On network error correction with limited feedback capacity

Tracey Ho*, Sukwon Kim†, Yanbo Yang*, Michelle Effros* and Salman Avestimehr‡
*California Institute of Technology, †Samsung Electronics, ‡Cornell University
Email: *{tho, yanbo, effros}@caltech.edu, †sw921.kim@samsung.com, ‡avestimehr@ece.cornell.edu

***Abstract*—We consider the problem of characterizing network capacity in the presence of adversarial errors on network links, focusing in particular on the effect of low-capacity feedback links across network cuts. We give a new outer bound as well as a new achievable strategy, and show a family of networks where the inner and outer bounds coincide.**

I. Introduction

The problem of reliable network communication in the presence of adversarial link errors was first considered by Yeung and Cai [1], [2] for the case of networks with equal capacity links. In this problem, an adversary can arbitrarily corrupt information on a set of z network links whose locations are unknown to the network user. Yeung and Cai [1], [2] showed that the multicast capacity is given by $m-2z$, where m is the minimum source-sink cut capacity, and that the capacity can be achieved by linear network coding.

In our recent work [3]–[5], we considered the case of networks with unequal link capacities. We showed that, unlike the case of equal link capacities, feedback across network cuts can increase the error correction capacity. We provided upper bounds on capacity, and coding strategies that achieve the upper bounds in a family of four-node networks with feedback links of sufficient capacity. The related problem of network error correction with adversarial nodes was considered by [6], [7].

In this paper, we consider the case of small-capacity feedback links, for which our previous bounds in [5] are not tight. We provide a new upper bound and a new coding strategy that achieves the upper bound in a family of four-node networks with small feedback capacity.

II. Preliminaries

We consider a communication network represented by a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Source node $s \in \mathcal{V}$ transmits information to the sink node $t \in \mathcal{U}$. We use $r(a,b)$ to denote the capacity of edge $(a,b) \in \mathcal{E}$.

Let q be the size of the code alphabet $\mathcal{A}$, which can be represented as the integers $0, 1, \ldots, q-1$. A link $l \in \mathcal{E}$ transmits $r(l)$ symbols in $\mathcal{A}$ per time unit. We can view an error vector on a link $l \in \mathcal{E}$ as set of $r(l)$ symbols in $\mathcal{A}$, where the output y_l of link l equals the mod q sum of the input x_l to link l and the error e_l applied to link l, i.e.

$$y_l = x_l + e_l \mod q.$$

We say that there are τ error links in the network if $e_l \neq 0$ on τ links.

Definition 1. *A network code is z-error link-correcting if it can correct any τ adversarial links for $\tau \leqslant z$. That is, if the total number of adversarial links in the network is at most z, then the source message can be recovered by all the sink nodes $u \in \mathcal{U}$.*

Let (A,B) be a partition of $\mathcal{V}$, and define the cut for the partition (A,B) of $\mathcal{V}$ by

$$cut(A,B) = \{(a,b) \in \mathcal{E} : a \in A, b \in B\}.$$

$cut(A,B)$ is called a cut between two nodes a and b if $a \in A$ and $b \in B$. Let $CS(a,b)$ denote the set of cuts between a and b. The links in $cut(A,B)$ are called the forward links of the cut. The links (a,b) for which $a \in B$, $b \in A$ are called the feedback links of the $cut(A,B)$. The capacity of a cut is the sum of the capacities of the forward links of the cut.

For any cut $Q = cut(P, \mathcal{V}\backslash P)$, let Q^R denote the set of feedback links across cut Q. We say that a feedback link $l \in Q^R$ is directly downstream of a forward link $l' \in Q$ (and that l' is directly upstream of l) if there is a directed path starting from l' and ending with l that does not include other links in Q or Q^R.

III. New upper bound

We introduce a new upper bounding approach which, for some source-sink cut, considers confusion between two possible sets of z forward adversarial links, when there exist two codewords that differ in these forward links but coincide in the values of their directly downstream feedback links and the remaining forward links. The bound is then the sum of the capacities of these feedback links plus the capacities of the remaining forward links. Therefore, this bound is useful when the feedback link capacities are sufficiently small.

