Truly Efficient 2-Round Perfectly Secure Message Transmission Scheme

Kaoru Kurosawa[1] and Kazuhiro Suzuki[2]

[1] Department of Computer and Information Sciences, Ibaraki University
kurosawa@mx.ibaraki.ac.jp
[2] Venture Business Laboratory, Ibaraki University, Hitachi, Ibaraki 316-8511, Japan
tutetuti@dream.com

Abstract. In the model of perfectly secure message transmission schemes (PSMTs), there are n channels between a sender and a receiver. An infinitely powerful adversary $\mathbf{A}$ may corrupt (observe and forge) the messages sent through t out of n channels. The sender wishes to send a secret s to the receiver perfectly privately and perfectly reliably without sharing any key with the receiver.
In this paper, we show the first 2-round PSMT for $n = 2t + 1$ such that not only the transmission rate is $O(n)$ but also the computational costs of the sender and the receiver are both polynomial in n. This means that we solve the open problem raised by Agarwal, Cramer and de Haan at CRYPTO 2006.

Keywords: Perfectly secure message transmission, information theoretic security, efficiency

1 Introduction

In the model of (r-round, n-channel) message transmission schemes [2], there are n channels between a sender and a receiver. An infinitely powerful adversary $\mathbf{A}$ may corrupt (observe and forge) the messages sent through t out of n channels. The sender wishes to send a secret s to the receiver in r-rounds without sharing any key with the receiver.

We say that a message transmission scheme is perfectly secure if it satisfies perfect privacy and perfect reliability. The perfect privacy means that the adversary $\mathbf{A}$ learns no information on s, and the perfect reliability means that the receiver can output $\hat{s} = s$ correctly.

For $r = 1$, Dolev et al. showed that there exists a 1-round perfectly secure message transmission scheme (PSMT) if and only if $n \geq 3t + 1$ [2]. They also showed an efficient 1-round PSMT [2].

For $r \geq 2$, it is known that there exists a 2-round PSMT if and only if $n \geq 2t + 1$ [2]. However, it is very difficult to construct an efficient scheme for $n = 2t + 1$. Dolev et al. [2] showed a 3-round PSMT such that the transmission rate is $O(n^5)$, where the transmission rate is defined as

$$\frac{\text{the total number of bits transmitted}}{\text{the size of the secrets}}.$$

Sayeed et al. [7] showed a 2-round PSMT such that the transmission rate is $O(n^3)$.

Recently, Srinathan et al. showed that n is a lower bound on the transmission rate of 2-round PSMT [8]. Then Agarwal, Cramer and de Haan [1] showed a 2-round PSMT such that the transmission rate is $O(n)$ at CRYPTO 2006 based on the work of Srinathan et al. [8] [1]. However, the communication complexity is exponential because the sender must broadcast consistency check vectors of size [2]

$$w = \binom{n-1}{t+1} = \binom{2t}{t+1}.$$

In other words, Agarwal et al. [1] achieved the transmission rate of $O(n)$ by sending exponentially many secrets. Therefore, the computational costs of the sender and the receiver are both exponential. Indeed, the authors wrote [1, Sec.6] that:

"We do not know whether a similar protocol can exist where sender and receiver restricted to polynomial time (in terms of the number of channels n) only".

In this paper, we solve this open problem. That is, we show the first 2-round PSMT for $n = 2t + 1$ such that not only the transmission rate is $O(n)$ but also the computational costs of the sender and the receiver are both polynomial in n.

Table 1. 2-Round PSMT for $n = 2t + 1$

	Trans. rate	com. complexity	Receiver	Sender
Agarwal et al. [1]	$O(n)$	exponential	exponential	exponential
This paper	$O(n)$	$O(n^3)$	poly	poly

The main novelty of our approach is to introduce a notion of *pseudo-basis* to the coding theory. Let $\mathcal{C}$ be a linear code of length n over a finite field F with the minimum Hamming distance $d = t+1$. Consider a message transmission scheme such that the sender chooses a codeword $X_i = (x_{i1}, \cdots, x_{in})$ of $\mathcal{C}$ randomly and sends x_{ij} through channel j for $j = 1, \cdots, n$. Note that the receiver can detect t errors, but cannot correct them because $d = t + 1$.

If the sender sends many codewords, however, then we can do something better. Suppose that the sender sent X_i as shown above, and the receiver received $Y_i = X_i + E_i$ for $i = 1, \cdots, m$, where E_i is an error vector caused by the adversary. We now observe that the dimension of the space $\mathcal{E}$ spanned by the error vectors

[1] Srinathan et al. claimed that they constructed a 2-round PSMT such that the transmission rate is $O(n)$ in [8]. However, Agarwal et al. pointed out that it has a flaw in [1].

[2] Indeed, in [1, page 407], it is written that "at most $O(w)$ indices and field elements are broadcast ....", where w is defined in [1, page 403] as shown above.

$E_1, \cdots, E_m$ is at most t because the adversary corrupts at most t channels. Suppose that $\{E_{i_1}, \cdots, E_{i_k}\}$ is such a basis, where $k \leq t$. For the same indices, we say that $\mathcal{B} = \{Y_{i_1}, \cdots, Y_{i_k}\}$ is a *pseudo-basis* of $\mathcal{Y} = \{Y_1, \cdots, Y_m\}$. We then show that a receiver can find a pseudo-basis $\mathcal{B}$ of $\mathcal{Y}$ in polynomial time.

By using this algorithm, we first show a 3-round PSMT for $n = 2t + 1$ such that the transmission rate is $O(n)$ and the computational cost of the sender and the receiver are both polynomial in n. (See Fig.3.) Then combining the technique of [8, 1], we show a 2-round PSMT such that not only the transmission rate is $O(n)$ but also the computational cost of the sender and the receiver are both polynomial in n.

(Remark) Recently, Fitzi et al. showed an efficient 2-round PSMT for $n \geq (2+\epsilon)t$ for any constant $\epsilon > 0$ [4], but not for $n = 2t + 1$.

