{8p^2}{\epsilon(2p-\epsilon)}-1\right)(p-1)+1}\right)$
weighted p-MRCT	2	$O\left(n^3\right)$	[Wu et al., 2000a]	**PTAS**	$O\left(n^{\left(\frac{8\xi^2}{\epsilon(2\xi-\epsilon)}-1\right)(p-1)+1}\right)$
p-WDOCT	2	$O\left(n^{p-1}\right)$	[Wu, 2004]	**PTAS**	$O\left(n^{\left(\frac{8p^2}{\epsilon(2p-\epsilon)}-1\right)(p-1)+1}\right)$
fixed parameter p-WSDOCT	2	$O\left(n^{p-1}\right)$	[Wu, 2004]	**PTAS**	$O\left(n^{\left(\frac{8\xi^2}{\epsilon(2\xi-\epsilon)}-1\right)(p-1)+1}\right)$
p-USCOCT	2	$O\left(n^{p-1}\right)$	[Wu, 2004]	**PTAS**	$O\left(n^{\left(4log_{\frac{2}{2+\epsilon}}(r)-1\right)(p-1)+1}\right)$

Another particular case of p-OCT in which we can apply our results is the p-Uniform Source Connecting Optimum Communication spanning Tree problem (p-USCOCT). In this problem the minimum routing requirement between two sources must be at least a fixed ratio of the sum of requirements between all the pairs of nodes[2]:

Problem 5. p-USCOCT - p-Uniform Source Connecting Optimum Communication spanning Tree problem.

Input: A graph G, a non-negative length function over the edges of G, $\omega : E \to \mathbb{Q}_+$, a set of p sources $S \subseteq V$, a fixed ratio $r > 0$ and a non-negative routing requirement function between each node of S and each node of V, $\psi : S \times V \to \mathbb{Q}_+$, where each pair of sources $u, v \in S$ satisfies $\psi(u,v) \geq r \sum_{u \in S} \sum_{v \in V} \psi(u,v)$.

Output: A spanning tree T of G that minimizes the total requirement routing cost: $C(T) = \sum_{u \in S} \sum_{v \in V} \psi(u,v) d(T,u,v)$.

Similar ideas allow us to prove that for $0 < \delta \leq \frac{1}{2}$ we can divide any connecting-source path P of a spanning tree of G in an instance of metric p-USCOCT with no more than $2log_{\frac{1}{1+\delta}}(r)$ internally disjoint δ-balanced-paths. Then by applying lemma 3 we obtain:

[2] Metric p-USCOCT is NP-hard, a reduction from SAT is given in the full version of the paper.

Theorem 2. *There exists a* PTAS *for the metric* p-USCOCT*. Given* $0 < \delta \leq \frac{1}{2}$*, the algorithm guarantees a* $(1 + 2\delta)$*-approximation of the optimum solution in time complexity* $O\left(n^{\left(4log_{\frac{1}{1+\delta}}(r)-1\right)(p-1)+1}\right)$.[3]

6 Conclusions

In this work we consider different NP-hard variants of p-OCT: metric case of p-MRCT, p-WDOCT fixed parameter WSDOCTand p-USCOCT. We prove a lemma that allows us to present PTAS's for these problems, being possible to use that result in order to obtain other approximations for some other particular cases of p-OCT. Also we prove that metric USCOCT is NP-hard. In table 1 we summarize our approaches and compare with previous results in literature. Many questions remain open regarding p-OCT and related problems. For example, no PTAS for metric p-OCT is known. Also, when we do not consider metric problems much is still to be researched.

References

[Bartal, 1996] 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)

[Hu, 1974] Hu, T.C.: Optimum communication spanning trees. SIAM J. Comput. 3(3), 188–195 (1974)

[Johnson et al., 1978] Johnson, D.S., Lenstra, J.K., Kan, A.H.G.R.: The complexity of the network design problem. Networks 8, 279–285 (1978)

[Talwar et al., 2003] 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)

[Wu, 2002] Wu, B.Y.: A polynomial time approximation scheme for the two-source minimum routing cost spanning trees. J. Algorithms 44, 359–378 (2002)

[Wu, 2004] Wu, B.Y.: Approximation algorithms for the optimal p-source communication spanning tree. Discrete and Applied Mathematics 143, 31–42 (2004)

[Wu and Chao, 2004] Wu, B.Y., Chao, K.M.: Spanning Trees and Optimization Problems. Chapman & Hall / CRC (2004) ISBN: 1584884363

[Wu et al., 2000a] 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)

[Wu et al., 2000b] 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)

[Wu et al., 2000c] 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)

[3] The complete proof of this theorem can be found in the full version of the paper.