Secure Multiplex Coding and Its Application to Secure Network Coding

Ryutaroh MATSUMOTO[1], Masahito HAYASHI[2]

[1]Tokyo Institute of Technology [2]Tohoku University / National University of Singapore

March 23, 2011
Institute of Network Coding
Chinese University of Hong Kong

Structure of this talk

Part 1: Secure multiplex coding

1. What is the secure multiplex coding
2. Problems in the existing research
3. Our improvements
4. Key ideas in the proof

Part 2: Secure multiplex network coding

1. Applying the same idea to the secure network coding
2. Relation to the existing results

Conclusion (if there is enough time)

Review of the wiretap channel

$$
\begin{tikzcd}
S \arrow[r] & \boxed{\text{Alice's encoder}} \arrow[r, "X"] & \boxed{\text{Channel}} \arrow[r] \arrow[d] & \boxed{\text{Bob's decoder}} \arrow[r] & S \\
& D \arrow[u] & Z & & \\
& & \boxed{\text{Eve}} & &
\end{tikzcd}
$$

- The secret message S should be sent reliably.
- S should be kept secret from Eve.

D: dummy random message statistically independent of S

(S, D) is encoded by an ordinary channel encoder.

Initiated by Wyner (1975), Csiszár-Körner (1978).

- Information rate is decreased by $\log |D|/n$.
- When the rate loss $\log |D|/n > I(X;Z)$, one can make S almost statistically independent of Eve's received signal.

If you cannot tolerate the rate loss...

Secure Multiplex Coding (Yamamoto et al. (ITW 2005))

$$S \to \boxed{\text{Alice's encoder}} \xrightarrow{X} \boxed{\text{Channel}} \xrightarrow{Y} \boxed{\text{Bob's decoder}} \to S\&S_2$$

$$\begin{array}{ccc} \uparrow & & \downarrow \\ \cancel{D}S_2 & & Z \\ & & \boxed{\text{Eve}} \end{array}$$

S: secret message

S_2: another secret message, statistically independent of S

- Substitution of D with meaningful S_2 remove the rate loss.
- To completely hide S and S_2 from Eve, rates of S and S_2 has to be $> I(X;Z)$.
- When $I(X;Z) > I(X;Y)/2$, we need three or more secret messages.
- Each message has to be distributed uniformly.

Problems in the existing research

$S_1, S_2, \ldots, S_T$: T independent secret messages

1. (*) They proved $I(S_i; Z) \simeq 0$ for $\forall i$ but did not evaluated $I(S_{i_1}, S_{i_2}, \ldots, S_{i_j}; Z)$.
2. The evaluation of $I(S_i; Z)$ could be improved (we failed to improve).
3. (*) Alice cannot send a common message destined for both Bob and Eve as "Broadcast Channel with Confidential Messages"

We improved two of them with (*) marks. For simplicity I assume no common message.

Proposed encoding procedure for secure multiplex coding

1. Fix a nonsingular (binary) matrix L randomly and Alice and Bob agree on its choice. Eve is allowed to know L.
2. Set the (binary) vector $\vec{m}$ from T secret messages $\vec{s}_1, \ldots, \vec{s}_T$ by

$$\vec{m} = L \begin{pmatrix} \vec{s}_1 \\ \vdots \\ \vec{s}_T \end{pmatrix}$$

3. Encode $\vec{m}$ by an ordinary channel encoder

Bob reverses the above procedure for his decoding.

The decoding probability is not more than the underlying channel code. Evaluation of the mutual information is remaining. . . .

