Research Article

A Computational Perspective on Network Coding

Qin Guo,[1,2,3] Mingxing Luo,[1,2,3] Lixiang Li,[1,2,3] and Yixian Yang[1,2,3]

[1] *Information Security Center, State Key Laboratory of Networking and Switching Technology,*
  Beijing University of Posts and Telecommunications, Beijing 100876, China
[2] *Key Laboratory of Network and Information Attack and Defence Technology of MOE,*
  Beijing University of Posts and Telecommunications, Beijing 100876, China
[3] *National Engineering Laboratory for Disaster Backup and Recovery,*
  Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Lixiang Li, lixiang@bupt.edu.cn

From the perspectives of graph theory and combinatorics theory we obtain some new upper bounds on the number of encoding nodes, which can characterize the coding complexity of the network coding, both in feasible acyclic and cyclic multicast networks. In contrast to previous work, during our analysis we first investigate the simple multicast network with source rate $h = 2$, and then $h \geq 2$. We find that for feasible acyclic multicast networks our upper bound is exactly the lower bound given by M. Langberg et al. in 2006. So the gap between their lower and upper bounds for feasible acyclic multicast networks does not exist. Based on the new upper bound, we improve the computational complexity given by M. Langberg et al. in 2009. Moreover, these results further support the feasibility of signatures for network coding.

1. Introduction

When network coding was firstly used by Ahlswede et al. [1], the node produced each of its outgoing packets as an arbitrary combination of its incomings, which is referred to as encoding node. Those functions applied by all nodes in the network specify the different network codes, such as linear network codes and random linear network codes via linear function and random linear function, respectively [1–3]. In [3], Li et al. showed that linear network codes are sufficient for achieving the capacity of the network. And in a subsequent work, Koetter and Médard [4] developed an algebraic framework for network coding and studied linear network codes for cyclic networks. Based on this framework, Ho et al. [2] showed that linear network codes can be efficiently constructed through a randomized

algorithm. Jaggi et al. [5] presented a deterministic polynomial-time algorithm for finding a feasible network codes in multicast networks. Errors introduced into even a single packet transmitted on the way can propagate and pollute multiple packets making their way to the destination. To prevent the spread of the error packets, signatures for network coding are proposed [6, 7]. Nevertheless, whatever the kind of network codes will be used, the number of encoding nodes plays an important role on the encoding complexity of network coding. In [8, 9], Langberg et al. studied the design of multicast coding networks with a limited number of encoding nodes. And they showed that in a directed acyclic coding network, the number of encoding nodes required to achieve the capacity of the network is bounded by $O(h^3 k^2)$ (h is the source rate and k is the number of terminal nodes) which is independent of the size of the network. And for the general networks, in which there exist cycles, the number of encoding nodes is limited by the size of the minimum feedback link set and bounded by $(2B + 1)h^3 k^2$, where B is the minimum size of a feedback link set.

In this paper, from the perspective of graph theory [10] and combinatorics theory [11] we mainly reinvestigate the upper bounds of encoding nodes in simple multicast networks and the computational complexity of constructing network coding on feasible acyclic multicast networks. Similar with [8], we put different colors on the paths from source to different terminal nodes. However, in this paper we investigate the number of *encoding nodes* instead of *nodes* on the paths. And we find that our upper bounds are their lower bounds in [8], so our upper bounds are optimal for simple multicast networks. Using this result, we can answer the open question raised in the end of [8]; that is, there donot exist such feasible acyclic multicast networks that the number of encoding nodes are between the lower bound and the upper bound in [8]. Based on these results for acyclic multicast networks, we improve the computational complexity of the deterministic algorithm of constructing network codes given by Langberg et al. in [9].

The rest parts of this paper are organized as follows. In Section 2 we mainly give some necessary notations and definitions. In Section 3 we reinvestigate the upper bounds of encoding nodes for simple acyclic and cyclic multicast networks, respectively. Some new results on computational complexity of constructing network codes for feasible multicast networks are also given. In the final section, Section 4, some conclusions and future topics are presented.

2. Basic Notions and Definitions

Let $\mathbb{N}(G, s, T, h)$ be a multicast network with a directed graph G, a source node s, a set of terminal nodes T, and source rate h. All links in G are of unit capacity. A network code for a multicast network $\mathbb{N}(G, s, T, h)$ is said to be feasible if it allows communication at rate h between s and each terminal node $t \in T$. We say that $\mathbb{N}(G, s, T, h)$ is a feasible multicast network if there exists a feasible network code for $\mathbb{N}(G, s, T, h)$.

Definition 2.1 (minimal multicast network [8]). A feasible multicast network $\mathbb{N}(G, s, T, h)$ is said to be minimal with respect to link removal if any network $\widehat{\mathbb{N}}(\widehat{G}, s, T, h)$ formed from $\mathbb{N}$ by deleting a link e from G is no longer feasible.

