The Single Source Two Terminal Network with Network Coding

Aditya Ramamoorthy
Department of Electrical Engineering
University of California
Los Angeles, CA 90095
Email: adityar@ee.ucla.edu

Richard D. Wesel
Department of Electrical Engineering
University of California
Los Angeles, CA 90095
Email: wesel@ee.ucla.edu

Abstract—We consider a communication network with a single source that has a set of messages and two terminals where each terminal is interested in an arbitrary subset of messages at the source. A tight capacity region for this problem is demonstrated. We show by a simple graph-theoretic procedure that any such problem can be solved by performing network coding on the subset of messages that are requested by both the terminals and that routing is sufficient for transferring the remaining messages.

I. INTRODUCTION

The seminal work of Ahlswede et al. [1] established that for the single-source multiple-terminal multicast problem the achievable rate was the minimum of the maximum flows to each terminal from the source. They showed that in general, it is necessary to perform network coding to achieve this capacity. The basic idea is to give the nodes in the network the flexibility of performing operations on the data rather than simply replicating and/or forwarding it. Li et al. [2] showed that linear network coding is sufficient for achieving the capacity of the transmission of a single source to multiple terminals. Subsequent work by Koetter and Médard [3] and Jaggi et al. [4] presented constructions of linear multicast network codes. A randomized construction of multicast codes was demonstrated by Ho et al. [5].

It is important to realize that the multicast capacity result of [1] assumes that all the terminals are interested in the same data. The general network coding problem with multiple sources and terminals and an arbitrary set of connections is much harder and not much is known about it. In fact it has been shown in [6] that non-linear network codes are necessary in certain non-multicast problems. Network coding has also been considered from a lossless compression point of view in [7][8][9][10].

In this paper we study a specific example of a non-multicast problem with a single source and two sinks. We find a tight capacity region for this problem. This problem was independently considered by Ngai and Yeung [11] and Erez and Feder [12]. However our method of proof is very different and is based on a simple graph-theoretic procedure that may be of independent interest. This procedure was also utilized in [10].

We became aware of this work after the submission of the current paper.

II. PROBLEM FORMULATION

Consider a communication network modelled as a directed graph $G$, with a specified source node $S$ and two terminal nodes $T_1$ and $T_2$. We assume that the links are noiseless and that each edge in $G$ has unit capacity. This assumption can be realized by picking a suitably large time unit, assuming sufficient error-correction at the lower layers of the network and splitting edges of higher capacity into parallel unit capacity edges.

Suppose that the source node $S$ observes three independent processes $X_0, X_1$ and $X_2$ such that terminal $T_1$ is interested in $(X_0, X_1)$ and terminal $T_2$ is interested in $(X_0, X_2)$. Let the entropy rates of the processes be $H_0, H_1$ and $H_2$ respectively. We show the necessary and sufficient conditions for the feasibility of this connection. Furthermore it is shown that this problem can be solved by a combination of pure routing and network coding, where the sources $X_1$ and $X_2$ can be simply routed to $T_1$ and $T_2$ whereas the source $X_0$ may need network coding. The case of connections between terminal nodes is handled more naturally in our framework as compared to [11].

In the sequel the capacity assignment to an edge $a \rightarrow b$ is denoted by $ca (a \rightarrow b)$ and the minimum cut between nodes $V_1$ and $V_2$ is denoted by $min-cut(V_1, V_2)$. By the max-flow min-cut theorem [13], the minimum cut is also the maximum rate that can be transmitted from $V_1$ to $V_2$. By a solution to a given problem we mean an assignment of appropriate coding vectors to each edge so that the required network connection can be supported.

III. RESULTS

The following theorem is the main result of this paper.

