IEICE TRANS. FUNDAMENTALS, VOL.E91–A, NO.10 OCTOBER 2008

PAPER Special Section on Information Theory and Its Applications

Strongly Secure Linear Network Coding

Kunihiko HARADA[†a)], *Nonmember and* **Hirosuke YAMAMOTO**[††b)], *Fellow*

SUMMARY In a network with capacity h for multicast, information $X^h = (X_1, X_2, \cdots, X_h)$ can be transmitted from a source node to sink nodes without error by a linear network code. Furthermore, secret information $S^r = (S_1, S_2, \cdots, S_r)$ can be transmitted securely against wiretappers by k-secure network coding for $k \leq h-r$. In this case, no information of the secret leaks out even if an adversary wiretaps k edges, i.e. channels. However, if an adversary wiretaps $k+1$ edges, some S_i may leak out explicitly. In this paper, we propose strongly k-secure network coding based on strongly secure ramp secret sharing schemes. In this coding, no information leaks out for every $(S_{i_1}, S_{i_2}, \cdots, S_{i_{r-j}})$ even if an adversary wiretaps $k+j$ channels. We also give an algorithm to construct a strongly k-secure network code directly and a transform to convert a nonsecure network code to a strongly k-secure network code. Furthermore, some sufficient conditions of alphabet size to realize the strongly k-secure network coding are derived for the case of $k < h-r$.

key words: *network coding, secure network coding, linear network coding, secret sharing schemes, strong ramp secret sharing schemes*

1. Introduction

A communication network like the Internet can be modeled by a graph, in which each edge and each node correspond to a channel and a computer, respectively. Each node encodes information received from its incoming edges and sends the encoded information to other nodes via outgoing edges. Ahlswede-Cai-Li-Yeung [1] treated a multicast coding problem such that a source node in a network sends the same information to several sink nodes without error. They showed that network coding can increase the capacity of the information than the case of routing, and the capacity is given by the minimum of the max-flows from a source node to sink nodes. Furthermore, Li-Yeung-Cai [2] showed that a linear code can attain the multicast capacity of a network.

In network coding, secure transmission against wiretappers was treated by Cai and Yeung [3]. They showed that if a network has capacity h, we can realize k-secure network coding such that secret information $S^r = (S_1, S_2, \cdots, S_r)$, $r = h-k$, can be transmitted to every sink node, which has at least h edges, without error and no information of S^r leaks out even if an adversary wiretaps any k edges. Furthermore, Feldman-Malkin-Servedio-Stein(FMSS) [4] showed that a nonsecure linear network code can be transformed linearly to a k-secure network code.

We note that the above secure network coding is closely related to (h, r, N)-threshold ramp secret sharing schemes (SSSs), which were studied by Yamamoto [5] and Blakley-Meadows [6] independently. In the (h, r, N)-threshold ramp SSS, secret information S^r is encoded to N shares such that S^r can be recovered from any h shares, but no information of S^r leaks out from any $k(= h-r)$ shares.

In [5], Yamamoto classified the ramp SSSs into weakly secure ramp SSSs and strongly secure ramp SSSs. In the case of weakly secure ramp SSSs, some S_i might leak out explicitly if an adversary gets more than k shares. But, in the case of strongly secure ramp SSSs, no information leaks out for every $(S_{i_1}, S_{i_2}, \cdots, S_{i_{r-j}})$ even if an adversary gets $k+j$ shares. Hence, the strongly secure ramp SSSs are more secure than the weakly secure ramp SSSs.

The k-secure network coding treated in [3] and [4] corresponds to the weakly secure ramp SSSs. But, in this paper, we propose strongly k-secure network coding based on the strongly secure ramp SSSs.

In Sect. 2, we review the results of network coding and ramp SSSs. Some notations are also given in Sect. 2. In Sect. 3, we define the strongly k-secure network code such that no information leaks out for every $(S_{i_1}, S_{i_2}, \cdots, S_{i_{r-j}})$ even if an adversary wiretaps $k+j$ channels. We also derive an algorithm to construct a strongly k-secure network code directly for a given network. Furthermore, in Sect. 4, we give a transform to convert a nonsecure network code to a strongly k-secure network code. These construction algorithm and transform can always realize the strongly k-secure network coding for a given multicast network without cycle if the alphabet size of information is sufficiently large. Some sufficient conditions of the alphabet size to realize the strongly k-secure network coding are derived for the case of $k < h-r$ in Sect. 6. Some examples of the strongly k-secure network codes are given in Sect. 5.

2. Notations and Preliminaries

In this section, we define some notations used in this paper and we review some known results for network coding and ramp secret sharing schemes.

Manuscript received January 25, 2008.
Manuscript revised April 17, 2008.

[†]The author is with the Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8656 Japan.
[††]The author is with the Department of Complexity Science and Engineering, Graduate School of Frontier Science, The University of Tokyo, Tokyo, 277-8561 Japan.
a) E-mail: harada@it.k.u-tokyo.ac.jp
b) E-mail: Hirosuke@ieee.org

DOI: 10.1093/ietfec/e91–a.10.2720

2.1 Network Coding

A communication network can be represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ are the sets of all nodes and all edges, respectively. Each node corresponds to an encoder in the network and each edge represents a channel between two nodes. Hence, $\mathcal{E}$ satisfies that $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ and

$$\mathcal{E} = \{(u, v) | u \in \mathcal{V} \text{ and } v \in \mathcal{V} \text{ such that there is an edge from } u \text{ to } v\}.$$

The cardinality of $\mathcal{V}$ and $\mathcal{E}$ are represented by $|\mathcal{V}|$ and $|\mathcal{E}|$, respectively. We denote starting and ending nodes of edge e by tail(e) and head(e), respectively. For edge $e = (u, v)$, e is said to be an incoming edge of v and an outgoing edge of u. We assume that every edge is a noiseless channel which can transmit any symbol in a finite alphabet $\mathcal{X}$ without error.

If an edge can transmit c symbols at once, we divide the edge into c parallel edges, each of which can transmit one symbol at once. Then, the capacity of every edge can be normalized as one without loss of generality. We assume in this paper that $\mathcal{G}$ is normalized in this way.

We also assume for simplicity that $\mathcal{G}$ is *acyclic*, that is, $\mathcal{G}$ does not have a directed cycles $(v_1, v_2)(v_2, v_3) \cdots (v_{i-1}, v_i)$ (v_i, v_1) nor a loop (v, v). In this paper, we consider multicast networks such that only one *source node* $s \in \mathcal{V}$ generates information $X^n = (X_1, X_2, \cdots, X_n)$, where X_i is an independent, identically and uniformly distributed random variable over an alphabet $\mathcal{X}$, and every sink node $t_1, t_2, \cdots t_L \in \mathcal{V}$ must recover the information X^n without error. A main problem in the multicast network coding is to clarify how large n is achievable for an arbitrarily given network $\mathcal{G}$.

The network coding must satisfy the following natural requirements.

Definition 1 (network code): A network code is a code that satisfies the following.

- The outputs of a source node are encoded from the information generated at the node, and they are sent via its outgoing edges.
- Each sink node decodes X^n from the information received via its incoming edges.
- At a general node, its outputs are encoded from the information received via its incoming edges, and the outputs are sent via its outgoing edges.

We assume that every encoding at every node and every transmission on every edge have no delay. Then, Ahlswede-Cai-Li-Yeung [1] showed that the achievable n is determined by the max-flow bound as shown in the following theorem.