Definition 2.2 (simple multicast network [8]). A multicast network $\mathbb{N}(G, s, T, h)$ is said to be simple if and only if (a) $\mathbb{N}$ is feasible; (b) $\mathbb{N}$ is minimal with respect to link removal; (c) the total degree of G each node in G is at most 3 (excluding the source and terminal nodes); (d) the terminal nodes in T have no outgoing links.
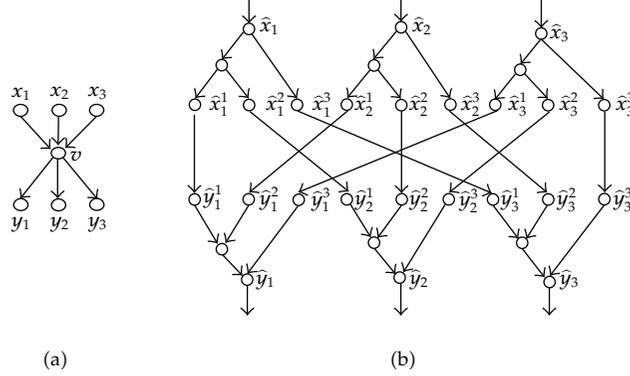
Figure 1: (a) A node $v \in G_1$ with degree larger than 3. (b) The subgraph Γ_v for the node v in (a).

In fact, for each general multicast network $\mathbb{N}(G, s, T, h)$ there exists a simple multicast network $\widehat{\mathbb{N}}(\widehat{G}, s, \widehat{T}, h)$ corresponding to it. And the construction is computationally efficient and includes the following three steps.

Step 1 (replacing terminal nodes). For each terminal node $t_i \in T$, we add a new node $\widehat{t}_i$ to G and connect t_i to $\widehat{t}_i$ by h parallel links. Denote the resulting graph by G_1 and the new set $\{t_i \mid 1 \leq i \leq k\}$ of terminal nodes by $\widehat{T}$. Then, the terminal nodes in $\widehat{T}$ have no outgoing links.

Step 2 (reducing degrees). Suppose G_2 is the graph formed from G_1 by replacing each node $v \in G_1$, $v \neq s$, $v \notin \widehat{T}$ whose degree is more than 3 by a subgraph Γ_v, constructed as Figure 1. So in G_2 there do not exist the nodes with degrees more than 3.

Step 3 (removing links). Let $\widehat{G}$ be any subgraph of G_2 such that $\widehat{\mathbb{N}}(\widehat{G}, s, \widehat{T}, h)$ is minimal with respect to link removal.

By Step 2, we can learn that the number of encoding nodes in a feasible multicast network is no more than that of the corresponding simple multicast network. In this paper, we let $\alpha(\mathbb{N})$ denote the number of encoding nodes in $\mathbb{N}$.

In general communication networks with cycles, there exist feedback links. And we will show that the value of $\alpha(\mathbb{N})$ in a cyclic network $\mathbb{N}(G, s, T, h)$ depends on the size of the minimum feedback link set.

Definition 2.3 (minimum feedback link set [12]). Let $G(V, E)$ be a directed graph. A subset $\widehat{E} \subseteq E$ is referred to as a feedback link set if the graph G' formed from G by removing all links in $\widehat{E}$ is acyclic. A feedback link set of minimum size is referred to as the minimum feedback link set.

Given a network $\mathbb{N}(G, s, T, h)$, let B denote the minimum size of feedback link set of underlying graph G.

3. Main Results

Let e_i, $1 \leq i \leq h$, be the h out links of source node s. Consider that P_r^i and P_b^i ($1 \leq i \leq h$) have the common starting link e_i. And paths P_r^i and P_b^j ($i \neq j$) have different starting links. Let
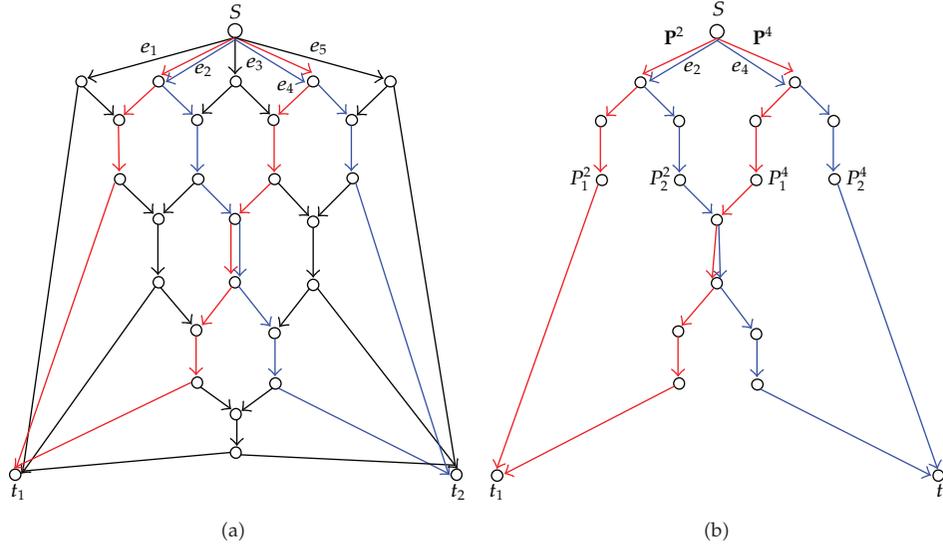
Figure 2: (a) A simple multicast network. (b) When investigating the number of encoding nodes on the paths in the sets $\mathbf{P}^2$ and $\mathbf{P}^4$, we ignore other paths which donot start at the links e_2 and e_4. So the only four paths P_1^2, P_1^4, P_2^2, and P_2^4 are left.

