Cryptanalysis of 2-layer Nonlinear Piece In Hand Method

No Author Given

No Institute Given

Abstract. Piece in Hand method is a security enhancement method for Multivariate Public Key Cryptosystems (MPKCs). Since 2004, many types of this method have been proposed. In this paper, we consider the 2-layer nonlinear Piece in Hand method as proposed by Tsuji et al. in 2009. The key point of this method is to introduce an invertible quadratic polynomial map on the linear combination of plaintext variables to construct perturbation of the original MPKC. Through our analysis we find that the security of the enhanced scheme is mainly relying on the quadratic polynomials of the auxiliary map. The two examples used for this map by Tsuji can not resist the Linearization Equation attack. Given a valid ciphertext, we can easily get a public key which is equivalent to the original MPKC. If there is an algorithm that can recover the plaintext corresponding to a valid ciphertext of the original MPKC, we can construct an algorithm that can recover the plaintext corresponding to a valid ciphertext of the enhanced MPKC.

Keywords: Multivariate Public Key Cryptosystems, Quadratic Polynomials, Algebraic Cryptanalysis, Linearization Equations, Piece in Hand.

1 Introduction

Multivariate public key cryptosystems (MPKCs) are promising candidates to resist the quantum computer attack. Their security is based on the difficulty of finding solutions of systems of multivariate quadratic (MQ) equations over a finite field, which is an NP-hard problem in general.

Since 1988, many MPKCs have been proposed, such as MI [MI88], HFE [Pat96], MFE [WYH06], TTM [Moh99], Rainbow [DS05b] etc. However, many of these schemes have shown to be insecure [Pat95,DHN07,NHL06]. In order to enhance the security of MPKCs, many enhancement methods were proposed. There are plus/minus [PGC98], internal perturbation [Ding04], Extended Multivariate public key Cryptosystems (EMC) [WZW11] etc. SFlash [PCG01], which combined the MI scheme with the minus method, was broken by Dubois et al. using differential attacks [DFSS07]. PMI [Ding04] and IPHFE [DS05a], which combined internal perturbation with the MI and HFE scheme, respectively, were also broken by differential attacks [FGS05,DGS07]. Extended Multivariate public key Cryptosystems (EMC) can not really enhance the security of the original

MPKC because we can find a public key equivalent to the original one from the public key of EMC.
Piece in Hand (PH) method is another security enhancement method introduced and studied in a series of papers [TTF04,TTF06,FTT08,TTF09]. In [TTF09], Tsuji et al. proposed the 2-layer nonlinear Piece in Hand method. For this, they introduced two vectors of polynomials: an auxiliary polynomial vector $\mathbf{H}$ and a perturbation polynomial vector $\mathbf{J}$. The perturbation polynomial vector is used to add perturbation to the original MPKC. And the auxiliary polynomial vector is constructed to be efficiently invertible which will be used in decryption.
Because the information of the auxiliary polynomial vector is part of the public key, the security of the whole scheme relies on the structure of this vector. In the paper, the authors give two examples of this vector, called $\mathbf{H_1}$ and $\mathbf{H_2}$.

In this paper we show that both $\mathbf{H_1}$ and $\mathbf{H_2}$ satisfy Linearization Equations (LEs) of the form

$$\sum a_{ij}x_iy_j + \sum b_ix_i + \sum c_jy_j + d = 0, \tag{1}$$

where x_i are the plaintext variables and y_j are the ciphertext variables.

After finding all the LEs and substituting a valid ciphertext into these equations, we can get a system of linear equations in the plaintext variables. By solving this system, we can represent some plaintext variables by linear combinations of other plaintext variables. Hence, we can do elimination on the public key using these representations. And we can perform a similar analysis on the eliminated public key to check whether or not there are some LEs satisfied by the simplified public key. In the case of $\mathbf{H_1}$, given a valid ciphertext, we can, after two eliminations on the public key, find a public key equivalent to the original MPKC. In the case of $\mathbf{H_2}$, given a valid ciphertext, we can achieve the same goal using three eliminations on the public key. This means that Piece in Hand method using these two auxiliary polynomial vectors can not enhance the security of the original MPKCs. So, we must be very careful when designing the auxiliary polynomial vector of PH method.

The paper is organized as follows. In Section 2 we give a brief description of MPKCs and Linearization Equations. Section 3 introduces the 2-layer nonlinear Piece in Hand method. In section 4, we present our cryptanalysis. Finally, in Section 5, we conclude the paper.

2 Preliminaries

2.1 Multivariate Public Key Cryptography

To build a multivariate public key cryptosystem, one starts with an easily invertible map $F : \mathbb{F}^n \to \mathbb{F}^m$ (central map). To hide the structure of F in the

public key, one combines it with two invertible affine maps T and U. Therefore the public key has the form

$$E : \mathbb{F}^n \to \mathbb{F}^m,\ \mathbf{y} = (y_1, \ldots, y_m) = E(x_1, \ldots, x_n) = T \circ F \circ U(x_1, \ldots, x_n) \quad (2)$$

