Public Key Cryptosystems Constructed Based on Reed-Solomon Codes, K(XV)SE(2)PKC, Realizing Coding Rate of Exactly 1.0

Masao KASAHARA
Faculty of Informatics, Osaka Gakuin University, Suita-shi, 564-8511 Japan.
kasahara@ogu.ac.jp

Abstract
In this paper, we present a new class of public-key cryptosystems, K(XV)SE(2)PKC realizing the coding rate of exactly 1.0, based on Reed-Solomon codes (RS codes). We show that K(XV)SE(2)PKC is secure against the various attacks including the attacks based on the Gröbner basis calculation (Gröbner basis attack, GB attack) and a linear transformation attack.

Keyword
Public key cryptosystem, PQC, Reed-Solomon code, Code based PKC, Multivariate PKC, Gröbner basis.

1 Introduction
Most of the multivariate PKC are constructed by the simultaneous equations of degree larger than or equal to 2 [1]∼[14]. All these proposed schemes are very interesting and important. However unfortunately, some of these schemes have been proved not necessarily secure against the conventional attacks such as Patarin’s attack[3], the attack based on the Gröbner basis calculation (GB Attack)[11-13] and Braeken-Wolf-Preneel (BWP) attack[14].

The author recently proposed several classes of multivariate PKC’s that are constructed by many sets of linear equations[15-20]. It should be noted that McEliece PKC[21] presented in 1978 can be regarded as a member of the class of linear multivariate PKC.

In 2011 the author presented a new class of public key cryptosystem, K(XV)SE(2)PKC based on error-correcting codes, realizing the coding rate of exactly 1.0. K(XV)SE(2)PKC is constructed on the basis of K(X)SE(1)PKC[23] which is not secure against a linear transformation attack[24]. We show that K(XV)SE(2)PKC is secure against the attacks based on a linear transformation attack and the GB Attack[11-13].

Throughout this paper, when the variable \(v_i\) takes on a value \(\tilde{v}_i\), we shall denote the corresponding vector \(\tilde{v} = (\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_n)\) as

\[
\tilde{v} = (\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_n).
\] (1)

The vector \(v = (v_1, v_2, \cdots, v_n)\) will be represented by the polynomial as

\[
v(x) = v_1 + v_2 x + \cdots + v_n x^{n-1}.
\] (2)

The \(\tilde{u}, \tilde{u}(x)\) et al. will be similarly defined.

2 k(XV)SE(1)PKC
In this section we present a simple version of K(XV)SE(2)PKC referred to as k(XV)SE(2)PKC. Generalization of k(XV)SE(2)PKC to K(XV)SE(2)PKC is straightforward.

2.1 Preliminaries
Let us define several symbols:

- \(g_i(x)\): Generator polynomial of RS code of degree 2 over \(F_{2^m}; i = 1, 2\).
- \(L_i\): Location, \(L_i \geq 2, i = 1, 2, L_1 \neq L_2\).
- \(m_i\): Single error whose error value is 1 that occurred at the location \(L_i; i = 1, 2\).
- \(P_C[g_i(x)]\): Probability that \(g_i(x)\) is estimated correctly, \(i = 1, 2\).
- \(P_C[L_i]\): Probability that \(L_i\) is estimated correctly, \(i = 1, 2\).
Let the message vector $A$ over $\mathbb{F}_2$ be represented by
$$A = (A_1, A_2, \ldots, A_N).$$

Throughout this paper we assume that the messages $A_1, A_2, \ldots, A_N$ are mutually independent and equally likely.

Let $A$ be transformed into
$$A \cdot H_1 = a$$
$$= (a_1, a_2, \ldots, a_N),$$
where $H_1$ is an $N \times N$ non-singular random matrix over $\mathbb{F}_2$.

Let $a$ be partitioned into
$$a = (m_1, m_2, m_3, m_4),$$
where $m_1$'s are $m$-tuples over $\mathbb{F}_2$.