paths set $\mathbf{P}^i = \{P_j^i \mid j = 1, 2, \ldots, k\}$. In addition, all the paths in the sets $\{\mathbf{P}^i\}_{i=1}^n$ i can cover the whole network of G. If some link e belongs to G but any path P_n^i, then $\mathbb{N}(G, s, T, h)$ will not be a simple multicast network, because this contradicts the minimality of $\mathbb{N}(G, s, T, h)$.

Similar to [8], we also put different colors on the paths from source node s to different terminal nodes. While the color of a link is the same to the color of the path which passes on it. Of course, one link can have several different colors if there are several paths to different terminal nodes passing on it. When all the colors of a link are blotted out, we consider that this link disappears.

3.1. New Upper Bounds

In this subsection, we establish the number of encoding nodes instead of nodes on a pair of link-disjoint paths. Since the source rate h may be very large, we investigate pairs of link-disjoint paths to terminal nodes one by one. For example, when investigating the pair of path sets $\mathbf{P}^i$ and $\mathbf{P}^j$ ($1 \leq i \neq j \leq h$), we only leave all the paths in the two sets $\mathbf{P}^i$ and $\mathbf{P}^j$ in which the paths have the same starting link e_i and e_j, respectively, and temporarily ignore other paths in the paths set $\mathbf{P}^l$ in which the paths have other starting links e_l, $l \neq i, j$, see Figure 2. When investigating the number of encoding nodes on the paths in the sets $\mathbf{P}^2$ and $\mathbf{P}^4$ which have the starting links e_2 and e_4, respectively, we only leave the paths P_1^2, P_1^4, P_2^2, and P_2^4 (see Figure 2(b)).

During the following analysis, we first investigate the simple multicast networks with $h = 2$, and then generalize this result to the networks with $h \geq 2$. So our method is very different from that given by Langberg et al. in [8].
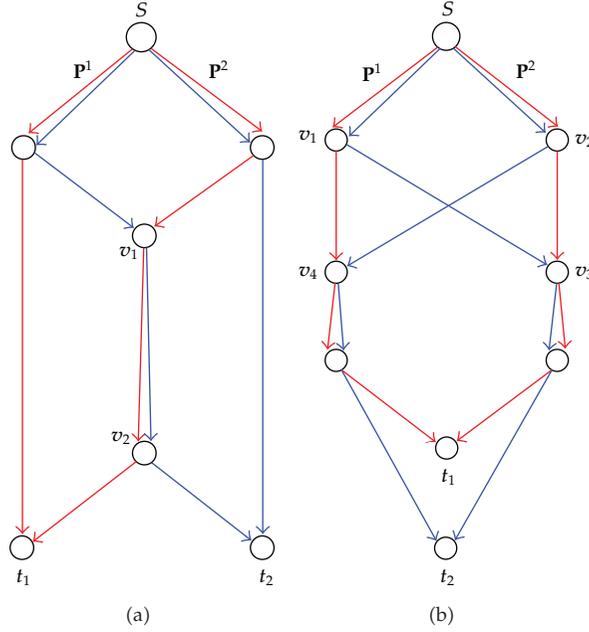
Figure 3: (a) A simple multicast network with two terminals has only one encoding node; (b) A feasible multicast network with two terminals has two encoding nodes. But it is not a simple multicast network, because if we delete (v_1, v_3) and (v_2, v_4), this multicast network is also feasible.

