Cryptanalysis of the TRMC-4 Public Key Cryptosystem

Xuyun Nie[1], Lei Hu[1,⋆], Jintai Ding[2,3,⋆⋆], Jianyu Li[1], and John Wagner[2]

[1] State Key Laboratory of Information Security, Graduate School of Chinese Academy of Sciences, Beijing 100049, China
[2] Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH, 45220, USA
[3] Fachbereich Informatik, Technische Universität Darmstadt, Germany
{nxy04b,hu}@is.ac.cn, ding@math.uc.edu, ljy@is.ac.cn, wagnerjh@email.uc.edu

Abstract. In 2006, the inventors of TRMC public key cryptosystem proposed a new variant of TRMC, TRMC-4, which can resist the existing attack, in particular, the Joux et al attack. In this paper, we show that the new version is vulnerable to attack via the linearization equations (LE) method. For any given valid ciphertext and its corresponding TRMC-4 public key, we can derive the corresponding plaintext within 2^{24} $\mathbb{F}_{2^8}$-operations, after performing once for the public key a computation of complexity less than 2^{34}. Our results are confirmed by computer experiments.

Keywords: multivariate public key cryptosystem, quadratic polynomial, algebraic cryptanalysis, linearization equation, TRMC.

1 Introduction

For the last three decades, public key cryptosystems (PKC) become an indispensable part of our modern communication system. The security of traditional PKC, such as RSA and ElGamal, depends on hard number theory based problems such as factoring or discrete logarithms. However, due to the quantum computer attack by Shor [Sho97], and demand for more efficient cryptosystems for small devices, there is a need to search for alternatives which are based on other classes of problems.

Multivariate public key cryptosystem (MPKC) is a promising alternative. Different from traditional PKC, the public key of MPKC is usually a set of quadratic polynomials. The security of MPKC relies on the difficulty of solving systems of nonlinear polynomial equations with many variables, and the latter is an NP-hard problem in general. Compared with RSA public key cryptosystems,

⋆ The work of this author is supported by NSFC (60573053) and National 863 Project of China (2006AA01Z416).
⋆⋆ The work of this author is partially supported by the Charles Phelps Taft Research Center and the Alexander von Humboldt Foundation.

J. Katz and M. Yung (Eds.): ACNS 2007, LNCS 4521, pp. 104–115, 2007.

the computation in MPKC can be very fast because it is operated on a small finite field.

The first promising construction of MPKC is the Matsumoto-Imai (MI) scheme [MI88] proposed in 1988. Unfortunately, it was defeated by Patarin in 1995 with the linearization method [Pat95].

Tractable rational map cryptosystem (TRMC) is a family of MPKC. It is a type of stepwise triangular system (STS) [Wo05]. There are some STS schemes such as TTM cryptosystems [Moh99] and TTS signature schemes [YC05]. All existing instances of TTM have a common defect: their plaintext and ciphertext variables always satisfy some linearization equations. Hence, they are all insecure [GC00], [DH03], [DS03], [NHLCD06]. Compared to TTM, the construction of TRMC is more systematic. Its central map is a so-called tractable rational map.

A previous version of TRMC is TRMC-2. The decryption of TRMC-2 involves solving a sub-system of equations. Joux et al pointed out that the existence of the sub-system turned out to be a weakness [JKMR05]. Utilizing this weakness, Joux et al introduced a variant of the XL algorithm and built a pseudo-private key equivalent to the original private key for a given valid ciphertext. With this pseudo-private key, they find the corresponding plaintext.

To avoid this attack, the inventors of TRMC proposed TRMC-4 [WC04] recently. But unfortunately, we find there exist some linearization equations satisfied by plaintext variables m_i and ciphertext variables w_j, namely

$$\sum_{i=1,j=1}^{n,m} a_{ij} m_i w_j + \sum_{i=1}^{n} b_i m_i + \sum_{j=1}^{m} c_j w_j + d = 0.$$

Linearization equation attack was proposed first by Patarin in 1995 to defeat the MI scheme [Pat95]. The linearization equation is also called the Patarin relation. The authors claimed that it would be computationally infeasible if one carefully designs the tractable rational maps [WC04]. But for TRMC-4, we find that there are some Paratin relations in TRMC-4 and we can find all linearization equations in 2^{34} operations. Then for a given valid ciphertext, via three eliminations, we can find the corresponding plaintext in 2^{24} operations.

This paper is organized as follows. We introduce tractable rational map and TRMC-4 encryption scheme in Section 2. In Section 3, we describe how to attack TRMC-4, present a practical attack procedure, and calculate the complexity of our attack. Finally, in Section 4, we conclude the paper.

2 TRMC Cryptosystems

2.1 Tractable Rational Map

TRMC is an MPKC. Its central map is a so-called tractable rational map, which is different from other MPKCs such as TTM etc..