Theorem 1 (max-flow bound [1]): For any given network $\mathcal{G}$, there exists a network code that can transmit X^n generated at a source node s to all sink nodes $t_1, t_2, \cdots, t_L$ if and only if n satisfies that

$$n \le \min_{t \in \mathcal{T}} \text{maxflow}(s, t), \tag{1}$$

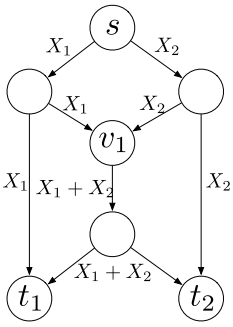

Fig. 1 An example of linear network coding.

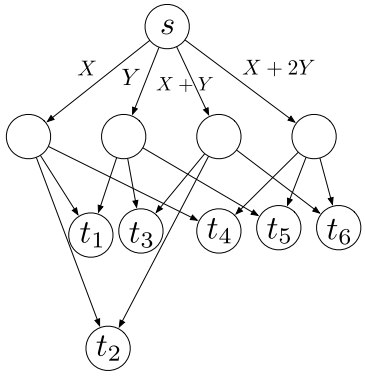

Fig. 2 An example that requires $q \ge 3$.

where $\mathcal{T} = \{t_1, t_2, \cdots, t_L\}$ and maxflow(s, t) stands for the maximum flow from node s to node t in $\mathcal{G}$, which is equal to the total capacity given by the minimum cut between s and t [7].

If all encoding in a network is implemented by linear operations on $\mathcal{X} = \mathbb{F}_q$ which is a finite field with cardinality q, then it is called linear network coding.

Figure 1 shows a well-known example of a linear network code, in which source node s generates (X_1, X_2) and two sink nodes t_1 and t_2 must decode (X_1, X_2) without error. In this network, we note that operation $X_1 + X_2$ at node v_1 is a linear operation. It is shown in [2] that a linear network code can achieve the bound of Eq. (1) with equality if q is sufficiently large. But, if q is not large, the equality in Eq. (1) cannot always be attained. For example, consider a network given in Fig. 2, in which $X + 2Y$ must be transmitted on one of the outgoing edges of the source node to send (X, Y) to all the sink nodes. This means that the cardinality of the finite field, q, must be larger than 2 because $\mathbb{F}_q$ must include $\{0, 1, 2\}$ at least. Therefore, the necessary alphabet size is an important factor in the linear network coding, and it is shown that the linear network coding can be realized if $q \ge |\mathcal{T}|$ [8], [9].

In this paper, we treat the multicast linear network coding such that a source output X^h with $h = \min_{t \in \mathcal{T}} \text{maxflow}(s, t)$ is transmitted to all the sink nodes in $\mathcal{T}$. The information flowing on each node is a linear transform of X^h, and it can be represented by a *coding vector*,

which is defined as follows.

Definition 2 (coding vector): A coding vector $b(e)$ assigned to an edge e is an h-dimensional column vector which satisfies the following conditions.

1. A flow on edge e, say W_e, is given by $W_e = X^h b(e)$.
2. Let $v = \mathrm{tail}(e)$ be not the source node s. Then, $b(e)$ is generated by a linear combination of the coding vectors assigned to the incoming edges of the node v.
3. For each sink node t, all the coding vectors assigned to the incoming edges of t must span the h-dimensional vector space.

For a subset of edges $\mathcal{A} \subseteq \mathcal{E}$, let $Z_{\mathcal{A}}$ be the matrix obtained by concatenating all the coding vectors of $\mathcal{A}$. Furthermore, let $Z \equiv Z_{\mathcal{E}}$ for simplicity. Note that Z describes all encoding in a network. It is known that Z can be derived by a randomized algorithm [10] and deterministic algorithms [9], [11], [12]. In this paper, we use the deterministic algorithm given by Jaggi-Sanders-Chau-Tolhuizen [9] to construct strongly secure network codes.

2.2 Ramp Secret Sharing Schemes

A secret sharing scheme (SSS) is a method to encrypt a secret information $S^r = (S_1, S_2, \cdots, S_r)$ into N shares, $V_1, V_2, \cdots, V_N$. In the case of (h, r, N) ramp threshold schemes, S^r can be decoded from any h shares, but no information of S^r leaks out from any k shares for $k = h - r$. More precisely, the ramp SSS satisfies that for any $\{i_1, \cdots, i_{k+m}\} \subseteq \{1, 2, \cdots, N\}$ and any $0 \le m \le r$,

$$H(S^r | V_{i_1}, V_{i_2}, \cdots, V_{i_{k+m}}) = \frac{r-m}{r} H(S^r), \tag{2}$$

where $H(\cdot)$ and $H(\cdot|\cdot)$ are the entropy and conditional entropy functions, respectively.

In [5], the ramp SSS satisfying Eq.(2) is called a weakly secure ramp SSS, and a strongly secure ramp SSS is defined as the ramp SSS that satisfies, instead of Eq.(2),

$$H(S_{j_1}, S_{j_2}, \cdots, S_{j_{r-m}} | V_{i_1}, V_{i_2}, \cdots, V_{i_{k+m}}) = H(S_{j_1}, \cdots, S_{j_{r-m}}), \tag{3}$$

for any $\{i_1, \cdots, i_{k+m}\} \subseteq \{1, \cdots, N\}$ and any $\{j_1, \cdots, j_{r-m}\} \subseteq \{1, \cdots, r\}, 0 \le m \le r$.

In the case of $0 < m < r$, some part of S^r might leak out from $V_{i_1}, V_{i_2}, \cdots, V_{i_{k+m}}$ in the weakly secure ramp SSS. But, no information leaks out for every $(S_{j_1}, S_{j_2}, \cdots, S_{j_{r-m}})$ from any $V_{i_1}, V_{i_2}, \cdots, V_{i_{k+m}}$ in the strongly secure ramp SSS. So, from the viewpoint of security, the strongly secure ramp SSS is better than the weakly secure ramp SSS.

In the case of linear ramp SSSs, shares are constructed from a secret $S^r = (S_1, S_2, \cdots, S_r)$ and random numbers $(R_1, R_2, \cdots, R_k)$ by

$$(V_1, V_2, \cdots, V_N) = (S_1, S_2, \cdots, S_r, R_1, R_2, \cdots, R_k)G,$$

where G is a matrix, and all S_i and R_j are independent and each of them is uniformly distributed over $\mathbb{F}_q$. The conditions of G to realize a weakly or strongly secure ramp SSS is given by the following theorem [5, Theorem 2].

Theorem 2: Let $b_{h,j}$, $j = 1, 2, \cdots, h$ be the j-th column vector of the identity matrix with rank h, and let g_j, $j = 1, 2, \cdots, N$, be the j-th column vector of G. Then, the following hold.

(a) The SSS generated by G is a weakly secure ramp SSS if and only if $b_{h,1}, b_{h,2}, \cdots, b_{h,r}$ and any $h - r$ column vectors in $\{g_1, g_2, \cdots, g_N\}$ are linearly independent.
(b) The SSS is a strongly secure ramp SSS if and only if any h column vectors obtained by picking up i vectors from $\{b_{h,1}, b_{h,2}, \cdots, b_{h,r}\}$ and $h - i$ vectors from $\{g_1, g_2, \cdots, g_N\}$ are linearly independent for any $0 \le i \le h$.

We note that the matrix G of a ramp SSS is closely related to the matrix Z of a linear network code. Based on this similarity, we define strongly secure network coding in the following section.