2.2 Linearization Equations

Let

$$\boldsymbol{y} = (y_1, y_2, \ldots, y_m) = E(x_1, x_2, \ldots, x_n)$$

be the encryption function of an MPKC, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$ be the plaintext variables and the ciphertext variables in the MPKC, respectively. A Linearization Equation (LE) is an equation in the $n + m$ variables $x_1, x_2, \ldots, x_n$, $y_1, y_2, \ldots, y_m$ of the form

$$\sum_{i=1,j=1}^{n,t} a_{ij} x_i g_j(y_1, y_2, \ldots, y_m) + \sum_{k=1}^{l} c_k f_k(y_1, y_2, \ldots, y_m) + d = 0. \quad (3)$$

where f_k, $1 \le k \le l$, g_j, $1 \le j \le t$, are polynomial functions in the ciphertext variables. The highest degree of g_j, $1 \le j \le l$ is called the order of the LE.
For example, a first order Linearization Equation (FOLE) looks like

$$\sum_{i=1,j=1}^{n,m} a_{ij} x_i y_j + \sum_{i=1}^{n} b_i x_i + \sum_{j=1}^{m} c_j y_j + d = 0. \quad (4)$$

Note that, given a valid ciphertext $\mathbf{y}' = (y'_1, y'_2, \ldots, y'_m)$, we can substitute it into equation (3) to get a linear equation in the plaintext variables. By finding all these equations we get a linear system which can be solved by Gaussian Elimination. After having found a solution, we can do elimination on the public key.

3 2-layer Piece In Hand Method

In this section we describe the general construction of the 2-layer nonlinear Piece in Hand method. Hereby we use the same notation as in [TTF09].
Let $E : \mathbb{F}^n \to \mathbb{F}^m$ be the public map of a multivariate public key encryption scheme and l be a positive integer.
To enhance the security of the original MPKC, the inventors of the 2-layer nonlinear Piece in Hand method introduced an auxiliary polynomial vector $\boldsymbol{H}$ and a perturbation polynomial vector $\boldsymbol{J}$. The auxiliary polynomial vector $\boldsymbol{H}$ is constructed with the products of two random linear polynomials h_i and h_j, where the functions h_i can be expressed as $h_i = \sum_{j=1}^{n} a_{ij} x_j$ $(i = 1, \ldots, l)$ with $a_{ij} \in_R \mathbb{F}$. The perturbation polynomial vector $\boldsymbol{J}$ is a vector with $l(l-1)/2$ components constructed from the polynomials $h_i h_j$ $(1 \le i < j \le l)$. Note that the polynomial components of the vector $\boldsymbol{H}$ are designed to be easily invertible for decryption. Therefore, one can use the vector $\boldsymbol{H}$ to compute the values of h_i $(i = 1, \ldots, l)$

and sequentially calculate the value of the vector $\boldsymbol{J}$. By doing so, one gets an enhanced public key $\tilde{E} : \mathbb{F}^n \to \mathbb{F}^{m+l}$ of the form

$$\tilde{E}(x_1, \ldots, x_n) = B \begin{pmatrix} E(x_1, \ldots, x_n) + D\boldsymbol{J} \\ C\boldsymbol{H} \end{pmatrix} \tag{5}$$

where B is an $(m+l) \times (m+l)$ invertible matrix over $\mathbb{F}$, D is an $m \times \frac{l(l-1)}{2}$ matrix over $\mathbb{F}$, and C is an $l \times l$ invertible matrix over $\mathbb{F}$.

Secret key: The secret key includes the following:

- the secret key of the underlying MPKC;
- the matrices B, C and D;
- the auxiliary polynomial vector $\boldsymbol{H}$;
- the perturbation polynomial vector $\boldsymbol{J}$.

Public key: The expression of function $\tilde{E}$.

Encryption: Given a plaintext vector $\boldsymbol{x}' = (x'_1, \ldots, x'_n)^T$, calculate

$$\boldsymbol{y}' = (y'_1, \ldots, y'_{m+l})^T = \tilde{E}(x'_1, \ldots, x'_n)^T.$$

Decryption: Given a valid ciphertext $\boldsymbol{y}' = (y'_1, \ldots, y'_{m+l})^T$, decryption includes the following steps:

(1) Compute $\boldsymbol{v}' = (v'_1, \ldots, v'_{m+l})^T = B^{-1}(y'_1, \ldots, y'_{m+l})^T$;
(2) Compute $\boldsymbol{H} = C^{-1}(v'_{m+1}, \ldots, v'_{m+l})^T$ and get the values of h_i $(i = 1, \ldots, l)$;
(3) Compute the value of $\boldsymbol{J}$ by substituting the values of h_i $(i = 1, \ldots, l)$ into its components;
(4) Compute $\boldsymbol{x}' = (x'_1, \ldots, x'_n)^T = E^{-1}(v'_1 - dj_1, \ldots, v'_m - dj_m)^T$, where $(dj_1, \ldots, dj_m)^T = D\boldsymbol{J}$.