Let K be a finite field. A tractable rational map is a map on K of following form:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_j \\ \vdots \\ y_n \end{pmatrix} = \phi \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} r_1(x_1) \\ r_2(x_2) \cdot \frac{p_2(x_1)}{q_2(x_1)} + \frac{f_2(x_1)}{g_2(x_1)} \\ \vdots \\ r_j(x_j) \cdot \frac{p_j(x_1,x_2,\cdots,x_{j-1})}{q_j(x_1,x_2,\cdots,x_{j-1})} + \frac{f_j(x_1,x_2,\cdots,x_{j-1})}{g_j(x_1,x_2,\cdots,x_{j-1})} \\ \vdots \\ r_n(x_n) \cdot \frac{p_n(x_1,x_2,\cdots,x_{n-1})}{q_n(x_1,x_2,\cdots,x_{n-1})} + \frac{f_n(x_1,x_2,\cdots,x_{n-1})}{g_n(x_1,x_2,\cdots,x_{n-1})} \end{pmatrix}$$

where f_j, g_j, p_j, q_j are polynomials on K, r_j is an invertible polynomial over K whose inverse can easily be computed.

The inverse process is very simple. One can derive $x_1 = r_1^{-1}(y_1)$ from $y_1 = r_1(x_1)$, then compute x_2 from x_1 and y_2. By iteration, we can obtain the values of $x_3, \cdots, x_n$ in turn. So TRMC can be regarded as a triangular system.

2.2 TRMC-4

Let $\mathbb{K} = \mathbb{F}_{2^8}$ be a finite field with 2^8 elements. Map $F : \mathbb{K}^{45} \to \mathbb{K}^{50}$ is a composition of 4 maps $\phi_1, \phi_2, \phi_3, \phi_4$. Let

$$(x_1, \cdots, x_{45}) = \phi_1(m_1, \cdots, m_{45}), (y_1, \cdots, y_{50}) = \phi_2(x_1, \cdots, x_{45}),$$

$$(z_1, \cdots, z_{50}) = \phi_3(y_1, \cdots, y_{50}), (w_1, \cdots, w_{50}) = \phi_4(z_1, \cdots, z_{50}),$$

where ϕ_1 and ϕ_4 are invertible affine maps, ϕ_2 and ϕ_3 are tractable rational maps. Note that the central map of TRMC-4 is the composition of two tractable rational maps.

The expressions of ϕ_2 and ϕ_3, except for a few parameters, are public information in the TRMC-4. ϕ_1 and ϕ_4 are taken as the private key, while the expression of the map $(w_0, \cdots, w_{50}) = F(m_0, ..., m_{45})$ is the public key. The public key $F(m_1, \cdots, m_{45})$ is 50 quadratic equations in 45 variables. Denote by F_j the j-th component function of F.

$$\begin{aligned} (w_1, \cdots, w_{50}) &= F(m_1, \cdots, m_{45}) \\ &= \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(m_1, \cdots, m_{45}) \\ &= (F_1(m_1, \cdots, m_{45}), \cdots, F_{50}(m_1, \cdots, m_{45})) \end{aligned}$$

To list ϕ_2 and ϕ_3, we firstly fix some notation.

Let $\mathbb{E} = \mathbb{F}_{2^{48}}$ be a degree 6 extension field of $\mathbb{K}$. $\pi : \mathbb{E} \to \mathbb{K}^6$ is a natural $\mathbb{K}$-linear isomorphism. Namely we take a basis of $\mathbb{E}$ over $\mathbb{K}$, $\{\theta_1, \cdots, \theta_6\}$, and define π by $\pi(a_1\theta_1 + \cdots + a_6\theta_6) = (a_1, \cdots, a_6)$ for any $a_1, \cdots, a_6 \in \mathbb{K}$. It is natural to regard π as a $\mathbb{K}$-linear isomorphism from $\mathbb{E}^8$ to $\mathbb{K}^{48}$.

In TRMC-4, the intermediate variables $x_1, \cdots, x_{45}, y_1, \cdots, y_{48}$ and $z_1, \cdots, z_{48}$ are grouped into elements in $\mathbb{E}$, shown in Table 1. Here the second and the forth column are the images of entries in the first and the third column, respectively. For example, $\pi(X_1) = c_1\theta_1 + x_1\theta_2 + \cdots + x_5\theta_6$. The $c_1, \cdots, c_6 \in \mathbb{K}$ are constants, such that $c_1, c_4, c_5 \neq 0$ to avoid decryption failure.

