Albert-Ludwigs-Universität Freiburg
Dept. of Computer Science
Algorithms and Complexity

Distributed MST Computation in the LogP Model

Bachelor or Master Thesis (or Project)

Project Description

The problem of computing a minimum spanning tree (MST) of a weighted graph $G = (V, E, w)$ is one of the most widely studied graph problems in the context of distributed algorithms. When studying distributed algorithms for the MST problem, one typically assume that the graph G defines the network. The nodes V of the G are the individual computing devices of the network and an edge $\{u, v\} \in E$ represents a bidirectional communication channel between the nodes u and v. At the end of a distributed MST algorithm, each node $v \in V$ needs to know with of its incident edges are in the computed spanning tree. The standard model in which this is studied is the so-called CONGEST model. In this model, time is divided into synchronous time steps called rounds and in each round, each node can send a (possibly different) $O(\log n)$-bit message to each of its neighbors, where n is the number of nodes of the graph G. The running time of an algorithm in the CONGEST model is the number of rounds that are needed for all nodes to terminate. The worst-case complexity of the MST problem is well understood in the CONGEST model. We know that $\tilde{\Theta}(D + \sqrt{n})$ rounds are necessary and also sufficient, where D is the diameter of the graph G (the $\tilde{\Theta}(\cdot)$-notation hides $O(\log n)$-factors).

While the CONGEST model is standard in the theory literature, in the more practical networking literature, people often work with the so-called LogP model. LogP stands for L: latency, o: overhead, g: gap, and P number of processors. In this project, we want to apply the LogP model to distributed graph computations. Applied to a graph topology, the model roughly means the following. Each node can send or receive one new message every o time units (i.e., the time for accessing the network stack is o time units for each message transmission or receival). The time for a message to be transmitted over a single edge is given by the latency L, and the gap g specifies the minimum time required between two messages over the same edge. The goal of the project is to understand the complexity of the MST problem as a function of the parameters o, L, and g, the number of nodes n, the diameter D, and maybe the number of bits B that can be packed into a single message. We believe that by adapting the techniques that are used for the CONGEST model, it should be possible to obtain some interesting results.

Requirements

- mathematical maturity, interest in mathematical questions
- algorithm theory and/or distributed systems (network algorithms) lecture (or comparable lectures) are an advantage

Contact

- **Fabian Kuhn**, kuhn@cs.uni-freiburg.de