2 Main Idea

Suppose that there are n channels between the sender and the receiver, and an adversary may corrupt t out of n channels. We use F to denote $GF(p)$, where p is a prime such that $p > n$. [3] Let $\mathcal{C}$ be a linear code of length n such that a codeword is $X = (f(1), \cdots, f(n))$, where $f(x)$ is a polynomial over F with $\deg f(x) \leq t$.

2.1 Difference from Random t Errors

Consider a message transmission scheme such that the sender chooses a codeword $X = (f(1), \cdots, f(n))$ of $\mathcal{C}$ randomly, and sends $f(i)$ through channel i for $i = 1, \cdots, n$. Then the adversary learns no information on $f(0)$ even if she observes t channels because $\deg f(x) \leq t$. Thus perfect privacy is satisfied.

If $n = 3t+1$, then the minimum Hamming distance of $\mathcal{C}$ is $d = n-t = 2t+1$. Hence the receiver can correct t errors caused by the adversary. Thus perfect reliability is also satisfied. Therefore we can obtain a 1-round PSMT easily.

If $n = 2t + 1$, however, the minimum Hamming distance of $\mathcal{C}$ is $d = n - t = t + 1$. Hence the receiver can only detect t errors, but cannot correct them. This is the main reason why the construction of PSMT for $n = 2t + 1$ is difficult.

What is a difference between usual error correction and PSMTs ? If the sender sends a single codeword $X \in \mathcal{C}$ only, then the adversary causes t errors randomly. Hence there is no difference. If the sender sends many codewords $X_1, \cdots, X_m \in \mathcal{C}$, however, the errors are not totally random. This is because the errors always occur at the same t (or less) places !

To see this more precisely, suppose that the receiver received

$$Y_i = X_i + E_i, \tag{1}$$

[3] We adopt $GF(p)$ only to make the presentation simpler, where the elements are denoted by $0, 1, 2, \cdots$. But in general, our results hold for any finite field F whose size is larger than n.

where $E_i = (e_{i1}, \cdots, e_{in})$ is an error vector caused by the adversary. Define

$$support(E_i) = \{j \mid e_{ij} \neq 0\}.$$

Then there exist some t-subset $\{j_1, \cdots, j_t\}$ of n channels such that each error vector E_i satisfies

$$support(E_i) \subseteq \{j_1, \cdots, j_t\}, \tag{2}$$

where $\{j_1, \cdots, j_t\}$ is the set of channels that the adversary forged.

This means that the space $\mathcal{E}$ spanned by $E_1, \cdots, E_m$ has dimension at most t. We will exploit this fact extensively.

2.2 Pseudo-Basis and Pseudo-Dimension

Let $\mathcal{V}$ denote the n-dimensional vector space over F . For two vectors $Y, E \in \mathcal{V}$, we write

$$Y = E \bmod \mathcal{C}$$

if $Y - E \in \mathcal{C}$.

For $i = 1, \cdots, m$, suppose that the receiver received Y_i such that

$$Y_i = X_i + E_i,$$

where $X_i \in \mathcal{C}$ is a codeword that the sender sent and E_i is the error vector caused by the adversary. From now on, (Y_i, X_i, E_i) has this meaning. Then we have that

$$Y_i = E_i \bmod \mathcal{C} \tag{3}$$

for each i. Let $\mathcal{E}$ be a subspace spanned by $E_1, \cdots, E_m$.

We first define a notion of *pseudo-span*.

Definition 1. *We say that* $\{Y_{j1}, \cdots, Y_{jk}\} \subset \mathcal{Y}$ *pseudo-spans* $\mathcal{Y}$ *if each* $Y_i \in \mathcal{Y}$ *can be written as*

$$Y_i = a_1 Y_{j1} + \cdots + a_k Y_{jk} \bmod \mathcal{C}$$

for some $a_i \in \mathsf{F}$.

We next define a *pseudo-basis* and the *pseudo-dimension* of $\mathcal{Y}$.

Definition 2.
- *Let* k *be the dimension of* $\mathcal{E}$. *We then say that* $\mathcal{Y}$ *has the pseudo-dimension* k.
- *Let* $\{E_{j1}, \cdots, E_{jk}\}$ *be a basis of* $\mathcal{E}$. *For the same indices, we say that* $\{Y_{j1}, \cdots, Y_{jk}\}$ *is a pseudo-basis of* $\mathcal{Y}$.

The following theorem is clear since the adversary forges at most t channels.

Theorem 1. *The pseudo-dimension of* $\mathcal{Y}$ *is at most* t.

Suppose that $\{Y_{j1}, \cdots, Y_{jk}\}$ is a pseudo-basis of $\mathcal{Y}$. Define

$$\mathsf{FORGED} = \bigcup_{i=1}^{k} support(E_{ji}). \tag{4}$$

It is then clear that FORGED is the set of all channels that the adversary forged. Therefore, the following theorem holds.

Theorem 2. *For each j,*

$$support(E_j) \subseteq \mathsf{FORGED}.$$

We finally prove the following theorem.

Theorem 3. $\mathcal{B} = \{Y_{j1}, \cdots, Y_{jk}\}$ *is a pseudo-basis of $\mathcal{Y}$ if and only if $\mathcal{B}$ is a minimal subset of $\mathcal{Y}$ which pseudo-spans $\mathcal{Y}$.*

(Proof) (I) Suppose that $\mathcal{B}$ is a minimal subset of $\mathcal{Y}$ which pseudo-spans $\mathcal{Y}$. Then each $Y_i \in \mathcal{Y}$ can be written as

$$Y_i = a_1 Y_{j1} + \cdots + a_k Y_{jk} \bmod \mathcal{C}$$

for some $a_i \in \mathsf{F}$. From eq.(3), we obtain that

$$E_i = a_1 E_{j1} + \cdots + a_k E_{jk} \bmod \mathcal{C}.$$

Hence

$$E_i - a_1 E_{j1} - \cdots - a_k E_{jk} \in \mathcal{C}.$$

The Hamming weight of the left hand side is at most t while the minimum Hamming weight of $\mathcal{C}$ is $t+1$. Therefore, $E_i - a_1 E_{j1} - \cdots - a_k E_{jk}$ is a zero-vector. Hence we obtain that

$$E_i = a_1 E_{j1} + \cdots + a_k E_{jk}.$$

This means that $\{E_{j1}, \cdots, E_{jk}\}$ spans $\mathcal{E}$. Further the minimality of $\mathcal{B}$ implies that $\{E_{j1}, \cdots, E_{jk}\}$ is a basis of $\mathcal{E}$. Therefore, from Def.2, $\mathcal{B} = \{Y_{j1}, \cdots, Y_{jk}\}$ is is a pseudo-basis of $\mathcal{Y}$.