Let us regard $m_1$ as an element of $\mathbb{F}_{2^m}$.

Let $m_1$ and $m_2$ be represented by
$$m_1 = (a_{11}, \ldots, a_{1m})$$
and
$$m_2 = (a_{21}, \ldots, a_{2m}).$$

### 2.2 Construction

**Figure 1**: Schematic diagram of principle of k(XV)SE(2)PKC

- **$P_{C}[g_i(x) \cap L_i]$**: Probability that both $g_i(x)$ and $L_i(x)$ are estimated correctly, $i = 1, 2$.

- **$H_i$**: Random matrix whose component takes on 0 or 1 equally likely, $i = 1, 2$.

- **$g_F(x)$**: Random primitive polynomial of degree $2m$ whose coefficients except those of $x^m$ and $x^{2m}$ take on 0 or 1 equally likely.

- **$G_i(x)$**: Generator polynomial of RS code of degree $g$ over $\mathbb{F}_{2^m}$; $i = 1, 2, \ldots, E$.

Let $r_1(x)$ be obtained by
$$m_2 x^{r_1} \equiv r_1(x) = m_2 r_2 x + m_2 r_1 \mod g_1(x),$$
yielding a code word, $F_1(x)$:
$$F_1(x) = m_2 x^{r_1} + m_2 r_2 x + m_2 r_1 \equiv 0 \mod g_1(x).$$

Let $m_3$ and $m_4$ be represented by
$$m_3 = (a_{31}, \ldots, a_{3m})$$
and
$$m_4 = (a_{41}, \ldots, a_{4m})$$
respectively.

Let $r_2(x)$ be given by
$$m_4 x^{r_2} \equiv r_2(x) = m_4 r_4 x + m_4 r_3 \mod g_2(x),$$
yielding a code word, $F_2(x)$:
$$F_2(x) = m_4 x^{r_2} + m_4 r_4 x + m_4 r_3 \equiv 0 \mod g_2(x).$$

Regarding $(m_2 r_2, m_4 r_4)$ as a $2m$-tuple over $\mathbb{F}_2$, it is transformed into
$$(m_2 r_2, m_4 r_4) H_2 = s = (s_1, s_2, \ldots, s_{2m}),$$
where $H_2$ is a $2m \times 2m$ non-singular random matrix over $\mathbb{F}_2$.

**Remark 1**: The component $s_i; i = 1, 2, \ldots, 2m$, is a linear equation in the variables $A_1, A_2, \ldots, A_N$ over $\mathbb{F}_2$. □
Regarding \((m_2r_1, m_4r_3)\) as a 2m-tuple over \(F_2\), it is transformed into
\[
(m_2r_1, m_4r_3)P_t = t
= (t_1, t_2, \ldots, t_{2m}),
\]
where \(P_t\) is a 2m \(\times\) 2m random permutation matrix.

Letting \(t(x)\) be represented by \(t_1 + t_2x + \cdots + t_{2m}x^{2m-1}\), it is transformed into
\[
\{t(x)\}^2 \equiv \gamma(x)
\equiv \gamma_1 + \gamma_2x + \cdots + \gamma_{2m}x^{2m-1} \text{ mod } g_F(x),
\]
where \(g_F(x)\) is a random primitive polynomial of degree 2m over \(F_2\).

Let \(\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_{2m})\) be transformed into
\[
\gamma H_1 = \delta
= (\delta_1, \delta_2, \ldots, \delta_{2m}),
\]
where \(H_1\) is a 2m \(\times\) 2m matrix over \(F_2\) which is not necessarily non-singular.

Let \(\delta\) be represented by
\[
\delta = (\delta_1, \delta_2, \ldots, \delta_{2m}),
\]
where \(\delta_1\) and \(\delta_2\) over \(F_2\) are defined by
\[
\delta_1 = (\delta_1, \delta_2, \ldots, \delta_m)
\]
and
\[
\delta_2 = (\delta_{m+1}, \delta_{m+2}, \ldots, \delta_{2m})
\]
respectively.

