Network Function Computation With Different Secure Conditions

Min Xu, Gennian Ge and Minqian Liu

Abstract

In this paper, we investigate function computation problems under different secure conditions over a network with multiple source nodes and a single sink node which desires a function of all source messages without error. A wiretapper has access to some edges of the network. Based on different practical requirements, we consider two secure conditions named as secure and user secure respectively. The main parameter concerned here is the computing rate, which is the average times of the target function that can be computed securely or user securely without error for one use of the network. In the secure case, a new upper bound which is tighter than the previous one is provided for arithmetic sum functions and arbitrary networks. Moreover, we show that the improved upper bound is strictly tight for tree-like networks. In the user secure case, we give a sufficient and necessary condition for the existence of user secure network codes and obtain an upper bound for the computation capacity.

Keywords: network function computation, secure, user secure, capacity

I. INTRODUCTION

In this paper, we study the problem of function computation over a communication network with two different secure conditions. The general setup of secure network function computation can be modeled as follows. In a communication network characterized by a directed acyclic graph, there is a sink node which wants to compute a function \( f \) about the messages generated at source nodes independently and uniformly. Meanwhile, there is a wiretapper who has access to some edge subsets \( W \) of the graph, we use \( \mathcal{W} \) to denote all possible wiretap sets. Each link of the network has unit transmission capacity and each node has unlimited computation capacity. We hope the sink node can compute the function without error and the wiretapper cannot get any information about the source message or the result of the function.

In last few years, network function computation has attracted considerable attention due to its application in sensor networks [1] and data processing [2], etc. The security problem is necessary to be studied when the network function computation is applied to practice.

A. Related Works

The main parameter concerned in this paper is the computation capacity for the sink node to reliably compute the target function over the network under the secure conditions. It’s obvious that when \( \mathcal{W} = \emptyset \), the secure network function computation problem reduces to non-secure case, which has been studied in recent years [3]–[7]. When the target function is linear, given the network, the capacity of the network is completely determined and the construction of achievable network codes is given [4]. However, for general target functions, there are some cut-set type upper bounds of the computation capacity [5]–[7], but these upper bounds are not always tight. Thus, we will focus on linear target function over finite field in this paper.

There have been extensive studies and applications of the security problem in network coding theory. In 2002, Cai and Yeung [12] first investigated the existence of secure network codes. In network coding problem, there is a single source node and multiple sink nodes, each wants a copy of the source message. The secure condition means the wiretapper cannot get any information about the source message. And some other secure conditions have also been studied [9], [10]. For secure network function computing, limited work has been done. Guang et al. [13] developed an upper bound of the capacity of secure function computing for linear target functions in 2021. And Li et al. [11] considered another secure condition which they called secret sum.

B. Contributions and Organization

Our main results are as follows. On one hand, we show that Guang’s upper bound in [13] is not tight by an example. From this example, we establish a new upper bound of the capacity under the secure condition. Moreover, we show that for a special class of networks, our new upper bound is tight. On the other hand, we define a new secure condition based on practical
requirements, and investigate the same problem under the new secure condition. A sufficient and necessary condition for the existence of secure network codes is addressed and an upper bound of the computing capacity is obtained.

This paper is organized as follows. Section III introduces network function computation problems under two different secure conditions and reviews some existing bounds in related research works. In Sections III and IV, we focus on the capacity under the secure condition in [13]. We establish our new upper bound from an example and prove it in Section III. In Section IV, we derive the condition of the linear transformation which turns a non-secure network code into a secure one, and show that for a special class of networks, our new upper bound is tight. For the new security, we address a sufficient and necessary condition for the existence of secure network codes and establish an upper bound by information-theoretical approach in Section V. Finally, we give a conclusion of all results in Section VI.

II. MODEL AND PRELIMINARIES

A. Model and Secure Conditions

Let \( \mathcal{N} = (V, E) \) be a directed acyclic graph where \( V \) is the node set and \( E \) is the edge set. For any edge \( e \in E \), we use tail\( (e) \) and head\( (e) \) to denote the tail and head of \( e \). For any vertex \( v \in V \), let In\( (v) = \{ e \in E : \text{head}(e) = v \} \) and Out\( (v) = \{ e \in E : \text{tail}(e) = v \} \), respectively. An edge sequence \( \{e_1, e_2, \cdots, e_n\} \) is a path from node \( u \) to node \( v \) if tail\( (e_1) = u \), head\( (e_n) = v \) and tail\( (e_{i+1}) = \text{head}(e_i) \) for \( i = 1, 2, \cdots, n - 1 \). For two nodes \( u, v \in V \), a cut of them is an edge set \( C \) such that every path from \( u \) to \( v \) contains some edges in \( C \). If \( U \subset V \) is a set of nodes, the cut of \( U \) and \( V \) is the edge set separating any vertex \( u \in U \) from \( v \). The minimum size of a cut of \( u \) and \( v \) is denoted as mincut\( (u, v) \). There are two types of special nodes in the graph. Source nodes are the nodes which have no incoming edges and sink nodes are nodes which have no outgoing edges. In this paper, we focus on multi-source single terminal network, denote the source nodes as \( S = \{ \sigma_1, \sigma_2, \cdots, \sigma_l \} \) and the sink node as \( \rho \).