(II) Suppose that $\mathcal{B} = \{Y_{j1}, \cdots, Y_{jk}\}$ is a pseudo-basis of $\mathcal{Y}$. Then $\{E_{j1}, \cdots, E_{jk}\}$ is a basis of $\mathcal{E}$. Therefore each E_i can be written as

$$E_i = a_1 E_{j1} + \cdots + a_k E_{jk}$$

for some $a_i \in \mathsf{F}$. This means that each Y_i is written as

$$Y_i = a_1 Y_{j1} + \cdots + a_k Y_{jk} \bmod \mathcal{C}$$

from eq.(3). Hence $\mathcal{B}$ pseudo-spans $\mathcal{Y}$. If $\mathcal{B}$ is not minimal, then we can show that a smaller subset of $\{E_{j1}, \cdots, E_{jk}\}$ is a basis of $\mathcal{E}$. This is a contradiction. Therefore, $\mathcal{B}$ is a minimal subset of $\mathcal{Y}$ which pseudo-spans $\mathcal{Y}$.

Q.E.D.

2.3 How to Find Pseudo-Basis

In this subsection, we show a polynomial time algorithm which finds the pseudo-dimension k and a pseudo-basis $\mathcal{B} = \{B_1, \cdots, B_k\}$ of $\mathcal{Y} = \{Y_1, \cdots, Y_m\}$. We begin with a definition of *linearly pseudo-express*.

Definition 3. *We say that Y is linearly pseudo-expressed by $\{B_1, \cdots, B_k\}$ if*

$$Y = a_1 B_1 + \cdots + a_k B_k \bmod \mathcal{C}$$

for some $a_1 \cdots, a_k \in \mathsf{F}$.

We first show in Fig.1 a polynomial time algorithm which checks if Y is linearly pseudo-expressed by $\{B_1, \cdots, B_k\}$. For a parameter $\alpha = (a_1 \cdots, a_k)$, define $X(\alpha)$ as

$$\begin{aligned} X(\alpha) &= Y - (a_1 B_1 + \cdots + a_k B_k) \\ &= (x_1(\alpha), \cdots, x_n(\alpha)). \end{aligned} \tag{5}$$

From the definition, Y is linearly pseudo-expressed by $\{B_1, \cdots, B_k\}$ if and only if there exists some α such that $X(\alpha) \in \mathcal{C}$. It is clear that $x_j(\alpha)$ is a linear expression of $(a_1 \cdots, a_k)$ from eq.(5). In Fig.1, it is also easy to see that each coefficient of $f_\alpha(x)$ is a linear expression of $(a_1 \cdots, a_k)$. Hence $f_\alpha(j) = x_j(\alpha)$ is a linear equation on $(a_1 \cdots, a_k)$ at step 3.

It is now clear that the algorithm of Fig.1 outputs YES if and only if $X(\alpha) \in \mathcal{C}$ for some α. Hence it outputs YES if and only if Y is linearly pseudo-expressed by $\{B_1, \cdots, B_k\}$.

Fig. 1. How to Check if Y is linearly pseudo-expressed by $\mathcal{B}$

Input: Y and $\mathcal{B} = \{B_1, \cdots, B_k\}$.

1. Construct $X(\alpha) = (x_1(\alpha), \cdots, x_n(\alpha))$ of eq.(5).
2. Construct a polynomial $f_\alpha(x)$ with $\deg f_\alpha(x) \le t$ such that
   $$f_\alpha(i) = x_i(\alpha)$$
   for $i = 1, \cdots, t+1$ by using Lagrange formula.
3. Output YES if the following set of linear equations has a solution α.
   $$\begin{aligned} f_\alpha(t+2) &= x_{t+2}(\alpha), \\ &\vdots \\ f_\alpha(n) &= x_n(\alpha). \end{aligned}$$
   Otherwise output NO.

We finally show in Fig.2 a polynomial time algorithm which finds the pseudo-dimension k and a pseudo-basis $\mathcal{B} = \{B_1, \cdots, B_k\}$ of $\mathcal{Y} = \{Y_1, \cdots, Y_m\}$. The correctness of the algorithm is guaranteed by Theorem 3.

Fig. 2. How to Find a Pseudo-Basis $\mathcal{B}$ of $\mathcal{Y}$

Input: $\mathcal{Y} = \{Y_1, \cdots, Y_m\}$.

1. Let $i = 1$ and $\mathcal{B} = \emptyset$.
2. While $i \leq m$ and $|\mathcal{B}| < t$, do:
(a) Check if Y_i is linearly pseudo-expressed by $\mathcal{B}$ by using Fig.1.
If NO, then add Y_i to $\mathcal{B}$.
(b) Let $i \leftarrow i + 1$.
3. Output $\mathcal{B}$ as a pseudo-basis and $k = |\mathcal{B}|$ as the pseudo-dimension.

2.4 Broadcast

We say that a sender (receiver) broadcasts x if it she sends x over all n channels. Since the adversary corrupts at most t out of $n = 2t + 1$ channels, the receiver (sender) receives x correctly from at least $t + 1$ channels. Therefore, the receiver (sender) can accept x correctly by taking the majority vote.

2.5 How to Apply to 3-Round PSMT

We now present an efficient 3-round PSMT for $n = 2t + 1$ in Fig.3.

Fig. 3. Our 3-round PSMT for $n = 2t + 1$

The sender wishes to send $\ell = nt$ secrets $s_1, \cdots, s_\ell \in \mathsf{F}$ to the receiver.