Examples for the auxiliary vector $\boldsymbol{H}$ and the perturbation vector $\boldsymbol{J}$

In [TTF09], the authors gave two examples for the choice of the auxiliary vector $\boldsymbol{H}$, denoted by $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, respectively.

$$\boldsymbol{H}_1 = (u_1, \ldots, u_l)^T = \begin{pmatrix} h_1h_2 + \alpha_1 \\ h_2h_3 + \alpha_2 \\ h_3h_1 + \alpha_3 \\ h_1h_4 + \alpha_4 \\ h_1h_5 + \alpha_5 \\ \vdots \\ h_1h_{l-1} + \alpha_{l-1} \\ h_1h_l + \alpha_l \end{pmatrix} \tag{6}$$

with randomly chosen field elements α_i $(i = 1, \ldots, l)$.
Apparently, given the value of the vector $(u_1, \ldots, u_l)$, we can use the first three equations to get

$$h_1 = \left(\frac{(u_1 - \alpha_1)(u_3 - \alpha_3)}{(u_2 - \alpha_2)} \right)^{\frac{1}{2}} \tag{7}$$

and then get the values of $h_2, h_3, \ldots, h_l$ in turn.
For the auxiliary map $\boldsymbol{H}_2$, the value l is fixed to 15. We have

$$\boldsymbol{H}_2 = (u_1, \ldots, u_{15})^T = \begin{pmatrix} h_1 h_2 + \alpha_1 \\ h_2 h_3 + \alpha_2 \\ h_3 h_4 + \alpha_3 \\ h_4 h_5 + \alpha_4 \\ h_5 h_1 + \alpha_5 \\ h_6^2 + h_1 h_3 + \alpha_6 \\ h_7^2 + h_3 h_5 + \alpha_7 \\ h_8^2 + h_5 h_2 + \alpha_8 \\ h_9^2 + h_2 h_4 + \alpha_9 \\ h_{10}^2 + h_4 h_1 + \alpha_{10} \\ h_1 h_{10} + h_6 h_{11} + \alpha_{11} \\ h_2 h_9 + h_7 h_{12} + \alpha_{12} \\ h_3 h_8 + h_8 h_{13} + \alpha_{13} \\ h_4 h_7 + h_9 h_{14} + \alpha_{14} \\ h_5 h_6 + h_{10} h_{15} + \alpha_{15} \end{pmatrix} \tag{8}$$

where $\alpha_i \in_R \mathbb{F}$ $(i = 1, 2, \ldots, l)$. Similar to $\boldsymbol{H}_1$, $\boldsymbol{H}_2$ can be easily inverted.
The perturbation vector $\boldsymbol{J}$ used in [TTF09] is given as follows:

$$\boldsymbol{J} = (j_1, j_2, \ldots, j_{l(l-1)/2}) = \begin{pmatrix} h_1 h_2 + \beta_1 \\ h_1 h_3 + \beta_2 \\ \vdots \\ h_1 h_l + \beta_{l-1} \\ h_2 h_3 + \beta_l \\ \vdots \\ h_2 h_l + \beta_{2l-3} \\ h_3 h_4 + \beta_{2l-2} \\ \vdots \\ h_{l-1} h_l + \beta_{l(l-1)/2} \end{pmatrix} \tag{9}$$

where $\beta_i \in_R \mathbb{F}$ $(i = 1, 2, \ldots, l(l-1)/2)$.

4 Cryptanalysis of the 2-layer PH Method

Although the perturbation polynomial vector $\boldsymbol{J}$ can hide the weak point of the underlying MPKC scheme, the security of the enhanced scheme depends mainly on the design of the vector $\boldsymbol{H}$. Bad design of the vector $\boldsymbol{H}$ will bring some new security problems to the scheme. Both vectors $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are not properly

chosen to enhance the security of the underlying scheme, since they satisfy Linearization Equations. In this section, we present the cryptanalysis of the 2-layer nonlinear Piece in Hand (PH) method with auxiliary polynomial vector $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, respectively.
Given a valid ciphertext $\boldsymbol{y}' = (y'_1, \ldots, y'_{m+l})^T$, our goal is to find its corresponding plaintext. Namely, we have to solve the following system:

$$\begin{cases} y'_1 = \tilde{E}_1(x_1, \ldots, x_n) \\ \vdots \\ y'_{m+l} = \tilde{E}_{m+l}(x_1, \ldots, x_n) \end{cases} \tag{10}$$

4.1 Case of $\boldsymbol{H_1}$

Through theoretical analysis, we find that the system $\tilde{E}$ satisfies Linearization Equations, which are brought in by the vector $\boldsymbol{H}_1$. Given a valid ciphertext, after finding all FOLEs, we can find the corresponding plaintext easily.

4.1.1 Linearization Equations

In the expression of the polynomial vector $\boldsymbol{H}_1$, we have

$$u_1 = h_1 h_2 + \alpha_1, u_2 = h_2 h_3 + \alpha_2,$$

hence,

$$h_3(u_1 - \alpha_1) = h_1(u_2 - \alpha_2). \tag{11}$$

Since the matrices B and C are invertible, the elements u_i $(i = 1, \ldots, l)$ can be expressed by linear equations in the ciphertext variables, namely $u_i = \sum_{j=1}^{m+l} t_{ij} y_j$ $(i = 1, \ldots, l)$. Analogously we get $h_i = \sum_{j=1}^{n} a_{ij} x_j$ $(i = 1, \ldots, l)$. Hence equation (11) implies that the plaintext variables $x_1, \ldots, x_n$ and ciphertext variables $y_1, \ldots, y_{m+l}$ satisfy an equation of the form:

$$\sum_{i=1,j=1}^{n,m+l} a_{ij} x_i y_j + \sum_{i=1}^{n} b_i x_i + \sum_{j=1}^{m+l} c_j y_j + d = 0. \tag{12}$$

This equation is exactly a FOLE. Similarly, from each of the pairs $h_j(u_i - \alpha_i) = h_i(u_j - \alpha_j)$ $(1 \le i < j \le l, i \ne 2)$ and the pair $h_1(u_2 - \alpha_2) = h_2(u_3 - \alpha_3)$, we can get an additional FOLE. Hence there exist at least $(l-2)(l-1)/2+1$ linear independent Linearization Equations of type (12).