Table 1. Intermediate variables and their corresponding entries in $\mathbb{E}$

X_1	$(c_1, x_1, x_2, x_3, x_4, x_5)$	Y_1	$(y_1, y_2, y_3, y_4, y_5, y_6)$
X_2	$(c_2, x_6, x_7, x_8, x_9, x_{10})$	Y_2	$(y_7, y_8, y_9, y_{10}, y_{11}, y_{12})$
X_3	$(c_3, x_{11}, x_{12}, x_{13}, x_{14}, x_{15})$	Y_3	$(y_{13}, y_{14}, y_{15}, y_{16}, y_{17}, y_{18})$
X_4	$(x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{21})$	Y_4	$(y_{19}, y_{20}, y_{21}, y_{22}, y_{23}, y_{24})$
X_5	$(x_{22}, x_{23}, x_{24}, x_{25}, x_{26}, x_{27})$	Y_5	$(y_{25}, y_{26}, y_{27}, y_{28}, y_{29}, y_{30})$
X_6	$(x_{28}, x_{29}, x_{30}, x_{31}, x_{32}, x_{33})$	Y_6	$(y_{31}, y_{32}, y_{33}, y_{34}, y_{35}, y_{36})$
X_7	$(x_{34}, x_{35}, x_{36}, x_{37}, x_{38}, x_{39})$	Y_7	$(y_{37}, y_{38}, y_{39}, y_{40}, y_{41}, y_{42})$
X_8	$(x_{40}, x_{41}, x_{42}, x_{43}, x_{44}, x_{45})$	Y_8	$(y_{43}, y_{44}, y_{45}, y_{46}, y_{47}, y_{48})$
$\tilde{X}_1$	$(c_4, x_1, x_4, x_7, x_{10}, x_{13})$	Z_1	$(z_1, z_2, z_3, z_4, z_5, z_6)$
$\tilde{X}_2$	$(c_5, x_2, x_5, x_8, x_{11}, x_{14})$	Z_2	$(z_7, z_8, z_9, z_{10}, z_{11}, z_{12})$
$\tilde{X}_3$	$(c_6, x_3, x_6, x_9, x_{12}, x_{15})$	Z_3	$(z_{13}, z_{14}, z_{15}, z_{16}, z_{17}, z_{18})$
		Z_4	$(z_{19}, z_{20}, z_{21}, z_{22}, z_{23}, z_{24})$
		Z_5	$(z_{25}, z_{26}, z_{27}, z_{28}, z_{29}, z_{30})$
		Z_6	$(z_{31}, z_{32}, z_{33}, z_{34}, z_{35}, z_{36})$
		Z_7	$(z_{37}, z_{38}, z_{39}, z_{40}, z_{41}, z_{42})$
		Z_8	$(z_{43}, z_{44}, z_{45}, z_{46}, z_{47}, z_{48})$

ϕ_2 is defined as follows:

$$\begin{cases} Y_1 = \tilde{X}_1; \\ Y_2 = \tilde{X}_2\tilde{X}_1; \\ Y_3 = \tilde{X}_3\tilde{X}_2; \\ Y_4 = X_4X_1 + X_3X_2 \\ \begin{pmatrix} Y_5 & Y_6 \\ Y_7 & Y_8 \end{pmatrix} = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} \begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix} = \begin{pmatrix} X_1X_5 + X_2X_7 & X_1X_6 + X_2X_8 \\ X_3X_5 + X_4X_7 & X_3X_6 + X_4X_8 \end{pmatrix}; \\ y_{49} = L_1L_6 + L_2L_7 + L_3L_8 + L_4L_9 + L_5L_{10}; \\ y_{50} = L_1L_{11} + L_2L_{12} + L_3L_{13} + L_4L_{14} + L_5L_{15}. \end{cases} \tag{2.1}$$

where $L_1, \cdots, L_{15}$ are randomly chosen linear maps in $x_1, \cdots, x_{45}$.

ϕ_3 is defined as follows:

$$\begin{cases} Z_1 = Y_1^2\frac{Y_3}{Y_2} + g_1(\frac{Y_5Y_8+Y_6Y_7}{Y_4}) = \tilde{X}_1\tilde{X}_3 + g_1(X_5X_8 + X_6X_7); \\ Z_2 = Y_2 + g_2(\frac{Y_5Y_8+Y_6Y_7}{Y_4}) = \tilde{X}_2\tilde{X}_1 + g_2(X_5X_8 + X_6X_7); \\ Z_3 = Y_3 + g_3(\frac{Y_5Y_8+Y_6Y_7}{Y_4}) = \tilde{X}_3\tilde{X}_2 + g_3(X_5X_8 + X_6X_7); \\ Z_4 = Y_4 = X_4X_1 + X_3X_2; \\ Z_5 = Y_5 = X_1X_5 + X_2X_7; \\ Z_6 = Y_6 = X_1X_6 + X_2X_8; \\ Z_7 = Y_7 = X_3X_5 + X_4X_7; \\ Z_8 = Y_8 = X_3X_6 + X_4X_8; \\ z_{49} = y_{49}; \\ z_{50} = y_{50}. \end{cases} \tag{2.2}$$

where g_i, $i = 1, 2, 3$, are maps from $\mathbb{E}$ to $\mathbb{E}$, each of them corresponds to a map f_i, where $f_i = \pi \circ g_i \circ \pi^{-1}$, is a $\mathbb{K}$-linear transformation from $\mathbb{K}^6$ to $\mathbb{K}^6$.

The inverting process of TRMC-4 is very simple. Applying ϕ_4^{-1} on $w_1, \cdots, w_{50}$, one can derive the $z_1, \cdots, z_{50}$, then the $Z_1, \cdots, Z_8$ and $Y_4, \cdots, Y_8$. One can

compute the value of Y_1, Y_2, Y_3 from the first three formulas of (2.2) and Z_1, Z_2, Z_3. Then from the first three formulas of (2.1), one can obtain $\tilde{X_1}, \tilde{X_2}, \tilde{X_3}$ hence X_1, X_2, X_3. Then one can derive X_4, X_5, X_6, X_7, X_8 from the matrix equation and the fourth equation. So one obtains all the $(x_1, \cdots, x_{45})$. Finally, applying ϕ_1^{-1} on $(x_1, \cdots, x_{45})$, one derives all plaintext $(m_1, \cdots, m_{45})$. Note that if $Z_4 = Y_4 = 0$, the decryption mentioned above will not work.