1. The sender sends a random codeword $X_i = (f_i(1), \cdots, f_i(n))$, and the receiver receives $Y_i = X_i + E_i$ for $i = 1, \cdots, \ell + t$, where $\deg f_i(x) \leq t$ and E_i is the error vector caused by the adversary.
2. The receiver finds a pseudo-basis $\mathcal{B} = \{Y_{j1}, \cdots, Y_{jk}\}$, where $k \leq t$, by using the algorithm of Fig.2. He then broadcasts $\mathcal{B}$ and $\Lambda_{\mathcal{B}} = \{j_1, \cdots, j_k\}$.
3. The sender constructs FORGED of eq.(4) from $\{E_j = Y_j - X_j \mid j \in \Lambda_{\mathcal{B}}\}$, encrypts $s_1, \cdots, s_\ell$ by using $\{f_i(0) \mid i \notin \Lambda_{\mathcal{B}}\}$ as the key of one-time pad, and then broadcasts FORGED and the ciphertexts.
4. The receiver reconstructs $f_i(x)$ by ignoring all channels of FORGED, and applying Lagrange formula to the remaining elements of Y_i. He then decrypts the ciphertexts by using $\{f_i(0) \mid i \notin \Lambda_{\mathcal{B}}\}$.

Further by combining the technique of [8, 1], we can construct a 2-round PSMT such that not only the transmission rate is $O(n)$, but also the computational cost of the sender and the receiver are both polynomial in n. The details will be given in the following sections.

3 Details of Our 3-Round PSMT

In this section, we describe the details of our 3-round PSMT for $n = 2t + 1$ which was outlined in Sec.2.5, and prove its security. We also show that the transmission rate is $O(n)$ and the computational cost of the sender and the receiver are both polynomial in n.

Remember that FORGED is the set of all channels which the adversary forged, and "broadcast" is defined in Sec.2.4.

3.1 3-round Protocol for $n = 2t + 1$

The sender wishes to send $\ell = nt$ secrets $s_1, \cdots, s_\ell \in \mathsf{F}$ to the receiver.

Step 1. The sender does the following for $i = 1, 2, \cdots, t + \ell$.

1. She chooses a polynomial $f_i(x)$ over F such that $\deg f_i(x) \leq t$ randomly.
   Let $X_i = (f_i(1), \cdots, f_i(n))$.
2. She send $f_i(j)$ through channel j for $j = 1, \cdots, n$.
   The receiver then receives $Y_i = X_i + E_i$, where E_i is the error vector caused by the adversary.

Step 2. The receiver does the following.

1. Find the pseudo-dimension k and a pseudo-basis $\mathcal{B} = \{Y_{j1}, \cdots, Y_{jk}\}$ of $\{Y_1, \cdots, Y_{t+\ell}\}$ by using the algorithm of Fig.2.
2. Broadcast k, $\mathcal{B}$ and $\Lambda_{\mathcal{B}} = \{j_1, \cdots, j_k\}$. where $\Lambda_{\mathcal{B}}$ is the set of indices of $\mathcal{B}$.

Step 3. The sender does the following.

1. Construct FORGED of eq.(4) from $\{E_j = Y_j - X_j \mid j \in \Lambda_{\mathcal{B}}\}$.
2. Compute $c_1 = s_1 + f_{i_1}(0), \cdots, c_\ell = s_\ell + f_{i_\ell}(0)$ for $i_1, \cdots, i_\ell \notin \Lambda_{\mathcal{B}}$.
3. Broadcast FORGED and $(c_1, \cdots, c_\ell)$.

Step 4. The receiver does the following. Let $Y_i = (y_{i1}, \cdots, y_{in})$.

1. For each $i \notin \Lambda_{\mathcal{B}}$, find a polynomial $f'_i(x)$ with $\deg f'_i(x) \leq t$ such that
   $$f'_i(j) = y_{i,j}$$
   for all $j \notin \mathsf{FORGED}$.
2. Compute $s'_1 = c_1 - f'_{i_1}(0), \cdots, s'_\ell = c_\ell - f'_{i_\ell}(0)$ for $i_1, \cdots, i_\ell \notin \Lambda_{\mathcal{B}}$.
3. Output $(s'_1, \cdots, s'_\ell)$.

3.2 Security

We first prove the perfect privacy. Consider $f_i(x)$ such that $i \notin \Lambda_{\mathcal{B}}$. For such i, Y_i is not broadcast at step 2-2. Hence the adversary observes at most t elements of $(f_i(1), \cdots, f_i(n))$. This means that she has no information on $f_i(0)$ because $\deg f_i(x) \leq t$. Therefore since $\{f_i(0) \mid i \notin \Lambda_{\mathcal{B}}\}$ is used as the key of one-time-pad, the adversary learns no information on $s_1, \cdots, s_\ell$.

We next prove the perfect reliability. We first show that there exist ℓ indices $i_1, i_2, \cdots, i_\ell$ such that

$$\{i_1, i_2, \cdots, i_\ell\} \subseteq \{1, 2, \cdots, t + \ell\} \setminus \Lambda_{\mathcal{B}}.$$

This is because

$$t + \ell - |\Lambda_{\mathcal{B}}| \geq t + \ell - t = \ell.$$

from Theorem 1. We next show that $f_i'(x) = f_i(x)$ for each $i \notin \Lambda_{\mathcal{B}}$ at Step 4. This is because

$$f_i'(j) = y_{i,j} = x_{i,j} = f_i(j)$$

for all $j \notin \mathsf{FORGED}$, and

$$n - |\mathsf{FORGED}| \geq 2t + 1 - t \geq t + 1.$$

Also note that $\deg f_i(x) \leq t$ and $\deg f_i'(x) \leq t$. Therefore $s_i' = s_i$ for $i = 1, \cdots, \ell$.