It should be noted that the component of \(\delta_i\) : \(i = 1, 2, \ldots, 2m\), is a quadratic equation in the variables \(A_1, A_2, \ldots, A_N\) over \(F_2\).

Let \(\delta_1'\) and \(\delta_2'\) be defined by
\[
\delta P_t^{-1} = (\delta_1', \delta_2').
\]

Regarding \((m_1, m_3)\) over \(F_2\), as an element of \(F_{2^m}\), \((m_1, m_3)\) is transformed into
\[
(m_1, m_3)\theta = \omega
= (\omega_1, \omega_2, \ldots, \omega_{2m}) \in F_{2^m},
\]
where \(\theta\) is given by
\[
\theta = 2^0 + 2^1 + \cdots + 2^{2m-1} < 2^{2m} - 1.
\]

Let \(\omega_1, \omega_2, \ldots, \omega_{2m}\) be defined by
\[
\omega = (\omega_1, \omega_2, \ldots, \omega_{2m})
\]
and
\[
\omega_1 = (\omega_{m+1}, \omega_{m+2}, \ldots, \omega_{2m}).
\]

Let \(\omega_1'\) and \(\omega_2'\) be defined by
\[
\omega P_t^{-1} = (\omega_1', \omega_2').
\]

At the sending end, after calculating \(\omega\) by Eq.(22), \(t + \delta + \omega\) is obtained. The ciphertext \(C\) is given by
\[
C = (s, t + \delta + \omega).
\]

We have the following set of keys.

| Public key | \(m_1, m_3, \{s_i\}, \{t_i + \delta_i\}, \theta, F_{2^m}, F_{2^m}^r, r_1, r_2, r_3, r_4, g_1(x), g_2(x), g_F(x)\) |
| Secret key | \(H_1, H_2, H_3, P_1, L_1, L_2, r_1, r_2, r_3, r_4\) |

### 2.3 Encryption and Decryption

**Encryption:**

**Step 1:** The \(\tilde{s}\) is calculated from the public keys whose component is represented by the variables \(A_1, A_2, \ldots, A_N\).

**Step 2:** The \(\tilde{t} + \tilde{\delta}\) is calculated from the public key \(t + \delta\) whose component is represented by the variables \(A_1, A_2, \ldots, A_N\).

**Step 3:** The \(\tilde{\omega}\) is calculated from \(\tilde{m}_1\) and \(\tilde{m}_3\) by Eq.(22).

**Step 4:** The ciphertext \(\tilde{C}\) is given by
\[
\tilde{C} = (\tilde{s}, \tilde{t} + \tilde{\delta} + \tilde{\omega}).
\]

**Decryption:**

**Step 1:** The \((\tilde{m}_2r_2, \tilde{m}_4r_4)\) is decoded by \((s_1, s_2, \ldots, s_{2m})H_2^{-1}\).

**Step 2:** The \((\tilde{m}_2r_1 + \tilde{\delta}_1', \tilde{m}_4r_3 + \tilde{\delta}_2', \tilde{\omega}_1, \tilde{\omega}_2)\) is decoded by \((\tilde{t} + \tilde{\delta} + \tilde{\omega})P_t^{-1}\).

**Step 3:** From \((\tilde{m}_2r_2, \tilde{m}_2r_1 + \tilde{\delta}_1' + \tilde{\omega}_1')\) and \((\tilde{m}_4r_4, \tilde{m}_4r_3 + \tilde{\delta}_2' + \tilde{\omega}_2')\), the sets \((\tilde{m}_2r_2, \tilde{\delta}_1' + \tilde{\omega}_1')\) and \((\tilde{m}_4r_4, \tilde{\delta}_2' + \tilde{\omega}_2')\) are decoded, for example, by Euclidean decoding[25], yielding \((\tilde{m}_2, \tilde{\delta}_1' + \tilde{\omega}_1')\) and \((\tilde{m}_4, \tilde{\delta}_2' + \tilde{\omega}_2')\).