Informal description of the upper bound on the mutual information

R_i: information rate of the i-th secret message $\vec{s}_i$
n: code length, X: transmitted symbol, Z: Eve's received symbol

$I(\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_j; Z^n)/n$

$$\begin{array}{ll} \to 0 (\text{exponentially of } n) & \text{if } \sum_{i=j+1}^{T} R_i > I(X;Z) \\ \to I(X;Z) - \sum_{i=j+1}^{T} R_i & \text{otherwise (not evaluated by Yamamoto et al.)} \end{array}$$

- $\vec{S}_{j+1}, \ldots, \vec{S}_T$ serve as a dummy random message making $(\vec{S}_1, \ldots, \vec{S}_j)$ secret to Eve.
- If the capacity $I(X;Z)$ is filled by $\vec{S}_{j+1}, \ldots, \vec{S}_T$, then $I(\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_j; Z^n)$ converges to zero exponentially with n.
- Otherwise $I(\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_j; Z^n)$ converges to ∞ linearly of n.
- We cannot asymptotically decrease $I(\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_j; Z^n)$ without decreasing rates R_i, i.e., the presented coding scheme is optimal, i.e., the capacity region is determined with this problem formulation.

Formal definition of the capacity region (w/o common message)

R_i: the rate of i-th secret message ($1 \le i \le T$)
$\mathcal{I} \subseteq \{1, \dots, T\}$
$R_{e,\mathcal{I}}$: minimum requirement on $H(S_{i,n} : i \in \mathcal{I} | Z^n)/n$
The tuple $(R_1, \dots, R_T)$ and $(R_{e,\mathcal{I}})$ is achievable if

$$\begin{aligned}
\lim_{n\to\infty} \text{decoding error prob.} &= 0, \\
\liminf_{n\to\infty} H(S_{\mathcal{I},n}|Z^n)/n &\ge R_{e,\mathcal{I}}, \\
\lim_{n\to\infty} I(S_{\mathcal{I},n}; Z^n) &= 0\Big(\text{if } R_{e,\mathcal{I}} = \sum_{i\in\mathcal{I}} R_i\Big), \\
\liminf_{n\to\infty} \frac{\log |\mathcal{S}_{i,n}|}{n} &\ge R_i,
\end{aligned}$$

for all $i = 1, \dots, T$ and all $\mathcal{I} \subseteq \{1, \dots, T\}$.
The capacity region is defined to be the closure of the achievable rate tuples.

Single-letterized formula for the capacity region

$P_{Y|X}$: channel from Alice to Bob
$P_{Z|X}$: channel from Alice to Eve
$V \to X \to YZ$.

$$
\begin{aligned}
\sum_{i=1}^{T} R_i &\leq I(V;Y) \\
R_{e,\mathcal{I}} &\leq I(V;Y) - I(V;Z) \text{ for all } \emptyset \neq \mathcal{I} \subseteq \{1,\ldots,T\}, \\
R_{e,\mathcal{I}} &\leq \sum_{i\in\mathcal{I}} R_i.
\end{aligned}
$$

Key ideas in the proof

Converse part can be borrowed from that for the broadcast channel with confidential messages (BCC) without change, by regarding $(S_{i,n} : i \in \mathcal{I})$ as the secret message in the BCC.

Direct part

- The decoding error probability is not worse than the underlying channel code.
- We have to evaluate the mutual information $I([S_{i,n} : i \in \mathcal{I}]; Z^n)$.

We need to introduce the two-universal hash functions and the privacy amplification theorem for the final task.

Family of two-universal hash functions

Definition

Let $\mathcal{F}$ be a set of functions from a set $\mathcal{S}_1$ to $\mathcal{S}_2$, F a (uniform) RV on $\mathcal{F}$. If for all $x_1 \neq x_2 \in \mathcal{S}_1$ we have

$$\Pr[F(x_1) = F(x_2)] \leq \frac{1}{|\mathcal{S}_2|},$$

then $\mathcal{F}$ is said to be a family of two-universal hash functions.

New privacy amplification (PA) theorem

(M, Z): discrete RVs
$\mathcal{F}$: a family of two-universal hash functions from $\mathcal{M}$ to $\mathcal{S}$
F: an RV on $\mathcal{F}$ statistically independent of (M, Z).

