Network Coding for Undirected Information Exchange

Jasper Goseling, Christina Fragouli, Suhas N. Diggavi

***Abstract*—We consider the information exchange problem where each in a set of terminals transmits information to all other terminals in the set, over an undirected network. We show that the design of only a single network code for multicasting is sufficient to achieve an arbitrary point in the achievable rate region. We also provide an alternative proof for the set of achievable rate tuples.**

I. Introduction

Consider a group of users in a network represented as an undirected graph that want to exchange information, *i.e.*, each user has information that needs to be received by all other users. This model arises in multimedia file exchange applications such as video-conferencing and internet games.

In this work we consider the use of network coding for such information exchange conferences. Network coding admits a larger rate region than routing, at polynomial complexity [1]. We start by reviewing the characterization of the achievable rate region provided in [1]. We give an alternative proof for this result based on information flow decomposition techniques [2] and graph-theoretic properties.

A main challenge that deployment of network coding for such applications faces, stems from the fact that during the conference duration, the rates at which users transmit will naturally vary. Employing a network coding solution that uses different sets of global coding vectors depending on the rates of the users, would result in unacceptable complexity. For example, it would require the users to update their decoding operations accordingly.

Our main contribution is to prove that a common set of global coding vectors can be used to support all achievable rate tuples, thus allowing a smooth operation at different rates. Moreover, we show that identifying such vectors is in fact exactly equivalent to identifying the network coding solution for a single multicast session, where the source is any one of the participating nodes in the conference.

II. Model

We are given a network represented as an undirected graph $G = (\mathbb{V}, \mathbb{E})$, and a set of terminals $\mathbb{T} \subseteq \mathbb{V}$ of size $N \triangleq |\mathbb{T}|$. These terminals act as sources and receivers in an information exchange conference. Each terminal needs to receive all information transmitted by the other terminals. We will call $\{G = (\mathbb{V}, \mathbb{E}), \mathbb{T}\}$ an information exchange configuration.

Assume time is slotted, and let rate R_i be the number of symbols that are generated per time-slot by the source located at terminal $T_i \in \mathbb{T}$, $i = 1, \ldots, N$. We consider only integral rates.

Definition 1 (Achievable Rate Region)**.** *The achievable rate region for an exchange configuration $\{G, \mathbb{T}\}$ consists of all tuples $(R_1, R_2, \ldots, R_N)$ for which there exists a valid linear network code that supports these rates.*

A code consists of an orientation of the graph and sets of global and local coding vectors. Throughout the paper we assume that we can impose a partial order on the edges. Furthermore, we assume that each edge in the graph has unit capacity, *i.e.*, it can carry one symbol per time-slot.

Definition 2 (Uniform Pairwise Min-cut)**.** *A set of terminals $\mathbb{T} \subseteq \mathbb{V}$ has uniform pairwise min-cut $h \in \mathbb{N}^+$ over a graph $G(\mathbb{V}, \mathbb{E})$ if $\text{min-cut}(T_1, T_2) = h$, $\forall\, T_1, T_2 \in \mathbb{T}$.*

We are interested in minimal configurations, in the sense that removing any network edge would reduce the min-cut to smaller than h for at least one pair of terminals. Minimal configurations are desirable as they allow the conservation of network resources. Such configurations can be identified in polynomial time by first selecting edge disjoint paths between pairs of terminals, and then removing redundant edges [2].

III. Rate Region

Theorem 1. *A rate tuple $(R_1, \ldots, R_N)$ is achievable if and only if for all $i = 1, \ldots, N$, a multicast session with only terminal T_i acting as a source of rate $R^* \triangleq \sum_{j=1}^{N} R_j$ is achievable.*

This theorem shows that the achievable rate region can be completely characterized in terms of the maximum achievable sum-rate R^*. This result was first obtained in [1]. We provide an alternative proof in Section IV, but first we connect R^* to the pairwise min-cut between the terminals thereby giving a graph-theoretic interpretation.

We start with a graph-theoretic result for networks that have a uniform pairwise min-cut. Let $\text{min-cut}\,(T, \mathbb{T} \setminus \{T\})$, $T \in \mathbb{T}$, be the value of a minimum cut that separates T from all terminals in $\mathbb{T} \setminus \{T\}$.

