Low Complexity Algebraic Multicast Network Codes

Sidharth Jaggi[1]
Dept. of Electrical Engineering
California institute of Technology
Pasadena, CA, USA 91125
e-mail: jaggi@caltech.edu

Philip A. Chou, and Kamal Jain
Microsoft Corporation
One Microsoft Way
Redmond, WA, USA 98052-6399
{pachou, kamalj}@microsoft.com

Abstract — **We present a low complexity algorithm for designing algebraic codes that achieve the information theoretic capacity for the multicast problem on directed acyclic networks. These codes operate over field sizes which are significantly smaller than those previously known, leading to significantly lower design and implementation complexity, and network link usage. These codes can be extended for networks with cycles and delays, and for robustness properties.**

A network is represented by a directed graph $\mathcal{G} = (V, E)$ with a vertex set V and edge set E. Multiple edges are allowed between vertices, and $\mathcal{G}$ is acyclic. A source node $s \in V$ generating information at rate R, and a set of N receiver nodes $\mathcal{R} = \{r_1, r_2, \ldots, r_N\} \subset V$, are specified. Nodes are allowed to transmit on outgoing edges any causal function of information on incoming edges. The *multicast capacity* is defined as the maximum rate at which identical information can be decoded by each node in $\mathcal{R}$. The multicast network coding problem is the 4-tuple $(\mathcal{G}, s, \mathcal{R}, R)$, and it is said to have a *solution* if there exist causal functions for each node in V such that the multicast capacity equals R. The solution comprises these functions at each node in V and the decoders at nodes in $\mathcal{R}$. Let $C \leq \min_i\{C_i\}$, where C_i is the min-cut max-flow capacity of the network between s and r_i. For any $\epsilon > 0$, the multicast network coding problem $(\mathcal{G}, s, \mathcal{R}, C - \epsilon)$ was shown by a random coding argument in [1] to have a solution. Further, [2] defined an $\mathbb{F}_{2^m}$*-linear network* as follows. Source bits are concatenated in blocks of length m to form symbols in the finite field $\mathbb{F}_{2^m}$. Nodes in V perform linear combinations (over $\mathbb{F}_{2^m}$) of the symbols on incoming edges to obtain symbols on outgoing edges. They show that for $2^m > NC$ there exist $\mathbb{F}_{2^m}$-linear networks which solve the network coding problem $(\mathcal{G}, s, \mathcal{R}, C)$.

Our main results, many of which were independently and concurrently discovered by [3], are as follows

Theorem 1 (Network Multicast Algorithm) *For* $m = \lceil \log_2 N \rceil$ *and any* $R \leq C$, *there exist* $\mathbb{F}_{2^m}$*-linear networks which solve the multicast network coding problem* $(\mathcal{G}, s, \mathcal{R}, R)$.

Outline of Proof:- We provide a polynomial time construction for the solution. The network is decomposed into R edge-disjoint paths to each r_i. These paths are then merged using the directed acyclic structure of the network to form the *network flow.* Linear combinations are designed for this network flow one edge at a time. One ensures that at every step of this design algorithm, for each receiver there exists a set of R edges on edge-disjoint paths to that receiver, such that these edges carry linearly independent combinations of the R symbols generated by s in any coding interval. More details are given in [4]. Decoding at each r_i is then the inverse transformation of the linear map taking the symbols generated by the source to the symbols on incoming edges to r_i. □

Theorem 2 (Lower bound on field size) *For any* k *there exist networks* $\mathcal{G}$ *with multicast capacity* 2 *with* $\binom{k+1}{2}$ *receivers, such that for any* $2^m < k$ *there exist no* $\mathbb{F}_{2^m}$*-linear networks which solve the network coding problem* $(\mathcal{G}, s, \mathcal{R}, 2)$.

The proof is constructive (details in [4]). This provides a lower bound to the finite field size required.

Theorem 3 (Robustness) *For* $m = \lceil \log_2(N\binom{C-1}{R-1}) \rceil$, *there exist* $\mathbb{F}_{2^m}$*-linear networks which solve the multicast network coding problem* $(\mathcal{G}', s, \mathcal{R}, R)$, *for any network* $\mathcal{G}'$ *obtained by less than* $C - R + 1$ *edge failures in* $\mathcal{G}$.

This generalizes the concept of MDS codes to multicast networks and shows the existence of codes which can tolerate up to $C - R$ network link failures while leaving the linear combinations at the source and internal nodes the same.

We now define the more general concept of $(\mathbb{F}_{2^m})^n$-linear networks, wherein each edge transmits an n-length vector of symbols from $\mathbb{F}_{2^m}$, vectors on incoming edges to a node are concatenated to form a single *concatenated input vector*, and each node performs a different *local linear transformation* to its concatenated input vector to obtain an n-length vector on each outgoing edge. Using random coding arguments one can prove that randomly chosen $(\mathbb{F}_{2^m})^m$-linear networks have a high probability of being solvable, and of having strong robustness properties.

Theorem 4 (Randomized Algorithm) *Under iid choices of all local linear transformations the probability that the resulting* $(\mathbb{F}_{2^m})^n$*-linear network solves the multicast network coding problem* $(\mathcal{G}, s, \mathcal{R}, R) > 1 - 2^{\log_2 N + 2 - nm(C-R)}$.

Theorem 5 (Randomized Robust Algorithm) *Under iid choices of all local linear transformations the probability that the resulting* $(\mathbb{F}_{2^m})^n$*-linear network solves the multicast network coding problem* $(\mathcal{G}', s, \mathcal{R}, R)$ $> 1 - 2^{\log_2 N + 2 + |E| - nm(C' - R)}$ *(where* C' *is the multicast capacity of any* $\mathcal{G}'$ *whose edge-set is a subset of that of* $\mathcal{G}$*).*

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network Information Flow", *IEEE Transactions on Information Theory*, IT-46, pp. 1204-1216, 2000.

[2] R. Koetter, M. Medard, "An Algebraic Approach to Network Coding", presented at *INFOCOM 2002.*

[3] P. Sanders, S. Egner, L. Tolhuizen, "Polynomial Time Algorithms for Network Information Flow", Proc. of the 15th ACM Symposium on Parallelism in Algorithms and Architectures.

[4] P. A. Chou, M. Effros, S. Egner, S. Jaggi, K. Jain, P. Sanders, L. Tolhuizen, "Linear Multicast Network Coding Algorithms", in preparation for *IEEE Transactions on Information Theory.*

[1] The first author was supported by Microsoft Corporation.

0-7803-7728-1/03/$17.00 ©2003 IEEE.