3.3 Efficiency

Let $|\mathsf{F}|$ denote the bit length of the field elements. Let $\mathsf{COM(i)}$ denote the communication complexity of Step i for $i = 1, 2, 3$. Then

$$\begin{aligned}
\mathsf{COM}(1) &= O(n(t + \ell))|\mathsf{F}|) = O(n\ell|\mathsf{F}|), \\
\mathsf{COM}(2) &= O(n^2 t|\mathsf{F}|) = O(n\ell|\mathsf{F}|), \\
\mathsf{COM}(3) &= O(n\ell|\mathsf{F}| + tn \log_2 n) = O(n\ell|\mathsf{F}|)
\end{aligned}$$

since $\ell = nt$. Hence the total communication complexity is $O(n\ell|\mathsf{F}|) = O(n^3|\mathsf{F}|)$. Further the sender sends ℓ secrets $s_1, \cdots, s_\ell \in \mathsf{F}$. Therefore, the transmission rate is $O(n)$ because

$$\frac{n\ell|\mathsf{F}|}{\ell|\mathsf{F}|} = n.$$

It is easy to see that the computational costs of the sender and the receiver are both polynomial in n.

4 Our Basic 2-Round PSMT

In this section, we show our basic 2-round PSMT for $n = 2t + 1$ such that the transmission rate is $O(n^2 t)$ and the computational costs of the sender and the receiver are both polynomial in n.

For two vectors $U = (u_1, \cdots, u_n)$ and $Y = (y_1, \cdots, y_n)$, define

$$d_u(U, Y) = \{u_j \mid u_j \neq y_j\}$$
$$d_I(U, Y) = \{j \mid u_j \neq y_j\}.$$

Remember that $\mathcal{C}$ is the set of all $(f(1), \cdots, f(n))$ such that $\deg f(x) \leq t$.

4.1 Randomness Extractor

Suppose that the adversary has no information on ℓ out of m random elements $r_1, \cdots, r_m \in \mathsf{F}$. In this case, let $R(x)$ be a polynomial with $\deg R(x) \leq m - 1$ such that $R(i) = r_i$ for $i = 1, \cdots, m$. Then it is well known [1, Sec.2.4] that the adversary has no information on

$$z_1 = R(m+1), \cdots, z_\ell = R(m+\ell).$$

4.2 Basic 2-round Protocol

The sender wishes to send a secret $s \in \mathsf{F}$ to the receiver.

Step 1. The receiver does the following for $i = 1, 2, \ldots, n$.

1. He chooses a random polynomial $f_i(x)$ such that $\deg f_i(x) \leq t$.
2. He sends
$$X_i = (f_i(1), \cdots, f_i(n))$$
through channel i, and the sender receives
$$U_i = (u_{i1}, \ldots, u_{in}).$$
3. Through each channel j, he sends $f_i(j)$ and the sender receives
$$y_{ij} = f_i(j) + e_{ij},$$
where e_{ij} is the error caused by the adversary. Let
$$Y_i = (y_{i1}, \cdots, y_{in}),\ E_i = (e_{i1}, \cdots, e_{in}).$$

Step 2. The sender does the following.

1. For $i = 1, \cdots, n$,
   (a) If $u_{ii} \neq y_{ii}$ or $|d_u(U_i, Y_i)| \geq t + 1$ or $U_i \notin \mathcal{C}$,
   then broadcast "ignore channel i". [4]
   This channel will be ignored from now on because it is forged clearly.
   (b) Else define r_i as
$$r_i = u_{ii} = y_{ii}. \tag{6}$$

[4] For simplicity, we assume that there are no such channels in what follows.

2. Find a polynomial $R(x)$ with $\deg R(x) \le n-1$ such that

$$R(i) = r_i$$

for each i.

3. Compute $R(n+1)$ and broadcast

$$c = s + R(n+1).$$

4. Find the pseudo-dimension k and a pseudo-basis $\mathcal{B} = \{Y_{j1}, \cdots, Y_{jk}\}$ of $\{Y_1, \cdots, Y_n\}$ by using the algorithm of Fig.2.
Broadcast $k, \mathcal{B}$ and $\Lambda_{\mathcal{B}} = \{j_1, \cdots, j_k\}$.
5. Broadcast $d_u(U_i, Y_i)$ and $d_I(U_i, Y_i)$ for each i.

Step 3. The receiver does the following.

1. Construct FORGED of eq.(4) from $\{E_i = Y_i - X_i \mid i \in \Lambda_{\mathcal{B}}\}$.
2. For each i, find a polynomial $u_i(x)$ with $\deg u_i(x) \le t$ such that

$$u_i(j) = u_{ij} \text{ for all } j \in d_I(U_i, Y_i),$$
$$u_i(j) = f_i(j) \text{ for all } j \text{ such that } j \notin d_I(U_i, Y_i) \text{ and } j \notin \mathsf{FORGED}$$

3. Find a polynomial $R'(x)$ with $\deg R'(x) \le n-1$ such that

$$R'(i) = u_i(i)$$

for each i. [5]

4. Compute $R'(n+1)$ and output

$$s' = c - R'(n+1).$$

4.3 Security

We first prove the perfect privacy.

Lemma 1. *There is at least one r_i on which the adversary has no information.*

Proof. Consider a non-corrupted channel i such that $i \notin \Lambda_{\mathcal{B}}$. First the sender does not broadcast r_i at step 2-4 because $i \notin \Lambda_{\mathcal{B}}$. Next because $f_i(i)$ is sent through channel i that the adversary does not corrupt, we have

$$r_i = u_{ii} = f_i(i).$$

Further the adversary observes at most t values of $(f_i(1), \cdots, f_i(n))$. Hence the adversary has no information on $r_i = f_i(i)$ because $\deg f_i(x) \le t$.

Finally there exists at least one non-corrupted channel i such that $i \notin \Lambda_{\mathcal{B}}$ because

$$n - t - |\Lambda_{\mathcal{B}}| \ge n - 2t = 1.$$

□

[5] "for each i" can be replaced by "for each $i \notin \Lambda_{\mathcal{B}}$" at step 2-2 and step 3-3.

Therefore, the adversary has no information on $R(n+1)$ from Sec.4.1. Hence she learns no information on s from $c = s + R(n+1)$.