Lemma 3.1. *Let $\mathbb{N}(G, s, T, 2)$ be a simple acyclic multicast network. Then the number of encoding nodes $\alpha(\mathbb{N})$ of $\mathbb{N}$ is at most $k - 1$, where $k = |T|$.*

Proof. We prove this result by induction on k. For the base step, we note that a simple multicast network $\mathbb{N}(G, s, \{t_1, t_2\}, 2)$ with $k = 2$ has at most one encoding node (see, e.g., Figure 3(a)). If there exist two encoding nodes, then $\mathbb{N}(G, s, \{t_1, t_2\}, 2)$ can only be characterized by Figure 3(b) (i.e., Figure 3(b) as a subnetwork of $\mathbb{N}(G, s, \{t_1, t_2\}, 2)$). However, we find that the two links (v_1, v_3) and (v_2, v_4) are redundant, since when we delete them, the network is also feasible. It means that Figure 3(b) is not a simple multicast network, and then $\mathbb{N}(G, s, \{t_1, t_2\}, 2)$ is not simple which contradicts the assumption above. Similarly, we can prove that more than two encoding nodes are impossible for $\mathbb{N}(G, s, \{t_1, t_2\}, 2)$.

For the induction step, we assume that $\alpha(\mathbb{N}) \leq k - 1$ for $k = 3, \ldots, n$. Now, we need to prove $\alpha(\mathbb{N}) \leq n$ for $k = n + 1$. By reduction to absurdity, suppose that $\alpha(\mathbb{N}) = n + 1$ for $\mathbb{N}(G, s, T, 2)$ with $k = n + 1$. Notice that in $\mathbb{N}(G, s, T, 2)$ there does not exist such a color of pair paths to any terminal: when we blot out this color, two encoding nodes disappear. Assume there exists such a color. Then we can obtain two classes of subnetworks shown in Figure 4. In Figure 4(a), v_5 and v_6 are two such encoding nodes for terminal t_1. For each of v_5 and v_6, one of the two incoming links has only one red color, such as (v_1, v_5) and (v_4, v_6). Otherwise, when we blot out red color, both of v_5 and v_6 are not encoding nodes anymore at the same time. But we find that the two links (v_1, v_5) and (v_4, v_6) are redundant, that is to say this network is not a simple multicast network. Because the messages can be transmitted along the paths (s, v_2), (v_2, v_5), (v_5, v_7), (v_7, t_1) and (s, v_3), (v_3, v_6), (v_6, v_8), (v_8, t_1) to terminal t_1. Similarly, in Figure 4(b), v_1 and v_3 are two encoding nodes. We find that the blue incoming
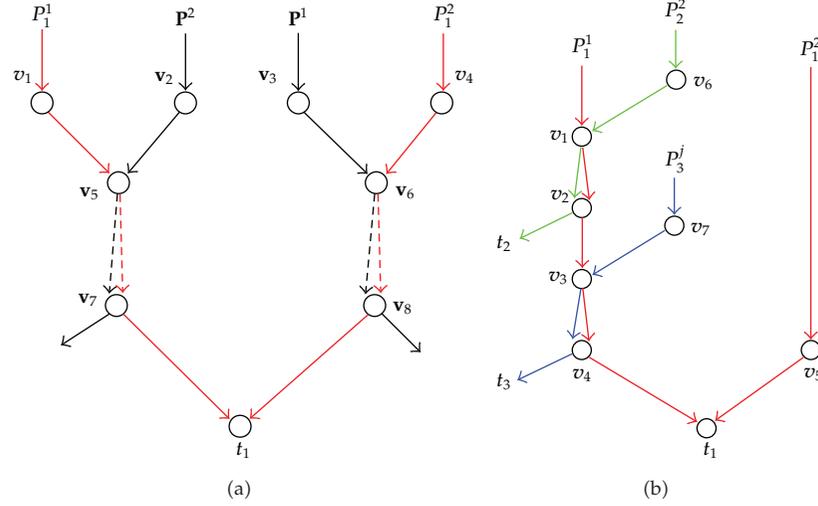
Figure 4: (a) One subnetwork of a feasible acyclic multicast network, in which if we blot out the red color, two encoding nodes v_5 and v_6 disappear. Here, at least one color path passes through the bold black links; (b) The other subnetwork in which if we blot out red color, two encoding nodes v_1 and v_3 disappear.

link of v_3 is redundant for this multicast network, because the path P_3^2 can go along the path (s, v_6), (v_6, v_1), (v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, t_3). So, when we blot out one color, only one encoding node will disappear, and the corresponding terminal is no longer a terminal. Then in the rest simple multicast network there exist n terminals and $\alpha(\mathbb{N}) = n$. This contradicts the assumption $\alpha(\mathbb{N}) \leq n - 1$ for $k = n$.

Therefore, the number of encoding nodes in any simple acyclic multicast network $\mathbb{N}(G, s, T, 2)$ is no more than $k - 1$. □

By the decomposition and combination of a simple acyclic multicast network, we can generalize this result to the multicast networks with any $h \geq 2$.