To state this result formally, consider any cut $Q = cut(P, \mathcal{V}\backslash P)$, and two disjoint sets of forward links $Z_1, Z_2 \subset Q$ where $|Z_i| \leqslant z$ for $i = 1, 2$. Let W_1 be the set of links in Q^R which are directly downstream of a link in Z_1 and upstream of a link in $Q\backslash Z_1$. Let W_2 be the set of links in Q^R which are directly downstream of a link in Z_2 and upstream of a link in $Q\backslash Z_2$.

Theorem 1. *Let*

$$M = \sum_{(a,b) \in ((Q\backslash Z_1)\backslash Z_2) \cup W_1 \cup W_2} r(a,b)$$

The first three authors are listed in alphabetical order.

denote the sum of capacities of forward links in $(Q\backslash\ Z_1)\backslash Z_2$ and feedback links in W_1 and W_2. If no link in Z_2 is directly upstream of any link in W_1, and no link in Z_1 is directly upstream of any link in W_2, then the capacity is at most M.

Proof: We assume that the codebook X contains more than q^M codewords, and show that this leads to a contradiction. Let k denote the number of links on the cut Q, and let $m = |Z_1|$, $n = |Z_2|$.

Since $|X| > q^M$, from the definition of M, there exist two distinct codewords $x, x' \in X$ such that the error-free outputs on the links in $(Q\backslash Z_1)\backslash Z_2$ and $W_1 \cup W_2$ are the same. The corresponding observations on the sink side of the cut are

$$\begin{aligned} O(x) &= \{y_1, .., y_{k-m-n}, u_1, .., u_m, w_1, .., w_n\} \\ O(x') &= \{y_1, .., y_{k-m-n}, u'_1, .., u'_m, w'_1, .., w'_n\}, \end{aligned}$$

where $(y_1, .., y_{k-m-n})$ denote the error-free outputs on the links in $(Q\backslash Z_1)\backslash Z_2$ for x and x'; $(u_1, .., u_m)$ and $(u'_1, .., u'_m)$ denote the error-free outputs on the links in Z_1 for x and x' respectively; and $(w_1, .., w_n)$ and $(w'_1, .., w'_n)$ denote the error-free outputs on the links in Z_2 for x and x' respectively.

Since no link in Z_2 is directly upstream of any link in W_1, the values on W_1 are determined by the values on $Q\backslash Z_2$. Similarly, since no link in Z_1 is directly upstream of any link in W_2, the values on W_2 are determined by $Q\backslash Z_1$.

We will show that it is possible for the adversary to produce exactly the same outputs on all the channels in Q under x and x' when errors occur on at most z links.

Assume the input of network is x. The adversary could choose forward links set Z_1 as its z adversarial links, and apply errors on Z_1 to change the output from u_i to u'_i $\forall 1 \leqslant i \leqslant m$. Note that the values on W_1 are determined by the values on $Q\backslash Z_2$, which, under these errors, are same as for x', and that the values on W_1 are the same for x and x'. Therefore, the values on W_1 are not changed, and thus the values on $Q\backslash Z_1$ are not affected. The observations on the sink side of the cut are $\{y_1, .., y_{k-m-n}, u'_1, .., u'_m, w_1, .., w_n\}$.

When codeword x' is transmitted, the adversary could choose forward links set Z_2 as its z adversarial links, and apply errors on them to change $(w'_1, .., w'_n)$ to $(w_1, .., w_n)$. Similarly, since the values on W_2 are determined by the values on $Q\backslash Z_1$, which, under these errors, are the same as for x, and since the values on W_1 are the same for x and x', the values on W_2 and on $Q\backslash Z_2$ are not affected. The observations on the sink side of the cut are $\{y_1, .., y_{k-m-n}, u'_1, .., u'_m, w_1, .., w_n\}$, the same output as before. ■

A number of variations of this result are possible. For instance, if there are no links in Q^R which are directly downstream of a link in Z_2 and upstream of a link in $Q\backslash Z_2$ (i.e. W_2 is empty), we can redefine W_1 to be the set of links in Q^R which are directly downstream of a link in Z_1 and upstream of a link in $(Q\backslash Z_1)\backslash Z_2$, which is smaller compared to the previous definition.