To find these FOLEs, we randomly generate $D_1 \ge n(m+l) + n + m + l + 1$ plaintext/ciphertext pairs and substitute them into equation (12). By doing so we get a system of D_1 linear equations in the $n(m+l)+n+m+l+1$ unknowns a_{ij}, b_i, c_j and d which can be solved by Gaussian Elimination. We denote the solution space by V and its dimension by D. Hence, we derive D linearly independent equations of type (12) in the plaintext and ciphertext variables.

The work above depends only on the public key and can be done once for a given public key.

By substituting the given ciphertext $\mathbf{y}' = (y'_1, \ldots, y'_{m+l})$ into the Linearization Equations found above we get D linear equations in the plaintext variables. Let's assume that t_1 of these equations are linearly independent and denote the $n - t_1$ dimensional solution space by S_1.

4.1.2 First Elimination

We can substitute the t_1 equations found above into the public key $\tilde{E}$ of the 2-layer nonlinear PH scheme. By doing so, we can eliminate t_1 equations from $\tilde{E}$. Therefore we get a simplified public key $\tilde{E}'$ of the form

$$\begin{cases} y'_j = \tilde{E}'_j(w_1, \ldots, w_{n-t_1}) \\ 1 \leq j \leq m + l \end{cases} \tag{13}$$

4.1.3 Second Elimination

In the practical instance of the paper [TTF09], the characteristic of the underlying field $\mathbb{F}$ was chosen to be 2. Using this property, we can find another type of Linearization Equations satisfied by all plaintext/ciphertext pairs in S_1.

Firstly, we denote by u'_i, $i = 1, \ldots, l$ the value of u_i corresponding to the given ciphertext $\mathbf{y}' = (y'_1, \ldots, y'_{m+l})$.
Applying this to (6), we have

$$\begin{cases} u'_1 = h_1 h_2 + \alpha_1 \\ u'_2 = h_2 h_3 + \alpha_2 \\ u'_3 = h_3 h_1 + \alpha_3 \\ u'_4 = h_1 h_4 + \alpha_4 \\ u'_5 = h_1 h_5 + \alpha_5 \\ \quad \vdots \\ u'_{l-1} = h_1 h_{l-1} + \alpha_{l-1} \\ u'_l = h_1 h_l + \alpha_l \end{cases} \tag{14}$$

Note that the Linearization Equations found during the previous step are, regardless of this substitution, still satisfied.
According to relations similar to equation (11), we have

$$\begin{cases} h_2 = \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1 \\ h_3 = \frac{u_2'-\alpha_2}{u_1'-\alpha_1} h_1 \\ h_4 = \frac{u_4'-\alpha_4}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1 \\ h_5 = \frac{u_5'-\alpha_5}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1 \\ \vdots \\ h_l = \frac{u_l'-\alpha_l}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1 \end{cases} \tag{15}$$

By substituting (15) into (6), we get

$$\begin{cases} u_1 = \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1^2 + \alpha_1 \\ u_2 = \frac{u_2'-\alpha_2}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1^2 + \alpha_2 \\ u_3 = \frac{u_2'-\alpha_2}{u_1'-\alpha_1} h_1^2 + \alpha_3 \\ u_4 = \frac{u_4'-\alpha_4}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1^2 + \alpha_4 \\ u_5 = \frac{u_5'-\alpha_5}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1^2 + \alpha_5 \\ \vdots \\ u_l = \frac{u_l'-\alpha_l}{u_1'-\alpha_1} \cdot \frac{u_2'-\alpha_2}{u_3'-\alpha_3} h_1^2 + \alpha_l \end{cases} \tag{16}$$

Due to $u_i = \sum_{j=1}^{m+l} t_{ij} y_j$ $(i = 1, \ldots, l)$ and $h_i = \sum_{j=1}^{n} a_{ij} x_j$ $(i = 1, \ldots, l)$ and using the fact that squaring is a linear operation on a field of characteristic 2, we have at least one equation of the following form satisfied by ciphertext variables and the remaining plaintext variables.

$$\begin{cases} \sum_{j=1}^{m+l} \tilde{a_j} \cdot y_j' + \sum_{i=1}^{n-t_1} \tilde{b}_i \cdot w_i^2 + \tilde{c} = 0 \\ \forall w_1, \ldots, w_{n-t_1} \in \mathbb{F} \end{cases} \tag{17}$$

It is easy to solve the above linear system for the $\tilde{a}_i$, $\tilde{b}_j$ and $\tilde{c}$. Let $(\tilde{a}_1^{(\rho)}, \cdots, \tilde{a}_{m+l}^{(\rho)}$, $\tilde{b}_1^{(\rho)}, \cdots, \tilde{b}_{n-t}^{(\rho)}, \tilde{c}^{(\rho)})$, $1 \le \rho \le l$ be a basis of the solution space of the system (17). Set

$$\begin{cases} \sum_{j=1}^{n-t_1} (\tilde{b}_j^{(\rho)})^{1/2} \cdot w_j + \left(\sum_{i=1}^{m+l} \tilde{a_i}^{(\rho)} \cdot y_i' + \tilde{c}^{(\rho)} \right)^{1/2} = 0 \\ 1 \le \rho \le l \end{cases} . \tag{18}$$

For any $(x_1, \ldots, x_n) \in S_1$, the corresponding vector $(w_1, \ldots, w_{n-t_1})$ satisfies (18). Therefore we represent at least one variable of $w_1, \ldots, w_{n-t_1}$ as a linear equation in the remaining variables. Denote the remaining variables by $v_1, \ldots, v_{n-t_1-1}$.