3. Strongly Secure Network Coding

In the following, we assume that there exist several adversaries in a network. They wiretap several edges cooperatively. The number of wiretapped edges may vary depending on the situation of adversaries. In order to transmit $S^r = (S_1, S_2, \cdots, S_r)$ securely against adversaries from a source node s to all sink nodes $t \in \mathcal{T}$ in the network, we use a linear network code Z for $X^h = (S_1, S_2, \cdots, S_r, R_1, R_2, \cdots, R_{h-r})$, where h is defined by $h = \min_{t \in \mathcal{T}} \mathrm{maxflow}(s, t)$, and R_i are random numbers. S_i and R_i take values in $\mathcal{X} = \mathbb{F}_q$, where q is assumed to be sufficiently large in this section. But, the size of q will be evaluated in Sect. 6.

We also assume that adversaries can know the linear network code Z, i.e. they can know what linear transform of X^h flows on each edge, and the source node and every sink node share no secret key in advance.

We now define two types of security conditions for network coding with adversaries in the same way as ramp SSSs.

Definition 3 (k-secure [3]): A linear network code Z is called k-secure if Z satisfies that for any $\mathcal{A} \subseteq \mathcal{E}$ with $\mathrm{rank}Z_{\mathcal{A}} \le k$

$$H(S^r | X^h Z_{\mathcal{A}}) = H(S^r). \tag{4}$$

In the case of linear network codes, the amount of leaked secrets increases linearly as $\mathrm{rank}Z_{\mathcal{A}}$ increases. Hence, Eq. (4) means that for $k \le \mathrm{rank}Z_{\mathcal{A}} \le h$,

$$H(S^r | X^h Z_{\mathcal{A}}) = \frac{[r + k - \mathrm{rank}Z_{\mathcal{A}}]_+}{r} H(S^r), \tag{5}$$

where $[a]_+ = \max\{a, 0\}$.

Definition 4 (strongly k-secure): A linear network code Z is called strongly k-secure if Z is k-secure and it satisfies that for any $\mathcal{A} \subseteq \mathcal{E}$ with $k \le \mathrm{rank}Z_{\mathcal{A}} \le h$,

$$H(S_{i_1}, S_{i_2}, \cdots, S_{i_{r+k-\mathrm{rank}Z_{\mathcal{A}}}} | X^h Z_{\mathcal{A}})$$
$$= \frac{[r+k-\mathrm{rank}Z_{\mathcal{A}}]_+}{r} H(S^r), \quad (6)$$
$$= H(S_{i_1}, S_{i_2}, \cdots, S_{i_{r+k-\mathrm{rank}Z_{\mathcal{A}}}}) \quad (7)$$
$$\text{for } \forall\{i_1, i_2, \cdots, i_{r+k-\mathrm{rank}Z_{\mathcal{A}}}\} \subseteq \{1, 2, \cdots r\}.$$

Note that k-secure and strongly k-secure network codes are closely related to weakly and strongly secure ramp SSSs, respectively. In the case of k-secure network codes, some S_i in S^r might leak out explicitly if adversaries wiretap more than k edges. But, in the case of strongly k secure network codes, no information leaks out for every $(S_{i_1}, S_{i_2}, \cdots, S_{i_{r+k-\mathrm{rank}Z_{\mathcal{A}}}})$ in S^r even if they wiretap more than k edges. Hence, the strongly k-secure network codes are more preferable than (weakly) k-secure network codes.

Remark 1: In the case of $r + k = h$, the operation $[\cdot]_+$ in Eqs. (5) and (6) is not necessary. But, a general case of $r + k \leq h$ is considered in this paper because, as shown in [4] and Sect. 6 of this paper, the size of q can be decreased as $h - (r + k)$ becomes larger.

Remark 2: Bhattad-Narayanan [13] defined weakly secure network coding in a different sense. They considered a network code Z satisfying that for any given edge set $\mathcal{A}$ with $\mathrm{rank}Z_{\mathcal{A}} \leq h - 1$,

$$H(S_i | X^h Z_{\mathcal{A}}) = H(S_i) \quad \text{for all } i. \quad (8)$$

They also extended the above case of a single S_i to the case of given subsets of secrets $S_1, S_2, \cdots, S_r$, and they called their schemes *weakly secure network coding*.

Compared with the Bhattad-Narayanan scheme, our scheme has ramp threshold security for S^r, and hence, we need not specify the edges that will be wiretapped by the adversaries in advance. Furthermore, our network code Z can easily be constructed as shown in this paper.

Cai-Yeung [3] showed that a k-secure network code can be realized if and only if $r \leq h - k$ is satisfied. The same also holds for the strongly k-secure codes as shown in the following theorem.

Theorem 3: For sufficiently large q, there exists a strongly k-secure code for any given network $\mathcal{G}$ if and only if $r \leq h - k$.

Proof. The part of "only if" holds obviously from Cai-Yeung's results because a strongly k-secure code is k-secure. The part of "if" holds because a strongly k-secure network code can be constructed by the following Algorithm 1. □

Algorithm 1 is based on the Jaggi-Sanders-Chau-Tolhuizen scheme [9] which can give coding vectors for the case with no adversaries. We first explain briefly their idea to construct a linear network code. For each sink node $t \in \mathcal{T}$ of a given digraph $\mathcal{G}$, there exist h edge-disjoint paths $p_{t,1}, p_{t,2}, \cdots, p_{t,h}$ from the source node s to each sink node t from Menger's theorem[†]. For each t, let $\mathcal{B}_t$ be the set of coding vectors, each of which is assigned to the latest considered edge in each path $p_{t,j}$. Then we select a node v in a topological order of $\mathcal{G}$ and we assign a coding vector to each outgoing edge of v such that all the coding vectors in $\mathcal{B}_t$ are linearly independent for every $t \in \mathcal{T}$.

In the case of a strongly k-secure network code, the coding vectors in $\mathcal{B}_t$ must satisfy more conditions. In order to describe the algorithm to construct a strongly k-secure network code, we use the following notations.

Let $b_{h,j}$, $j = 1, 2, \cdots, h$, represent the j-th column vector of the identity matrix with rank h. For an edge e on a path $p_{t,j}$, $f_t^{\leftarrow}(e)$ stands for the adjacent preceding edge of e on $p_{t,j}$, i.e., $\mathrm{tail}(e) = \mathrm{head}(f_t^{\leftarrow}(e))$, and $b(f_t^{\leftarrow}(e))$ represents the coding vector of edge $f_t^{\leftarrow}(e)$. In the case that e is an outgoing edge of the source s, $f_t^{\leftarrow}(e)$ does not exist. But, for the sake of convenience to simplify the description of algorithm, we define that $b(f_t^{\leftarrow}(e)) = b_{h,j}$ if e on path $p_{t,j}$ is an outgoing edge of s.

For a node e in $\mathcal{E}$, $\mathcal{T}(e)$ represents the set of sink nodes t such that the node e is on the path $p_{t,j}$ for some j. Let $\hat{\mathcal{V}} \equiv \{v \in \mathcal{V} \mid v \text{ is on a path } p_{t,j} \text{ for some } t \text{ and } j\}$ and $\hat{\mathcal{E}} \equiv \{e \in \mathcal{E} \mid e \text{ is on a path } p_{t,j} \text{ for some } t \text{ and } j\}$. For $\hat{\mathcal{V}}$, let $Q_{\hat{\mathcal{V}}}$ be the queue of nodes in $\hat{\mathcal{V}}$ such that the order is determined by a topological order of $\mathcal{G}$. Note that the first node in $Q_{\hat{\mathcal{V}}}$ is the source node s.