We next prove the perfect reliability. If $j \notin \mathsf{FORGED}$ and $j \notin d_I(U_i, Y_i)$, then $f_i(j) = y_{ij} = u_{ij}$ from the definition of $d_I(U_i, Y_i)$. Therefore, at step 3-2,

$$u_i(j) = u_{ij}$$

for all $j \in d_I(U_i, Y_i)$, and for all j such that $j \notin d_I(U_i, Y_i)$ and $j \notin \mathsf{FORGED}$. This means that $u_i(j) = u_{ij}$ for each $j \in (\overline{\mathsf{FORGED}} \cup d_I(U_i, Y_i))$, where

$$|\overline{\mathsf{FORGED}} \cup d_I(U_i, Y_i))| \geq |\overline{\mathsf{FORGED}}| \geq n - t = (2t+1) - t = t + 1.$$

Further since $\deg u_i(x) \leq t$ and $U_i \in \mathcal{C}$, it holds that

$$(u_i(1), \cdots, u_i(n)) = (u_{i1}, \cdots, u_{in}).$$

In particular, $u_i(i) = u_{ii}$. Therefore from eq.(6), we have that

$$R(i) = r_i = u_{ii} = u_i(i) = R'(i)$$

for each i. Hence we obtain that $R'(x) = R(x)$ because $\deg R'(x) \leq n-1$ and $\deg R(x) \leq n - 1$. Consequently,

$$s' = c - R'(n+1) = c - R(n+1) = s.$$

Thus the receiver can compute $s' = s$ correctly.

4.4 Efficiency

Let $\mathsf{COM(i)}$ denote the communication complexity of Step i for $i = 1, 2$. Note that $|d_u(U_i, Y_i)| = |d_I(U_i, Y_i)| \leq t$ for each i. Then

$$\begin{aligned}
\mathsf{COM}(1) &= O(n(n+n)|\mathsf{F}|) = O(n^2|\mathsf{F}|), \\
\mathsf{COM}(2) &= O((|d_I(U_i, Y_i)| \log_2 n + |d_u(U_i, Y_i)||\mathsf{F}|)n^2 \\
&\quad +(\log_2 n + n|\mathcal{B}||\mathsf{F}| + |\Lambda_{\mathcal{B}}| \log_2 n)n + |\mathsf{F}|n) \\
&= O(tn^2 \log_2 n + tn^2|\mathsf{F}| + n \log_2 n + n^2 t|\mathsf{F}| + tn \log_2 n + |\mathsf{F}|n) \\
&= O(n^2 t|\mathsf{F}|)
\end{aligned}$$

because $|\mathcal{B}| = |\Lambda_{\mathcal{B}}| \leq t$. Hence the total communication complexity is $O(n^2 t|\mathsf{F}|)$. The transmission rate is $O(n^2 t)$ because the sender sends one secret.

It is easy to see that the computational cost of the sender and the receiver are polynomial in n.

5 More Efficient 2-Round Protocol

In our basic 2-round protocol, the transmission rate was $O(n^2 t)$. In this section, we reduce it to $O(n^2)$. We will use nt codewords $X_i \in \mathcal{C}$ to send t^2 secrets in this section while n codewords were used to send a single secret in the basic 2-round PSMT.

5.1 Protocol

The sender wishes to send $\ell = t^2$ secrets $s_1, s_2, \ldots, s_\ell \in \mathsf{F}$ to the receiver.

Step 1. The receiver does the following for each channel i.
For $h = 0, 1, \cdots, t-1$;

1. He chooses a random polynomial $f_{i+hn}(x)$ such that $\deg f_{i+hn}(x) \le t$.
2. He sends
$$X_{i+hn} = (f_{i+hn}(1), \cdots, f_{i+hn}(n))$$
through channel i, and the sender receives
$$U_{i+hn} = (u_{i+hn,1}, \cdots, u_{i+hn,n})$$
3. Through each channel j, he sends $f_{i+hn}(j)$ and the sender receives
$$y_{i+hn,j} = f_{i+hn}(j) + e_{i+hn,j},$$
where $e_{i+hn,j}$ is the error caused by the adversary. Let
$$Y_{i+hn} = (y_{i+hn,1}, \cdots, y_{i+hn,n}),\ E_{i+hn} = (e_{i+hn,1}, \cdots, e_{i+hn,n}).$$

Step 2. The sender does the following.

1. Find the pseudo-dimension k and a pseudo-basis $\mathcal{B} = \{Y_{j1}, \ldots, Y_{jk}\}$ of $\{Y_1, \cdots, Y_{tn}\}$ by using the algorithm of Fig.2.
Broadcast $k, \mathcal{B}$ and $\Lambda_{\mathcal{B}} = \{j_1, \cdots, j_k\}$.
2. For $i = 1, \cdots, n$,
   (a) If $u_{i+hn,i} \ne y_{i+hn,i}$ or $|d_u(U_{i+hn}, Y_{i+hn})| \ge k+1$ [6]
   or $U_{i+hn} \notin \mathcal{C}$ for some h, then broadcast "ignore channel i". [7]
   This channel will be ignored from now on because it is forged clearly.
   (b) Else define r_{i+hn} as
$$r_{i+hn} = u_{i+hn,i} = y_{i+hn,i} \tag{7}$$
   for $h = 0, \cdots, t-1$.
3. Find a polynomial $R(x)$ with $\deg R(x) \le nt - 1$ such that
$$R(i + hn) = r_{i+hn}$$
for each $i + hn$.
4. Compute $R(nt+1), \cdots, R(nt+\ell)$ and broadcast
$$c_1 = s_1 + R(nt+1), \cdots, c_\ell = s_\ell + R(nt+\ell).$$
5. Broadcast $d_u(U_{i+hn}, Y_{i+hn})$ and $d_I(U_{i+hn}, Y_{i+hn})$ for each $i + hn$.

[6] k is the number of channels that the adversary forged on $\{Y_{i+hn}\}$.
[7] For simplicity, we assume that there are no such channels in what follows.

Step 3. The receiver does the following.