Theorem 3.2. *Let $\mathbb{N}(G, s, T, h)$ be a simple acyclic multicast network. Then,*

$$\alpha(\mathbb{N}) \leq \frac{(k-1)h(h-1)}{2} = O\left(h^2 k\right). \tag{3.1}$$

Proof. We investigate a pair of path sets $\mathbf{P}^i$ and $\mathbf{P}^j$ for all terminals. And all of the paths in these two sets compose a subnetwork $\mathbb{N}^{i,j}(G^{i,j}, s, T, 2)$ of $\mathbb{N}(G, s, T, h)$ which is also a simple multicast network. By Lemma 3.1, there exist at most $k - 1$ encoding nodes in $\mathbb{N}^{i,j}(G^{i,j}, s, T, 2)$. And each encoding node in $\mathbb{N}(G, s, T, h)$ will be counted in some subnetwork $\mathbb{N}^{i,j}(G^{i,j}, s, T, 2)$. Suppose that an encoding node v is not counted and its two incoming links are e_1 and e_2. By the minimality of $\mathbb{N}(G, s, T, h)$, e_1 and e_2 must be on some paths, which belong to the path sets $\mathbf{P}^m$ and $\mathbf{P}^n$, respectively. Then v, as an encoding node, is in the corresponding subnetwork $\mathbb{N}^{m,n}(G^{m,n}, s, T, 2)$. This contradicts the assumption above. In $\mathbb{N}(G, s, T, h)$ there are $h(h-1)/2$

pairs of paths; that is, there are $h(h-1)/2$ such subnetworks in all, so

$$\alpha(\mathbb{N}) \leq \frac{(k-1)h(h-1)}{2} = O\left(h^2 k\right). \tag{3.2}$$

$\square$

Being different from the results that the numbers of encoding nodes for $k = 2$ and $k > 2$ have different descriptions in [8], this theorem is a unified presentation for all simple acyclic multicast networks. Notice that by Lemma 12 of [8], each feasible multicast network $\mathbb{N}(G, s, T, h)$ corresponds to a simple multicast network $\widehat{\mathbb{N}}(\widehat{G}, s, \widehat{T}, h)$. Moreover, the minimum number of encoding nodes required for $\mathbb{N}(G, s, T, h)$ is no more than that for $\widehat{\mathbb{N}}(\widehat{G}, s, \widehat{T}, h)$. Therefore, by Lemma 12 of [8] and Theorem 3.2 above we can show the existence of network codes with at most $(k-1)h(h-1)/2$ encoding nodes for any feasible acyclic multicast network.

Theorem 3.3. *Let G be an acyclic graph and $\mathbb{N}(G, s, T, h)$ be a feasible multicast network. Then, there exists a feasible network code with at most $(k-1)h(h-1)/2$ encoding nodes.*

From the investigation above, we can find that our method is very different from that given by Langberg et al. in [8]. On one hand, they investigated the simple multicast networks based on the number of terminal nodes, classified by $k = 2$ and $k > 2$, and got two different results for these two kinds of networks. In this paper, we classify the simple multicast networks by source rate h. One is with $h = 2$, and the other is with $h > 2$. We first investigate the simple multicast networks with $h = 2$, then generalize this result to the networks with any $h \geq 2$.

On the other hand, by scaling law Langberg et al. estimated the number of *nodes* instead of *encoding nodes* on the three kinds of paths: red path P_r, blue path P_b and green path P_g (defined in [8]). And they proved that there exists at most one node that belongs to all three paths P_r, P_b, and P_g. Though each encoding node belongs to all these color paths, there exist such nonencoding nodes that belong to all three color paths. For example, see Figure 3 (a). By the definitions of three color paths in [8], paths $P_{r_1} = \{(s, v_1), (v_1, t_1)\}$ and $P_{r_2} = \{(s, v_2), (v_2, v_3), (v_3, v_4), (v_4, t_1)\}$ are two red paths, paths $P_{b_1} = \{(s, v_1), (v_1, v_3), (v_3, v_4), (v_4, t_2)\}$ and $P_{b_2} = \{(s, v_2), (v_2, t_2)\}$ are two blue paths, and paths $P_{g_1} = \{(t_1, v_1), (v_1, v_3), (v_3, v_2), (v_2, t_2)\}$ and $P_{g_2} = \{(t_1, v_4), (v_4, t_2)\}$ are two green paths. v_1 belongs to three color paths P_{r_1}, P_{b_1}, and P_{g_1}, v_2 belongs to P_{r_2}, P_{b_2}, and P_{g_1}, v_3 belongs to P_{b_1}, P_{r_2}, and P_{g_1}, and v_4 belongs to P_{b_1}, P_{r_2}, and P_{g_2}. Obviously, only v_3 is an encoding node, but v_1, v_2, and v_4. This shows that there exist nonencoding nodes that belong to all three paths of different colors. So they get a higher upper bound. In our paper, based on the properties of encoding nodes and simple acyclic multicast networks, we investigate the number of *encoding nodes* in simple acyclic multicast networks directly and obtain the new upper bound. And we find that our upper bound is just their lower bound. Therefore, the gap of hk between their lower and upper bounds for acyclic networks is inexistent which is the answer to the open question in the end of [8].