For the example network in Fig. 1, with two adversarial links, our previous result in [5] gives a bound of 8, whereas Theorem 1 gives a bound of $5+c$, which is tighter when $c < 3$.

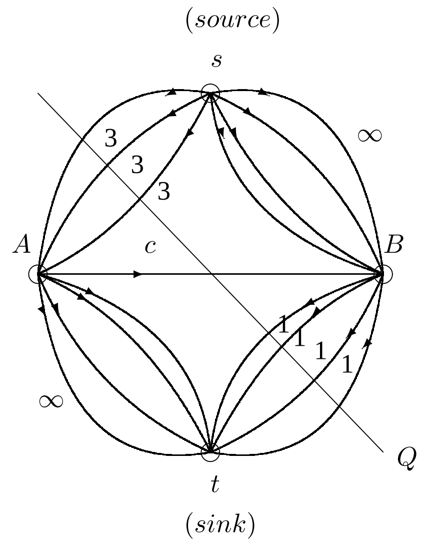

Fig. 1. Four node acyclic networks: Given the cut $Q = cut(\{s, B\}, \{A, t\})$, unbounded reliable communication is allowed from source s to its neighbor B on one side of the cut and from node A to sink t on the other side of the cut, respectively. There is a feedback link from A to B with capacity c.

IV. Achievability with small feedback link capacity

In this section, we present a new achievable code construction which exploits small-capacity feedback links more efficiently than our previous constructions in [5], which were designed for sufficiently large feedback link capacities. We show that this construction achieves the upper bound of Theorem 1 in a family of four-node networks with small feedback link capacity.

Consider the network of Figure 2, in which there are n links between source s and node A, of capacities a_i $(i = 1, \ldots, n)$, one feedback link from node A to node B of capacity b, and m links from node B to sink t, of capacities c_i $(i = 1, \ldots, m)$. Unbounded reliable communication is allowed from source s to B and from node A to sink t. Assume there are at most z adversarial links. For simplicity, we consider the case where $a_1 = a_2 = \ldots = a_n = a$, $c_1 = c_2 = \ldots = c_m = c < a$, $b \leqslant a - c$, and $n \geqslant z, m > z$. A more general construction is given in the longer version of this paper [8]. Note that if $a \leqslant c$ or $m \leqslant z$, the capacity is given by the generalized Singleton Bound in [3], and achieved without requiring the feedback link, using the construction in [3].

A. Codebook construction

Let the code alphabet be $GF(q)$ for some prime power q. Consider an $(n, n-z)$ MDS code in $GF(q^{a-c})$ mapping symbols $(g_1, g_2, \ldots, g_{n-z})$ to $(x_1, x_2, \ldots, x_n)$, and an $(n+m, n+m-2z)$ MDS code in $GF(q^c)$ mapping symbols $(g'_1, \ldots, g'_{n+m-z})$ to $(x'_1, x'_2, \ldots, x'_{n+m})$. Let

$$w_i = x_i + \alpha_i f,\ 1 \leqslant i \leqslant z,$$

where $w_i \in GF(q^{a-c})$, $\alpha_i \in GF(q^{a-c})$ and f is an element of $GF(q^b)$ but viewed as an element of $GF(q^{a-c})$. The ith

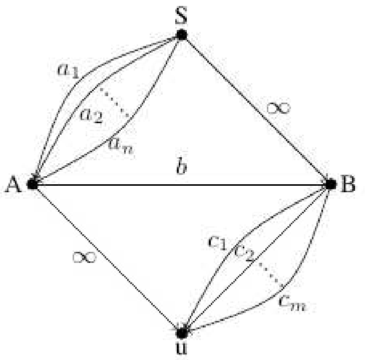

Fig. 2. Four node acyclic networks: there are n links between source s and node A, of capacities a_i $(i = 1, \ldots, n)$, one feedback link from node A to node B of capacity b, and m links from node B to sink t, of capacities $c_i(i = 1, \ldots, m)$. Unbounded reliable communication is allowed from source s to B and from node A to sink t.

link in (s, A) carries

$$\begin{array}{ll} (w_i, x'_i) & 1 \leqslant i \leqslant z \\ (x_i, x'_i) & z+1 \leqslant i \leqslant n, \end{array}$$

the feedback link carries f, and the m links in (B, u) carry

$$(x'_{n+1}, \ldots, x'_{n+m})$$

respectively. The field size q, the two MDS codes and the coefficients $\alpha_i, i \leqslant n$, are chosen such that the overall code is generic, in that any subset of links carries the maximum possible number of independent functions of the variables $g_1, g_2, \ldots, g_{n-z}, g'_1, \ldots, g'_{n+m-z}, f$.