Generally, a target function \( f : \mathcal{A}^* \rightarrow \mathcal{O} \) is needed to be computed at sink node \( \rho \), where \( \mathcal{A}, \mathcal{O} \) are finite fields. We assume that each source node in the network \( \mathcal{N} \) generates a sequence of random symbols over \( \mathcal{A} \) independently and uniformly, and each edge of the network transmits a symbol of \( \mathcal{B} \) once a time, where \( \mathcal{B} \) is also a finite field. Let the message generated at source node \( \sigma_i \) be a random variable \( M_i \), where \( M_1, M_2, \cdots, M_s \) are mutually independent. Computing the function \( f \) times implies source \( \sigma_i \) generates a random vector \( M_i = (M_{i1}, \cdots, M_{il}) \), and denote all the source messages by \( M_S = (M_1, \cdots, M_s) \).

In secure network function computation, there is a wiretapper who wants information about messages in the network. Let \( \mathcal{W} \) be all possible wiretap sets, and the wiretapper can access exactly one edge subset \( W \in \mathcal{W} \). \( \mathcal{W} \) is known by the source nodes. The sink node need to compute target function

\[
f(M_S) = (f(M_{j1}, M_{j2}, \cdots, M_{js}) : j = 1, 2, \cdots, l)
\]

without error while the wiretapper cannot get any useful information. In network coding case, different secure conditions such as secure [9], weakly secure [9], strongly secure [10] are considered. In network function computation case, Guang et al. [13] focused on protecting all source messages from the wiretapper. Sometimes, in sensor network, the original messages generated at sources are not important, while the computation result is. Li et al. considered protecting the arithmetic sum of all source messages from the wiretapper [11]. They assumed each edge has unlimited transmission capacity, i.e. each edge can send unlimited symbols of \( \mathcal{B} \) once a time. We are concerned in this paper with the following two secure conditions:

\[
(\text{secure}) \quad I(M_S; Y_W) = 0,
\]

\[
(\text{user secure}) \quad I(f(M_S); Y_W) = 0,
\]

where \( Y_W \) is the message transmitted by the edges in \( W \) and observed by the wiretapper. In network coding, the message sent by each node is a function of all the messages the node has. Therefore, random keys are needed under the secure condition. Otherwise, any edge could leakage the information about \( M_S \). But for the user secure case, random keys are not necessary.

An \((l, n)\) network code for the secure (or user secure) problem \((\mathcal{N}, f, \mathcal{W})\) is defined as follows. First, we let \( m_i \in \mathcal{A}^l \) and \( k_i \in \mathcal{K}_i \) be the outputs of the source message \( M_i \) and the key \( K_i \) respectively, \( \mathcal{K}_i = \emptyset \) if random keys are not necessary. An \((l, n)\) network code consists of a local encoding function for each edge \( e \) which is defined as follows,

\[
\hat{\theta}_e : \mathcal{A}^l \times \mathcal{K}_i \mapsto \mathcal{B}^n, \quad \text{if} \ \text{tail}(e) = \sigma_i \ \text{for some} \ i;
\]

\[
\prod_{e \in \text{In}(\text{tail}(e))} \mathcal{B}^n \mapsto \mathcal{B}^n, \quad \text{otherwise},
\]

and a decoding function \( \hat{\varphi} : \prod_{\text{In}(\rho)} \mathcal{B}^n \mapsto \mathcal{O}^l \). By the definition of local encoding function, the message transmitted by edge \( e \) denoted as \( U_e \) is a function of all source messages and random keys, i.e. for each \( e \in E \),

\[
\hat{g}_e(m_S, k_S) = \begin{cases} 
\hat{\theta}_e(m_i, k_i), & \text{if} \ \text{tail}(e) = \sigma_i \ \text{for some} \ i; \\
\hat{\theta}_e(\hat{\varphi}(\text{tail}(e))(m_S, k_S)), & \text{otherwise}.
\end{cases}
\]
We say a rate $R$ is achievable if there exists a valid network code with rate $R$ with respect to $(N, f, W)$, which implies that by using the network $n$ times, the sink node can compute target function securely (or user securely) $l$ times. The capacity of the network $N$ to compute target function $f$ with possible wiretap set $W$ is the maximum achievable rate, i.e.

$$\hat{C}(N, f, W) = \max \left\{ \frac{l}{n} \mid \text{there exists a valid } (l, n) \text{ network code} \right\}.$$ 

In this paper, suppose $A = B = O = F_q$, and $f$ is a linear function over $F_q$, i.e.

$$f(x_1, x_2, \ldots, x_s) = \sum_{i=1}^s a_i \cdot x_i,$$

where $a_i \in F_q$ for all $i = 1, 2, \ldots, s$. Without loss of generality, let $a_i \neq 0$, $\forall i = 1, 2, \ldots, s$. Otherwise, we can remove the source nodes $\sigma_i$ for those $i$ such that $a_i = 0$ and consider computing the same target function over this new network. Moreover, let $y_i = a_i \cdot x_i$, and consider the sum function $f'(y_1, y_2, \ldots, y_s) = \sum_{i=1}^s y_i$ over $F_q$. A secure (or user secure) network code for $(f, N, W)$ is equivalent to a secure (or user secure) network code for $(f', N, W)$. From this observation, we only need to consider the computing capacity for the arithmetic sum function over $F_q$.

Particularly, we consider the wiretap set with the form

$$W_r = \{W \subset E : |W| \leq r \},$$

where $r$ is defined as the security level. In the rest of this paper, we use $\hat{C}(N, f, r)$ to denote $\hat{C}(N, f, W_r)$.