Next, we consider the networks with cycles. Based on Lemma 3.1, we also establish the number of encoding nodes in simple cyclic multicast networks. Before investigating general simple multicast networks, we first study the ones with $h = 2$.

Lemma 3.4. *Let $\mathbb{N}(G, s, T, 2)$ be a simple cyclic multicast network and B be the size of the minimal feedback link set in $\mathbb{N}$. Then, the number of encoding nodes $\alpha(\mathbb{N}) \leq (k-1)(B+1)$.*
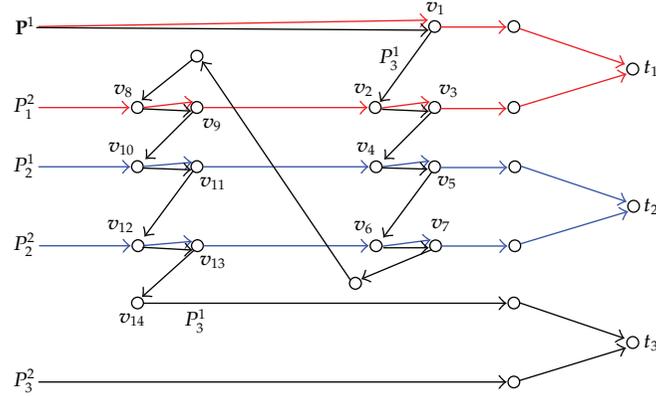
Figure 5: A simple cyclic multicast network in which paths P_3^1 and P_2^1 have common nodes.

Proof. By Lemma 3.1, if there are no feedback links, then there exist at most $k - 1$ encoding nodes in $\mathbb{N}(G, s, T, 2)$, and if we add one feedback link in $\mathbb{N}(G, s, T, 2)$, then there will exist at most $B(k - 1)$ additional encoding nodes. Because the path on which the feedback link passes does not intersect with the paths with the same starting edge. Otherwise, there will exist redundant links for the simple multicast networks. See Figure 5, for example, in which the black path from v_1 to v_{12} is a feedback link. The paths P_3^1 and P_2^1 have common starting link e_1, and they have the nodes of intersection: v_4 and v_{10}. Since the path from source node s to t_3 can go along with the path $\{(s, v_{10}), (v_{10}, v_{11}), (v_{11}, v_{12}), (v_{12}, v_{13}), (v_{13}, v_{14}), (v_{14}, t_3)\}$, the black link from v_1 to v_{10} is redundant. And this contradicts the minimality of the simple multicast network. □

In the following theorem, we establish the upper bounds on the size of $\alpha(\mathbb{N})$ for simple cyclic multicast networks with source rate $h \geq 2$.

Theorem 3.5. *Let* $\mathbb{N}(G, s, T, h)$ *be a simple cyclic multicast network. Then*

$$\alpha(\mathbb{N}) \leq \frac{(B+1)(k-1)h(h-1)}{2} = O\big((B+1)h^2 k\big). \tag{3.3}$$

Proof. By Lemma 3.4, the proof is similar with that of Theorem 3.2. □

We can find that if $B = 0$, that is to say there do not exist feedback links in $\mathbb{N}(G, s, T, h)$, in Lemma 3.4 and Theorem 3.5, then they will degenerate into Lemma 3.1 and Theorem 3.2, respectively.

By comparing our Theorem 3.2 with Theorem 6 in [8], we find that their lower bound $O(h^2 k)$ is indeed the real upper bound of the number of encoding nodes in simple acyclic multicast networks. So our upper bound is the optimal one. Moreover, by using Lemma 3.4 and Theorem 3.2 we have established a much tighter upper bound for cyclic network, shown in Theorem 3.5. In addition, for each encoding node in an arbitrary simple multicast network there exists at least one corresponding nonencoding node at which the joint flows are split. So it is impossible for us to construct a cyclic minimal multicast network with the number of encoding nodes more than $|V|/2$, where $|V|$ is the total number of nodes in the network.

Therefore, their lower bound of $O(\max(|V|/2, Bh, h^2k))$ in [8] can also be considered as another upper bound of encoding nodes in cyclic networks.

In fact, our these results are further in favor of the feasibility of signatures for network coding. In the scheme of signature for network coding, when an intermediate node is a nonencoding node, the operations are in two steps at this node: (i) verify the signature for incoming packet; (ii) transmit the uncorrupted packet or discard corrupted packet. While at an encoding node, the operations are in four steps: (i) verify the two signatures for incoming packets; (ii) encode the two incoming packet if they are not corrupted, or discard it (or them) when corrupted. (iii) sign the encoded packet; (iv) transmit the signed packet. So the highly consuming signature operations are at the encoding nodes. While, by the results above, the number of encoding node is independent of the size of the network and its upper bound is $(B + 1)h^2k/2$ ($B = 0$ when the network is acyclic).