**Step 4:** From \(\tilde{m}_2\) and \(\tilde{m}_4\), the \((\tilde{\delta}_1', \tilde{\delta}_2')\) is decoded by Eqs.(15) \sim (21), yielding \(\tilde{\omega}_1'\) and \(\tilde{\omega}_2'\).

**Step 5:** The \(\tilde{\omega}\) is decoded by \((\tilde{\omega}_1', \tilde{\omega}_2')P_t\).

**Step 6:** The \((\tilde{m}_1, \tilde{m}_3)\) an element of \(F_{2^m}\), is decoded by \(\tilde{\omega}^{1/\theta}\).

**Step 7:** The original message \(\tilde{A}\) is decoded by \((\tilde{m}_1, \tilde{m}_2, \tilde{m}_3, \tilde{m}_4)H_1^{-1} = \tilde{A} = (\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_N)\).
Let us define several symbols:

- \( N_V \): Total number of message variables, \( N = 4m \).
- \( N_{ESI} \): Total number of 2m linear equations representing the components of \( s, s_1, s_2, \ldots, s_{2m} \).
- \( N_{ESQ} \): Total number of 2m quadratic equations representing the components of \( t + \delta \).
- \( \text{SE}(s) \): Set of linear equations related to \( s \) in the variables \( A_1, A_2, \ldots, A_N \).
- \( \text{SE}(t + \delta) \): Set of quadratic equations related to \( t + \delta \) in the variables \( A_1, A_2, \ldots, A_N \).
- \( \text{SE}(m_i) \): Set of linear equations related to \( m_1 \) or \( m_3 \) in the variables \( A_1, A_2, \ldots, A_N \).

The sizes of \( \text{SE}(s), \text{SE}(t + \delta), \text{SE}(m_1) \) and \( \text{SE}(m_3) \) are given by

\[
|\text{SE}(s)| = N_V \cdot N_{ESI}, \quad |\text{SE}(t + \delta)| = N_V \cdot H_2 \cdot N_{ESQ},
\]

and

\[
|\text{SE}(m_1)| = |\text{SE}(m_3)| = N_V \cdot m,
\]

respectively.

The size of the public key for \( m_1, m_3, \{s\} \) and \( \{t + \delta\} \), \( S_{PK} \) is given by

\[
S_{PK} = N_V \cdot H_2 \cdot 2m + 4N_V \cdot m.
\]

The coding rate \( \rho \) is given by

\[
\rho = \frac{|M|}{|C|} = \frac{|m_1| + |m_2| + |m_3| + |m_4|}{|m_2r_2| + |m_4r_4| + |m_2r_1 + \omega \lambda + \delta_k| + |m_4r_3 + \omega \mu + \delta_l|} = 1.0.
\]

We see that the coding rate is given by exactly 1.0.

The probability that the location is estimated correctly, \( P_C[\hat{L}_i] \), is given by

\[
P_C[\hat{L}_i] = (2^m - 3)^{-1} \approx 2^{-m}; i = 1, 2.
\]

The probability that \( g_i(x) \) is estimated correctly, \( P_C[\hat{g}_i(x)] \), is given by

\[
P_C[\hat{g}_i(x)] = \left\{2^{m(g-2)} + (2^m - 1)\right\}^{-1} \approx 2^{-m(g-1)}; i = 1, 2.
\]

### 2.5 Examples

In Table 1, we show three examples of k(XV)SE(2)PKC.