Let $\mathcal{I}_r$ be the set of the first r column vectors of the identity matrix with rank h, i.e. $\mathcal{I}_r \equiv \{b_{h,1}, b_{h,2}, \cdots, b_{h,r}\}$. Then, a subset of $\mathcal{I}_r$ is represented by $\mathcal{I}_r^{(\ell)}$ if the subset has ℓ column vectors. Similarly, for $\mathcal{Z}$ which is a set to store coding vectors assigned to edges in the following algorithm, a subset of $\mathcal{Z}$ is represented by $\mathcal{Z}^{(\ell)}$ if the subset has ℓ coding vectors. Furthermore, for a set of column vectors $\mathcal{A}$, $\dim\mathcal{A}$ stands for the dimension of the vector space spanned by $\mathcal{A}$.

Algorithm 1 (strongly k-secure network coding):

1. For a given graph $\mathcal{G}$, obtain h edge-disjoint paths $p_{t,1}, p_{t,2}, \cdots, p_{t,h}$ for each $t \in \mathcal{T}$ and $Q_{\hat{\mathcal{V}}}$.
2. Initialize $\mathcal{B}_t$ as $\mathcal{B}_t = \{b_{h,1}, b_{h,2}, \cdots, b_{h,h}\}$ for each $t \in \mathcal{T}$, and let $\mathcal{Z} = \emptyset$.
3. Let v be the first node of $Q_{\hat{\mathcal{V}}}$. Then, repeat the following for each outgoing edge $e \in \hat{\mathcal{E}}$ of the node v.
   a. To the edge e, assign a column vector $b(e)$ that satisfies the following three conditions.
      i. $b(e)$ can be generated by a linear combination of coding vectors $\{b(f_t^{\leftarrow}(e)), t \in \mathcal{T}(e)\}$.
      ii. For each $t \in \mathcal{T}(e)$, $b(e)$ must be linearly independent from the coding vectors included in $\mathcal{B}_t \setminus \{b(f_t^{\leftarrow}(e))\}$.
      iii. $b(e)$ must satisfy for any $\mathcal{I}_r^{(i)}$ and $\mathcal{Z}^{(h-i-1)}$, $0 \leq i \leq r$, that
      $$\dim(\mathcal{I}_r^{(i)} \cup \{b(e)\} \cup \mathcal{Z}^{(h-i-1)}) = i + \dim(\{b(e)\} \cup \mathcal{Z}^{(h-i-1)}).$$
   b. For each $t \in \mathcal{T}(e)$, update $\mathcal{B}_t$ as $\mathcal{B}_t = (\mathcal{B}_t \setminus$

[†] h edge-disjoint paths can be easily obtained by, for instance, Ford-Fulkerson's algorithm. [14, Sect. 9.2].

$\{b(f_t^{\leftarrow}(e))\} \cup \{b(e)\}$. Also update $\mathcal{Z}$ as $\mathcal{Z} = \mathcal{Z} \cup \{b(e)\}$.

4. Remove v from $\mathcal{Q}_{\hat{\mathcal{V}}}$. If $\mathcal{Q}_{\hat{\mathcal{V}}}$ is not empty, go to step 3. Otherwise, assign the zero column vector to all edges in $\mathcal{E} \setminus \hat{\mathcal{E}}$, and terminate this algorithm.

Note that the conditions i and ii in Step 3-a correspond to the conditions 2 and 3 in Definition 2, respectively. Furthermore, the condition iii in Step 3-a corresponds to the condition (b) in Theorem 2, which can guarantee the property of strongly k-secure network coding given in Definition 4. We also note that in step 3, $b(f_t^{\leftarrow}(e))$ is always included in $\mathcal{B}_t$ because a node v is treated in a topological order and $\mathcal{G}$ has no loop. We also note that step 3 works in any oder of the outgoing edges of v.

In the case that the cardinality of $\mathbb{F}_q$ is sufficiently large and $r \leq h - k$, we can always give a coding vector $b(e)$ that satisfies the three conditions shown in step 3. For instance, generate $b(e)$ as a random linear combination of $\{b(f_t^{\leftarrow}(e)), t \in \mathcal{T}(e)\}$, and check whether it satisfies the conditions ii and iii. If the conditions are not satisfied, regenerate another $b(e)$. If q is very large compared with the minimum required size of q, a desired $b(e)$ can be obtained by one or a few trials. On the other hand, in the case that q is not large, we can exhaustively check all vectors $b(e)$ that can be genrated from $\{b(f_t^{\leftarrow}(e)), t \in \mathcal{T}(e)\}$. In this case, the number of trials can be bounded by $q^{|\mathcal{T}(e)|}$.

If we want to get a weakly k-secure network code rather than a strongly k-secure network code, we can easily obtain it by modifying only the step 3-a-iii in Algorithm 1 as follows.

Algorithm 2 (weakly k-secure network coding):

3.-a.-iii. $b(e)$ must satisfy for any $\mathcal{Z}^{(h-r-1)}$ that

$$\dim(\mathcal{I}_r \cup \{b(e)\} \cup \mathcal{Z}^{(h-r-1)}) = r + \dim(\{b(e)\} \cup \mathcal{Z}^{(h-r-1)}).$$

This condition also corresponds to the condition (a) shown in Theorem 2 for the weakly secure ramp SSS. Clearly the condition in Algorithm 2 is weaker than the condition in Algorithm 1.

Remark 3: In this paper, we assume that adversaries can wiretap any edges in $\mathcal{E}$. But, in the case that adversaries can wiretap any edges only in a set $\mathcal{A} \subset \mathcal{E}$, it suffices to apply the step 3-a-iii and the update $\mathcal{Z} = \mathcal{Z} \cup \{b(e)\}$ in the step 3-b only for the edges included in $\mathcal{A} \cap \hat{\mathcal{E}}$ in Algorithms 1 and 2.

Some examples of strongly k-secure network coding will be shown in Sect. 5.

4. Transform from a Nonsecure Network Code to a Strongly Secure Network Code

In the previous section, we showed how to construct strongly (or weakly) k-secure network code. But, in this section, we consider how to realize strongly k-secure transmission by using a given nonsecure network code. Our scheme is based on the FMSS scheme [4], which can transform a nonsecure network code to a k-secure network code.

For a given network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, assume that a linear network code Z is already given, but this code may not be secure. The problem is to find a method to realize the strongly k-secure transmission of a secret $S^r = (S_1, S_2, \cdots, S_r)$ by using the network code Z.

Let $Y^h = (S_1, S_2, \cdots, S_r, R_1, R_2, \cdots, R_l)$, $h = r + l$, where $R^l = (R_1, R_2, \cdots, R_l)$ is a tuple of random numbers and all symbols of Y^h are independent and uniformly distributed over $\mathcal{X} = \mathbb{F}_q$. Instead of Y^h, we send a linearly transformed message $X^h = Y^h M^{-1}$ from a source node s to the sink nodes by using the network code Z. M is a nonsingular matrix with size $h \times h$ and we assume that the information of M is public. Since X^h can be decoded at each sink node for the network code Z, S^r can be recovered by calculating $Y^h = X^h M$ at the sink node.

The informations on all edges are given by $X^h Z = Y^h M^{-1} Z$. Hence, in this case, the matrix $M^{-1}Z$, instead of Z, randomizes the secret S^r to attain secure network coding. The following theorem gives the necessary and sufficient conditions such that M attains the strongly k-secure network coding.

Theorem 4: The matrix M can attain the strongly k-secure network coding if and only if any $k+r$ column vectors, which are obtained by picking up $k+j$ linearly independent column vectors of Z and $r - j$ column vectors of the first r column vectors of M, are linearly independent for any j, $0 \leq j \leq r - 1$.