\section{B. Existing Bounds}

First, let’s review some known results about secure network function computation. For a network $N$, let $C_{min}$ be the size of minimum cut of $N$. We call a cut as $\sigma_i$-cut if it is an edge subset separating $\sigma_i$ from $\rho$ and there doesn’t exist any path from other sources to this cut. For each source node $\sigma_i$, the minimum size of a $\sigma_i$-cut is denoted as $D_{min}^i$.

An important result for linear network computation over finite field is that the capacity of network $N$ is $C_{min}$ [4]. For secure case, Guang et al. established an upper bound [13]

$$\hat{C}(N, f, r) \leq \min \{C_{min}, D_{min} - r\},$$

where $D_{min} = \min_{\sigma_i \in S} \{D_{min}^i\}$. On the other hand, similar to the construction of secure linear network codes, Guang et al. gave a lower bound [13]

$$\hat{C}(N, f, r) \geq C_{min} - r.$$ 

Whether the upper bound is tight is not known yet. In the scenario of user secure, Li et al. investigated the problem under the assumption of unlimited edge transmission capacity [11]. A sufficient and necessary condition for the existence of user secure network codes was provided based on the network topology.

\section{III. An Improved Upper Bound}

We begin this section with an example which shows Guang’s upper bound is not always tight.

\textbf{Example 3.1:} For the secure network function computation problem shown in Fig. [11] the minimum cut and minimum $\sigma_i$-cut in this network have size 1 and 3 respectively, i.e., $C_{min} = 1, D_{min} = 3$. When the security level $r = 2$, which means the wiretapper has access to any two edges, we have $\hat{C}(N, f, r) \leq 1$ from (6). In fact, there doesn’t exist a secure code in this case. Suppose the wiretapper gets the messages transmitted by $e_4$ and $e_7$ which are denoted as $U_{e_4}$ and $U_{e_7}$. If there exists a secure network code, then $U_{e_4}$ is a function of the source message and the random key generated at $\sigma_2$, i.e. $U_{e_4} = g_1(y, k')$. Similarly, $U_{e_7} = g_2(x, k, U_{e_4})$. Since the sink node can complete the computation and $U_{e_5}, U_{e_6}$ contain no information about $k$, $U_{e_7}$ cannot contain any information about $k$ either, i.e. $U_{e_7} = g_2(x, U_{e_4})$. Now, after getting $U_{e_4}$ and $U_{e_7}$ the wiretapper can get a function of $x$, which contradicts the secure condition [11].

From the example, we know that Guang’s upper bound is not tight although it is tight over some special $(N, f, r)$, see examples in [13]. For any cut $C$ in a network, let $I_C$ be the source nodes separated from $\rho$ by $C$, i.e.

$$I_C = \{\sigma_i \in S : \text{all paths from } \sigma_i \text{ to } \rho \text{ pass through } C\}.$$
From a series of papers in network function computation [5]-[7], we know that each of those sources which can attach to C, i.e. there exists a path from the source to some edges in C, also matters to the capacity. Define

\[ J_C = \{ \sigma_i \in S \setminus I_C : \text{there exists a path from } \sigma_i \text{ to } C \} . \]

Now we define a new edge set for each cut \( C \) as follows. Define

\[ \hat{B} = \min \{ B \subset E : \text{any path from } \sigma_i \text{ to } e \text{ pass through, } \forall \sigma_i \in J_C, \forall e \in C \}, \]

where the minimum is taken with respect to the size of \( B \). For any cut \( C \), let \( A = C \cup \hat{B} \), and define \( A_{\min} = \min_{C \subseteq \Lambda} \{|A| : A = C \cup \hat{B}\} \), where \( \Lambda \) is the collection of all cut sets of the network.

We present our upper bound in the following theorem.

**Theorem 3.2:** Let \( \mathcal{N} \) be a network and \( f \) be an algebraic sum over a finite field \( \mathbb{F}_q \). If the security level \( r \leq A_{\min} \), then

\[ \hat{C}(\mathcal{N}, f, r) \leq \min\{C_{\min}, A_{\min} - r\}, \tag{8} \]

and \( \hat{C}(\mathcal{N}, f, r) = 0 \) otherwise.

**Proof.** For the secure network function computation problem \( (\mathcal{N}, f, r) \), suppose there exists an \((l, n)\) secure code. To prove the theorem, we only need to show \( l/n \leq A_{\min} - r \).

Let the message generated at source node \( \sigma_i \) be a random variable \( X_i = (M_i, K_i) = (M_{i1}, \ldots, M_{il}, K_{i1}, \ldots, K_{ir}) \), where \( M_{ij}, K_{ij}, i \in [s], j \in [l], j' \in [r] \) are generated independently and uniformly at random over \( \mathbb{F}_q \). Then by the definition of entropy function, we have

\[ H(M_i) = l \cdot H(M_{ij}) = l \cdot \log q, \]

where the logarithm base is two. The message transmitted by any edge \( e \) is denoted as \( U_e \in \mathbb{F}_q^n \).