1. Construct FORGED of eq.(4) from $\{E_i = Y_i - X_i \mid i \in \Lambda_{\mathcal{B}}\}$.
2. For each $i+hn$, find a polynomial $u_{i+hn}(x)$ with $\deg u_{i+hn}(x) \le t$ such that

   $u_{i+hn}(j) = u_{i+hn,j}$ for all $j \in d_I(U_{i+hn}, Y_{i+hn})$
   $u_{i+hn}(j) = f_{i+hn}(j)$ for all j such that $j \notin d_I(U_{i+hn}, Y_{i+hn})$ and $j \notin \mathsf{FORGED}$

3. Find a polynomial $R'(x)$ with $\deg R'(x) \le nt-1$ such that

$$R'(i+hn) = u_{i+hn}(i)$$

   for each $i+hn$. [8]
4. Compute $R'(nt+1), \cdots, R'(nt+\ell)$ and output

$$s'_1 = c_1 - R'(nt+1), \cdots, s'_\ell = c_\ell - R'(nt+\ell).$$

5.2 Security

We first prove the perfect privacy.

Lemma 2. *There exists a subset $A \subset \{r_1, \cdots, r_{tn}\}$ such that $|A| \ge \ell$ and the adversary has no information on A.*

Proof. Consider a non-corrupted channel i such that $i+hn \notin \Lambda_{\mathcal{B}}$. First the sender does not broadcast r_{i+hn} at step 2-1 because $i+hn \notin \Lambda_{\mathcal{B}}$. Next since $f_{i+hn}(i)$ is sent through channel i that the adversary does not corrupt, we have

$$r_{i+hn} = u_{i+hn,i} = f_{i+hn}(i).$$

Further the adversary observes at most t values of $(f_{i+hn}(1), \cdots, f_{i+hn}(n))$. Hence the adversary has no information on $r_{i+hn} = f_{i+hn}(i)$ because $\deg f_{i+hn}(x) \le t$.

Note that the adversary corrupts at most t channels and for each corrupted channel i, the adversary gets $r_i, r_{i+n}, \ldots, r_{i+(t-1)n}$. Therefore, there exists a subset $A \subset \{r_1, \cdots, r_{tn}\}$ such that

$$|A| \ge nt - |\Lambda_{\mathcal{B}}| - t^2 = nt - k - t^2$$

and the adversary has no information on A. Finally

$$nt - k - t^2 \ge (2t+1)t - t - t^2 = t^2 = \ell.$$

□

[8] "for each $i+hn$" can be replaced by "for each $i+hn \notin \Lambda_{\mathcal{B}}$" at step 2-3 and step 3-3.

Therefore, the adversary has no information on $R(nt+1)$, …, $R(nt+\ell)$ from Sec.4.1. Hence she learns no information on s_i for $i=1,\cdots,\ell$.

We next prove the perfect reliability. If $j \notin \mathsf{FORGED}$ and $j \notin d_I(U_{i+hn}, Y_{i+hn})$, then $f_{i+hn}(j) = y_{i+hn,j} = u_{i+hn,j}$ from the definition of $d_I(U_{i+hn}, Y_{i+hn})$. Therefore,

$$u_{i+hn}(j) = u_{i+hn,j}$$

for all $j \in d_I(U_{i+hn}, Y_{i+hn})$, and for all j such that $j \notin d_I(U_{i+hn}, Y_{i+hn})$ and $j \notin \mathsf{FORGED}$. This means that $u_{i+hn}(j) = u_{i+hn,j}$ for each $j \in (\overline{\mathsf{FORGED}} \cup d_I(U_{i+hn}, Y_{i+hn}))$, where

$$|\overline{\mathsf{FORGED}} \cup d_I(U_{i+hn}, Y_{i+hn}))| \geq |\overline{\mathsf{FORGED}}| \geq n-t = 2t+1-t = t+1.$$

Further since deg $u_{i+hn}(x) \leq t$ and $U_{i+hn} \in \mathcal{C}$, it holds that

$$(u_{i+hn}(1), \cdots, u_{i+hn}(n)) = (u_{i+hn,1}, \cdots, u_{i+hn,n}).$$

In particular, $u_{i+hn}(i) = u_{i+hn,i}$. Therefore from eq.(7), we have that

$$R(i+hn) = r_{i+hn} = u_{i+hn,i} = u_{i+hn}(i) = R'(i+hn)$$

for each $i+hn$. Hence we obtain that $R'(x) = R(x)$ because $\deg R'(x) \leq nt-1$ and $\deg R(x) \leq nt-1$. Consequently,

$$s'_i = c_i - R'(nt+i) = c_i - R(nt+i) = s_i.$$

Thus the receiver can compute $s'_i = s_i$ correctly for $i=1,\cdots,\ell$.

5.3 Efficiency

Let $\mathsf{COM(i)}$ denote the communication complexity of Step i for $i=1,2$. Note that $|d_u(U_{i+hn}, Y_{i+hn})| = |d_I(U_{i+hn}, Y_{i+hn})| \leq t$ for each $i+hn$. Then

$$\begin{aligned}
\mathsf{COM}(1) &= O(tn(n+n)|\mathsf{F}|) = O(tn^2|\mathsf{F}|),\\
\mathsf{COM}(2) &= O((|d_I(U_{i+hn}, Y_{i+hn})| \log_2 n + |d_u(U_{i+hn}, Y_{i+hn})||\mathsf{F}|)tn \times n\\
&\qquad + (\log_2 n + n|\mathcal{B}||\mathsf{F}| + |\Lambda_{\mathcal{B}}| \log_2 n)n + t^2|\mathsf{F}|n)\\
&= O(n^2t^2 \log_2 n + n^2t^2|\mathsf{F}| + n\log_2 n + n^2t|\mathsf{F}| + tn\log_2 n + t^2|\mathsf{F}|n)\\
&= O(n^2t^2|\mathsf{F}|)
\end{aligned}$$

because $|\mathcal{B}| = |\Lambda_{\mathcal{B}}| \leq t$. Hence, the total communication complexity is $O(n^2t^2|\mathsf{F}|)$, and the transmission rate is $O(n^2)$ because the sender sends t^2 secrets.