Theorem 1: Consider a communication network modelled by a directed graph $G = (V, E)$ with one source node $S$ and two terminal nodes $T_1$ and $T_2$. Three independent processes $X_0, X_1$ and $X_2$ are observed at $S$ such that $H(X_0) = H_0, H(X_1) = H_1$ and $H(X_2) = H_2$. $T_1$ is interested in receiving $(X_0, X_1)$ and $T_2$ is interested in receiving $(X_0, X_2)$. If

$$
\text{min-cut}(S, T_1) \geq H_0 + H_1, \quad (1)
$$

$$
\text{min-cut}(S, T_2) \geq H_0 + H_2 \quad \text{and,} \quad (2)
$$

$$
\text{min-cut}(S, (T_1, T_2)) \geq H_0 + H_1 + H_2 \quad (3)
$$
there exists a solution where \( X_1 \) can be routed to \( T_1 \), \( X_2 \) can be routed to \( T_2 \) and \( X_0 \) can be sent to both \( T_1 \) and \( T_2 \) via network coding. Conversely if any of the inequalities \((\overline{1}) - (\overline{3})\) are violated then the connection cannot be supported.

We defer the proof of this theorem until we have established a lemma that is required. We start by defining an augmented graph \( G_1 = (V_1, E_1) \) as depicted in Fig. 1.

1) The new vertex set is \( V_1 = V \cup \{ T_1', T_2', Y_1, Y_2 \} \) as shown in Fig. 1. \( T_1' \) and \( T_2' \) can be regarded as virtual terminals, where the data is actually decoded. \( Y_1 \) and \( Y_2 \) are virtual nodes introduced for the purposes of our proof.

2) The capacity assignments of the new edges are

\[
\begin{align*}
\text{cap} (T_1 \to T_1') &= H_0 + H_1, \quad \text{cap} (T_1' \to Y_1) = H_0 + H_1, \\
\text{cap} (T_1' \to Y_2) &= H_1, \quad \text{cap} (T_2 \to T_2') = H_0 + H_2, \\
\text{cap} (T_2' \to Y_2) &= H_2 \\
\end{align*}
\]

Lemma 1: For the augmented graph \( G_1 \) the following is true :-

\[
\begin{align*}
\text{min-cut}(S, T_1') &= H_0 + H_1 \quad (4) \\
\text{min-cut}(S, T_2') &= H_0 + H_2 \quad (5) \\
\text{min-cut}(S, (T_1', T_2')) &= H_0 + H_1 + H_2 \quad (6) \\
\text{min-cut}(S, Y_1) &= H_0 + H_1 + H_2 \quad (7) \\
\text{min-cut}(S, Y_2) &= H_0 + H_1 + H_2 \quad (8)
\end{align*}
\]

Proof :- The first two equalities are obviously true. To see that \( \text{min-cut}(S, Y_1) = H_0 + H_1 + H_2 \) note that all cuts between \( S \) and \( Y_1 \) can be divided into four types:

a) The cut \((C, C^c)\) such that \( S, T_1, T_2 \in C \) and \( Y_1 \in C^c \). By inspection such a cut has capacity larger than or equal to \( H_0 + H_1 + H_2 \).

b) \( S, T_1 \in C \) and \( T_2, Y_1 \in C^c \).

The min-cut \((S, T_2) \geq H_0 + H_2 \) and min-cut \((T_1, Y_1) = H_0 + H_1 \) and the edges connecting \( T_1 \) and \( Y_1 \) are independent of the edges connecting \( S \) and \( Y_1 \). This means that such a cut has capacity at least \( 2H_0 + H_1 + H_2 \).

c) \( S, T_2 \in C \) and \( T_1, Y_1 \in C^c \).

The min-cut \((S, T_1) \geq H_0 + H_1 \) and min-cut \((T_2, Y_1) = H_2 \) and the edges connecting \( S \) to \( T_1 \) are independent of the edges connecting \( T_2 \) to \( Y_1 \). This means that such a cut has capacity at least \( H_0 + H_1 + H_2 \).

d) \( S \in C \) and \( T_1, T_2, Y_1 \in C^c \).

Since the min-cut \((S, (T_1, T_2)) \geq H_0 + H_1 + H_2 \), therefore any such cut has capacity at least \( H_0 + H_1 + H_2 \).

Finally, the sum of the capacities on the incoming edges of \( Y_1 \) is exactly \( H_0 + H_1 + H_2 \). This means that \( \text{min-cut}(S, Y_1) = H_0 + H_1 + H_2 \). The other statements in the lemma can be shown to be true in a similar manner.