Let \( A = \mathcal{C} \cup B \) be the edge set with size \( A_{\min} \), where \( C \) is a cut of the network and \( B \) is the minimum cut separating \( J_C \) and \( C \). Without loss of generality, we assume \( J_C = \{\sigma_1\} \), and consider a wiretap set \( W \subset A \) with size \( r \). Let \( Y_W = \{U_e : e \in W\} \), and \( Y_A = \{U_e : e \in A\} \). Since each source generates message independently, we have

\[ l \cdot \log q = H(M_1) = H(M_1 | M_{S \setminus D}), \tag{9} \]

where \( D = J_C \cup \{\sigma_1\} \). The secure condition \([1]\) implies

\[ H(M_1, M_{S \setminus D} | Y_W) = H(M_1, M_{S \setminus D}), \]

\[ H(M_{S \setminus D} | Y_W) = H(M_{S \setminus D}). \]

Expanding both sides of the first equation, we get

\[ H(M_1 | M_{S \setminus D}, Y_W) + H(M_{S \setminus D} | Y_W) = H(M_1, M_{S \setminus D} | Y_W) = H(M_1, M_{S \setminus D}) = H(M_1 | M_{S \setminus D}) + H(M_{S \setminus D}), \]

which shows

\[ H(M_1 | M_{S \setminus D}, Y_W) = H(M_1 | M_{S \setminus D}). \tag{10} \]
By the definition of edge set $A$, we know that every edge in $A$ transmits a function of messages generated at sources in $D = J_C \cup \{\sigma_1\}$, which means $Y_W$ is a function of $M_D$ and $K_D$. Then,

$$H(M_1|\{MS\}_{D}, Y_W) = H(M_1|\{MS\}_{D}, Y_W, K_{S\setminus D}).$$  \hspace{1cm} (11)

We claim that

$$H(M_1|\{MS\}_{D}, Y_A, K_{S\setminus D}) = 0.$$  \hspace{1cm} (12)

This claim can be shown as follows

$$H(M_1|\{MS\}_{D}, Y_A, K_{S\setminus D}) = H(M_1|\{MS\}_{D}, Y_A, K_{S\setminus A})$$

$$= H(M_1|\{MS\}_{D}, Y_A, K_{S\setminus A}, Y_{C'})$$

$$= H(M_1|\{MS\}_{D}, K_{S\setminus A}, f(M_1, \cdots, M_s))$$

$$\leq H(M_1|\{MS\}_{D}, f(M_1, \cdots, M_s)) = 0,$$

where $C' = \bigcup_{i \in S \setminus \{A\}} \text{Out}(\sigma_i)$. The first equality holds for the independence of source messages and keys. By the definition of $C'$, $Y_{C'}$ is a function of $\{MS\}_{S \setminus \{A\}}$ and $K_{S \setminus A}$, then the second equality follows. The third equation holds since that $A \cup C'$ is a global cut of the network, the target function can be computed for any global cut messages. And the last equation is from that $f(M_1, \cdots, M_s) = M_1 + \cdots + M_s$ and $\{MS\}_{S \setminus \{A\}}$ is known.

Combining (9), (10), (11), (12), we get

$$l \cdot \log q = H(M_1|\{MS\}_{D}, Y_W, K_{S\setminus D})$$

$$= H(M_1|\{MS\}_{D}, Y_A, K_{S\setminus D}) - H(M_1|\{MS\}_{D}, Y_A, K_{S\setminus D})$$

$$= I(M_1; Y_{A\setminus W}|MS_{D}, K_{S\setminus D})$$

$$\leq H(Y_{A\setminus W}) \leq (|A| - r) \cdot n \cdot \log q.$$

Then we can conclude that

$$\frac{l}{n} \leq |A| - r = A_{min} - r.$$

**Remark.** Comparing with Guang’s upper bound [5], we only change $D_{min}$ to $A_{min}$. By the definition of $A_{min}$, if $C$ is a $\sigma_{i\ast}$-cut of the network, then $B = \emptyset$ since $J_C = \emptyset$, then $A = C \cup B = C$. Thus $D_{min}$ is the minimum size of $A$, where the minimum is taken over all $\sigma_{i\ast}$-cut, and $A_{min}$ is the minimum taken over all cut. Therefore, $A_{min} \leq D_{min}$ in general, which implies our new upper bound is tighter than the previous one.

**Example 3.3:** Consider the network function computation problem shown in Fig. [3] For $r = 1$, our new upper bound shows $\tilde{C}(N, f, 1) \leq 1$, which is achievable and the secure code is the same as the example in [13]. For $r = 2$, applying Theorem 3.2, we have $\tilde{C}(N, f, 2) \leq 0$, since $A_{min} = 2$ for the network. This also matches our analysis in Example 3.1.

**IV. LOWER BOUND**

For the lower bound of the secure computing capacity over a network, we need to construct an algorithm to generate a secure $(l, n)$-code for any given network, target function and security level, which is difficult in general. For non-secure case, most known results about lower bound are obtained by connecting the problem with equivalent network coding problem [3], [14].

Similar to the construction of secure network coding problem, Guang et al. [13] provided a lower bound with rate $C_{min} - r$ by a linear transformation at the source nodes. In the following, we address a sufficient and necessary condition of the linear transformation which turns a non-secure network code into a secure one, and prove that for multi-edge tree graph, the upper bound in Theorem 3.2 is tight.

Recall that an $(l, n)$ network code can be represented by global encoding functions for all edges in the network, i.e., \{\hat{g}_e : e \in E\}.

We denote the message generated at each source node $\sigma_i$ as $M_i \in \mathbb{F}_q^{n\times r}$, and denote the random key as $K_i \in \mathbb{F}_q^{r}$ in this section. Each edge $e \in E$ transmits a message of length 1 over $\mathbb{F}_q$. Such a linear secure code has rate $n - r$, and each global encoding function can be represented by a column vector $g_e \in \mathbb{F}_q^{rn}$, i.e.

$$\hat{g}_e(M_1, K_1, \cdots, M_s, K_s) = (M_1, K_1, \cdots, M_s, K_s) \cdot g_e.$$

A linear transformation at all source nodes is an $sn \times sn$ matrix $A$ which is a block diagonal matrix,

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_s \end{pmatrix},$$
where $A_t$ is an $n \times n$ transformation matrix at source node $\sigma_i$. For $T \subset [sn]$, let $A_T$ be the matrix obtained by selecting the columns of $A$ corresponding to $T$. For instance, $A_{\{1\}}$ is the first column of $A$. Then, we can give a sufficient and necessary condition for the linear transformation transforming a non-secure code into a secure one.