Theorem 4 holds for any q. But, if q is not large enough, there might not exist a matrix M that satisfies the conditions shown in Theorem 4. Since the dimension of each column vector is h, we can conclude from Theorem 4 that a strongly k-secure M exists if and only if $k + r \leq h$, i.e. $k \leq l = h - r$, in the case that q is sufficiently large. We remark that the size of q becomes smaller as k becomes smaller as shown in Sect. 6. This property coincides with the result shown in [4] which treats k-secure network coding.

In order to prove Theorem 4, we first show the next Lemma 1.

Lemma 1: Assume that adversaries wiretap all edges in $\mathcal{A} \subset \mathcal{E}$ for $Y^h = X^h M$. Then, the following conditions are equivalent.

A : $H(S^r | Y^h M^{-1} Z_{\mathcal{A}}) = H(S^r)$.
B : $\mathrm{spn} Z_{\mathcal{A}} \cap \mathrm{spn}\{M^{(r)}\} = \{\mathbf{0}\}$.

In the above, $M^{(r)}$ represents the matrix that consists of the first r columns of M, "spn" stands for the space spanned by the columns of a matrix, and $\mathbf{0}$ is the zero column vector.

Proof. Instead of the equivalence between A and B, we prove the equivalence of the following conditions, which are the negatives of A and B, respectively.

$\bar{\mathrm{A}}$: $H(S^r | Y^h M^{-1} Z_{\mathcal{A}}) = H(S^r | X^h Z_{\mathcal{A}}) < H(S^r)$.
$\bar{\mathrm{B}}$: $\mathrm{spn} Z_{\mathcal{A}} \cap \mathrm{spn}\{M^{(r)}\} \neq \{\mathbf{0}\}$.

Since we consider only linear network coding, the condition $\bar{\mathrm{A}}$ occurs if and only if there exist nonzero row vectors $a \in \mathbb{F}_q^r$ and $b \in \mathbb{F}_q^{|A|}$ such that

$$S^r a^{\mathrm{T}} + X^h Z_{\mathcal{A}} b^{\mathrm{T}} = 0 \tag{9}$$

where a^{T} represents the transpose of vector a. Because of $S^r = X^h M^{(r)}$, Eq. (9) means that

$$X^h M^{(r)} a^{\mathrm{T}} + X^h Z_{\mathcal{A}} b^{\mathrm{T}} = 0 \tag{10}$$

Since the above relation must hold for any X^h, we obtain $M^{(r)} a^{\mathrm{T}} + Z_{\mathcal{A}} b^{\mathrm{T}} = \mathbf{0}$, which means condition $\bar{\mathrm{B}}$.

The next corollary can be derived easily from Lemma 1.

Corollary 1 ([4]): Matrix M can attain k-secure transmission if and only if any k linearly independent column vectors of Z and $M_1, \cdots, M_r$ are linearly independent.

Now we show the proof of Theorem 4.

The proof of Theorem 4. Assume that adversaries wiretap $\mathcal{A}$ with $\mathrm{rank} Z_{\mathcal{A}} = k + j$. Then, in order to realize the strongly k-secure network coding, Eq. (7) must be satisfied. Hence, from Lemma 1, we have that $\mathrm{spn} Z_{\mathcal{A}} \cap \mathrm{spn}\{M_{i_1}, \cdots, M_{i_{r-j}}\} = \{\mathbf{0}\}$, which coincides with the statement in the theorem. □

When the cardinality q of $\mathbb{F}_q$ is sufficiently large, a strongly k-secure M can easily be constructed. But, in the case that the cardinality q is not large, a strongly k-secure M may not exist. The necessary size of q will be discussed in Sect. 6 for the case of $k < l$.

5. Some Examples of Strongly Secure Network Codes

In this section, we show some examples of k-secure network codes. Since coding vector Z in Sect. 3 corresponds to $M^{-1}Z$ in Sect. 4, we show only the examples of $M^{-1}Z$.

Consider a nonsecure linear network code Z shown in Fig. 3, where the column vector attached at each edge is the coding vector of the edge. When we transmit $S^3 = (S_1, S_2, S_3)$ in this network, this code Z is not secure because, for instance, S_1 leaks out if an edge with coding vector $[1, 0, 0]^T$ is wiretapped.

Now we convert this nonsecure network code Z to a k-secure network code $M^{-1}Z$ with $k = l = 1$ and $r = 2$. Although Z has 15 columns for this network, we describe only the columns that have different values for simplicity in the following.

$$Z = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Furthermore, we assume that the cardinality q is given by $q = 5$, i.e. $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$

Let M_s and M_w be the matrices to realize strongly and weakly k-secure network coding, respectively. Then M_s and M_w can be obtained from Theorem 4 and Corollary 1, respectively, for instance, as follows.

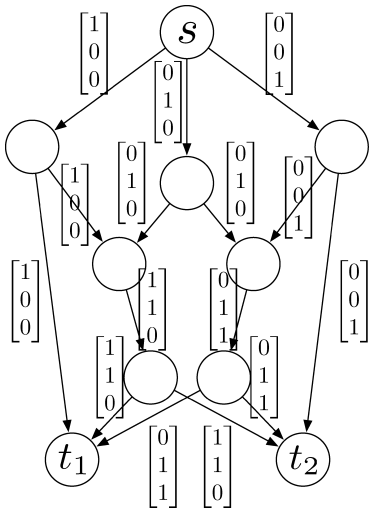

Fig. 3 An example: A non-secure network code.

$$M_s = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \quad M_s^{-1} = \begin{pmatrix} 3 & 0 & 4 \\ 4 & 0 & 3 \\ 3 & 1 & 3 \end{pmatrix} \tag{11}$$

$$M_w = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 0 \\ 1 & 3 & 0 \end{pmatrix}, \quad M_w^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 3 & 0 \\ 1 & 3 & 4 \end{pmatrix} \tag{12}$$

In the case of weakly 1-secure transmission, the symbol on each edge is given by

$$(S_1, S_2, R_1) M_w^{-1} Z = (S_1, S_2, R_1) \begin{pmatrix} 0 & 1 & 1 & 1 & 2 \\ 0 & 3 & 0 & 3 & 3 \\ 1 & 3 & 4 & 4 & 2 \end{pmatrix}. \tag{13}$$

We can easily check that adversaries cannot get any information about (S_1, S_2) even if they wiretap any one edge. But, if they wiretap two edges with coding vectors $[0, 0, 1]^{\mathrm{T}}$ and $[1, 0, 4]^{\mathrm{T}}$, S_1 leaks out perfectly.

In the case of strongly 1-secure transmission, the symbol on each edge is given by

$$(S_1, S_2, R_1) M_s^{-1} Z = (S_1, S_2, R_1) \begin{pmatrix} 3 & 0 & 4 & 3 & 4 \\ 4 & 0 & 3 & 4 & 3 \\ 3 & 1 & 3 & 4 & 4 \end{pmatrix} \tag{14}$$

In this case, adversaries cannot get any information about S_1 nor S_2 even if they wiretap any two edges.