3.2. New Computational Complexity of Network Coding Construction

In this subsection, we mainly extend the results on the computational complexity of network coding constructions, shown in [9]. Firstly, we briefly show some notations used in this subsection.

For two multicast coding networks $\mathbb{N}(G(V, E), s, T, h)$ and $\widehat{\mathbb{N}}(\widehat{G}(\widehat{V}, \widehat{E}), \widehat{s}, \widehat{T}, \widehat{h})$, we say that $\widehat{\mathbb{N}}$ models $\mathbb{N}$ if the following three conditions hold.

(i) $\mathbb{N}$ is feasible if and only if $\widehat{\mathbb{N}}$ is feasible.

(ii) For any feasible network code $\widehat{\mathbb{F}}(\widehat{\mathbb{N}})$ for $\widehat{\mathbb{N}}$, there exists a corresponding network code $\mathbb{F}(\mathbb{N})$ for $\mathbb{N}$ that includes the same number of encoding nodes or less.

(iii) Given a feasible network code $\widehat{\mathbb{F}}(\widehat{\mathbb{N}})$ for $\widehat{\mathbb{N}}$, the corresponding network code $\mathbb{F}(\mathbb{N})$ for $\mathbb{N}$ can be found through an efficient procedure whose running time is bounded by $O(|E| + |\widehat{E}|)$.

In addition, we use $\mathbb{N}'(G', s, T, h)$, $\mathbb{N}^*(G^*, s, T, h)$, and $\widehat{\mathbb{N}}(\widehat{G}, s, T, h)$ to denote three auxiliary coding networks, all of which model $\mathbb{N}(G, s, T, h)$. They are constructed by algorithms Procedure EXPAND, Algorithm MIN-GLOBAL, and Procedure SHRINK (shown in [9] for details) as the following steps.

Step 1. $\mathbb{N}'(G', s, T, h)$ = Pro. EXPAND ($\mathbb{N}(G, s, T, h)$).

Step 2. $\mathbb{N}^*(G^*, s, T, h)$ = Alg. MIN-GLOBAL ($\mathbb{N}'(G', s, T, h)$).

Step 3. $\widehat{\mathbb{N}}(\widehat{G}, s, T, h)$ = Pro. SHRINK ($\mathbb{N}^*(G^*, s, T, h)$).

From the proofs of Theorems 6 and 7 in [9], we find that their upper bound $O(h^3)$ for an acyclic multicast network with two terminal nodes has greatly influenced these two theorems. Based on our new results on the upper bound of encoding nodes in feasible multicast networks, we give two farther results as follows.

Theorem 3.6. *Let $\mathbb{N}^*(G^*(V^*, E^*), s, T, h)$ be the coding network returned by Algorithm Min-Global($\mathbb{N}'(G', s, T, h)$). Let $\overline{V}^*$ be the subset of $V^* \setminus T$ that includes nodes of in-degree two and $\overline{E}^*$ be the set of incoming links of nodes in $\overline{V}^*$. Then, it holds that $|\overline{E}^*| = O(h^2k)$.*

Proof. By Theorems 1, the proof follows the same line as that of Theorem 6 in [9]. □

Theorem 3.7. *Let $\mathbb{N}(G, s, T, h)$ be an acyclic feasible coding network. Then there exists a deterministic algorithm that computes a network code $\mathbb{F}(\mathbb{N})$ in time $O(|E|kh + |V|k^2h^2 + h^3k^2(k + h))$. Moreover, the number of encoding nodes in $\mathbb{F}(\mathbb{N})$ is bounded by $O(h^2k)$.*

Proof. By Theorems 3.2 and 3.6, the proof follows the same line as that of Theorem 7 in [9]. □

So the complexity $O(|E|kh + |V|k^2h^2 + h^3k^2(k + h))$ of computing a feasible network code $\mathbb{F}(\mathbb{N})$ in Theorem 7 of [9] is in fact $O(|E|kh + |V|k^2h^2 + h^4k^3(k + h))$ and their bound on the number of encoding nodes $O(h^3k^2)$ in this feasible network code is really $O(h^2k)$. This is the best known running time for constructing network code.