**Theorem 4.1:** Given the network $\mathcal{N}$, sum function $f$, and security level $r$, if there exists a non-secure network code $\{\hat{g}_e : e \in E\}$ with rate $R = n$, then it can be transformed into a secure network code with rate $n - r$ by a linear transformation $A$ if and only if $A$ satisfies

$$span(A_T^{-1} \cap \text{span}(F(W))) = 0,$$

where $T = \bigcup_{i=0}^{n-1} \{i + 1, i + 2, \ldots, i + n - r\}$.

**Proof.** Denote all source messages and random keys as a row vector of length $sn$ by $X = (M_1, K_1, \cdots, M_s, K_s)$. By the definition of global encoding functions, before transformation, edge $e \in E$ sends a message $U_e = X \cdot g_e$. After transformation, the message sent by edge $e$ is $U_e = X \cdot A \cdot g_e$. Let $W \subset E$ be a wiretap edge subset, $G(W) = \{g_e : e \in W\}$, the wiretapper can get

$$Y_W = X \cdot A \cdot G(W).$$

Here, the security condition

$$I(M_1, M_2, \cdots, M_s; Y_W) = 0$$

is equivalent to that there doesn’t exist any vector $a \in F_q^{(n-r)}$, $b \in F_q^{|W|}$ such that

$$(M_1, M_2, \cdots, M_s) \cdot a = Y_W \cdot b,$$

since

$$(M_1, M_2, \cdots, M_s) = X \cdot \begin{pmatrix} J_{n,n-r} & 0 & \cdots & 0 \\ 0 & J_{n,n-r} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n,n-r} \end{pmatrix},$$

where $J_{n,n-r}$ is the first $n - r$ columns of an identity matrix of order $n$. From above, Eq. (14) is equivalent to

$$X \cdot \begin{pmatrix} J_{n,n-r} & 0 & \cdots & 0 \\ 0 & J_{n,n-r} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n,n-r} \end{pmatrix} \cdot a = X \cdot A \cdot G(W) \cdot b,$$

then we have

$$A^{-1} \cdot \begin{pmatrix} J_{n,n-r} & 0 & \cdots & 0 \\ 0 & J_{n,n-r} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n,n-r} \end{pmatrix} \cdot a = G(W) \cdot b,$$

the lefthand side of Eq. (16) can be represented by $A_T^{-1}$, where $T = \bigcup_{i=0}^{n-1} \{i + 1, i + 2, \ldots, i + n - r\}$. This is equivalent to $span(A_T^{-1} \cap \text{span}(F(W))) = 0$.

For sum-network, the capacity of network $\mathcal{N}$ is equal to $C_{min}$ [3], [14]. Thus from Theorem 4.1 for secure network sum function computing problem, we have $C(\mathcal{N}, f, r) \geq C_{min} - r$.

We call a graph $\mathcal{N} = (V, E)$ as multi-edge tree if for any node $v \in V$, there exists a node $u$ such that all the out edges of $v$ are in edges of $u$, which means $Out(v) \subset In(u)$. We obtain the following corollary.

**Corollary 4.2:** For multi-edge tree graphs, the upper bound in Theorem 3.2 is tight.

**Proof.** In multi-edge tree graphs, for any cut $C$, $J_C = \emptyset$. Thus, $A_{min} = C_{min}$, and by Theorem 3.2, the capacity is $C_{min} - r$. And in Theorem 4.1 there exists a secure network code with rate $C_{min} - r$.

For general network which satisfies $A_{min} > C_{min}$, whether the upper bound (8) is tight or not is not known yet. But we can provide some necessary conditions for linear secure global encoding functions. Let $\hat{g}_e(M_1, K_1, \cdots, M_s, K_s) = (M_1, K_1, \cdots, M_s, K_s) \cdot g_e$, where $g_e$ is a column vector with length $sn$. For any source node $\sigma_i$, let $\hat{g}_{e|\{\sigma_i\}}$ be the global encoding function restricted on $M_i, K_i$, i.e.

$$\hat{g}_{e|\{\sigma_i\}}(M_i, K_i) = (M_i, K_i) \cdot g_{e|\{\sigma_i\}}.$$

**Theorem 4.3:** If there exists an $(n - r, 1)$ linear secure network code $\{\hat{g}_e : e \in E\}$ for sum-network problem $(\mathcal{N}, f, r)$, then for any cut $C$, the messages $\{\hat{g}_{e|C} : e \in C\}$ contain all information about $\sum_{\sigma_i \in I_C} M_i$. 
Suppose we cannot decode \( \sum_{e \in I_C} M_i \) from the truncated messages \( \{ \hat{g}_e : e \in C \} \). Then there must be some random keys from source nodes \( \sigma_i \in I_C \) we cannot eliminate. Let \( G = C \cup (\cup_{e \in I_C} Out(\sigma_i)) \) be a global cut, we can decode the sum from the messages transmitted on these edges. However, \( \cup_{e \in I_C} Out(\sigma_i) \) don’t contain any message about \( K_i \), for \( \sigma_i \in I_C \), which means the random keys in \( \{ \hat{g}_e : e \in C \} \) cannot be eliminated. Then it’s impossible to obtain the sum from the global cut. \( \blacksquare \)

For the case \( A_{\min} \geq C_{\min} \), if the wiretapper has access to a cut \( C \) of the network with size less than \( A_{\min} \), he cannot get any information about any source message from the secure condition. However, from above theorem, the sum \( \sum_{\sigma_i \in I_C} M_i \) can be obtained from truncated messages \( \{ \hat{g}_e \}_{e \in C} \), which indicates that the messages in the min-cut from \( J_C \) to \( C \) are served as random keys to protect the sum from the wiretapper.