Lemma 1. *Consider an information exchange configuration $\{G, \mathbb{T}\}$, with uniform pairwise min-cut h. The minimum min-*

J. Goseling is with Delft University of Technology, The Netherlands. C. Fragouli and S.N. Diggavi are with Ecole Polytechnique Fédérale de Lausanne, Switzerland. This work was done while J. Goseling was with Ecole Polytechnique Fédérale de Lausanne, Switzerland, and was supported in part by SNF NCCR-MICS grant 5005–67332.

cut from a terminal to all other terminals is h*:*

$$\min_{T\in\mathbb{T}} \text{min-cut}\,(T, \mathbb{T}\setminus\{T\}) = h.$$

Proof: Let H be a Gomory-Hu tree (see e.g. [3]) constructed only on the set of terminal vertices. By the uniform pairwise min-cut assumption, every edge in H has weight h. Let T be any terminal node that is a leaf in H. The edge in H incident on T separates T from all other terminals. By the properties of a Gomory-Hu tree there is a capacity h cut in the original network, separating T from all other terminals, hence $\text{min-cut}\,(T, \mathbb{T}\setminus\{T\}) = h$. ∎

Theorem 2. *Consider an exchange configuration* $\{G, \mathbb{T}\}$ *with uniform pairwise min-cut* h*. The maximum achievable sum rate satisfies* $h\frac{N}{2(N-1)} \leq R^* \leq h$*.*

Proof: We first show the upper bound. By Lemma 1 we know that there exists a terminal $T \in \mathbb{T}$ for which $\text{min-cut}(T, \mathbb{T}\setminus\{T\}) = h$. Consider any such minimum cut. The information from all sources needs to cross this cut.

For the lower bound we show that rate tuple $\left(\frac{h}{2(N-1)}, \ldots, \frac{h}{2(N-1)}\right)$ is achievable. Create a directed network by replacing each undirected edge with two oppositely directed edges of capacity $1/2$. Add a virtual source S' to the network and connect it to each terminal in $\mathbb{T}$ using directed edges of capacity $h/(2(N-1))$. We will show that

$$\text{min-cut}\,(S', T) \geq Nh/(2(N-1)), \quad \forall\, T \in \mathbb{T}. \tag{1}$$

Consider any cut that separates the virtual source S' from a terminal T. The edge that connects S' to T has to cross this cut. In fact, if all terminals are on T's side of the cut, all edges from the virtual source S' necessarily cross the cut which therefore has value at least $Nh/(2(N-1))$. If, on the other hand, a subset of terminals $\mathbb{T}_1 \subseteq \mathbb{T}$ are on the same part of the cut as the source, from construction $\text{min-cut}\,(\mathbb{T}_1, T) \geq h/2$. Adding the contribution of the direct edge from S' to T results in a cut value of at least $Nh/(2(N-1))$.

Since (1) is satisfied there exists a network code that allows to multicast rate $Nh/(2(N-1))$ from S' to all terminals. Any valid solution will require S' to send independent information to each terminal at rate $h/(2(N-1))$ and thus corresponds to a solution of the information exchange problem. ∎

Note that the ratio between the inner and outer bound is $2(N-1)/N$. In [1] it was shown that over undirected graphs, if the min-cut from a source to all receivers equals h, use of network coding allows to achieve rate R^* with $h/2 \leq R^* \leq h$. Thus we get a slightly smaller ratio.

IV. Preserving Global Coding Vectors

We here show that, given a linear network code (and its associated set of global coding vectors) that supports a specific rate tuple, we can in fact use the same global coding vectors to operate the network for all achievable rate tuples. Thus, the terminal nodes always have the same set of linear equations to solve, no matter what is the operating rate-tuple. We can achieve this by sequentially applying the following procedure.

Procedure 1 Preserving global coding vectors

Input: Network code at rate tuple $(R_\ell)_{\ell=1}^N$, $i, j \in \{1, \ldots, N\}$.
Output: Code using same global coding vectors, at rate tuple $(\tilde{R}_\ell)_{\ell=1}^N$, $\tilde{R}_i = R_i - 1$, $\tilde{R}_j = R_j + 1$, $\tilde{R}_k = R_k, k \neq i, j$.