The following two theorems are about integral and fractional network codes. In an integral network, codes packets cannot be split and have to be sent through the network in one piece. In fractional network codes, each packet can be split into a number of smaller packets, each of which is sent over different paths. We assume that all integral packets are elements in finite field $GF(2^n)$, which implies that each such packet can be represented by n bits. m-Fractional Network Code $\mathbb{F}_m(\mathbb{N})$), let $\mathbb{N}(G, s, T, h)$ be a feasible coding network, n be the size of the integral packet, and let m be a divisor of n. Suppose $\mathbb{N}_m(G_m, s, T, mh)$ be the coding network in which G_m is formed from G by splitting each link e in G of bit capacity c_e into m parallel links $e_1, \ldots, e_m$ of bit capacity c_e/m. Let $\mathbb{F}(\mathbb{N}_m)$ be a feasible integral network code for $\mathbb{N}_m(G_m, s, T, mh)$ over finite field $GF(2)^{n/m}$. We refer to $\mathbb{F}(\mathbb{N}_m)$ as a feasible m-fractional network code $\mathbb{F}_m(\mathbb{N})$ for $\mathbb{N}(G, s, T, h)$.

Theorem 3.8. *Let $\mathbb{N}(G, s, T, h)$ be a feasible acyclic network. Then a feasible integral network code $\mathbb{F}(\mathbb{N})$ can be constructed in time $|V|^{O(h^4k^2)}$.*

Proof. By Theorem 3.2, the proof follows the same line as that of Theorem 11 in [9]. □

By this theorem, we can find that the complexity $|V|^{O(h^6k^4)}$ of constructing a feasible integral network code $\mathbb{F}(\mathbb{N})$ in Theorem 11 of [9] is $|V|^{O(h^4k^2)}$.

Theorem 3.9. *Let $\mathbb{N}(G, s, T, h)$ be a feasible acyclic network. Then a feasible m-fractional network node $\mathbb{F}_m(\mathbb{N})$ for $\mathbb{N}$ can be constructed in time $|V|^{O(h^2k)}$.*

Proof. By Theorem 3.2, the proof is similar to that of Theorem 16 in [9]. □

From this result, we can find that the complexity $|V|^{O(h^3k^2)}$ of constructing a feasible m-fractional network code in Theorem 16 of [9] is actually $|V|^{O(h^2k)}$.

4. Conclusion

In this paper, we reinvestigated the upper bounds of encoding nodes in acyclic and cyclic networks and answered the open question in [8]: the gap of hk between their lower and upper bounds for acyclic networks does not exist. Then, we gave out some new upper bounds for cyclic networks, which are tighter than those in [8]. The number of encoding nodes required to achieve the capacity of the network is independent of the size of the network. This fact is in favor of the feasibility of signatures for network coding to a certain extent. Furthermore, we also give a new result on the computational complexity of deterministic algorithm given by Langberg et al. [9], $O(|E|kh + |V|k^2h^2 + h^3k^2(k + h))$. This is the best known running time for constructing network codes. At last, some new results on computational complexity

of feasible integral and fractional network codes constructions are given for feasible acyclic networks.

In this paper, we have established a much tighter upper bound for simple cyclic multicast networks, while this upper bound is not the optimal one. Therefore, like our result for simple acyclic multicast networks, to find the optimal upper bound for simple cyclic multicast networks is an interesting future topic. In addition, based on our current results whether a new algorithm of constructing network codes for cyclic multicast network with lower computational complexity can be found is an important further topic. Moreover, in future it is meaningful to discuss the relationships between the number of encoding nodes and the time-delay of network communication using network coding, between the number of encoding nodes and the optimizations of network resources (the encoding node can be viewed as the intelligent node in wireless sensor networks) and so on.

Acknowledgments

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] T. Ho, M. Médard, R. Koetter et al., "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.

[3] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 5, pp. 782–795, 2003.

[5] S. Jaggi, P. Sanders, P. A. Chou et al., "Polynomial time algorithms for multicast network code construction," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1973–1982, 2005.

[6] D. Charles, K. Jain, and K. Lauter, "Signatures for network coding," *International Journal of Information and Coding Theory*, vol. 1, pp. 3–14, 2009.

[7] D. Boneh, D. Freeman, J. Katz, and B. Waters, "Signing a linear subspace: signature schemes for network coding," in *Public Key Cryptography—PKC 2009*, vol. 5443 of *Lecture Notes in Computer Science*, pp. 68–87, Springer, Berlin, UK, 2009.

[8] M. Langberg, A. Sprintson, and J. Bruck, "The encoding complexity of network coding," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2386–2397, 2006.

[9] M. Langberg, A. Sprintson, and J. Bruck, "Network coding: a computational perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 147–157, 2009.

[10] J. Bang-Jensen and G. Gutin, *Digraphs. Theory, Algorithms and Applications*, Springer, London, UK, 2nd edition, 2008.

[11] V. Krishnamurthy, *Combinatorics: Theory and Applications*, Ellis Horwood Series in Mathematics and Its Applications: Statistics and Operational Research, Ellis Horwood, Chichester, UK, 1986.

[12] M. R. Garey and D. S. Johnson, *Computers and Intractability*, W. H. Freeman, San Francisco, Calif, USA, 1979.