Using the augmented graph \( G_1 \) we shall now demonstrate the existence of a certain number of paths from \( S \) to \( T'_1 \) and \( S \) to \( T'_2 \) over which data can be routed. Further, we shall show that it is possible to send the remaining data via network coding such that the demands of each sink are satisfied. The arguments proceed by utilizing the minimum cut conditions and performing a simple graph-theoretic procedure on the chosen paths in \( G_1 \). The details are given below.

Proof of Theorem \( \square \):-

First let us consider the paths from \( S \) to \( Y_1 \) and \( S \) to \( T'_2 \). Using Menger’s theorem (see the book by van Lint & Wilson [13]) we can conclude that :

- There exists a set of \((H_0 + H_1 + H_2)\) edge-disjoint paths from \( S \) to \( Y_1 \) from \( \square \). We call this set \( G \).
- There exists a set of \((H_0 + H_2)\) edge-disjoint paths from \( S \) to \( T'_2 \) from \( \square \). We call this set \( R \).

Now, we color the edges in paths \( \in G \), green and the edges in paths \( \in R \), red. At the end of this procedure some edges on these paths may have just one color while others may have two.

We claim that it is always possible to find \( H_2 \) exclusively green paths (i.e. paths that contain edges only having the color green) from \( S \) to \( T'_1 \). The technique of proof is similar to the one used in [10][14]. To prove this we define an algorithm \( A \) that shall be applied to a path \( P \in G \).

Algorithm A (P) :-

1) Traverse \( P \) starting at \( S \) and find the first edge \( e_1 \) that has color (green, red)

2) If no such \( e_1 \) is found then STOP.

3) ELSE

Suppose \( e_1 \in P' \) where \( P' \in R \) such that \( P' = P'_1 \).
$e_1 - P'_2$ where $P'_1$ is the portion of $P'$ from $S$ to $e_1$ and $P'_2$ is the portion of $P'$ from $e_1$ to $T'_2$. Color all edges on $P$ from $S$ to $e_1$, red in addition to their current color and remove red from the edges in $P'_1$. We now define a condition that each path $P \in \mathcal{G}$ needs to satisfy.

\[
\text{Cond}(P) = \{ \text{All edges in } P \text{ are } \text{green} \} \\
\text{or } \{ \text{First edge of } P \text{ is } \text{green, red} \}
\] (9)

We continue applying $A$ to each path of $\mathcal{G}$ until all paths in $\mathcal{G}$ satisfy $\text{Cond}$. It is easy to see that $A$ will eventually halt (for a proof see [10]).

At the end of this process we realize that there exist $H_1$ paths belonging to $\mathcal{G}$ that are exclusively green. This is true since if Algorithm $A$ re-routes a path $e \in \mathcal{R}$ it removes the color red from one outgoing edge of $S$ and places it on another outgoing edge. Therefore the total number of outgoing edges that are colored red remains constant at $H_0 + H_2$. It follows that $H_0 + H_1 + H_2 - (H_0 + H_2) = H_1$: outgoing edges are colored green and since the paths obey $\text{Cond}$ all those paths are exclusively green.

Next we note that all the exclusively green paths need to pass through $T'_1$ since $T'_2$ has exactly $(H_0 + H_2)$ incoming edges all of which have to be colored red. This proves the claim made above.

The critical point to be realized is that the re-routing of paths as above gives us $H_1$ paths from $S$ to $T'_1$ that are interference-free since these paths do not intersect with the paths from $S$ to $T'_2$. This means that data on these paths can be simply routed. Applying exactly the same procedure on the set of paths from $S$ to $Y_2$ and $S$ to $T'_1$ gives us $H_2$ paths from $S$ to $T'_2$ that are interference-free.

Now suppose that these paths ($H_1$ paths from $S$ to $T_1$ and $H_2$ paths from $S$ to $T_2$) are removed from $G_1$ to obtain a new graph $G_2$. Note that there still exist $H_0$ paths from $S$ to $T'_1$ and $H_0$ paths from $S$ to $T'_2$ in $G_2$. In other words, even after the removal of the interference-free paths the maximum flow from $S$ to $T'_1$ and $S$ to $T'_2$ in $G_2$ is $H_0$. Using the multicast result of [1] we can surely transmit the same $H_0$ bits from $S$ to $T'_1$ and $T'_2$ via network coding.

