

Secure Multiplex Network Coding

Ryutaroh MATSUMOTO¹, Masahito HAYASHI²

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

July 25, 2011
NETCOD 2011

www.leaderstudio.net

What is the secure network coding?

- **Goal:** Make the transmitted information secret from the adversary (Eve).
- (Usually) Single source multicast
- Linear network coding over \mathbf{F}_q
- n = the minimum of max flows
- Eve (eavesdropper) can wire-tap at most μ links per time slot.

Loss of information rate

If at least $n - \mu$ random dummy \mathbf{F}_q symbols are sent along with the information S , then $I(S; Z)$ can be made zero by a suitable network coding, where Z = Eve's observation.

If you don't like the rate loss...

The weakly and the strongly secure network coding

Let the secret information $S = (S_1, \dots, S_n) \in \mathbf{F}_q^n$.

μ : the number of wire-tapped links

Z : Eve's observation

The following codes have **no rate loss**.

Weakly secure network coding (Bhattad & Narayanan 2005)

For all $i = 1, \dots, n$, $I(S_i; Z) = 0$.

Practical construction was given by Silva and Kschischang (2009).

Strongly secure network coding (Harada & Yamamoto 2008)

For all $\mathcal{I} \subset \{1, \dots, n\}$ with $|\mathcal{I}| \leq n - \mu$, we have $I(S_{\mathcal{I}}; Z) = 0$. $S_{\mathcal{I}} = [S_i : i \in \mathcal{I}]$.

Remark: $I(S_1, \dots, S_n; Z) \approx \mu \log q$.

Harada & Yamamoto presented a construction algorithm with huge complexity...

RNC is strongly secure (Cai 2009)

Cai (2009) proved that RNC (Random linear Network Coding) gives the strongly secure network coding with arbitrary high probability over sufficiently large finite fields.

But...

- No explicit expression for the required field size
- The required field size seems to be much larger than the ordinary RNC...
- Structured coding at intermediate nodes is sometimes desirable.

www.leadersuniversity.net

Random linear coding at the source node realizes the strong security

Features and assumptions in the proposed method

- Vector linear network coding using m time slots
- Arbitrarily fixed field size
- T secret messages
- the i -th message is $S_i \in \mathbf{F}_q^{k_i}$

- $nm = \sum_{i=1}^T k_i$

- The set of wire-tapped links are constant during m time slots.

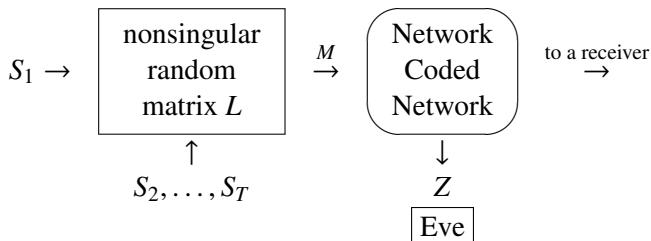
$$\sum_{i \in \mathcal{I}} k_i < m(n - \mu) \Rightarrow I(S_{\mathcal{I}}; Z) = 0 \text{ with high probability, where } S_{\mathcal{I}} = [S_i : i \in \mathcal{I}]$$

Comments on the proposed method

- The proposed method looks as a generalization of the strongly secure NC to multiple time slots.
- Yamamoto et al. (2005) applied almost the same idea to the wiretap channel and called their proposal as “secure multiplex coding.” The name of our proposal comes from it.

www.leaderstudio.net

Figure of the proposed method



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

Eve may know L . L is not a secret shared key between Alice and Bob.

- $I(S_I; Z) = 0$ with high probability of random choice of L , when $\sum_{i \in I} k_i < m(n - \mu)$.
- Encoding only at the source node is changed.
- Universal in the sense of Silva and Kschischang (2009).

I will show a proof sketch.

