A note on optimal communication spanning trees

Bang Ye Wu   Kun–Mao Chao

April 1, 2004

The *optimal communication spanning tree* (OCT) problem is defined as follows. Let $G = (V, E, w)$ be an undirected graph with nonnegative edge length function w. We are also given the requirements $\lambda(u, v)$ for each pair of vertices. For any spanning tree T of G, the communication cost between two vertices is defined to be the requirement multiplied by the path length of the two vertices on T, and the communication cost of T is the total communication cost summed over all pairs of vertices. Our goal is to construct a spanning tree with minimum communication cost. That is, we want to find a spanning tree T such that $\sum_{u,v \in V} \lambda(u, v) d_T(u, v)$ is minimized.

The requirements in the OCT problem are arbitrary nonnegative values. By restricting the requirements, several special cases of the problem were defined in the literature. We list the problems in the following, in which $r : V \to Z_0^+$ is a given vertex weight function and $S \subset V$ is a set of k vertices given as sources.

- $\lambda(u, v) = 1$ for each $u, v \in V$: This version is the Minimum Routing Cost Spanning Tree (MRCT) problem discussed in the previous chapter.
- $\lambda(u, v) = r(u)r(v)$ for each $u, v \in V$: This version is called the Optimal Product-Requirement Communication Spanning Tree (PROCT) problem.
- $\lambda(u, v) = r(u)+r(v)$ for each $u, v \in V$: This version is called the Optimal Sum-Requirement Communication Spanning Tree (SROCT) problem.
- $\lambda(u, v) = 0$ if $u \notin S$: This version is called the p-Source OCT (p-OCT) problem. In other words, the goal is to find a spanning tree minimizing $\sum_{u \in S} \sum_{v \in V} \lambda(u, v) d_T(u, v)$.
- $\lambda(u, v) = 1$ if $u \in S$, and $\lambda(u, v) = 0$ otherwise: This version is called the p-Source MRCT (p-MRCT) problem. In other words, the goal is to find a spanning tree minimizing $\sum_{u \in S} \sum_{v \in V} d_T(u, v)$.

We define two communication costs and notations for the PROCT and the SROCT problems.

Definition 1: The *product-requirement communication* (or p.r.c. in abbreviation) cost of a tree T is defined by $C_p(T) = \sum_{u,v} r(u)r(v)d_T(u, v)$.

When there are more than one vertex weight functions, we shall use $C_p(T, r)$ to indicate that the cost is with respect to weight r.

Definition 2: The *sum-requirement communication* (or s.r.c. in abbreviation) cost of a tree T is defined by $C_s(T) = \sum_{u,v}(r(u) + r(v))d_T(u, v)$.

Given a graph G, the PROCT (or SROCT) problem asks for a spanning tree T of G such that $C_p(T)$ (or $C_s(T)$ respectively) is minimum among all possible spanning trees.

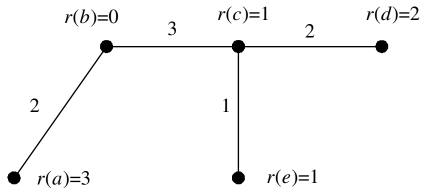

Figure 1: The product-requirement communication cost between vertices a and d is $3 \times 2 \times (2 + 3 + 2) = 42$, and the sum-requirement communication cost between vertices a and e is $(3 + 1) \times (2 + 3 + 1) = 24$.

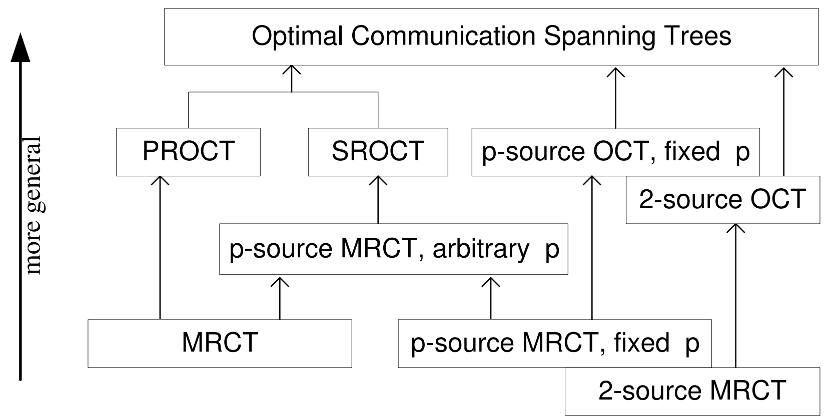

Figure 2: The relationship of the OCT problems.