Thus, the entire solution can be realized by an appropriate choice of paths such that,

1) $H_1$ bits (process $X_1$) can be routed from $S$ to $T'_1$ and $H_2$ bits (process $X_2$) can be routed from $S$ to $T'_2$.
2) $H_0$ bits (process $X_0$) can be sent to both $T'_1$ and $T'_2$ by linear network coding [2].

Finally we note that it is trivial to realize the virtual terminals $T'_1$ and $T'_2$ at the terminals.

The proof of the converse is easy to see since even if one of the inequalities (1) - (4) is violated then at least one terminal does not have enough capacity to support its demand. This completes the proof of Theorem 1.

It is possible to find networks where one needs to strictly perform network coding for transmitting $X_0$ (while routing $X_1$ and $X_2$) and hence our result is tight. A simple example that demonstrates this is provided in Fig. 2. Here we have $H_0 = 2$ and $H_1 = H_2 = 1$. The figure shows a network where it is necessary to send $X_0$ via network coding. All links have unit capacity.

![Fig. 2. The sources observed at $S$ are such that $H_0 = 2, H_1 = H_2 = 1$. The figure shows a network where it is necessary to send $X_0$ via network coding. All links have unit capacity.](image)

IV. CONCLUSION

We found the capacity region for a network information transfer problem with a single source and two terminals when the use of network coding is permitted by utilizing a simple graph-theoretic procedure that may be of independent interest. It is interesting to note that the use of network coding permits us to obtain a tight characterization of the capacity region of this problem. However the region for the general broadcast channel with two receivers is still unknown (this was also noted by [12]).

V. ACKNOWLEDGEMENT

This work was supported by the University of California and Texas Instruments through UC Discovery grant COM04-10155 and National Science Foundation grant CCR-0209110.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, “Network Information Flow,” IEEE Trans. on Info. Th., vol. 46, no. 4, pp. 1204–1216, 2000.

[2] S.-Y. Li, R. W. Yeung, and N. Cai, “Linear Network Coding,” IEEE Trans. on Info. Th., vol. 49, no. 2, pp. 371–381, 2003.

[3] R. Koetter and M. Medard, “Beyond Routing: An Algebraic Approach to Network Coding,” in IEEE Infocom, 2002.
[4] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, and L. Tolhuizen, “Polynomial time algorithms for multicast network code construction,” Submitted to IEEE Trans. on Info. Th.

[5] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger, “Towards a Random Operation of Networks,” Submitted to IEEE Trans. on Info. Th.

[6] R. Dougherty, C. Freiling, and K. Zeger, “Insufficiency of Linear Coding in Network Information Flow,” Submitted to IEEE Trans. on Info. Th.

[7] M. Effros, M. Medard, T. Ho, S. Ray, D. Karger, and R. Koetter, “Linear Network Codes: A Unified Framework for Source, Channel and Network Coding,” in Proceedings of the DIMACS Workshop on Network Information Theory, 2003.

[8] T. Ho, M. Médard, M. Effros, and R. Koetter, “Network Coding for Correlated Sources,” in CISS, 2004.

[9] A. Ramamoorthy, K. Jain, P. A. Chou, and M. Effros, “Separating Distributed Source Coding from Network Coding,” in Proceedings of the 42nd Allerton Conference on Communication, Control, and Computing, 2004.

[10] ——, “Separating Distributed Source Coding from Network Coding,” Submitted to IEEE Trans. on Info. Th.

[11] C. K. Ngai and R. W. Yeung, “MultiSource Network Coding with Two Sinks,” in IEEE ICCAS, 2004.

[12] E. Erez and M. Feder, “Capacity Region and Network Codes for Two Receivers Multicast with Private and Common Data,” in Workshop on Coding, Cryptography and Combinatorics, 2003.

[13] J. H. van Lint and R. M. Wilson, A Course in Combinatorics. Cambridge University Press, 2001.

[14] K. Jain, M. Mahdian, and M. R. Salavatipour, “Packing Steiner Trees,” in SODA, 2003.