The total number of codewords, corresponding to the number of possible values of $(g_1, g_2, \ldots, g_{n-z}, g'_1, \ldots, g'_{n+m-z}, f)$, is

$$q^{(a-c)(n-z)+c(n+m-2z)+b}.$$

Note that this matches the upper bound from Theorem 1.

B. guess-and-forward

We adapt the guess-and-forward strategy of [5] to this code construction.

If A's incoming links are not consistent with any codeword, A finds the largest subset of consistent links and send an error message to B, once for each distinct subset. If B receives an error message, or on the first time that the received feedback value $\tilde{f}$ doesn't match the correct value received from s, B sends an error message along with all the information received from s and A (equivalent to the claims in the description in [5]) to sink u using a repetition code.

Note that the code construction effectively decouples the network into two subnetworks as shown in Figure 3. The first subnetwork transmits $(g'_1, \ldots, g'_{n+m-z})$ reliably using an MDS code. In the second subnetwork, by the MDS properties of the code, for any codeword with some value of f, the adversary cannot produce another codeword with the same value of f. Also, by the generic properties of the code, the adversary must change z links in order to produce a codeword with another value $\tilde{f}$. If the adversary changes fewer than z links, the remaining links are identified by A as the largest subset of consistent links. Thus, A and B send error messages a finite number of times, each time allowing the sink to identify at least one more adversarial node, similarly to the scheme in [5].

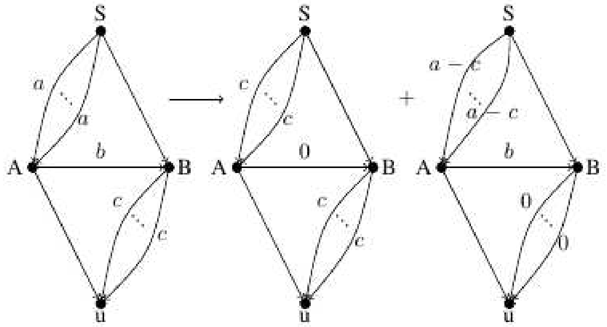

Fig. 3. Decoupling of four node network into two subnetworks.

ACKNOWLEDGEMENTS

This work is partly supported by the Air Force Office of Scientific Research under grant FA9550-10-1-0166 and Caltech's Lee Center for Advanced Networking.

REFERENCES

[1] R. Yeung and N. Cai, "Network error correction, part I: Basic concepts and upper bounds," *Communications in Information and Systems*, vol. 6, no. 1, pp. 19–36, 2006.

[2] N. Cai and R. Yeung, "Network error correction, part II: Lower bounds," *Communications in Information and Systems*, vol. 6, no. 1, pp. 37–54, 2006.

[3] S. Kim, T. Ho, M. Effros, and S. Avestimehr, "Network error correction with unequal link capacities," in *47th Annual Allerton Conference on Communication, Control, and Computing*, 2009.

[4] ——, "New results on network error correction: capacities and upper bounds," in *Information Theory and Applications Workshop (ITA), 2010*. IEEE, 2010, pp. 1–10.

[5] ——, "Network error correction with unequal link capacities," *IEEE Transactions on Information Theory*, 2011, to appear.

[6] O. Kosut, L. Tong, and D. Tse, "Nonlinear network coding is necessary to combat general byzantine attacks," in *47th Annual Allerton Conference on Communication, Control, and Computing*, 2009.

[7] ——, "Polytope codes against adversaries in networks," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*. IEEE, 2010, pp. 2423–2427.

[8] T. Ho, S. Kim, Y. Yang, M. Effros, and S. Avestimehr, "On network error correction with limited feedback capacity," 2011, in preparation.