3 Cryptanalysis on TRMC-4

The inventors of TRMC claimed [WC04] that searching the general Patarin relations would be computationally infeasible by carefully designing the tractable rational maps. But through theoretical analysis, we find that there still exist Patarin relations in TRMC-4 and we can find all Patarin relations in a short times. Given a valid ciphertext, starting from these equations, we can find the corresponding plaintext easily.

3.1 Linearization Equations

Firstly, set

$$M_1 = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}, M_2 = \begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix}, M = \begin{pmatrix} Y_5 & Y_6 \\ Y_7 & Y_8 \end{pmatrix} = \begin{pmatrix} Z_5 & Z_6 \\ Z_7 & Z_8 \end{pmatrix}.$$

Denote by A^* the associated matrix of a square matrix; for a second order matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, its associated matrix is $A^* = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

In TRMC-4, we have

$$M = M_1 M_2, \det(M_1) = Y_4 = Z_4,$$

Hence

$$M_2 \det(M_1) = M_1^* M,$$

namely,

$$\begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix} Z_4 = \begin{pmatrix} X_4 & X_2 \\ X_3 & X_1 \end{pmatrix} \begin{pmatrix} Z_5 & Z_6 \\ Z_7 & Z_8 \end{pmatrix}. \tag{3.1}$$

Expanding it, that is,

$$\begin{cases} X_4 Z_5 + X_2 Z_7 + X_5 Z_4 = 0; \\ X_2 Z_8 + X_4 Z_6 + X_6 Z_4 = 0; \\ X_1 Z_7 + X_3 Z_5 + X_7 Z_4 = 0; \\ X_1 Z_8 + X_3 Z_6 + X_8 Z_4 = 0. \end{cases} \tag{3.2}$$

Since F is derived from $\phi_3 \circ \phi_2$ by composing from the inner and outer sides by invertible affine maps ϕ_1 and ϕ_4. Hence equation (3.2) imply that for any $(m_1, \cdots, m_{45}) \in K^{45}$ satisfying the equation of the form:

$$\sum_{i=1,j=1}^{45,50} a_{ij} m_i F_j + \sum_{i=1}^{45} b_i m_i + \sum_{j=1}^{50} c_j F_j + d = 0 \tag{3.3}$$

Furthermore, the four equations in (3.2) are all linearly independent, therefore there exist at least 24 linearization equations such that their corresponding coefficient vectors are linearly independent. Actually, given the value of Z_i, the equations in X_i are also linearly independent. Hence, given a valid ciphertext, there still exist at least 24 linearly independent linear equations in $(m_1, \cdots, m_{45})$. Let V denote the $\mathbb{K}$-linear space composed of all linearization equations of the form (3.3), and let $D \geq 24$ be its dimension.

To find all equations in V is equivalent to find a basis of V. The equation (3.3) is equivalent to a system of equations on the coefficients a_{ij}, b_i, c_j, and d. The number of unknowns in these equations is equal to the number of monomials in m_i, F_j. So there are $2346 = 45 \times 50 + 45 + 50 + 1$ unknowns in these equations.

To find a basis of V, we randomly select slightly more than 2346, say 2500, plaintexts $(m_1, \cdots, m_{45})$, substitute them in (3.3) to get a system of 2500 linear equations, and solve the resulting system. Let $\{(a_{ij}^{(\rho)}, b_i^{(\rho)}, c_j^{(\rho)}, d^{(\rho)}), 1 \leq \rho \leq D\}$ be the coefficient vectors corresponding to a basis of V, where i, and j stand for $i = 1, \cdots, 45$, $1 \leq j \leq 50$, respectively. Hence, we derive D linearly independent equations in m_i and F_j. Let $E_\rho (1 \leq \rho \leq D)$ denote the equations:

$$\begin{cases} \sum\limits_{i=1,j=1}^{45,50} a_{ij}^{(\rho)} m_i F_j + \sum\limits_{i=1}^{45} b_i^{(\rho)} m_i + \sum\limits_{j=1}^{50} c_j^{(\rho)} F_j + d^{(\rho)} = 0 \\ (1 \leq \rho \leq D) \end{cases} \tag{3.4}$$

The work above depends only on any given public key, and it can be solved once for all cryptanalysis under that public key.

3.2 First Elimination

Let's assume we have a valid ciphertext $w' = (w'_1, \cdots, w'_{50})$. our goal is to find its corresponding plaintext $m' = (m'_1, \cdots, m'_{45})$.