1) Find a path from T_i to T_j.
2) Reverse the orientation of the edges along this path.
3) At each coding point find new local coding vectors so that the global coding vectors remain unaffected.

We prove feasibility of the procedure using the information flow decomposition techniques developed in [2]. The idea is to decompose the network into parts, each part being a tree, and thus called a subtree. We distinguish between source subtrees and coding subtrees. Each such subtree corresponds to one unit rate source, and thus a source emitting rate R_i results in R_i source subtrees. Each coding subtree starts with a coding point, an edge where linear combining occurs, and again represents a tree through which runs unit information rate. The linear combinations are represented by global coding vectors assigned to the subtrees. The local operations performed at the coding points are represented by local coding vectors.

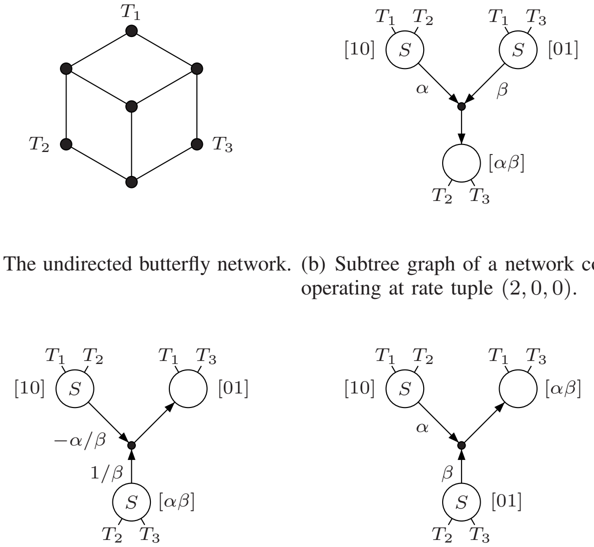
As an illustrating example consider the undirected butterfly network in Figure 1(a), on which terminals T_1, T_2 and T_3 exchange information, and the min-cut between any two terminals equals 2. Figure 1(b) gives the subtree graph, with two source subtrees and one coding subtree, corresponding to the case where terminal T_1 is acting as a source. That is, the operational rate tuple is $(2, 0, 0)$. A notational difference from [2] is that we explicitly represent the coding point.

Assume now that we would like to operate at the rate tuple $(1, 1, 0)$, where both T_1 and T_2 act as sources of rate one, while still retaining the same global coding vectors. The subtree graph after applying Procedure 1 is given in Figure 1(c).

We observe that applying Procedure 1 preserves the subtrees and coding points of the subtree graph. Step 2) amounts to repeatedly interchanging the role of a source and a coding subtree.

Lemma 2. *Consider a valid code over a minimal subtree graph and assume we exchange the role of a source subtree and one of its children. There exist local coding vectors resulting in the same global coding vectors as the original code.*

Proof: W.l.o.g. assume that in the original graph subtrees $C_1, \ldots, C_k$ are parents of C_{k+1}, and that after the procedure we have $C_2, \ldots, C_{k+1}$ parents of C_1. Let v_i denote the global coding vector associated with subtree C_i and $\langle v_1, \ldots, v_k\rangle$ the space spanned by vectors $v_1 \ldots, v_k$. Then $v_{k+1} \in \langle v_1, \ldots, v_k\rangle$ and from properties of minimal configurations $v_{k+1} \notin \langle v_2, \ldots, v_k\rangle$. This implies that $\langle v_1, \ldots, v_k\rangle = \langle v_2, \ldots, v_{k+1}\rangle$, i.e., $v_2, \ldots, v_{k+1}$ form a basis of the space $\langle v_1, \ldots, v_k\rangle$. Thus $v_1 \in \langle v_2, \ldots, v_{k+1}\rangle$, we can always find the required local coding vector. Note, that if we start from an orientation that allows a partial order of the edges and reverse edges according to Procedure 1, the resulting topology has a partial order of the edges. ∎



(a) The undirected butterfly network. (b) Subtree graph of a network code operating at rate tuple $(2, 0, 0)$.