Substituting this linear expression into the system (13), we can get a new public key with $(n - t_1 - 1)$ unknowns, denoted as

$$\begin{cases} y_j' = \tilde{E}_j''(v_1, \ldots, v_{n-t_1-1}) \\ 1 \le j \le m + l \end{cases} . \tag{19}$$

4.1.4 Eliminating Perturbation

Furthermore, after two eliminations on the public key, the vector $\boldsymbol{J}$ becomes a constant vector, namely, the perturbation of Piece in Hand method is eliminated. The reason for this is shown as follows. From (16), we have

$$h_1 = \left(\frac{(u'_1 - \alpha_1)(u'_3 - \alpha_3)}{u'_2 - \alpha_2} \right)^{1/2} \tag{20}$$

Substituting (20) and (15) into (9), the vector $\boldsymbol{J}$ becomes a constant vector on $\mathbb{F}$. For example,

$$j_1 = h_1 h_2 + \beta_1 = u'_1 - \alpha_1 + \beta_1,$$
$$j_{l+1} = h_2 h_4 + \beta_{l+1} = \left(\frac{(u'_2 - \alpha_2)(u'_4 - \alpha_4)}{u'_3 - \alpha_3} \right) + \beta_{l+1}.$$

Hence, the public key $\tilde{E}''$ of equation (19) is equivalent to the public key of the underlying MPKC scheme.
If there exists an algorithm which recovers the plaintext corresponding to a valid ciphertext for the underlying MPKC scheme, we can therefore find the values of the variables $v_1, \ldots, v_{n-t_1-1}$ corresponding to the valid ciphertext $\mathbf{y}' = (y'_1, \ldots, y'_{m+l})$. Using the linear equations found during the two eliminations above, we can recover the values of the remaining plaintexts variables.

4.2 Case of $\boldsymbol{H_2}$

Let $\mathbf{y}' = (y'_1, \ldots, y'_{m+15})$ be a valid ciphertext of the Piece in Hand MPKC with auxiliary map $\mathbf{H}_2$. Again we want to find the corresponding plaintext $\mathbf{x}' = (x'_1, \ldots, x'_n)$ by solving the system (10).
Similarly to the case of $\boldsymbol{H}_1$ we can get, from the first five equations in (8), five FOLEs between u_i and h_i $(1 \leq i \leq 5)$:

$$\begin{cases} h_3(u_1 - \alpha_1) = h_1(u_2 - \alpha_2) \\ h_4(u_2 - \alpha_2) = h_2(u_3 - \alpha_3) \\ h_5(u_3 - \alpha_3) = h_3(u_4 - \alpha_4) \\ h_1(u_4 - \alpha_4) = h_4(u_5 - \alpha_5) \\ h_2(u_5 - \alpha_5) = h_5(u_1 - \alpha_1) \end{cases}$$

Apparently, these five equations are linearly independent. Hence, we can get at least five Linearization Equations satisfied by plaintext variables and ciphertext variables of the form (12).
Using the same method as in Subsection 4.1, we can find all Linearization Equations of this form. And after substituting the valid ciphertext, we can get at least four linearly independent linear equations in the plaintext variables. Then, we do the first elimination on the system (10). Suppose we eliminated $t_1 \geq 4$ variables in the system. Denote the remaining plaintext variables by $w_1, \ldots w_{n-t_1}$ and let

$$\begin{cases} y'_j = \tilde{E}'_j(w_1, \ldots, w_{n-t_1}) \\ 1 \leq j \leq m + 15 \end{cases}. \tag{21}$$

be the simplified public key.
Using a similar method as in Subsection 4.1, we can perform two additional eliminations on the system (21). Due to the limitation of paper size, we omit the details of this part here. We will present them in the full version of this paper. But we should point out the following facts.
For the public key $\tilde{E}'_j(w_1, \ldots, w_{n-t_1})$, plain- and ciphertext variables satisfy equations of the form

$$\sum_{j=1}^{m+l} \tilde{a_j} y_j + \sum_{i=1}^{n-t_1} \tilde{b}_i w_i^2 + \tilde{c} = 0. \tag{22}$$

By substituting the ciphertext $\mathbf{y}'$ into these equations, we can find $t_2 \geq 6$ linear equations in the plaintext variables. We can therefore eliminate t_2 variables from the public key. After this elimination, the simplified public key has the form

$$\begin{cases} y'_j = \tilde{E}''_j(v_1, \ldots, v_{n-t_1-t_2}) \\ 1 \leq j \leq m+15 \end{cases}. \tag{23}$$

The public key $\tilde{E}''$ satisfies equations of the form

$$\sum_{j=1}^{m+15} \tilde{\tilde{a}}_j \cdot y_j + \sum_{i=1}^{n-t_1-t_2} \tilde{\tilde{b}}_i \cdot v_i + \tilde{\tilde{c}} = 0. \tag{24}$$