Example 1: The p.r.c. cost and s.r.c. cost between a pair of vertices are illustrated in Figure 1. The cost of the tree is the sum of the cost for all pairs of vertices.

The relationship of the different versions of the OCT problems is illustrated in Figure 2. Note that there are variants for the multi-source problems. By "arbitrary p," we mean there is no restriction on the number of sources in the input data, while by "fixed p," the number of sources is always equal to the constant p.

Table 1 summarizes the results.

Bibliographic Notes and Further Reading

The current best approximation ratio for the general OCT problem is due to Yair Bartal's algorithms which approximate arbitrary metrics by tree metrics. He first presented a randomized algorithm [1] and then derandomized it to a deterministic algorithm [2]. Its application to approximating the OCT problem was pointed out in [10].

The PROCT and SROCT problems were introduced in [8]. In that paper, Bang Ye Wu, Kun-Mao Chao, and Chuan Yi Tang gave a 1.577-approximation algorithm for the PROCT problem and a 2-approximation algorithm for the SROCT problem. The PTAS using the Scaling-and-Rounding technique for a PROCT problem was presented in [9] by the same authors. Scaling the input instances is a technique that has been used to balance the running time and the approximation ratio. For example, Oscar H. Ibarra and Chul E. Kim used the scaling technique to develop a

Table 1: The objectives and currently best ratios of the OCT problems.

Problem	Objective	Ratio
OCT	$\sum_{u,v} \lambda(u,v) d_T(u,v)$	$O(\log n \log \log n)$
PROCT	$\sum_{u,v} r(u)r(v)d_T(u,v)$	PTAS
SROCT	$\sum_{u,v}(r(u)+r(v))d_T(u,v)$	2
MRCT	$\sum_{u,v} d_T(u,v)$	PTAS
p-MRCT	$\sum_{u \in S} \sum_{v \in V} d_T(u,v)$	2
2-MRCT	$\sum_v (d_T(s_1,v)+d_T(s_2,v))$	PTAS

fully polynomial time approximation scheme (FPTAS) for the knapsack problem [5], and some improvement was made by Eugene L. Lawler [6]. A nice explanation of the technique can also be found in [4](pp. 134–137).

The NP-hardness of the 2-MRCT was shown by Bang Ye Wu [7], in which the reduction is from the Exact Cover By 3-Sets (X3C) problem ([SP2] in [4]). The transformation is simpler and easier to extend to the weighted case, which is designed to show the NP-hardness of the p-MRCT problem for any fixed p. A similar reduction (for 2-MRCT) was also shown by Harold Connamacher and Andrzej Proskurowski [3]. They showed that the 2-MRCT problem is NP-hard. The PTAS for the 2-MRCT problem also appeared in [7]. In addition to the PTAS for the 2-MRCT problem, there is also a PTAS for the weighted 2-MRCT problem. But the PTAS works only for metric inputs and the counterpart on general graphs was left as an open problem.

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*, pages 184–193, 1996.

[2] Y. Bartal. On approximating arbitrary metrics by tree metrics. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 161–168, 1998.

[3] H.S. Connamacher and A. Proskurowski. The complexity of minimizing certain cost metrics for k-source spanning trees. *Discrete Appl. Math.*, 131:113–127, 2003.

[4] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W.H. Freeman and Company, San Francisco, 1979.

[5] O.H. Ibarra and C.E. Kim. Fast approximation algorithms for the knapsack and sum of subset problems. *J. ACM*, 22:463–468, 1975.

[6] E.L. Lawler. Fast approximation algorithms for knapsack problems. *Math. Oper. Res.*, 4(4):339–356, 1979.

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

[8] B.Y. Wu, K.-M. Chao, and C.Y. Tang. Approximation algorithms for some optimum communication spanning tree problems. *Discrete Appl. Math.*, 102:245–266, 2000.

[9] B.Y. Wu, K.-M. Chao, and C.Y. Tang. A polynomial time approximation scheme for optimal product-requirement communication spanning trees. *J. Algorithms*, 36:182–204, 2000.

[10] B.Y. Wu, G. Lancia, V. Bafna, K.-M. Chao, R. Ravi, and C.Y. Tang. A polynomial time approximation scheme for minimum routing cost spanning trees. *SIAM J. Comput.*, 29:761–778, 2000.