V. Upper Bound Under User Secure

In this section, we consider the network function computation problem under the user secure condition (2). Li et al. referred this kind of security as secret sum when the target function is arithmetic sum over finite field \([11]\). In their model, each edge one source having a path to the sink node \( \rho \), the condition for a user secure network function computation problem is that after deleting all wiretap edges, there exists at least length once a time. Therefore, it is meaningless to bound the capacity under their model. In \([11]\), the sufficient and necessary condition for a user secure network function computation problem is that after deleting all wiretap edges, there exists at least one source having a path to the sink node \( \rho \).

Now, we assume the transmission capacity of each edge is unit, i.e. each edge transmits one symbol over \( \mathbb{F}_q \) with unlimited length once a time. Therefore, it is meaningless to bound the capacity under their model. In \([11]\), the sufficient and necessary condition for a user secure network function computation problem is that after deleting all wiretap edges, there exists at least one source having a path to the sink node \( \rho \).

We begin with the following illustrative examples.

![Network Diagram](image)

**Fig. 2:** The network has 2 sources and the target function is \( f = (x_1 + y_1, x_2 + y_2) \).

**Example 5.1:** Consider the user secure network function computation problem shown in Fig. 2. When the security level \( r = 1 \), i.e. the wiretapper has access to exactly one edge of the network, there exists a (2,1) network code over \( \mathbb{F}_2 \) defined as follows: \( U_{e_1} = x_1 + x_2, U_{e_2} = x_2, U_{e_3} = y_1, U_{e_4} = y_1 + y_2, \) and \( U_{e_5} = x_2 + y_1 \). The sink node \( \rho \) can decode \( x_1 + y_1 \) and \( x_2 + y_2 \) from \( U_{e_1}, U_{e_2}, U_{e_3}, U_{e_4}, \) and any one of \( U_{e_1}, \ldots, U_{e_5} \) cannot leakage any information about \( (x_1 + y_1, x_2 + y_2) \). This example shows that \( \hat{C}_{us}(N, f, r) \geq 2 \). On the other hand, for sum function computation task over network, the capacity cannot be greater than \( C_{\min} \). Otherwise, we will have a code with rate larger than \( C_{\min} \) which contradicts with the result in \([3]\). Thus, in this example, \( \hat{C}_{us}(N, f, r) = 2 \) is tight.

The fundamental reason why random keys are not needed is that the user secure condition (2) means the leakage messages \( Y_W \) should not be a function of target result \( X_1 + \cdots + X_n \). In above example, even the wiretapper knows \( x_2 \), he cannot get any information about \( x_2 + y_2 \) since \( y_2 \) is uniformly distributed on \( \mathbb{F}_q \). Therefore, we focus on the capacity without random keys. In fact, if we allow random keys, we may have larger capacity since if a code is secure then it is also user secure.

**Example 5.2:** In Example 5.1, we claim that when the security level \( r = 2 \), there doesn’t exist any user secure network code such that \( \rho \) can compute \( \hat{X} + \hat{Y} \) without error. Suppose there exists an \((l, n)\) user secure code, i.e. source messages
$X, Y \in \mathbb{F}_q^l$, edge $e$ transmits a message $U_e \in \mathbb{F}_q^n$, and from solvability and user security, we have

$$H(X + Y | U_{e_1}, U_{e_4}, U_{e_5}) = 0,$$

$$I(X + Y; U_{e_1}, U_{e_4}) = 0,$$

$$I(X + Y; U_{e_1}, U_{e_5}) = 0,$$

where $U_{e_1}$ is a function of $X$, and $U_{e_4}$ is a function of $Y$. Then we will have

$$I(U_{e_1}, U_{e_5}; X + Y) \geq I(U_{e_1}, U_{e_5}; X + Y | Y)$$

$$= H(U_{e_1}, U_{e_5} | Y) + H(X + Y | Y) - H(X + Y, U_{e_1}, U_{e_5} | Y)$$

$$= H(U_{e_1}) + H(X) - H(X + Y, U_{e_1}, U_{e_5} | Y)$$

$$= H(U_{e_1}) + H(X) - H(U_{e_1}, U_{e_5})$$

$$= I(X; U_{e_1}) = H(U_{e_1}) > 0,$$

which also contradicts the user security. The first equation is from the definition of mutual information, and the second equation holds since $X$ and $Y$ are independent and uniformly distributed over $\mathbb{F}_q^l$ and $U_{e_4}$ is a function of $Y$. From another view, consider $I(X + Y; U_{e_1}, U_{e_4})$, we can prove

$$I(X + Y; U_{e_1}, U_{e_4}) \geq I(X + Y; U_{e_1}, U_{e_4} | Y)$$

$$= H(U_{e_1}, U_{e_4} | Y) + H(X + Y | Y) - H(X + Y, U_{e_1}, U_{e_4} | Y)$$

$$= H(U_{e_1}) + H(X) - H(X + Y, U_{e_1}, U_{e_4} | Y)$$

$$= H(U_{e_1}) + H(X) - H(U_{e_1}, U_{e_4})$$

$$= I(X; U_{e_1}) = H(U_{e_1}) > 0,$$

which also contradicts with user security. The second equation holds since $U_{e_4}$ is determined by $Y$.