Substituting $(F_1, \cdots, F_{50}) = (w'_1, \cdots, w'_{50})$ into $E_\rho (1 \leq \rho \leq D)$, we can derive D linear equations in m_i. Reducing these D equations, we can derive a system of linearly independent linear equations. Let l $(l \geq 24)$ denote the number of linearly independent equations in these system. Let $E'_1, \cdots, E'_l$ denote these equations. Doing a simple Gaussian elimination, from these l equations we can represent l variables of $x_1, \cdots, x_{45}$ by linear combinations of other $45 - l$. That is, we can find two disjoint subsets of $\{1, \cdots, 45\}$, $A'_1 = \{u'_1, \cdots, u'_l\}$ and $A_1 = \{u_1, \cdots, u_{45-l}\}$, and linear expressions

$$m_{u'_j} = h_j(m_{u_1}, \cdots, m_{u_{45-l}}), 1 \leq j \leq l \tag{3.5}$$

such that $E'_1, \cdots, E'_l$ holds when (3.5) are substituted into them.

Let S denote a $(45 - l)$-dimensional affine subspace of $\mathbb{K}^{45}$ defined by (3.5); the component $m_{u'_j}$ of any vector $(m_0, \cdots, m_{45})$ in S is $h_j(m_{u_1}, \cdots, m_{u_{45-l}})$.

Now substitute (3.5) into $F_j(m_1, \cdots, m_{45})$ and derive 50 new quadratic functions $\hat{F}_j(m_{u_1}, \cdots, m_{u_{45-l}})$ $(1 \leq j \leq 50)$.

3.3 Second Elimination

Furthermore, through theoretical analysis, we find there still exist linearization equations on S.

Firstly, we denote by Z'_i, $i = 1, \cdots, 8$, the value of Z_i corresponding to a given valid ciphertext $w' = (w'_1, \ldots, w'_{50})$. Similar notations Y'_i, X'_i, $\tilde{X}'_i$, x'_i and m'_i are denoted for Y_i, X_i, $\tilde{X}_i$, x_i and m_i, respectively.

Since we have found a basis of all linearization equations and each linearization equation is a linear combination of this basis, this fact holds when the variables F_j in the equations are substituted by w'_j. Applying this fact to (3.1), we know

$$\begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix} = \begin{pmatrix} X_4 & X_2 \\ X_3 & X_1 \end{pmatrix} \begin{pmatrix} Z'_5 & Z'_6 \\ Z'_7 & Z'_8 \end{pmatrix} {Z'_4}^{-1} \tag{3.6}$$

namely,

$$\begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix} = \begin{pmatrix} (X_4Z'_5 + X_2Z'_7){Z'_4}^{-1} & (X_2Z'_8 + X_4Z'_6){Z'_4}^{-1} \\ (X_1Z'_7 + X_3Z'_5){Z'_4}^{-1} & (X_1Z'_8 + X_3Z'_6){Z'_4}^{-1} \end{pmatrix} \tag{3.7}$$

The linear equations in m_i derived from (3.6), (3.7) are all linear combinations of the equations $E'_1, \cdots, E'_l$, in other words, (3.6), (3.7) holds on S.

Calculate the determinants of matrixes in two sides of matrix equation (3.6), then

$$X_5X_8 + X_6X_7 = C'Z_4 \tag{3.8}$$

where $C' = (Z'_5Z'_8 + Z'_6Z'_7){Z'_4}^{-2}$.

Substitute (3.7) (3.8) into (2.2), then

$$\begin{cases} Z_1 = \tilde{X}_1\tilde{X}_3 + g_1(C'Z_4) \\ Z_2 = \tilde{X}_2\tilde{X}_1 + g_2(C'Z_4) \\ Z_3 = \tilde{X}_3\tilde{X}_2 + g_3(C'Z_4) \\ Z_4 = X_1X_4 + X_2X_3 \\ Z_5 = Z'_5{Z'_4}^{-1}Z_4 \\ Z_6 = Z'_6{Z'_4}^{-1}Z_4 \\ Z_7 = Z'_7{Z'_4}^{-1}Z_4 \\ Z_8 = Z'_8{Z'_4}^{-1}Z_4 \end{cases} \tag{3.9}$$

From the first three equations of (3.9), we can derive:

$$\begin{cases} \tilde{X}_3(Z_2 + g_2(C'Z_4)) = \tilde{X}_1(Z_3 + g_3(C'Z_4)) \\ \tilde{X}_2(Z_1 + g_1(C'Z_4)) = \tilde{X}_1(Z_3 + g_3(C'Z_4)) \end{cases} \tag{3.10}$$

Equation (3.10) implies that there exist at least 10 to 12 linearly independent linearization equations for remaining $45-l$ plaintext variables and the new public key polynomials, that is:

$$\sum_{i=1,j=1}^{45-l,50} \hat{a}_{ij} m_{u_i} \hat{F}_j + \sum_{i=1}^{45-l} \hat{b}_i m_{u_i} + \sum_{j=1}^{50} \hat{c}_j \hat{F}_j + \hat{d} = 0 \tag{3.11}$$

And these equations are still linearly independent when the value of $\hat{F}_i$ is given.

Additionally, from the last five equations, we find that some (at least 24) polynomials of the new public key polynomials can be linearly expressed by other polynomials.