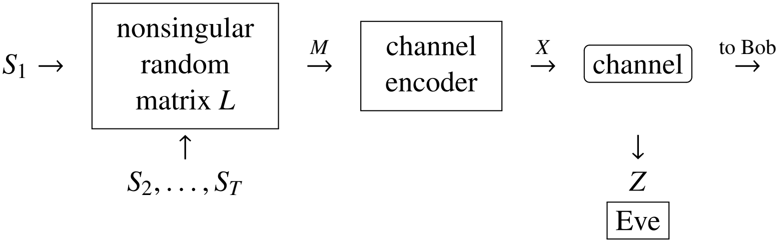
$$\begin{aligned} I(F(M); Z|F) &\leq \frac{|\mathcal{S}|^\rho \mathbf{E}[P_{M|Z}(M|Z)^\rho]}{\rho} \text{(Hayashi 2011)}, \\ \mathbf{E}_f \exp[\rho I(F(M); Z|F = f)] &\leq 1 + |\mathcal{S}|^\rho \mathbf{E}[P_{M|Z}(M|Z)^\rho] \text{(new)}, \end{aligned}$$

for all $0 < \rho \leq 1$.

The new inequality is stronger than Hayashi's latest work. The old one cannot tightly evaluate the mutual information when it goes to ∞.

How to apply the PA theorem

The artificial noise $P_{X|V}$ is ignored.



$M = L \times (S_1, \ldots, S_T)$.

Evaluate $I(S_1; Z)$ by applying the PA theorem to M and Z.

To do so, the function $M \mapsto S_1$ must be two-universal hashing.

To evaluate $I([S_{i,n} : i \in \mathcal{I}]; Z^n)$, the function $M \mapsto [S_{i,n} : i \in \mathcal{I}]$ must also be two-universal hashing.

Random matrix makes two-universal hashing

$\mathcal{B} = \prod_{i=1}^{T} \mathcal{S}_i$
$\mathcal{I} \subseteq \{1, \ldots, T\}$
$\alpha_{\mathcal{I}}$: projection from $\mathcal{B}$ to $\prod_{i \in \mathcal{I}} \mathcal{S}_i$

Proposition

If $\mathcal{L}$ is the set of all bijective linear maps on $\mathcal{B}$, then $\{\alpha_{\mathcal{I}} \circ L \mid L \in \mathcal{L}\}$ is a family of two-universal hash functions.

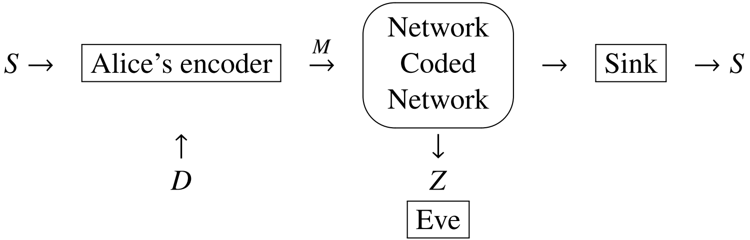
Structure of the mutual information evaluation

1. By the PA theorem, obtain an upper bound on $I([S_{i,n} : i \in \mathcal{I}]; Z^n)$.
2. Modify the upper bound so that the upper bound become concave w.r.t. the message distribution with fixed channel conditional distribution.
3. Moving the averaging of random coding (of the channel code) into the upper bound.
4. Single-letterize the upper bound.
5. Evaluate the speed of the convergence of the upper bound to zero or ∞.

For the detail, please refer to arXiv:1101.4036.

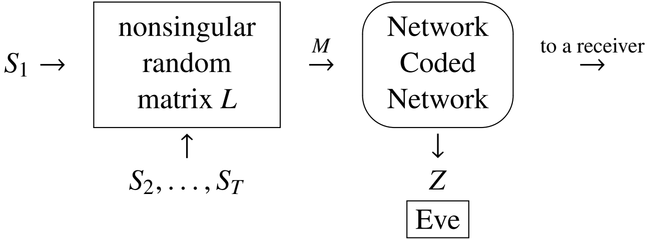
Applying the same idea to the secure network coding

Single source multicast is considered.