By substituting the ciphertext $\mathbf{y}'$ into these equations, we can find $t_3 \geq 5$ linear equations in the variables $v_1, \ldots, v_{n-t_1-t_2}$. Therefore, we can eliminate t_3 variables from the system (23) and get a new public key $\tilde{E}'''$ of the form

$$\begin{cases} y'_j = \tilde{E}'''_j(v_1, \ldots, v_{n-t_1-t_2-t_3}) \\ 1 \leq j \leq m+15 \end{cases}. \tag{25}$$

For the public key $\tilde{E}'''$, the vector $\boldsymbol{J}$ becomes a constant vector. Hence, $\tilde{E}'''$ is equivalent to the public key of the underlying MPKC. As in Subsection 4.1 we can therefore, under the assumption that there exists an algorithm which, for the underlying MPKC, finds for a given ciphertext the corresponding plaintext, construct an algorithm which, for any given ciphertext $\mathbf{y}' = (y'_1, \ldots, y'_{m+15})$ of the PH scheme, recovers the corresponding plaintext $\mathbf{x}' = (x'_1, \ldots, x'_n)$.

4.3 Complexity and Experimental Verification

In our concrete attack scenario we set $\mathbb{F} = GF(256)$ and $m = n = 25$. As the underlying MPKC we used the $C^\star$ scheme of Matsumoto and Imai [MI88]. We implemented the Piece in Hand cryptosystem in two different ways using $\mathbf{H}_1$ (with $l = 8$) and $\mathbf{H}_2$ as auxiliary map respectively. For our attack we chose randomly a valid ciphertext $\mathbf{y}' = (y'_1, \ldots, y'_{m+l}) \in \mathbb{F}^{m+l}$. Our goal was to find the corresponding plaintext $\mathbf{x}' = (x'_1, \ldots, x'_n) \in \mathbb{F}^n$.

Case of H_1 In the first step we computed 900 ($> n(m+l)+n+m+l+1 = 884$) plaintext/ciphertext pairs and substituted them into the Linearization Equation of type (12). We did Gaussian Elimination on a linear equation system with $n(m+l)+n+m+l+1$ unknowns and found a basis of all FOLEs. The complexity of Gaussian Elimination is equal to $(n(m+l)+n+m+l+1)^3$ operations on the finite field $\mathbb{F}$. In our experiments,

$$(n(m+l)+n+m+l+1)^3 = 884^3 \leq 2^{30}.$$

And we found that the dimension of the space spanned by all FOLEs is $D = (l-2)(l-1)/2 = 22$.
Computing the plaintext/ciphertext pairs and solving this large linear system proved to be the most time-consuming step of our attack. In our experiments, it took about 70 seconds, where it took about 68 seconds on generating the plaintext/ciphertext pairs and about 2 second on the Gaussian Elimination. This step has to be performed for each public key only once.
After substituting the ciphertext y' into these equations we obtained 7 linear equations in the plaintext variables. We denote the 18 dimensional subspace satisfying these equations by S_1.
In the second step we computed 100 plaintext/ciphertext pairs where the plaintext was randomly chosen in S_1 and substituted them into the Linearization Equation of type (17). By doing so, we got 15 linearly independent equations of the form (17). By evaluating equation (18), we got 1 linear equation in the plaintext variables. We denote the 17 dimensional subspace of S_1 satisfying this equation by S_2. For the elements of S_2, the map $\mathbf{J}$ became constant.
We substituted the 8 linear equations found in the previous steps into the public key and obtained a new public key $\tilde{E}''$ of 33 equations in 17 variables, which proved to be of the form of a $C^\star$ public key.
In the last step of the attack, we attacked the new public key $\tilde{E}''$ with the Linearization Equation attack of Patarin [Pat95]. We computed 500 plaintext/ciphertext pairs and substituted them into the Linearization Equation of type (12). By doing so, we got 25 linear independent equations of type (12). After substituting the ciphertext $\mathbf{y}'$ we obtained 17 linear equations in the plaintext variables which enabled us to reconstruct the plaintext $\mathbf{x}'$.
The running time of the whole attack was about 90 seconds.

Case of H_2 In the first step we computed 1100 ($> (n(m+15)+n+m+15+1) = 1066$) plaintext/ciphertext pairs and substituted them into the Linearization Equation of type (12). We solved the resulting linear system for the variables a_{ij}, b_i, c_j and d to find a basis of all FOLEs. The complexity of this step is equal to $1066^3 \leq 2^{31}$. It took about 104 seconds in our experiments, where it took about 102 seconds on generating the plaintext/ciphertext pairs and about 2 second on the Gaussian Elimination. This step has to be performed for each public key only once.
By doing so, we found 5 linear independent Linearization Equations. After substituting the ciphertext $\mathbf{y}'$ into these equations we obtained 4 linear equations in