From the above examples, we can conclude a sufficient and necessary condition for the existence of user secure network codes.

**Theorem 5.3:** Let $\mathcal{N}$ be a network and $f$ be a sum function over $\mathbb{F}_q$. Suppose the security level is $r$, then there exists a user secure network code if and only if $r < \min\{C_{min}, s\}$, where $s$ is the number of source nodes.

**Proof.** We first prove the only if part. Denote the source messages as $X_1, \ldots, X_s$, the sink node $\rho$ wants $X_1 + \cdots + X_s$, while the wiretapper has access to any $r$ edges and cannot get any information about the sum. We prove this part in two steps.

First, suppose $r \geq C_{min}$ and let $C$ be a cut with $|C| = C_{min}$. Without loss of generality, assume $I_C = \{\sigma_1\}$, let $C' = \cup_{i=2}^s \text{Out}(\sigma_i)$. Note that $G = C \cup C'$ is a global cut of the network, thus the sum can be decoded from messages transmitted by $G$, i.e.

$$H(X_1 + \cdots + X_s | \{U_e, e \in G\}) = 0.$$

If there exists a user secure code with security level $r \geq C_{min}$, then the sum is protected from the wiretapper who has all messages in cut $C$, which indicates

$$I(X_1 + \cdots + X_s; \{U_e, e \in C\}) = 0.$$

Similar with the process in Example 5.2,

$$I(\{U_e, e \in C\} ; X_1 + \cdots + X_s) \geq I(\{U_e, e \in C\} ; X_1 + \cdots + X_s | \{X_2, \ldots, X_s\})$$

$$= H(X_1 + \cdots + X_s | \{X_2, \ldots, X_s\}) - H(X_1 + \cdots + X_s | \{U_e, e \in C\}, \{X_2, \ldots, X_s\})$$

$$= H(X_1) - H(X_1 + \cdots + X_s | \{U_e, e \in C\}, \{X_2, \ldots, X_s\})$$

$$= H(X_1) - 0 > 0,$$

which contradicts with the secure condition above. The third equation holds since $\{U_e, e \in C'\}$ is determined by $\{X_2, \ldots, X_s\}$. And the last equation follows from that any global cut must carry all messages about $X_1 + \cdots + X_s$.

Similarly, suppose there is a user secure code with security level $r \geq s$, then for the edge set $E' = \{e_i : i \in [s], e_i \in \text{Out}(\sigma_i)\}$ with size $s$, the follow equation must be satisfied,

$$I(X_1 + \cdots + X_s; \{U_{e_i}, e_i \in E'\}) = 0.$$
Consider the conditional mutual information,
\[
I(\{U_{e_i}, e_i \in E\}; \sum_{i=1}^{s} X_i) \geq I(\{U_{e_i}, e_i \in E\}; \sum_{i=1}^{s} X_i|\{X_2, \cdots, X_s\})
\]
\[
= H(\sum_{i=1}^{s} X_i|\{X_2, \cdots, X_s\}) + H(\{U_{e_i}, e_i \in E\}|\{X_2, \cdots, X_s\})
\]
\[
- H(\sum_{i=1}^{s} X_i, \{U_{e_i}, e_i \in E\}|\{X_2, \cdots, X_s\})
\]
\[
= H(X_1) + H(U_{e_1}) - H(X_1, U_{e_1})
\]
\[
= I(X; U_{e_1}) = H(U_{e_1}) > 0,
\]
which is also a contradiction. Therefore, when \(r \geq s\) or \(r \geq C_{min}\), there doesn’t exist any user secure code.

For the if part, we claim that if \(r < \min\{C_{min}, s\}\), then there exists a user secure code with rate 1. The construction is as follows. For each source node \(s_i\), we select one path from \(s_i\) to \(\rho\). Each source node sends its message \(x_i\) directly and other nodes send the sum of messages they get. This network code has rate 1 and is user secure since \(r < C_{min}\) and \(r < s\), there exists at least one source and corresponding message cannot be observed by the wiretapper. ★

From Theorem 5.3, we know that there always exists a user secure network code when \(r < \min\{C_{min}, s\}\). The following example shows \(C_{min}\) is not always achievable, which means the study of user secure capacity is not trivial.

\[
\begin{align*}
I(X; U_{e_1}) &= H(U_{e_1}) > 0, \\
H(M_1 + M_2|Y_{e_8}) &= H(M_1 + M_2), \\
\end{align*}
\]

**Example 5.4:** Consider the reverse butterfly network shown in Fig. 3. We claim that there doesn’t exist a rate 2 network code for security level \(r = 1\). Suppose there exists a user secure network code with rate 2. Consider the global cut \(\{e_8, e_9\}\), and let the wiretap edge \(W = \{e_8\}\). By user secure condition 2,
\[
H(M_1 + M_2|Y_{e_8}) = H(M_1 + M_2).
\]

By the solvability of the code, we have
\[
H(M_1 + M_2|Y_{e_8}, Y_{e_9}) = 0.
\]

Combining the above equations,
\[
I(M_1 + M_2; Y_{e_9}|Y_{e_8}) = H(M_1 + M_2|Y_{e_8}) - H(M_1 + M_2|Y_{e_8}, Y_{e_9}) = H(M_1 + M_2) > 0,
\]
which means \(I(M_1 + M_2; Y_{e_9}) > 0\). This means if wiretap edge \(W = \{e_9\}\), the wiretapper will get some information about \(M_1 + M_2\), which is a contradiction.