Next we convert the nonsecure Z to a strongly 0-secure $M_s^{-1}Z$ with $r = 3$ and $\mathcal{X} = \mathbb{F}_7 = \{0, 1, 2, 3, 4, 5, 6\}$. Note that weakly 0-secure codes are nonsense because they are nonsecure, but strongly 0-secure codes are useful as shown below. A matrix M_s, which satisfies Theorem 4, is given by, for instance,

$$M_s = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & 5 & 1 \end{pmatrix}, \quad M_s^{-1} = \begin{pmatrix} 4 & 4 & 1 \\ 3 & 5 & 5 \\ 1 & 5 & 1 \end{pmatrix} \tag{15}$$

Then, the symbol on each edge is given by

$$(S_1, S_2, S_3)M_s^{-1}Z = (S_1, S_2, S_3)\begin{pmatrix} 4 & 4 & 1 & 1 & 5 \\ 3 & 5 & 5 & 1 & 3 \\ 1 & 5 & 1 & 6 & 6 \end{pmatrix} \tag{16}$$

In this case, when adversaries wiretap any one edge, no information leaks out for every pair of two S_i. Furthermore, even if adversaries wiretap any two edges, no information leaks out for every S_i. Note that the coding rate of strongly 0-secure $M_s^{-1}Z$ is the same as nonsecure Z because $r = 3$ is used. But, $M_s^{-1}Z$ is much more secure than Z. Therefore, strongly k-secure network coding is useful even in the case of $k = 0$.

6. Sufficient Alphabet Size for Realization

In Sects. 3 and 4, we showed that the strongly secure network coding can be realized if alphabet size q is sufficiently large. However, if q is too small, network coding cannot be realized as shown in the example of Fig. 2. In this section, based on the idea of FMSS [4], we clarify in the case of $k < l$ how large q is sufficient to construct strongly secure network codes proposed in Sects. 3 and 4.

First, we define some notations.

For a given linear network code Z, let $\mathcal{S}_Z$ be the kernel space of Z, which consists of all row vectors $x \in \mathbb{F}_q^{|\mathcal{E}|}$ satisfying $Zx^{\mathrm{T}} = \mathbf{0}$. Since Z is a matrix with size $h \times |\mathcal{E}|$, the dimension of $\mathcal{S}_Z$ is at most $|\mathcal{E}| - h$. hence, $\mathcal{S}_Z$ can be represented by $\mathcal{S}_Z = \{a\Lambda | a \in \mathbb{F}_q^{|\mathcal{E}|-h}\}$, where Λ is a generator matrix of the kernel space $\mathcal{S}_Z$, and the size of Λ is $(|\mathcal{E}| - h) \times |\mathcal{E}|$. We define a distance between two matrices with the same column size $|\mathcal{E}|$. Let P and Q are matrices with sizes $\alpha \times |\mathcal{E}|$ and $\beta \times |\mathcal{E}|$, respectively. Then the distance $\delta(P, Q)$ is defined by

$$\delta(P, Q) = \min_{\substack{a,b: \\ a \in \mathbb{F}_q^{\alpha}, a \neq \mathbf{0}, b \in \mathbb{F}_q^{\beta}}} d_H(aP, bQ)$$

where $d_H(\cdot, \cdot)$ is the Hamming distance between two vectors in $\mathbb{F}_q^{|\mathcal{E}|}$. In the space of $\mathbb{F}_q^{|\mathcal{E}|}$, let $\mathrm{Vol}_q(\ell, |\mathcal{E}|)$ be the volume of the sphere with radius ℓ which is measured by the Hamming distance. Note that $\mathrm{Vol}_q(\ell, |\mathcal{E}|)$ does not depend on a center vector.

The following theorems were proved by FMSS [4].

Theorem 5 ([4]): For a given Λ, there exists a matrix M^{-1} to realize the k-secure network coding for S^r if and only if there exists a matrix B with size $r \times |\mathcal{E}|$ satisfying $\delta(\Lambda, B) > k$.

This theorem means that the problem to obtain the matrix M^{-1} is equivalent to the problem to obtain the matrix B that satisfies $\delta(\Lambda, B) > k$ for a given Λ. Actually, M can be constructed as $M^{(r)} = ZB^{\mathrm{T}}$. For the latter problem, the next theorems hold.

Theorem 6 ([4]): Assume that a matrix B with size $r \times |\mathcal{E}|$ is selected randomly from all matrices in $\mathbb{F}_q^{r \times |\mathcal{E}|}$. Then, it holds for any $\epsilon > 0$ and $k = (h - r)/(1 + \epsilon)$ that

$$\Pr\{\delta(\Lambda, B) > k\} \geq 1 - P_{\mathrm{BAD}},$$

where P_{BAD} is defined by

$$P_{\mathrm{BAD}} = q^{-(1+\epsilon)k}\mathrm{Vol}_q(k, |\mathcal{E}|). \tag{17}$$

Theorem 7 ([4]): If $k + r < h$ and

$$h = \frac{\log |\mathcal{E}|}{\log q} + \frac{\log \mathrm{Vol}_q(k, |\mathcal{E}|)}{\log q} - 2\log |\mathcal{E}| - \log q - \log \ln q, \tag{18}$$

then, there exists a matrix Λ such that any matrix B with size $r \times |\mathcal{E}|$ does not satisfy $\delta(\Lambda, B) > k$.

A sufficient condition to realize k-secure network coding can be derived from Theorems 6 and 7. It is shown in [4] that in the case that $r = h - l$ for $k < l$, the sufficient condition of alphabet size q is given by $q > (k + 1)^{\frac{1}{l-k}} \cdot |\mathcal{E}|^{\frac{k}{l-k}}$. Hence, letting $l = 2k$, we can conclude that the sufficient size is given by $q \simeq |\mathcal{E}|$. Similarly, in the case of $k = \Theta(|\mathcal{E}|)$, it is given by $q = |\mathcal{E}|^{\Omega\left(\frac{k}{l-k}\right)}$. Furthermore, it can be obtained from Theorems 5 and 6 that for the case of $l = k$, the sufficient size is given by $q = |\mathcal{E}|^{\Omega\left(\sqrt{k/\log k}\right)}$.

Now we consider the case of the strongly k-secure network coding. Let $\mathcal{D}$ be a nonempty subset of $\{1, 2, \cdots, r\}$, i.e. $\mathcal{D} \subseteq \{1, 2, \cdots, r\}$ and $\mathcal{D} \neq \emptyset$, and let $B_{\mathcal{D}}$ be the matrix with size $|\mathcal{D}| \times |\mathcal{E}|$ which is constructed by concatenating every row of B corresponding $\mathcal{D}$. Then, the following theorem holds.

Theorem 8: For a given Λ, there exists a matrix M^{-1} to realize the strongly k-secure network coding for S^r if and only if there exists a matrix B with $r \times |\mathcal{E}|$ satisfying

$$\delta(\Lambda, B_{\mathcal{D}}) > k + r - |\mathcal{D}|, \quad \text{for any } \mathcal{D} \subseteq \{1, 2, \cdots, r\}, \mathcal{D} \neq \emptyset \tag{19}$$

Proof. This theorem holds obviously from Theorems 4 and 5. □

From Theorem 8, we can derive a sufficient condition of q to realize the strongly k-secure network coding by considering the condition given by Eq. (19).