the plaintext variables. We denote the 21 dimensional subspace satisfying these equations by S_1.
In the second step we computed 100 plaintext/ciphertext pairs where the plaintext was randomly chosen in S_1 and substituted them into the Linearization Equation of type (22). By doing so, we got 14 linear independent equations of form (22). After substituting the ciphertext $\mathbf{y}'$, we got 6 linear equations in the plaintext variables. We denote the 15 dimensional subspace of S_1 satisfying these equations by S_2.
In the third step we computed again 100 plaintext/ciphertext pairs, where the plaintext was chosen randomly in S_2 and substituted them into the Linearization Equation of type (24). We obtained 25 linear independent equations. By substituting the ciphertext $\mathbf{y}'$ into these equations, we got 5 linear equations in the plaintext variables. We denote the 10 dimensional subspace of S_2 satisfying these equations by S_3. For the elements of S_3, the map $\mathbf{J}$ became constant.
We substituted the 15 linear equations found in the previous steps into the public key and obtained a new public key $\tilde{E}'''$ of 40 equations in 10 variables, which proved to be of the form of a $C^\star$ public key.
In the last step of the attack, we attacked the new key $\tilde{E}'''$ with the Linearization Equation attack of Patarin. We computed 500 plaintext/ciphertext pairs and substituted them into the Linearization Equation of type (12). By doing so, we obtained 25 linear independent equations. After substituting the ciphertext $\mathbf{y}'$ we got 10 linear equations in the plaintext variables which enabled us to reconstruct the plaintext $\mathbf{x}'$.
The running time of the whole attack was about 127 seconds.

All experiments are performed on a server with 24 AMD Opteron processors and 128 GB RAM. However, for our experiments we used only a single core. The attack was programmed in Magma code and required about 120 MB of memory. The algorithms of the attack are listed in appendix of this paper.

Remark. Most of the time in our attack is consumed in generating the plaintext/ciphertext pairs. Therefore, with a more sophisticated implementation of the Piece in Hand scheme, the running time of our attack can be decreased drastically.

5 Conclusion

In this paper, we presented the cryptanalysis of two examples of the 2-layer nonlinear Piece in Hand method. Both examples do not really enhance the security of the underlying MPKC because they can not resist Linearization Equation attacks. From this paper, we find that the security of the 2-layer nonlinear Piece in Hand method depends mainly on the construction of the auxiliary polynomial vector $\boldsymbol{H}$. We should therefore design the auxiliary polynomial vector $\boldsymbol{H}$ in such a way that it resists existing attacks.

References

[Ding04] J. Ding. A new variant of the Matsumoto-Imai cryptosystem through perturbation. *Public key Cryptography, (PKC'04)*, LNCS, volume 2947, pages 305–318. Springer, 2004.

[DS05a] J. Ding, and D. Schmidt. Cryptanalysis of HFEV and the internal perturbation of HFE. *Public key Cryptography - (PKC'05)*, LNCS, volume 3386, pages 288-301 Springer, 2005.

[DS05b] J. Ding, and D. Schmidt. Rainbow, a new multivariate public key signature scheme. *The Third International Conference of Applied Cryptography and Network Security (ACNS 2005)*, LNCS, volume 3531, pages 164-175, Springer, 2005.

[DHN07] J. Ding, L. Hu, X. Nie, et al.. High Order Linearization Equation (HOLE) Attack on Multivariate Public Key Cryptosystems, Advance in PKC 2007, LNCS, volume 4450, pages 233-248. Springer, 2007.

[DFSS07] V. Dubois, P. Fouque, A. Shamir and J. Stern. Practical Cryptanalysis of SFLASH. *Advance in Cryptology - CRYPTO 2007*, LNCS, volume 4622, pages 1-12. Springer, 2007.

[DGS07] V. Dubois, L. Granboulan and J. Stern. Cryptanalysis of HFE with Internal Perturbation. *Public Key Cryptography - PKC 2007*, LNCS, volume 4450, pages 249-265. Springer, 2007.

[FGS05] P.-A. Fouque, L. Granboulan, and J. Stern. Differential Cryptanalysis for Multivariate Schemes Advances in Cryptology - EUROCRYPT 2005, LNCS, volume 3494, Page 341-353, Springer 2005, .

[FTT08] R. Fujita, K. Tadaki, and S. Tsujii. Nonlinear piece in hand perturbation vector method for enhancing security of multivariate public key cryptosystems. *Proc. PQCrypto 2008*, LNCS, Vol.5299, pp.148C164, Springer, 2008.

[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.

[NHL06] 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.

[Pat96] J. Patarin. Hidden field equations (HFE) and isomorphism of polynomials (IP): Two new families of asymmetric algorithms. In U. Maurer, editor, *Eurocrypt'96, LNCS*, volume 1070, pages 33–48. Springer, 1996.

[PCG01] J. Patarin, N. Courtois, and L. Goubin. Flash, a fast multivariate signature algorithm. *Progress in Cryptology, CT-RSA 2001*, LNCS, volume 2020, pages 298–307. Springer, 2001.

[PGC98] J. Patarin, L. Goubin and N. Courtois. C^*_{-+} and HM: variations around two schemes of T. Matsumoto and H.Imai. *Advances in Cryptology - ASIACRYPT'98*, LNCS, volume 1514, pages 35-50. Springer, 1998.

[TTF04] S. Tsujii, K. Tadaki, and R. Fujioka. Piece in Hand concept for enhanceing the security of multivariate type pulic key cryptosystem: public key without containing all the information of secret key. IACR eprint 2004/366, http://eprint.iacr.org.

[TTF06] S. Tsujii, K. Tadaki, and R. Fujioka. Proposal for piece in hand matrix ver.2: General concept for enhancing security of multivariate pulic key cryptosystem. IACR eprint 2006/051, http://eprint.iacr.org.

[TTF09] S. Tsujii, K. Tadaki,R. Fujita et al. Security Enhancement of Various MPKCs by 2-layer Nonlinear Piece in Hand Method. *IEICE TRANS. Fundamentals*, Vol. E92-A, NO. 10, pages 2438-2447, 2009.