Let \(G\) be a global cut of the network, and \(G_{min}\) be the smallest size of a global cut. From Example 5.4 we can conclude the following theorem.

**Theorem 5.5:** Let \(\mathcal{N}\) be a network and \(f\) be an algebraic sum over a finite field \(F_q\). If the security level \(r < \min\{C_{min}, s\}\), then
\[
\tilde{C}_{us}(\mathcal{N}, f, r) \leq \min\{C_{min}, G_{min} - r\},
\]
and \(\tilde{C}_{us}(\mathcal{N}, f, r) = 0\) otherwise.

\[
\begin{align*}
\text{Fig. 3: The network has 2 sources and the target function is } f = x + y
\end{align*}
\]
Proof. We only prove \( \hat{C}_{us}(N, f, r) \leq G_{\text{min}} - r \), since \( \hat{C}_{us}(N, f, r) \leq C_{\text{min}} \) can be obtained from [3]. In fact, there doesn’t exist a network code which can compute arithmetic sum on a network with rate larger than \( C_{\text{min}} \) [3]. Similar to Example 5.4 assume \( G \) is a global cut with size \( G_{\text{min}} \), and \( W \subset G \) is a wiretap set. Suppose there exists an \((l, n)\) user secure code. Since source messages are uniformly and independently distributed,

\[
H(M_1 + M_2 + \cdots + M_s) = l \cdot \log q.
\]

By user secure condition \([2]\),
\[
H(M_1 + \cdots + M_s|Y_W) = H(M_1 + \cdots + M_s).
\]

By the solvability of the code, we have
\[
H(M_1, \cdots, M_s|Y_G) = 0.
\]

Combining the above equations,
\[
(|G| - r) \cdot n \cdot \log q \geq H(Y_{G\setminus W})
\]
\[\geq I(M_1 + M_2; Y_{G\setminus W}|Y_W)
\]
\[= H(M_1 + M_2|Y_W) - H(M_1 + M_2|Y_G)
\]
\[= H(M_1 + M_2) = l \cdot \log q,
\]

which shows \( l/n \leq G_{\text{min}} - r \). \qed

VI. Conclusions

In this paper, we first present a new upper bound of the network computation capacity under the secure condition. The new upper bound is strictly tighter than the previous one. When the network is a multi-edge tree, we can prove that the new upper bound is tight. Meanwhile, we give a necessary condition for the existence of linear secure network codes for computing a sum function over network, which provides some insight of the construction of achievable network codes.

On the other hand, we consider the capacity problem under another new secure condition defined as user secure in this paper. We first give a sufficient and necessary condition for the existence of user secure network codes. Furthermore, an upper bound of the user secure computation capacity is addressed. However, whether the bound is tight or not remains unknown yet.

References

[1] A. Giridhar and P. Kumar, “Computing and communicating functions over sensor networks”, IEEE Journal on Selected Areas in Communications, vol. 23, no. 4, pp. 755-764, 2005.
[2] A. Destounis, G. S. Paschos, and I. Koutsopoulos, “Streaming big data meets backpressure in distributed network computation”, in Proceeding of the 37th Annual IEEE Symposium on Foundations of Computer Science, pp. 1-9, 2016.
[3] B. Rai and B. K. Dey, “On network coding for sum-networks”, IEEE Transactions on Information Theory, vol. 58, no. 1, pp 50-63, 2012.
[4] R. Appuswamy and M. Franceschetti, “Computing linear functions by linear coding over networks”, IEEE Transactions on Information Theory, vol. 60, no. 1, pp. 422-431, 2013.
[5] R. Appuswamy, M. Franceschetti, and N. Karamchandani, “Network coding for computing: Cut-set bounds”, IEEE Transactions on Information Theory, vol. 57, no. 2, pp. 1015-1030, 2011.
[6] X. Huang, Z. Tan, S. Yang, and X. Guang, “Comments on cut-set bounds on network function computation”, IEEE Transactions on Information Theory, vol. 64, no. 9, pp. 6454-6459, 2018.
[7] X. Guang, R. W. Yeung, S. Yang, and C. Li, “Comments on cut-set bounds on network function computation”, IEEE Transactions on Information Theory, vol. 65, no. 6, pp. 3790-3811, 2019.
[8] N. Cai and T. Chan, “Theory of secure network coding”, Proceedings of the IEEE, vol. 99, no. 3, pp. 421-437, 2011.
[9] K. Bhattacharyya, R. Poovendran, and B. S. Lee, “Weakly secure network coding”, NetCod, 2005.
[10] K. Harada and H. Yamamoto, “Strongly secure linear network coding”, IEICE transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. 91, no. 10, pp. 2720-2728, 2008.
[11] S. Li and C. T. Li, “Arithmetic network coding for secret sum computation”, arXiv preprint, arXiv:2201.03032, 2022.
[12] N. Cai and R. W. Yeung, “Secure network coding”, IEEE International Symposium on Information Theory (ISIT), IEEE, pp. 323, 2002.
[13] R. Koetter, M. Effros, and T. Ho, “Network codes as codes on graphs”, Proceeding of CISS, 2004.