In order to make our attack more efficient, we can do Gauss reduction first on the new public key polynomials. Note that here we must combine the given valid ciphertext with the new public key polynomials. Concretely, we consider the coefficients in each polynomial as a row vector and we concatenate w_i' and the coefficient vector corresponding to F_i. Therefore, we derive a $50 \times (\binom{45-l+2}{2} + 1)$ matrix. Doing Gauss reduction on this matrix, we can obtain a matrix whose order less than 26. Hence, we derive a new set of public key polynomials, denoted by $\hat{\hat{F}}_i$. Set there are $t \leq 26$ polynomials in this set. Denote the valid ciphertext corresponding to the new public key polynomials by $\hat{w}_i'$, $i = 1, \cdots, t$.

So the equation (3.11) can be changed into:

$$\sum_{i=1,j=1}^{45-l,t} \hat{a}_{ij} m_{u_i} \hat{\hat{F}}_j + \sum_{i=1}^{45-l} \hat{b}_i m_{u_i} + \sum_{j=1}^{t} \hat{c}_j \hat{\hat{F}}_j + \hat{d} = 0 \tag{3.12}$$

To find all equations of the form (3.12), we can use the same method as the one used for equations (3.3). Firstly, we must derive a system of linear equation in $\hat{a}_{ij}$, $\hat{b}_i$, $\hat{c}_j$ and $\hat{d}$. Since the number of public key polynomials decrease to t, these equation have only

$$(45-l)t + 45 - l + t + 1 \leq 594$$

unknowns. We randomly select 600 $m \in S$, and substitute them in (3.12) to get a system of 600 linear equations and then solve it.

Let $\hat{D}$ and $\{(\hat{a}_{ij}^{(\rho)}, \hat{b}_i^{(\rho)}, \hat{c}_j^{(\rho)}, \hat{d}^{(\rho)}) : 1 \leq \rho \leq \hat{D}\}$ be the dimension and a basis of solution space, respectively. So we derive $\hat{D}$ linearly independent quadratic equations in $m_{u_i}, i = 1, \cdots, 45-l$ and $\hat{\hat{F}}_j, j = 1, \cdots, t$. Then Substitute $\hat{\hat{F}}_j$ by $\hat{w}_j$ to get $\hat{D}$ linear equations in m_{u_i}. Assuming we can derive k ($k \geq 10$) linearly independent equations, denote these equations by $\hat{E}_1', \cdots, \hat{E}_k'$. Doing a simple Gaussian elimination, from these k equations we can represent k variables of $m_{u_1}, \cdots, m_{u_{45-l}}$ by linear combinations of other $45-l-k$. That is, we can find two disjoint subsets of $\{1, \ldots, 45-l\}$, $B_1' = \{v_1', \cdots, v_k'\}$ and $B_1 = \{v_1, \cdots, v_{45-l-k}\}$, and linear expressions

$$m_{v_j'} = \hat{h}_j(m_{v_1}, \cdots, m_{v_{45-l-k}}), 1 \leq j \leq k \tag{3.13}$$

such that $\hat{E}_1', \cdots, \hat{E}_k'$ holds when (3.13) are substituted into them. Let $\hat{S}$ denote $(45-l-k)$-dimensional affine subspace of S, where for each vector $(m_1, \cdots, m_{45})$ in $\hat{S}$, $m_{v_j'}$ is substituted by (3.13) for any $1 \leq i \leq k$.

Now substitute (3.13) into $F_j(m_1, \cdots, m_{45})$ and derive t new quadratic functions $\tilde{F}_j(m_{v_1}, \cdots, m_{v_{45-l-k}})$, $j = 1, \cdots, t$.

3.4 Third Elimination

Again, through theoretical analysis on $\tilde{F}_j(m_{v_1}, \cdots, m_{v_{45-l-k}})$, $j = 1, \cdots, t$, we find that we can do elimination once more.

Since we have found a basis of all linearization equations on S and each linearization equations is a linear combination of this basis, this fact of course holds when the variables $\hat{\tilde{F}}_j$ in the equations are substituted by $\hat{w}'_j$. Applying this fact to (3.10), we know

$$\begin{cases} \tilde{X}_3 = \tilde{X}_1 C'_1 \\ \tilde{X}_2 = \tilde{X}_1 C'_2 \end{cases} \tag{3.14}$$

where $C'_1 = (Z'_3 + g_3(C'Z'_4))(Z'_2 + g_2(C'Z'_4))^{-1}, C'_2 = (Z'_3 + g_3(C'Z_4))(Z'_1 + g_1(C'Z'_4))^{-1}$, in other words, (3.14) holds on $\hat{S}$.

Substitute (3.14) into the first three equations of (3.9), then

$$\begin{cases} Z_1 = \tilde{X}_1^2 C'_1 + g_1(C'Z_4) \\ Z_2 = \tilde{X}_1^2 C'_2 + g_2(C'Z_4) \\ Z_3 = \tilde{X}_1^2 C'_1 C'_2 + g_3(C'Z_4) \end{cases} \tag{3.15}$$

We find $\tilde{X}_1^2$ can be expressed as linear combinations of the Z_i. Utilizing the fact that squaring is a linear operation on a field of characteristic 2, we have, on $\hat{S}$, the 6 expressions corresponding to $\tilde{X}_1^2$ is of the form $\sum a'_i m_i^2 + b'$ and $\mathbb{K}$-linear combinations of $F_j(m_1, \cdots, m_{45})$ and 1 (constant). Thus, of linear combinations of $\tilde{F}_j(m_{v_1}, \cdots, m_{v_{45-l-k}})$, $j = 1, \cdots, t$, there must exist at least 6 expressions which all contain only squaring terms and a constant term and correspond to $\tilde{X}_1^2$.