- Originally, coding at intermediate nodes are carefully chosen. In this talk I do not change coding at intermediate nodes.
- Traditionally, exactly zero $I(S;Z)$ is required. In this talk, I regard nonzero but arbitrary small $I(S;Z)$ to be acceptable.

Applying the PA theorem to the secure network coding



$M = L \times (S_1, \ldots, S_T)$.
Eve may know L. L is not a secret shared key between Alice and Bob.

- Evaluate $I(S_1; Z)$ by applying the PA theorem to M and Z.
- Evaluate $I([S_{i,n} : i \in \mathcal{I}]; Z^n)$ by applying the PA theorem to M and Z.

The number of tapped link is at most μ per time slot. $I(S_i; Z)$ should be small with any choice of μ links.

Simplification of the upper bound

n: minimum of max flows to the legitimate receivers
m: number of time slots used for coding
$M \in \mathbf{F}_q^{mn}$
$Z \in \mathbf{F}_q^{m\mu}$

By the PA theorem

$$
\begin{aligned}
I(S_i; Z|L) &\leq \frac{|\mathcal{S}_i|^\rho \mathbf{E}[P_{M|Z}(M|Z)^\rho]}{\rho} \\
&\leq q^{-m\rho(n-\mu-\log_q |\mathcal{S}_i|/m)}/\rho \\
I(S_i; Z|L) &\leq q^{-m(n-\mu-\log_q |\mathcal{S}_i|/m)}
\end{aligned}
$$

This inequality shows that for a particular choice of μ links, almost all choices of random matrices L make $I(S_i; Z|L)$ small. But we need to ensure that almost every L makes $I(S_i; Z|L)$ small with all choices of μ links.

When the locations of wire-tapping are time-invariant

Let $Z = BM$, and the matrix B represents eavesdropping. Let B be drawn according to the uniform distribution of all possible eavesdropping.

$$I(S_i; Z|L, B) \leq q^{-m(n-\mu-\log_q |S_i|/m)}$$

For probability $1 - 1/C_E$, a realization b of B makes

$$I(S_i; Z|L, B) \leq C_E q^{-m(n-\mu-\log_q |S_i|/m)}.(**)$$

When the locations of tapped links do not change in time, the possible number of B is finite ($\leq q^{n\mu}$). By setting C_E larger than the number of B, we can see that (**) holds for every B.

The above argument breaks down when the locations of tapped links change in time. I made the same error in the talk presented at INC Sept. 22, 2010.

Relation to the existing research I

Bhattad and Nayaranan (NetCod 2005) proposed the weakly secure network coding, in which

- No dummy message nor rate loss,
- For each collection of $n - \mu$ or less messages, its mutual information to Eve is zero

The disadvantages are

- Code construction depends on the network topology,
- The computational complexity of code construction is huge.

The proposed construction does not have those disadvantages.

Relation to the existing research II

Silva and Kschischang (ITW 2009) proposed the universal weakly secure network coding, in which

- No dummy message nor rate loss,
- For each collection of 2 or less messages, its mutual information to Eve is zero,
- Low complexity of the code construction,
- Independent of network topology and coding at intermediate nodes.

The disadvantages are

- No explicit construction for collections of three or more messages.

Our construction ensures that for almost all choices of L, the mutual information of every collection of messages to Eve is below a specified value.

Conclusion

- We improved the analysis of secure multiplex coding by Yamamoto et al., in particular, mutual information of collections of messages and its optimal convergence speed to ∞ are clarified.
- The same idea is also applied to the secure network coding.

The security requirements are different from the traditional secure network coding:

- Arbitrary small mutual information instead of strictly zero.
- Almost all choices of random matrices ensures the small mutual information.

On multiple independent uniform messages

The secret messages $S_1, \ldots, S_T$ have to be uniform and independent. Otherwise the proof breaks down. The assumption is too restrictive.

- Optimal compression makes almost uniform distribution (w.r.t. the normalized KL divergence).
- Small deviation from the uniform distribution may increase the mutual information little.
- Compressed information could be used as $S_1, \ldots, S_T$.

We are working on formally prove the above.