It is easy to see that the computational costs of the sender and the receiver are both polynomial in n.

6 Final 2-Round PSMT

The transmission rate is still $O(n^2)$ in the 2-round PSMT shown in Sec.5. In this section, we show how to reduce it to $O(n)$ by using the technique of [1, page 406] and [8]. Then we can obtain the first 2-round PSMT for $n=2t+1$ such that not only the transmission rate is $O(n)$ but also the computational costs of the sender and the receiver are both polynomial in n.

6.1 Generalized Broadcast

Suppose that the receiver knows the locations of k ($\leq t$) channels that the adversary forged. For example, suppose that the receiver knows that channels $1, 2, \cdots, k$ are forged by the adversary. Then the adversary can corrupt at most $t-k$ channels among the remaining $n-k$ channels $k+1, \cdots, n$. In this case, it is well known that the sender can send $k+1$ field elements $u_1, u_2, \ldots, u_{k+1}$ reliably with the communication complexity $O(n|\mathsf{F}|)$ as shown below.

1. The sender finds a polynomial $p(x)$ with $\deg p(x) \leq k$ such that $p(1) = u_1$, $p(2) = u_2, \ldots, p(k+1) = u_{k+1}$.
2. She sends $p(i)$ through channel i for $i = 1, \cdots, n$.

Without loss of generality, suppose that the receiver knows that channels $1, \cdots, k$ are forged by the adversary. Then he consider a shortened code such that a codeword is $(p(k+1), \cdots, p(n))$. The minimum Hamming distance of this code is $(n-k)-k = 2t+1-2k = 2(t-k)+1$. Hence the receiver can correct the remaining $t-k$ errors.

This means that the receiver can decode $(p(k+1), \cdots, p(n))$ correctly. Then he can reconstruct $p(x)$ by using Lagrange formula because

$$n - k = 2t + 1 - k \geq 2k + 1 - k = k + 1 \geq \deg p(x) + 1.$$

Therefore he can obtain $u_1 = p(1), \ldots, u_{k+1} = p(k+1)$ correctly.

6.2 How to Improve Step 2-5

Step 2-5 is the most expensive part in the 2-round PSMT shown in Sec.5. In this subsection, we will show a method which reduces the communication complexity of step 2-5 from $O(n^2t^2|\mathsf{F}|)$ to $O(n^2t|\mathsf{F}|)$.

At step 2-5, the sender *broadcasts*

$$d_u(U_{i+hn}, Y_{i+hn}) \text{ and } d_I(U_{i+hn}, Y_{i+hn})$$

for each $i + hn$. Note that the size of all $d_u(U_{i+hn}, Y_{i+hn})$ is bounded by

$$\sum_{i=1}^{n} \sum_{h=0}^{t-1} |d_u(U_{i+hn}, Y_{i+hn})| \leq knt \tag{8}$$

because $|d_u(U_{i+hn}, Y_{i+hn})| \leq k$ from Step 2-2(a), where

$$k = |\mathcal{B}| = |\mathsf{FORGED}|$$

is the number of channels that the adversary forged on $\{Y_{i+hn}\}$. On the other hand, the following lemma holds.

Lemma 3. *The sender can send $k+1$ field elements reliably with the communication complexity $O(n|\mathsf{F}|)$ at step 2-5.*

Proof. The sender knows the value of k because she computes $\mathcal{B}$. The receiver knows the locations of k forged channels because he computes FORGED. Therefore, we can use the generalized broadcasting technique shown in Sec.6.1. □

Now from eq.(8) and Lemma 3, it is easy to see that the communication complexity of step 2-5 can be reduced to $O(n^2t|\mathsf{F}|)$.

6.3 Final Efficiency

Consequently, we obtain $\mathsf{COM}(2) = O(n^2t|\mathsf{F}|)$ because the communication complexity of step 2-5 is now reduced to $O(n^2t|\mathsf{F}|)$. On the other hand, $\mathsf{COM}(1) = O(n^2t|\mathsf{F}|)$ from Sec.5.3. To summarize,

$$\mathsf{COM}(1) = O(n^2t|\mathsf{F}|) \text{ and } \mathsf{COM}(2) = O(n^2t|\mathsf{F}|)$$

in our final 2-round PSMT. Hence, the total communication complexity is $O(n^3|\mathsf{F}|)$ because $n = 2t + 1$.

Now the transmission rate is $O(n)$ because the sender sends t^2 secrets which is $O(n^2|\mathsf{F}|)$. Finally, it is easy to see that the computational costs of the sender and the receiver are both polynomial in n.

References

1. S.Agarwal, R.Cramer and R.de Haan: Asymptotically Optimal Two-Round Perfectly Secure Message Transmission. CRYPTO 2006: pp.394–408 (2006)
2. D.Dolev, C.Dwork, O.Waarts, M.Yung: Perfectly Secure Message Transmission. J. ACM 40(1): pp.17–47 (1993)
3. Y.Desmedt, Y.Wang and M.Burmester: A Complete Characterization of Tolerable Adversary Structures for Secure Point-to-Point Transmissions Without Feedback. ISAAC 2005: pp.277–287 (2005)
4. M.Fitzi, M.Franklin, J.Garay, S.Vardhan: Towards Optimal and Efficient Perfectly Secure Message Transmission. TCC 2007: pp.311–322 (2007)
5. M.Hirt, U.Maurer: Player Simulation and General Adversary Structures in Perfect Multiparty Computation. J. Cryptology 13(1): pp.31–60 (2000)
6. M.V.N.A.Kumar, P.R.Goundan, K.Srinathan, C.P.Rangan: On perfectly secure communication over arbitrary networks. PODC 2002: pp.193–202 (2002)
7. H.Md.Sayeed and H.Abu-Amara: Efficient Perfectly Secure Message Transmission in Synchronous Networks. Inf. Comput. 126(1): pp.53–61 (1996)
8. K. Srinathan, Arvind Narayanan, C. Pandu Rangan: Optimal Perfectly Secure Message Transmission. CRYPTO 2004: pp.545–561 (2004)