It is easy to solve the following linear system on the $\tilde{a}_i$ and $\tilde{b}_j$:

$$\begin{cases} \sum\limits_{i=1}^{50} \tilde{a}_i \tilde{F}_i(m_{v_1}, \cdots, m_{v_{45-l-k}}) + \sum\limits_{j=1}^{45-l-k} \tilde{b}_j m_{v_j}^2 + \tilde{c} = 0 \\ \forall m_{v_1}, \cdots, m_{v_{45-l-k}} \in \hat{W} \end{cases} \tag{3.16}$$

Set $(\tilde{a}_1^{(\rho)}, \cdots, \tilde{a}_{50}^{(\rho)}, \tilde{b}_1^{(\rho)}, \cdots, \tilde{b}_{45-l-k}^{(\rho)}, \tilde{c}^{(\rho)})$, $1 \le \rho \le p$ (where p such that $p + k = 15$, because the vectors in $\mathbb{K}^6$ corresponding to $\tilde{X}_i$ have 15 variables), is a basis of solution space of system (3.16). Set

$$\begin{cases} \sum\limits_{j=1}^{45-l-k} (\tilde{b}_j^{(\rho)})^{1/2} m_{v_j} + (\sum\limits_{i=1}^{50} \tilde{a}_i w'_i)^{1/2} + \tilde{c}^{(\rho)} = 0 \\ 1 \le \rho \le p \end{cases} \tag{3.17}$$

For any $(m_1, \cdots, m_{45}) \in \hat{S}$, its corresponding $(m_{v_1}, \cdots, m_{v_{45-l-k}})$ satisfied (3.17). Therefore we can represent p variables of $m_{v_1}, \cdots, m_{v_{45-l-k}}$ as linear expressions of the remaining variables.

So far, we represent totally $l + k + p$ variables of $(m_1, \cdots, m_{45})$ as linear expressions of the remaining $45-l-k-p$ variables. In other words, we eliminated $l + k + p$ variables in public key polynomials.

3.5 Finding The Plaintext

Substitute the linear expressions derived from (3.17) into $\tilde{F}_j(m_{v_1}, \cdots, m_{v_{45-l-k}})$, $j = 1, \cdots, t$ to get t new public key polynomials. There are $45 - l - k - p(\leq 6)$ in these new polynomials. Denote them by $\tilde{\tilde{F}}_j(m_{v_1}, \cdots, m_{v_{45-l-k}})$, $j = 1, \cdots, t$. Since $45-l-k-p(\leq 6)$ is very small, in principle, we can use the Gröbner bases method or XL method to solve the system

$$\tilde{\tilde{F}}_j = \hat{w}'_j \tag{3.18}$$

very easily and to find the plaintext.

3.6 A Practical Attack Procedure, Its Complexity and Experimental Verification

Our attack can be further divide into the following five steps.

Step 1: *Find a basis of the linear space of the coefficient vectors* (a_{ij}, b_i, c_j, d) *of the linearization equations.*

As mentioned in subsection (3.1), we randomly select 2500 plaintexts (m_1, $\cdots, m_{45}$) and substitute them into equation (3.3) to get a linear system of 2500 equations on 2346 unknowns. The computational complexity to solve it is

$$2346^2 \times 2500 < 2500^3 < 2^{34}.$$

operations on the finite field $K = \mathbb{F}_{2^8}$.

This step is independent of the value of the ciphertext w' and can be done once for a given public key.

Our computer experiments show that indeed D is equal to 24.

Step 2: *For a given valid ciphertext* $(w'_1, \cdots, w'_{50})$*, we substitute it into (3.4) and solve the system of linear equations to get linear expression (3.5). Substitute (3.5) into the public key polynomials to derive a set of new public key polynomials* $\hat{F}_1, \cdots, \hat{F}_{50}$*. Then we combine the given valid cipheretext and the new public key polynomials and do Gauss reduction as subsection (3.3) described. At last, we derive t linearly independent public key polynomials and t new valid ciphertext components.*

The first part of this step is of computational complexity about

$$45^2 \cdot D < 45^3 < 2^{15},$$

and the second part is

$$\left(\binom{45-l+2}{2} + 1\right)^2 \times 50 < 2^{22}.$$

Our computer experiments show that the number of linear expression derived in this step is $l = 24$, and the number of the linearly independent public key polynomials is $t = 26$.