(c) Preserving the global coding vectors, operating at $(1, 1, 0)$. (d) Preserving the local coding vectors, operating at $(1, 1, 0)$.

Fig. 1. The undirected butterfly network, with different choices of sources and receivers.

Theorem 3. *Any valid code used as input for Procedure 1 results in a valid output code.*

Proof: Follows from the above observations and Lemma 2. ∎

Proof of Theorem 1: Procedure 1 allows, starting from any achievable rate tuple, to obtain a valid code in which a single terminal is acting as a source with a rate that is equal to the sum-rate of the original rate tuple, and vice versa. ∎

V. Preserving Local Coding Vectors

Preserving the same global coding vectors reduces the complexity for the terminal nodes. Alternatively, we may be interested in reducing the complexity at the coding points, by preserving the local coding vectors, while modifying the global coding vectors. Unlike the previous case, where starting from any valid code, for any achievable rate tuple we could maintain the same global coding vectors, we here need to calculate *in advance* specifically chosen *universal* local coding vectors that can be reused for all rate tuples. Note that the global coding vectors may now be different for each rate tuple.

We use Procedure 2, which is illustrated in Figure 1(d).

Theorem 4. *There exists a set of universal local coding vectors that can be repeatedly used in Procedure 2 and result in valid output codes.*

Proof: Start from any achievable rate tuple and an orientation of the graph supporting it. Assign orthonormal basis vectors as global coding vectors to the source subtrees. Following the algebraic framework [4], let the coefficients of the local coding vectors be variables. We show that there exists values for these variables that can be used as universal local coding vectors.

Procedure 2 Preserving local coding vectors

Input: Network code at rate tuple $(R_l)_{l=1}^N$, $i, j \in \{1, \ldots, N\}$.

Output: Code using same local coding vectors, at rate tuple $(\tilde{R}_l)_{l=1}^N$, $\tilde{R}_i = R_i - 1$, $\tilde{R}_j = R_j + 1$, $\tilde{R}_k = R_k, k \neq i, j$.

1) Find a path from T_i to T_j.
2) Reverse the orientation of the edges along this path.
3) At each coding point keep the local coding vector coefficients of the edges that are not reoriented. At affected coding points, to the new incoming edge assign the coefficient of the edge that is now outgoing.
4) Assign orthonormal basis vectors as global coding vectors to the source subtrees. The global coding vectors for the coding subtrees follow from the global coding vectors of the source subtrees and the local coding vectors.

Consider the transfer matrices to each terminal for each of the codes that can be obtained by Procedure 2. By Theorem 3 we know that for the orientations obtained after Step 2) there exists a valid network code. Therefore, there also exists a code using any set of orthonormal basis vectors as global coding vectors at the source subtrees. Hence, the determinants of all these matrices are non-identically zero polynomials.

Consider the polynomial formed as the product of the above determinants. This is again a non-identically zero polynomial. Any set of coefficients for which the product polynomial evaluates to a non-zero value, results in suitable local coding vectors. ∎

The theorem states that one can carefully construct a suitable code that allows to preserve the same local coding vectors.

VI. Discussion

This paper provides constructions to achieve arbitrary operating points in the achievable rate region for an undirected information exchange network, while using a single set of global or local coding vectors. Note that our approach lends itself to asynchronous network operation: the intermediate nodes in the network can deduce the coding operations they need to perform, based on the origin of the incoming packets, and without knowledge of the operating point. As a side result, we also get an alternative proof to the characterization of the exchange rate region. We also provide a proof for Theorem 2 based on graph-theoretic properties, that gives additional insight in our problem structure.

VII. Acknowledgements

We would like to thank Jos Weber and the reviewers for their helpful comments.

References

[1] Z. Li and B. Li, "Network coding in undirected networks," in *Proc. of the 38th Annual Conf. on Information Sciences and Systems (CISS)*, 2004.
[2] C. Fragouli and E. Soljanin, "Information flow decomposition for network coding," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 829–848, 2006.
[3] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*. John Wiley & Sons, Inc, 1998.
[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, vol. 11, no. 5, pp. 782–795, 2003.