Theorem 9: Assume that a matrix B with size $r \times |\mathcal{E}|$ is selected randomly from all matrices in $\mathbb{F}_q^{r \times |\mathcal{E}|}$. Then, it holds that

$$\Pr\left\{\delta(\Lambda, B_{\mathcal{D}}) > k + r - |\mathcal{D}|, \forall \mathcal{D} \in 2^{\{1,2,\cdots,r\}} \setminus \{\emptyset\}\right\} > 1 - \hat{P}_{\mathrm{BAD}}, \tag{20}$$

where $\hat{P}_{\mathrm{BAD}}$ is defined by

$$\hat{P}_{\mathrm{BAD}} = (k + r)2^r|\mathcal{E}|^{r+k-1}q^{r+k-h}. \tag{21}$$

Proof. For each $\mathcal{D}$, let $\mathrm{BAD}_{\mathcal{D}}$ be the set of vectors which are within $k + r - |\mathcal{D}|$ in the Hamming distance from the linear column space of Λ. Then, $\mathrm{BAD}_{\mathcal{D}}$ satisfies

$$|\mathrm{BAD}_{\mathcal{D}}| \leq q^{|\mathcal{E}|-h}\mathrm{Vol}_q(k + r - |\mathcal{D}|, |\mathcal{E}|) \tag{22}$$

Hence, the randomly selected B satisfies that

$$\begin{aligned}
&\Pr\{\exists x \in \mathbb{F}_q^{|\mathcal{D}|} \text{ s.t. } xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\} \\
&\quad \le q^{|\mathcal{D}|} \Pr\{xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\} \\
&\quad \le q^{|\mathcal{D}|} q^{|\mathcal{E}|-h} q^{-|\mathcal{E}|} \mathrm{Vol}_q(k+r-|\mathcal{D}|, |\mathcal{E}|) \\
&\quad = q^{|\mathcal{D}|-h} \mathrm{Vol}_q(k+r-|\mathcal{D}|, |\mathcal{E}|) \qquad (23) \\
&\quad =^1 q^{|\mathcal{D}|-h} \sum_{i=0}^{k+r-|\mathcal{D}|} (q-1)^i \binom{|\mathcal{E}|}{i} \\
&\quad <^2 q^{k+r-h}(k+r-|\mathcal{D}|+1) \binom{|\mathcal{E}|}{k+r-|\mathcal{D}|} \\
&\quad <^3 q^{k+r-h}(k+r-|\mathcal{D}|+1)|\mathcal{E}|^{k+r-|\mathcal{D}|}, \qquad (24)
\end{aligned}$$

where the numbered equality and inequalities hold from the following relations.

1. $\mathrm{Vol}_q(\ell, |\mathcal{E}|) = \sum_{i=0}^{\ell} (q-1)^i \binom{|\mathcal{E}|}{i}$
2. $\sum_{i=0}^{\ell} (q-1)^i \binom{|\mathcal{E}|}{i} \le (l+1)(q-1)^l \binom{|\mathcal{E}|}{l}$ for $q \ge 2$
3. $\binom{|\mathcal{E}|}{i} < |\mathcal{E}|^i$

If there exits a $\mathcal{D}$ such that $xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}$, then B does not satisfy the condition of $\Pr\{\cdots\}$ in Eq. (20). Hence, we have that

$$\begin{aligned}
&1 - \Pr\{\cdots\} \\
&\quad = \Pr\{\exists \mathcal{D}, \exists x \in \mathbb{F}_q^{|\mathcal{D}|} \text{s.t. } xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\} \\
&\quad = \Pr\left\{ \bigcup_{i=1}^{r} \bigcup_{\mathcal{D}:|\mathcal{D}|=i} \left[\exists x \in \mathbb{F}_q^{|\mathcal{D}|} \text{ s.t. } xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\right] \right\} \\
&\quad \le^4 \sum_{i=1}^{r} \sum_{\mathcal{D}:|\mathcal{D}|=i} \Pr\{\exists x \in \mathbb{F}_q^{|\mathcal{D}|} \text{ s.t. } xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\} \\
&\quad <^5 \sum_{i=1}^{r} \binom{r}{i} q^{k+r-h}(k+r-i+1)|\mathcal{E}|^{k+r-i} \\
&\quad =^6 \sum_{j=0}^{r-1} \binom{r}{j} q^{k+r-h}(k+j+1)|\mathcal{E}|^{k+j} \\
&\quad <^7 (k+r)2^r |\mathcal{E}|^{k+r-1} q^{k+r-h} \\
&\quad = \hat{P}_{\mathrm{BAD}} \qquad (25)
\end{aligned}$$

The numbered equality and inequalities 4–7 hold because

4. the union bound is applied,
5. the number of $\mathcal{D}$ with $|\mathcal{D}| = i$ is given by $\binom{r}{i}$, and Eq. (24) is substituted,
6. $j = r - i$ is substituted,
7. $\sum_{j=0}^{r} \binom{r}{j} = 2^r$.

□

We note from Eq. (21) that if

$$q > (k+r)^{\frac{1}{l-k}} 2^{\frac{r}{l-k}} |\mathcal{E}|^{\frac{k+r}{l-k}} \qquad (26)$$

holds for $l = h - r$ and $l > k$, we have that $\hat{P}_{\mathrm{BAD}} < 1$. This means from Theorem 9 that there exists at least one B that satisfies Eq. (19). Hence, we can conclude from Theorem 8 that if q satisfies Eq. (26) for $l > k$, then we can realize the strongly k-secure network coding.

In the case of $k = \Theta(|\mathcal{E}|)$, the next theorem also holds.

Theorem 10: Let $k = \gamma|\mathcal{E}|$ for $0 < \gamma < 1$ and let ϵ be the constant satisfying $k = \frac{h-r}{1+\epsilon}$. Then, if $q \ge 2^{\frac{\frac{r}{|\mathcal{E}|}+1}{\epsilon\gamma - \mathrm{o}(1)}}$ for $\mathrm{o}(1) = \frac{\log_q(1+\gamma|\mathcal{E}|)}{|\mathcal{E}|}$, there exists a matrix B that satisfies Eq. (19).

Proof. Let $H_q(w) = w\log_q(q-1) - w\log_q w - (1-w)\log_q(1-w)$. Then, it satisfies the following relations if $k = \gamma|\mathcal{E}|$ [15, Lemma 2.10.3].

$$H_q(w) \le w + \frac{1}{\log q} \qquad (27)$$

$$\frac{\log \mathrm{Vol}_q(k, |\mathcal{E}|)}{\log q} < \left(H_q\left(\frac{k}{|\mathcal{E}|}\right) + \mathrm{o}(1) \right) |\mathcal{E}| \qquad (28)$$

It holds from Eq. (23) that

$$\begin{aligned}
&\Pr\{\exists x \in \mathbb{F}_q^{|\mathcal{D}|} \text{ s.t. } xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\} \\
&\quad \le q^{|\mathcal{D}|-h} \mathrm{Vol}_q(k+r-|\mathcal{D}|, |\mathcal{E}|) \\
&\quad < q^{|\mathcal{D}|-h} q^{\left(H_q\left(\frac{k+r-|\mathcal{D}|}{|\mathcal{E}|}\right)+\mathrm{o}(1)\right)|\mathcal{E}|} \\
&\quad \le q^{|\mathcal{D}|-h} q^{\left(\frac{k+r-|\mathcal{D}|}{|\mathcal{E}|}+\frac{1}{\log q}+\mathrm{o}(1)\right)|\mathcal{E}|} \\
&\quad = q^{k+r-h} q^{\left(\frac{1}{\log q}+\mathrm{o}(1)\right)|\mathcal{E}|}, \qquad (29)
\end{aligned}$$

where the second and third inequalities follow from Eqs. (28) and (27), respectively.