[WYH06] L. Wang, B. Yang, Y. Hu, et al. A Medium-Field Multivariate Public key Encryption Scheme, CT-RSA 2006: The Cryptographers' Track at the RSA Conference 2006, LNCS, volume 3860, pages 132-149, Springer, 2006.

[WZW11] H. Wang, H. Zhang, Z. Wang and M. Tang. Extended multivariate public key cryptosystems with secure encryption function. SCIENCE CHINA Information Sciences, June 2011 Vol. 54 No. 6: 1161C1171.

Appendix: The algorithms of our attack

Algorithm 1 Cryptanalysis of 2-layer nonlinear Piece in Hand with auxiliary map of form H_1

Input: public key $\tilde{E}$ of Piece in Hand, ciphertext $\mathbf{y}' \in \mathbb{F}^{n+l}$
Output: corresponding plaintext $\mathbf{x}' \in \mathbb{F}^n$

1: Compute $D_1 \geq n \cdot (m+l) + n + m + l + 1$ plaintext/ciphertext pairs $(\mathbf{x}^{(i)}/\mathbf{y}^{(i)})$.
2: Compute all Linearization Equations of the form $\sum a_{ij}x_iy_j + \sum b_ix_i + \sum c_jy_j + d = 0$. This can be done once for a given public key.
3: Substitute the ciphertext $\mathbf{y}'$ into the linearization equations and find t_1 linear equations in the plaintext variables. Denote the solution space of these equations by S_1.
4: Compute $D_2 \geq (m+l)\cdot(n-t_1)$ plaintext/ciphertext pairs $(\mathbf{x}^{(i)}/\mathbf{y}^{(i)})$. The plaintexts are chosen from S_1.
5: Compute for $i = 1, \ldots, D_2$ equations of the form $\sum a_j y_j^{(i)} + \sum b_j (x_j^{(i)})^2 + d = 0$. Solve this system for a_j, b_j and d. Denote the basis of the solution space by B.
6: By substituting the ciphertext $\mathbf{y}'$ one gets for the elements of B linear equations in the plaintext variables of the form $\sum b_j^{1/2} x_i' + (\sum a_j y_j' + d)^{1/2} = 0$. There will be t_2 linear independent ones. Denote the solution space of these equations by S_2. For the elements of S_2 the maps $\mathbf{J}$ and $\mathbf{H}_1$ become constants.
7: Substitute the linear equations found in steps 3 and 6 into the public key. The simplified public key will have the form of a public key of the underlying multivariate scheme.
8: Use an attack against the underlying MPKC to recover the plaintext $\mathbf{x}'$.

Algorithm 2 Cryptanalysis of 2-layer nonlinear Piece in Hand with auxiliary map of form H_2

Input: public key $\tilde{E}$ of Piece in Hand, ciphertext $\mathbf{y}' \in \mathbb{F}^{n+l}$
Output: corresponding plaintext $\mathbf{x}' \in \mathbb{F}^n$

1: Compute $D_1 \geq n \cdot (m+l) + n + m + l + 1$ plaintext/ciphertext pairs $(\mathbf{x}^{(i)}/\mathbf{y}^{(i)})$.
2: Compute all linearization equations of the form $\sum a_{ij}x_iy_j + \sum b_ix_i + \sum c_jy_j + d = 0$. This can be done once for a given public key.
3: Substitute the ciphertext $\mathbf{y}'$ into the linearization equations and find t_1 linear equations in the plaintext variables. Denote the solution space of these equations by S_1.
4: Compute $D_2 \geq (m+l)\cdot(n-t_1)$ plaintext/ciphertext pairs $(\mathbf{x}^{(i)}/\mathbf{y}^{(i)})$. The plaintexts are chosen from S_1.
5: Compute for $i = 1, \ldots, D_2$ equations of the form $\sum a_j y_j^{(i)} + \sum b_j (x_j^{(i)})^2 + d = 0$. Solve this system for a_j, b_j and d. Denote the basis of the solution space by B_1.
6: By substituting the ciphertext $\mathbf{y}'$ one gets for the elements of B_1 linear equations in the plaintext variables of the form $\sum b_j^{1/2} x_i' + (\sum a_j y_j' + d)^{1/2} = 0$. There will be t_2 linear independent ones. Denote the solution space of these equations by S_2.
7: Compute $D_3 \geq (m+l) \cdot (n - t_1 - t_2)$ plaintext/ciphertext pairs $(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})$. The plaintexts are chosen from S_2.
8: Compute for $i = 1, \ldots, D_3$ equations of the form $\sum a_j y_j^{(i)} + \sum b_j x_j^{(i)} + d = 0$. Solve this system for a_j, b_j and d. Denote the basis of the solution space by B_2.
9: By substituting the ciphertext $\mathbf{y}'$ one gets for the elements of B_2 linear equations in the plaintext variables of the form $\sum b_j m_j + \sum a_j c_j + d = 0$. There will be t_3 linear independent ones. Denote the solution space of these equations by S_3. For the elements of S_3 the maps $\mathbf{J}$ and $\mathbf{H}_2$ become constant.
10: Substitute the linear equations found in steps 3, 6 and 9 into the public key. The simplified public key will have the form of a public key of the underlying multivariate scheme.
11: Use an attack against the underlying MPKC to recover the plaintext $\mathbf{x}'$.