A Time- and Message-Optimal Distributed Algorithm for Minimum Spanning Trees

Gopal Pandurangan (University of Houston), Peter Robinson (Royal Holloway, University of London), and Michele Scquizzato (University of Houston)

gopalpandurangan@gmail.com, peter.robinson@rhul.ac.uk, michele@cs.uh.edu

Problem

The distributed construction of a minimum spanning tree (MST) by a network where nodes can communicate by message passing.

Model of Computation

CONGEST, the standard model for distributed network computing. It consists of a communication network, modeled by a graph, where the n vertices represent computational entities and the m edges represent bidirectional communication links. Computation proceeds in synchronous rounds, and in every round each of the n nodes may send messages of $O(\log n)$ bits to each of its neighbors. Complexity measures:

- Time complexity: total number of rounds;
- Message complexity: total number of messages exchanged.

Definition

We say that a problem enjoys *singular optimality* if it admits a distributed algorithm whose time and message complexity are both optimal.

Question

Does MST enjoy singular optimality?

Lower Bounds

- $\tilde{\Omega}(D+\sqrt{n})$ rounds [3];
- $\Omega(m)$ messages [6].

Both apply to randomized Monte Carlo algorithms.

Previous Results

Reference	Time Complexity	Message Complexity
Gallager et al. [4]	$O(n \log n)$	$O(m + n \log n)$
Awerbuch [1]	$O(n)$	$O(m + n \log n)$
Garay et al. [5]	$O(D + n^{0.614} \log^* n)$	$O(m + n^{1.614})$
Kutten and Peleg [7]	$O(D + \sqrt{n} \log^* n)$	$O(m + n^{1.5})$
Elkin [2]	$\tilde{O}(\mu(G,w) + \sqrt{n})$	$O(m + n^{1.5})$

Main Result

A randomized Las Vegas distributed algorithm that constructs a minimum spanning tree in weighted networks in $\tilde{O}(D+\sqrt{n})$ rounds and exchanging $\tilde{O}(m)$ messages, with high probability.

The Algorithm, in a Nutshell

1. Simultaneously and independently, grow MST fragments by merging them through min-weight outgoing edges ("blue rule"), until at most $\sqrt{n}$ fragments, each of diameter $O(\sqrt{n})$, remain.
2. Keep merging fragments, using
   - an auxiliary BFS tree on the network if $D = O(\sqrt{n})$ or when the number of remaining fragments is $O(n/D)$;
   - a *hierarchy of sparse neighborhood covers* otherwise.

Key Ideas

Replace the 2nd phase of sublinear-time algorithms (which uses the "red rule", and which is not message-efficient) with a continuation of the 1st phase. This introduces the problem of fragments with diameter $> D$ (hence communication within fragments may require $> D$ time). Solution: use *neighborhood covers*, a collection of clusters with

- diameter smaller than D and smaller than that of the fragments they contain;
- small overlap, which implies low-congestion communication across clusters.