Furthermore, by taking the union bound for all $\mathcal{D}$, the following bound is obtained for the probability $\Pr\{\cdots\}$ in Eq. (20).

$$\begin{aligned}
&1 - \Pr\{\cdots\} \\
&\quad = \Pr\{\exists \mathcal{D}, \exists x \in \mathbb{F}_q^{|\mathcal{D}|} \text{ s.t. } xB_{\mathcal{D}} \in \mathrm{BAD}_{\mathcal{D}}\} \\
&\quad < \sum_{i=0}^{r-1} \binom{r}{i} q^{k+r-h} q^{\left(\frac{1}{\log q}+\mathrm{o}(1)\right)|\mathcal{E}|} \\
&\quad < 2^r q^{k+r-h} q^{\left(\frac{1}{\log q}+\mathrm{o}(1)\right)|\mathcal{E}|} \\
&\quad = q^{\frac{r}{\log q}} q^{-\epsilon\gamma|\mathcal{E}|} q^{\left(\frac{1}{\log q}+\mathrm{o}(1)\right)|\mathcal{E}|} \\
&\quad = q^{\left(\frac{\frac{r}{|\mathcal{E}|}+1}{\log q} - \epsilon\gamma + \mathrm{o}(1)\right)|\mathcal{E}|} \qquad (30)
\end{aligned}$$

We note that if $\frac{\frac{r}{|\mathcal{E}|}+1}{\log q} < \epsilon\gamma - \mathrm{o}(1)$, i.e. $q \ge 2^{\frac{\frac{r}{|\mathcal{E}|}+1}{\epsilon\gamma-\mathrm{o}(1)}}$, then Eq. (30) becomes less than 1. This means that this theorem holds from Theorems 8 and 9. □

From the relation $k = (h-r)/(1+\epsilon)$, ϵ is given by $\epsilon = (l-k)/k$ for $l = h - r$. Hence, from Theorem 10, there exists a strongly k-secure network code if $q = 2^{\Omega\left(\frac{k}{l-k}\left(1+\frac{r}{|\mathcal{E}|}\right)\right)}$.

Table 1 Sufficient alphabet size q.

	$k < l$		$k = l$
	$k = o(\|\mathcal{E}\|)$	$k = \Theta(\|\mathcal{E}\|)$	
k-secure	$\Theta\left(\|\mathcal{E}\|^{\frac{k}{l-k}}\right)$	$2^{\Omega\left(\frac{k}{l-k}\right)}$	$\|\mathcal{E}\|^{\Omega\left(\sqrt{\frac{k}{\log k}}\right)}$
strongly k-seruce	$\Theta\left(\|\mathcal{E}\|^{\frac{k+r}{l-k}}\right)$	$2^{\Omega\left(\frac{k}{l-k}\left(1+\frac{r}{\|\mathcal{E}\|}\right)\right)}$	open problem

7. Concluding Remarks

In this paper, based on the strongly secure ramp secret sharing schemes, we proposed the strongly k-secure linear network coding schemes, which are much more secure than previous known secure network schemes [3], [4]. We gave the direct construction method of the strongly k-secure linear network coding in Sect. 3. We also showed in Sect. 4 that a nonsecure linear network code can easily be transformed to a strongly k-secure network code if the alphabet size of information is sufficiently large. Furthermore, we derived some sufficient conditions of alphabet size to realize such transform for the case of $k < l$ in Sect. 6.

The sufficient conditions of alphabet size q can be summarized as Table 1, where the sufficient conditions of k-secure network coding are cited from [4]. We note from the table that in the case of $k = l$, a nontrivial sufficient condition has not been obtained yet for the strongly k-secure network coding. But, q must be larger than the case of k-secure network coding, i.e. $q = |\mathcal{E}|^{\Omega\left(\sqrt{\frac{k}{\log k}}\right)}$.

As shown in Sect. 3, we can assign the zero vector to all edges that are not included in the edge-disjoint paths of all sink nodes. Hence, in the evaluation of Sect. 6, the size $|\mathcal{E}|$ can be decreased to the number of edges included in all the edge-disjoint paths of all sink nodes.

References

[1] R. Ahlswede, N. Cai, S.Y.R. Li, and R.W. Yeung, "Network information flow," IEEE Trans. Inf. Theory, vol.46, no.4, pp.1204–1216, July 2000.
[2] S.Y.R. Li, R.W. Yeung, and N. Cai, "Linear network coding," IEEE Trans. Inf. Theory, vol.49, no.2, pp.371–381, Feb. 2003.
[3] N. Cai and R.W. Yeung, "Secure network coding," IEEE ISIT'02, p.323, June 2002.
[4] J. Feldman, T. Malkin, C. Stein, and R.A. Servedio, "On the capacity of secure network coding," 42nd Ann. Allerton Conf. on Comm., Control, and Comp., 2004.
[5] H. Yamamoto, "Secret sharing system using (k, L, n) threshold scheme," IECE Trans. Fundamentals (Japanese Edition), vol.J68-A, no.9, pp.945–952, Sept. 1985. (English Translation: Scripta Technica, Inc., Electronics and Comm. in Japan, Part 1, vol.69, no.9, pp.46–54, 1986.)
[6] G.R. Blakley and C. Meadows, "Security of ramp schemes," Advances in Cryptology — CRYPTO'84, LNCS 196, pp.242–269, Springer-Verlag, 1985.
[7] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network Flows: Theory, Algorithms, and Applications, Englewood Cliifs, NJ, Prentice-Hall, 1993.
[8] T. Ho, R. Koetter, and M. M'edard, "Network coding from a network flow perspective," IEEE ISIT'03, Yokohama, Japan, p.442, 2003.
[9] S. Jaggi, P. Sanders, C.P.A., and L. Tolhuizen, "Polynomial time algorithms for multicast network code construction," IEEE Trans. Inf. Theory, vol.51, no.6, pp.1973–1982, June 2005.
[10] T. Ho, R. Koetter, M. Médard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," IEEE ISIT'03, p.441, 2003.
[11] S. Jaggi, P.A. Chou, and J.K., "Low complexity algebraic multicast network codes," IEEE ISIT'03, p.441, 2003.
[12] P. Sanders, S. Egner, and L. Tolhuizen, "Polynomial time algorithms for network information flow," 15th ACM Symp. Parallel Alg. and Arch., pp.286–294, June 2003.
[13] K. Bhattad and K.R. Narayanan, "Weakly secure network coding," NetCod 2005, 2005.
[14] A. Schrijver, Combinatorial Optimization: Algorithms and Combinatorics, Springer Velag, 2003.
[15] W.C. Huffman and V. Pless, Fundamentals of Error Correcting Codes, Cambridge University Press, 2003.

Kunihiko Harada was born in Nagasaki, Japan, on December 24, 1981. He received the Bachelor's and Master's degrees from the University of Tokyo, in 2004 and 2006, respectively. He is now working toward the Ph.D. degree in information science and technology at the University of Tokyo. He won a prize of encouragement by the presentation in SITA2005 (in Japanese).

Hirosuke Yamamoto was born in Wakayama, Japan, on November 15, 1952, He received the B.E. degree from Shizuoka University, in 1975 and the M.E. and Ph.D. degrees from the University of Tokyo, in 1977 and 1980, respectively, all in electrical engineering. In 1980, he joined Tokushima University. He was an Associate Professor at Tokushima University, the University of Electro-Communications, and the University of Tokyo, from 1983 to 1987, from 1987 to 1993, and from 1993 to 1999, respectively. Since 1999, he has been a Professor at the University of Tokyo. He was with the School of Engineering and the School of Information Science and Technology from 1993 to 1999 and from 1999 to 2004, respectively, and is currently with the School of Frontier Sciences in the University of Tokyo. In 1989 and 1990, he was a Visiting Scholar at the Information Systems Laboratory, Stanford University. His research interests are in Shannon theory, data compression algorithms, and cryptology. Dr. Yamamoto is a member of the IEEE and the SITA (Society of Information Theory and its Applications). He served as the Chair of the IEEE Information Theory Society Japan Chapter in 2002 and 2003, and the TPC (Technical Program Committee) Co-Chair of ISITA2004 (the 2004 International Symposium on Information Theory and its Applications). He is currently the President of the SITA, an Associate Editor for Shannon Theory, IEEE Transactions on Information Theory, and the TPC Chair of ISITA2008.