Scheduling calls for multicasting in tree-networks

Johanne Cohen* Pierre Fraigniaud† Margarida Mitjana‡

Abstract

This paper is an extended abstract of the technical report [2]. In this report, we have shown that the multicast problem in trees can be expressed in term of arranging rows and columns of boolean matrices. The application of this result is that, given a directed tree T whose arcs are oriented from the root toward the leaves, and given a subset of nodes D, there exists a polynomial-time algorithm that computes an optimal multicast protocol from the root to all nodes of D.

1 Motivations

Multi-point protocols are of particular interest for group applications. Those groups involve more than two users (some may even involve thousands of users) sharing a common application, as video-conferences, distributed data-bases, media-spaces, games, etc. Several protocols have been proposed to handle and to control a large group of users. All of them are based on tree-networks, either a single tree connecting all the users (*e.g.*, Core-Based Tree), or several trees (*e.g.*, PIM). The traffic between the users is then routed along the edges of the tree(s).

One of the major communication problem related to multi-point applications consists to broadcast a message from one user to all the users of the application. This operation is called *broadcast* at the application level, though it is actually a *multicast* at the network level. The repetition of point-to-point connections between the source and the several destinations would significantly increase the traffic in the network, and it makes this solution not applicable in practice. Thence, the source must require the help of other nodes to relay messages. A broadcast message will then reach the destinations after having been relayed by several intermediate nodes. As the number of hops between the source and each destination is required to be as small as possible, we have derived a polynomial-time algorithm which returns, for any tree T, and for any source $u \in V(T)$, a multicast protocol from u to any arbitrary subset of nodes of T that minimizes the number of hops.

2 Models

We considered multicasting from the root to a set of destination nodes of a directed tree T whose arcs are oriented from the root toward the leaves. We have considered both 1-port and all-port models. In the 1-port model, we assume that, at any given time, each node of the tree can *call* at most one other node of the tree. In the all-port model, a node can call many other nodes simultaneously. Moreover, according to modern communication facilities (e.g., circuit-switched, wormhole, WDM, etc.), long-distance calls are allowed, in the sense that the receiver of a call is not necessarily a neighboring node of the initiator of the call, and a message crossing a node can cut-through the node if required. As a restriction though, we want the calls performed at the same time to not share any edge. This model is called the *line* model.

The set of all calls performed at the same time is called a *round*. The costs of our broadcast protocols are expressed in terms of number of rounds.

3 The contention-free matrix problem.

We have shown that the broadcast problem in directed trees under the 1-port line model gives rise to a matrix problem. Given a $p \times q$ matrix M with 0-1 entries, the *shadow* of M is the 1-dimensional boolean vector x of q entries such that $x_i = 0$ if and only if there is no 1-entry in the ith column of M, and $x_i = 0$ otherwise.

Definition 3.1. *Given a $p \times q$ matrix M with 0-1 entries, a* minimal contention-free version *of M is a matrix M^* such that:*

1. *M^* has at most one 1-entry per column;*
2. *every row r of M^* (viewed as the binary expression of an integer) is larger than the corresponding row r of M, $1 \leq r \leq q$; and*
3. *the shadow of M^* (viewed as the binary expression of an integer) is minimum.*

*Laboratoire de Recherche en Informatique, Bât. 490, Univ. Paris-Sud, 91405 Orsay cedex, France. Additional support by the DRET of the DGA, and by the AFFDU and the Editions Gauthier-Villars.

†Laboratoire de Recherche en Informatique, Bât. 490, Univ. Paris-Sud, 91405 Orsay cedex, France. Additional support by the CNRS.

‡Dept. Matematica i Telematica, Universitat Politecnica de Catalunya, 08034 Barcelona, Spain.

Note that the minimal contention-free version of a matrix is not necessarily unique, even up to a permutation of the rows. On the other hand, the shadow of a minimal contention-free version of a matrix is unique. As an example, let us consider

$$M = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \tag{3.1}$$

The reader can check that a minimal contention-free version of M is

$$M^* = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{3.2}$$

M^* has a shadow equal to $14 = (1110)_2$. We have shown:

THEOREM 3.1. *There is an $O(q(p+q))$-time algorithm that computes a minimal contention-free version of any $p \times q$ boolean matrix.*

4 Multicasting.

Given a broadcast protocol B from one extremity u of a path of length q, the *shadow* of B is the 1-dimensional boolean vector x of q entries such that $x_i = 1$ if and only if there is a call from u at round i of B, and $x_i = 1$ otherwise. B is *lexicographically optimal* if its shadow, considered as the binary expression of an integer, is minimum. We have established the following correspondence between multicast protocols and contention-free matrices.

LEMMA 4.1. *Let T be a star of p branches, each of length at most $2^q - 1$. Let M be the $p \times q$ matrix whose p rows are the p shadows of p broadcast algorithms B_i's from the root to the p branches of T. Assume that all B_i's are lexicographically optimal. Then any contention-free version M^* of M determines a broadcast protocol B from the root in the star under the 1-port line model. Moreover, if M^* is minimal, then B is lexicographically optimal, and conversely.*

Let us give an example of such correspondence. Let us consider a star of center v, and of three branches of 2, 4, and 3 nodes, respectively. The matrix M of the lemma is given in Equation 3.1. The matrix M^* given in Equation 3.2 is a minimal contention-free version of M of shadow $14 = (1110)_2$. M^* determines a broadcast protocol according to its 1-entries: at round 1, v calls the second branch; at round 2, v calls the third branch; and, at round 3, v calls the first branch. At round 4, v is idle.

The following result is a simple application of Lemma 4.1 and Theorem 3.1.

COROLLARY 4.1. *There exists a polynomial-time algorithm that computes an optimal multicast protocol from any source u to any destination set D in any star under the 1-port line model.*

More interestingly, using Lemma 4.1, Theorem 3.1, and the results in [1], we have shown that multicasting from the root of an arbitrary directed tree under the all-port line model can be solved in polynomial time.

COROLLARY 4.2. *There exists a polynomial-time algorithm that computes an optimal multicast protocol from any source u to any destination set D in any directed tree under the all-port line model.*

5 Further research.

We are currently working on an extension of Theorem 3.1 to make use of this result in the 1-port line model in arbitrary tree. The idea is to construct the protocol bottom-up from the leaves to the root. To make clear why Theorem 3.1 needs to be slightly adapted, let us consider the simple case of a *fork*, that is a particular type of directed tree in which the root u has a single child v which is the root of a star of p branches. Let X_i be the shadow of a broadcast algorithm from v to the ith branch, $i = 1, \ldots, p$, and let M be the $p \times q$ array whose ith row is X_i.

A non necessarily optimal broadcast protocol in the 1-port line model consists in two phases: first u informs v, then v informs the p branches according to a minimal contention-free version of M. This protocol may be suboptimal because it can be more efficient to have both u and v informing the p branches (in the 1-port line model, u and v can call two distinct branches simultaneously). The main question in this context is therefore to figure out how to make use optimaly of u. For a fork, this question can be easily solved. The generalization to an arbitrary tree still requires some work.

References

[1] J. Cohen. Broadcasting, multicasting and gossiping in trees under the all-port line model. In *10th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pages 164–171. ACM press, 1998.

[2] J. Cohen, P. Fraigniaud, and M. Mitjana. Scheduling calls for multicasting in tree-networks. Technical report 1189, LRI, Université Paris-Sud, 91405 Orsay cedex, France (http://www.lri.fr/~pierre), 1998.