### 2.6 Security Consideration

**Attack 1:** Attack on \( L_i \) and \( g_i(x) ; i = 1, 2 \).

The probability that \( g_i(x) \)'s and \( L_i \)'s are estimated correctly is given by

\[
\{P_C[\hat{g}_i(x)] : P_C[\hat{L}_i] \}^2; i = 1, 2,
\]

sufficiently small value for \( m \gtrsim 20 \). We conclude that k(XV)SE(2)PKC is secure against Attack 1.

**Attack 2:** GB Attack on ciphertext

The components of \( \omega \) over \( F_{2^m} \) can be represented by the set of equations over \( F_2 \) of very high degree. Sets of the components of \( (s, t + \delta + \omega) \) yield a set of 2m equations of degree \( \eta + 1 \) in the variables \( A_1, A_2, \ldots, A_N \) and a set of 2m linear equations in the variables \( A_1, A_2, \ldots, A_N \). The degrees take on a very high value of at least 63 as we see in the examples in Table 1. We thus conclude that our proposed scheme, k(XV)SE(2)PKC, can be secure against GB Attack.

**Attack 3:** Exhaustive attack on \( (\tilde{m}_1, \tilde{m}_3) \).

We see that when \( (\tilde{m}_1, \tilde{m}_3) \) is estimated correctly by an exhaustive manner, the ciphertext can be disclosed by the GB attack on the set of 2m quadratic equations and 2m linear equations. It is easy to see that the average number of times required for estimating \( (\tilde{m}_1, \tilde{m}_3) \) in an exhaustive manner is given by \( 2^{2m-1} \).

In order to be secure against this attack, 2m, the size of \( (\tilde{m}_1, \tilde{m}_3) \) is recommended to be sufficiently large \( (m \gtrsim 30) \).

**Attack 4:** Attack on \( (m_2r_1, m_4r_3) \) from \( s \).

It is apparent that \( (m_2r_1, m_4r_3) \) can be disclosed from \( (m_2r_2, m_4r_4) \) by a linear transformation.

---

| Example | | \(|A| N(\text{bit})\) | \(|m_1| m(\text{bit})\) | \(\eta + 1 = 2m - 1\) | \(P_C[\hat{g}_i(x) \cap L_i] \) | \(S_{PK}(\text{KB})\) | \(\rho\) |
|---------|---|----------------|----------------|----------------|---------------------------|----------------|-------|
| I       |   | 128            | 32             | 63             | 2.94 \times 10^{-39}     | 68.1           | 1.0   |
| II      |   | 160            | 40             | 79             | 6.84 \times 10^{-39}     | 132            | 1.0   |
| III     |   | 192            | 48             | 95             | 1.59 \times 10^{-38}     | 227            | 1.0   |

---

Table 1: Example of k(XV)SE(2)PKC (\(\rho = 1.0\)).
3 K(XV)SE(2)PKC

In this section we present a generalized version of k(XV)SE(2)PKC, referred to as K(XV)SE(2)PKC. As K(XV)SE(2)PKC is a straightforward generalization of k(XV)SE(2), we shall only present an outline of the construction of K(XV)SE(2)PKC.

3.1 Construction

Let $m_i$ be

$$m_i = (m_{i1}, \ldots, m_{im}) \quad ; i = 1, 2, \ldots, E. \quad (37)$$

In the followings let us regard $m_i$'s as the elements of $\mathbb{F}_{2^m}$. Let

$$m_1x^{L_1} \equiv r_1(x) = m_1r_{i1} + m_1r_{i2}x + \cdots + m_1r_{ig}x^{g-1} \mod G_1(X),$$

$$m_2x^{L_2} \equiv r_2(x) = m_2r_{i1} + m_2r_{i2}x + \cdots + m_2r_{ig}x^{g-1} \mod G_2(X),$$

$$\vdots$$

$$m_Ex^{L_E} \equiv r_{E}(x) = m_Er_{E1} + m_Er_{E2}x + \cdots + m_Er_{Eg}x^{g-1} \mod G_E(X), \quad (38)$$

where all the locations $L_i$'s are distinct and satisfy

$$g \leq L_i \leq 2^m - 2 \quad ; i = 1, 2, \ldots, E. \quad (39)$$

We also assume that the degree of $G_i(x); i = 1, 2, \ldots, E$, is given by $g$.