Neighborhood Covers for the 2nd Phase

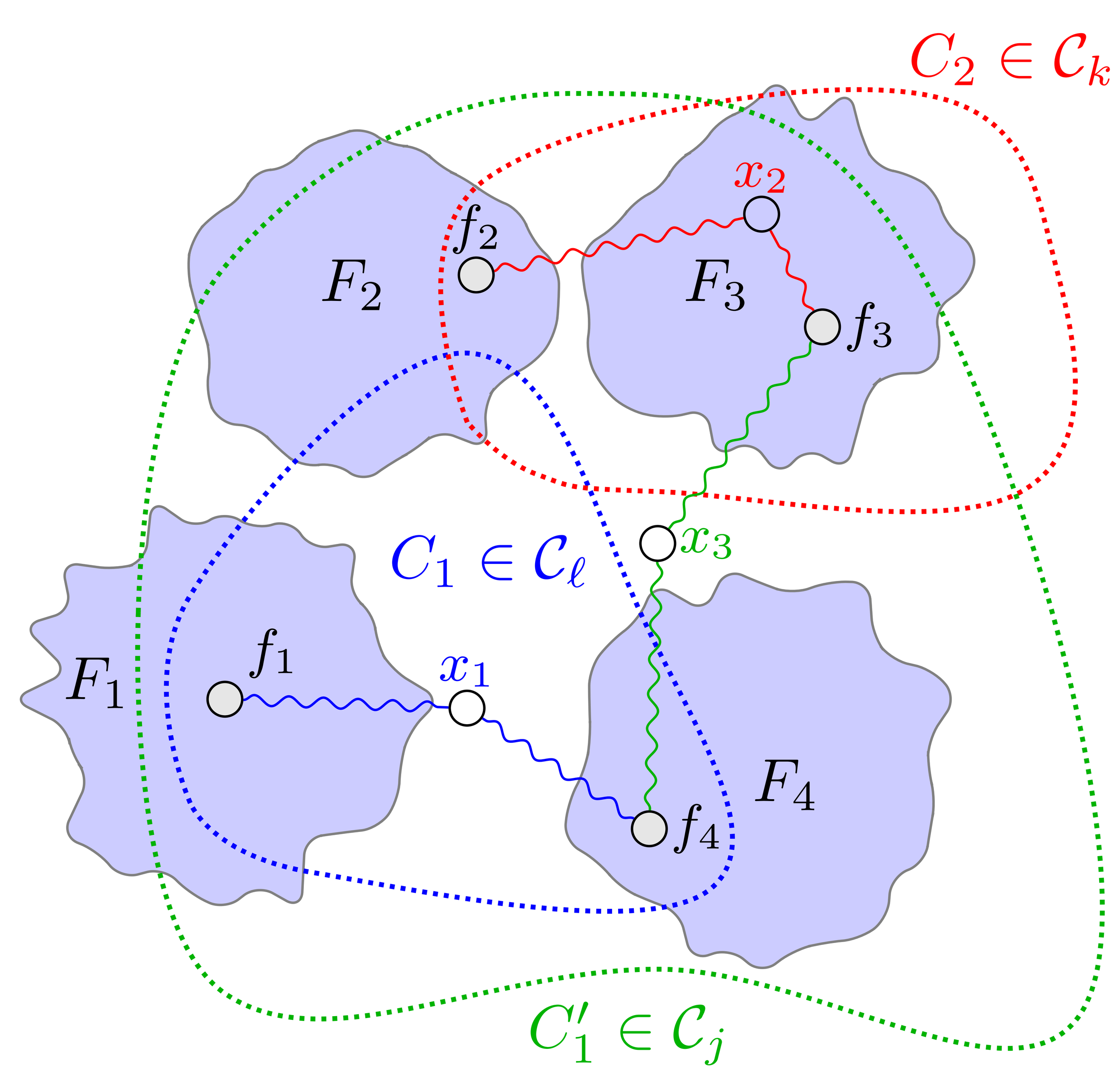

Figure 1: MST fragments $F_1, \ldots, F_4$, and communication-efficient paths within clusters C_i of covers $\mathcal{C}_j$. Used when $D = \omega(\sqrt{n})$ and the number of fragments is large.

Further Result

A graph construction for which every ϵ-error randomized distributed MST algorithm runs in $\tilde{\Omega}(D+\sqrt{n})$ rounds *and* exchanges $\Omega(m)$ messages in expectation.

Open Problems

- Investigate whether MST also enjoys singular optimality under the assumption that nodes initially have knowledge of the IDs of their neighbors (a.k.a. KT_1 variant).
- Investigate whether other fundamental problems, such as shortest paths, enjoy singular optimality.

References

[1] B. Awerbuch. Optimal distributed algorithms for minimum weight spanning tree, counting, leader election, and related problems. In *Proceedings of STOC 1987*, pages 230–240.

[2] M. Elkin. A faster distributed protocol for constructing a minimum spanning tree. *J. Comput. Syst. Sci.*, 72(8):1282–1308, 2006.

[3] M. Elkin. An unconditional lower bound on the time-approximation trade-off for the distributed minimum spanning tree problem. *SIAM J. Comput.*, 36(2):433–456, 2006.

[4] R. G. Gallager, P. A. Humblet, and P. M. Spira. A distributed algorithm for minimum-weight spanning trees. *ACM Trans. Program. Lang. Syst.*, 5(1):66–77, 1983.

[5] J. A. Garay, S. Kutten, and D. Peleg. A sublinear time distributed algorithm for minimum-weight spanning trees. *SIAM J. Comput.*, 27(1):302–316, 1998.

[6] S. Kutten, G. Pandurangan, D. Peleg, P. Robinson, and A. Trehan. On the complexity of universal leader election. *J. ACM*, 62(1), 2015.

[7] S. Kutten and D. Peleg. Fast distributed construction of small k-dominating sets and applications. *J. Algorithms*, 28(1):40–66, 1998.