Step 3: *Solve (3.12) to get a basis of solution space of (3.12),* $\{(\hat{a}_{ij}^{(\rho)}, \hat{b}_i^{(\rho)}, \hat{c}_j^{(\rho)}, \hat{d}^{(\rho)}) : 1 \leq \rho \leq \hat{D}\}$ *then substitute the given ciphertext into result system of equations to derive linear expression (3.13).*

The first part of this step is of computational complexity about

$$(594)^2 \times 600 < 2^{24},$$

and the second part is

$$(45 - l)^2 \cdot \hat{D} < 2^{11}.$$

Our computer experiments show $\hat{D} = k = 12$.

Substituting (3.13) into $\hat{\hat{F}}_j, j = 1, \cdots, t$, we can derive a set of new public key polynomials $\tilde{F}_j(m_{v_1}, \cdots, m_{v_{45-l-k}}), j = 1, \cdots, t$.

Step 4: *Solve (3.16) to get a basis of solution space of it and then solve the system of equations (3.17) to derive* p *linear expressions in remainder* $45-l-k-p$ *components.*

The first part of this step is of computational complexity about

$$(96 - l - k)^3 < 2^{18},$$

and the second part is

$$p(45 - l - k)^2 < 2^{10}.$$

Our computer experiments show $p = 3$,

Step 5: *Use the Gröbner basis method to solve the system of equations (3.18) to get* $45 - l - k - p$ *values of plaintext components and then collect all linear expressions between the variables derived in previous steps to get the values of remainder plaintext components.*

Our computer experiments show that there is 6 variables and 8 polynomials in the last new public key polynomials $\tilde{\tilde{F}}_j(m_{v_1}, \cdots, m_{v_{45-l-k}})$, $j = 1, \cdots, t$. The computational complexity in this step is

$$\left(\frac{6^3}{3!}\right)^3 < 2^{18}.$$

Hence, the total computational complexity of our attack is less then 2^{34} $\mathbb{F}_{2^8}$-operations.

We implement our attack on a Pentium IV 2.4Ghz PC with 256M memory, and we code the attack using VC++. For any given valid ciphertext, our experiments successfully find the corresponding plaintext less than 7 minutes, where 6 minutes were spent on the execution of the step 1 in subsection (3.6), and less than 1 minute was spent to execute the remaining steps.

4 Conclusion

In this paper, we present a very efficient attack on TRMC-4. We need to do precomputation first, which takes 6 minutes on a PC with a 2.4Ghz Pentium

IV processor. Our attack then recovers the corresponding plaintext of any valid ciphertext in less than 1 minute. The total computational complexity is less than 2^{34} $\mathbb{F}_{2^8}$-operations. The key point of the attack is finding all linearization equations in polynomial time. Therefore, TRMC-4 is totally insecure.

Although we break the TRMC-4, we still think the design of TRMC is a interesting idea; one can carefully design the tractable rational map to improve the security of TRMC.

References

[DH03] J.Ding and T.Hodges. Cryptanalysis of an Implementation Scheme of TTM. *J. Algebra Appl.*, pages 273-282, 2004. http://eprint.iacr.org/2003/084.

[DS03] J.Ding and D.Schmidt. The new TTM implementation is not secure. In H.Niederreiter K.Q.Feng and C.P. Xing, editors, *Proceedings of International Workshop on Coding, Cryptography and Combinatorics (CCC 2003)*, pages 106–121, 2003.

[GC00] L.Goubin and N.Courtois. Cryptanalysis of the TTM cryptosystem. *LNCS, Springer Verlag*, 1976:44–57, 2000.

[JKMR05] A. Joux, S. Kunz-Jacques, F. Muller, P.-M. Ricordel, Cryptanalysis of the Tractable Rational Map Cryptosystem, *PKC 2005, p258-274, Lecture Notes in Computer Sciences* 3386.

[MI88] T.Matsumoto and H.Imai. Public quadratic polynomial-tuples for efficient signature verification and message encryption. In C.G. Guenther, editor, *Advances in cryptology –EUROCRYPT'88, LNCS*, volume 330, pages 419–453. Springer, 1988.

[Moh99] T.Moh. A fast public key system with signature and master key functions. *Lecture Notes at EE department of Stanford University*, May 1999. http://www.usdsi.com/ttm.html.

[NHLCD06] X.Nie, L.Hu, J.Li, C.Updegrove and J.Ding. Breaking A New Instance of TTM Cryptosystem. *Advances in ACNS2006, LNCS*, volume 3989, Springer, 2006.

[Pat95] J.Patarin. Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt'88. In D. Coppersmith, editor, *Advances in Cryptology – Crypto'95, LNCS*, volume 963, pages 248–261, 1995.

[Sho97] P.Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing,* 26(5):1484-1509, October 1997.

[WC04] L.Wang and F.Chang. Tractable Rational Map Cryptosystem, available at http://eprint.iacr.org/2004/046, revised on February 3,2006.

[Wo05] C.Wolf, Multivariate Quadratic Polynomials in Public Key Cryptography, *Cryptology ePrint Archive, Report 2005/393*, 2005, avalable at http://eprint.iacr.org/.

[YC05] B.Yang and J.Chen. Building Secure Tame-like Multivariate Public key Cryptosystems–The New TTS. Information Security and Privacy: 10th Australasian Conference–ACISP 2005, LNCS 3574, 2005, Springer, P. 518-531.