Let the remainder $r_i(x) = m_1r_{i1} + m_1r_{i2}x + \cdots + m_1r_{ig}x^{g-1}$ given by Eq.(38) be partitioned into

$$r_i(x) = r_{IE}(x) + r_{IH}(x), \quad ; i = 1, 2, \ldots, E. \quad (40)$$

where $r_{IE}(x)$ and $r_{IH}(x)$ are given by

$$r_{IE}(x) = m_ir_{i1}H_{i1} + m_ir_{i1}H_{i2} + \cdots + m_ir_{ig}x^{g-1}$$

and

$$r_{IH}(x) = m_ir_{i1} + m_ir_{i2}x + \cdots + m_ir_{iH}x^{H-1} \quad (41)$$

$$\sum_{i=1}^{E} m_i x^{L_i} = \sum_{i=1}^{E} m_i r_{iH1} + \sum_{i=1}^{E} m_i r_{iH2} + \cdots + \sum_{i=1}^{E} m_i r_{ig}x^{g-1}. \quad (44)$$

Let the positive integers $E$ and $H$ satisfy

$$E + H = g. \quad (43)$$

respectively.

Given $m_1, m_2, \ldots, m_E$, we construct

$$V_E \cdot H_4 = s = (V_{H+1}, V_{H+2}, \ldots, V_g), \quad (45)$$

where $H_4$ is a random $E \times E$ non-singular matrix over $\mathbb{F}_{2^m}$. The vector $V_H = (\sum_{i=1}^{E} m_ir_{i1}, \cdots, \sum_{i=1}^{E} m_ir_{ig})$ is transformed into

$$V_H \cdot P_2 = T = (V_1, V_2, \ldots, V_H), \quad (46)$$

where $P_2$ is a $H \times H$ random permutation matrix over $\mathbb{F}_{2^m}$. The $T$ is transformed into

$$T^3 = \Gamma \quad (47)$$

in a similar manner as Eq.(16).

Messages $m_{E+1}, m_{E+2}, \ldots, m_N$, are publicized as

$$V_P = (m_{E+1}, m_{E+2}, \cdots, m_N), \quad (48)$$

In a similar manner as Eq.(22), $\Omega$ is given by

$$V^\theta_P = \Omega \quad (49)$$

The correspondence among the parameters for k(XV)SE(2)PKC and K(XV)SE(2)PKC is shown below:

$$\begin{array}{ll}
\text{k(XV)SE(2)PKC} & \text{K(XV)SE(2)PKC} \\
(m_1, m_3) & (m_{E}, m_{H}) \\
(m_1, m_3) & (m_{E}, m_{H})^\theta \\
(m_2r_{2}, m_4r_{4})H_2 & s & V_{E}H_4 = S \\
(m_2r_{1}, m_4r_{4})P_1 & t & V_{H}P_2 = T \\
\gamma & \delta & \Gamma = \Delta + \Omega \\
\end{array}$$

$$\begin{array}{ll}
k(XV)SE(2)PKC & \gamma \\
(m_2r_{1}, m_4r_{4}) & \delta \\
(m_2r_{1}, m_4r_{4}) & \omega \\
(m_1, m_3) & \theta \\
\end{array}$$

$$\begin{array}{ll}
C_I = (s, t + \delta + \omega) & C_H = (S, T + \Delta + \Omega) \\
\end{array}$$
4 Conclusion

In this paper we have presented k(XV)SE(2)PKC. We have briefly described K(XV)SE(2)PKC because it can be constructed by a straightforward generalization of k(XV)SE(2)PKC.

We have shown that k(XV)SE(2)PKC would be secure against the various attacks including GB Attack, the attack based on Gröbner basis calculation.

References