Privacy amplification (PA) theorem (special case)

(M, Z) : discrete RVs

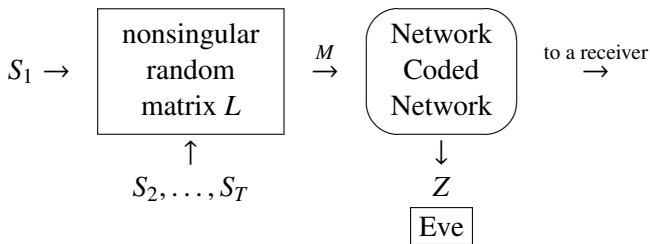
F : uniformly randomly chosen surjective linear maps \mathcal{M} to \mathcal{S} statistically independent of (M, Z) .

$$I(F(M); Z|F) \leq |\mathcal{S}| \mathbf{E}[P_{M|Z}(M|Z)] \text{ (Bennett et al. 1995).}$$

The base of all the logarithms are e .

www.leaderstudio.net

How to apply the PA theorem



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

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 assumptions in the PA theorem are satisfied.

www.leadrstudio.net

What we have proved so far??

Already proved

For a **fixed** set of μ wire-tapped links and **fixed** $\mathcal{I} \subset \{1, \dots, T\}$,

$$\mathbf{E}_\ell I(S_{\mathcal{I}}; Z|L = \ell) \leq \exp_q(-m(n - \mu - \sum_{i \in \mathcal{I}} k_i/m))$$

What we have to prove

For **every** set of μ wire-tapped links and **every** \mathcal{I} , with high probability we have

$$I(S_{\mathcal{I}}; Z|L = \ell) = 0$$

Final adjustments in proof

$I(S_{\mathcal{I}}; Z|L = \ell)$ is an integer multiple of $\log q$.

If $\mathbf{E}_{\ell} I(S_{\mathcal{I}}; Z|L = \ell) < (\log q)/C_1$ then $I(S_{\mathcal{I}}; Z|L = \ell) = 0$ with probability $> 1 - 1/C_1$.

The total number of wire-tapping patterns is $\binom{|E|}{\mu}$.

The total number of indices sets \mathcal{I} is $\leq 2^T - 1$. Random choice of ℓ makes $I(S_{\mathcal{I}}; Z|L = \ell) = 0$ for **every** set of μ wire-tapped links and **every** \mathcal{I} with probability

$$> 1 - \frac{2^T - 1}{C_1} \binom{|E|}{\mu}.$$

Recall $C_1 = \exp_q m(n - \mu - \sum_{i \in \mathcal{I}} k_i/m)$, and an explicit value of m is computable.

Remark: This argument depends on the fact that $\binom{|E|}{\mu}$ does not grow with m (coding length).

Analysis when $\sum_{i \in \mathcal{I}} k_i \geq m(n - \mu)$

We prove that for any $\delta > 0$

$$I(S_{\mathcal{I}}; Z|L = \ell) \leq m \left(\sum_{i \in \mathcal{I}} k_i/m + \mu - n + \delta \right)$$

for every wire-tapping pattern and every \mathcal{I} with high probability for large m .

For its proof, we need to strengthen the privacy amplification theorem as in our proc. paper.

www.leaderstudio.net

Concluding remarks

- Our proc. paper does not prove $I(S_I; Z) = 0$. Please refer to the arxiv eprint.
- Removal of the random dummy message and rate loss can be done for the interference channels (in preparation) and broadcast channels with confidential messages (ISIT 2011).
- \Rightarrow The dummy message might be unnecessary in other contexts...

www.leaderstudio.net

On multiple independent uniform messages

The secret messages S_1, \dots, S_T have to be uniform and independent. Otherwise the proof breaks down. The same assumption is made in all the previous paper. 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, \dots, S_T .

We are working on formally prove the above.

www.leaderstudio.net

How individual mutual information can be zero?

S_1, S_2 : uniform and independent over \mathbf{F}_q .

$$Z = S_1 + S_2$$

$$I(S_1; Z) = I(S_2; Z) = 0, \quad I(S_1, S_2; Z) = \log q.$$

Eve (having Z) has no information on S_1 or S_2 , while she knows possible combination of S_1 and S_2 .

When each secret message is generated by a different information source (person/organization), leakage of possible combination of unrelated messages is acceptable (in my view).

www.leaderstudio.net