A DEFECT OF THE IMPLEMENTATION SCHEMES OF THE TTM CRYPTOSYSTEM

ABSTRACT. We show all the existing TTM implementation schemes have a defect that there exist linearization equations

$$\sum_{i=1,j=1}^{n,m} a_{ij}x_iy_j(x_1,\ldots,x_n)+\sum_{i=1}^{n}b_ix_i+\sum_{j=1}^{m}c_jy_j(x_1,\ldots,x_n)+d=0,$$

which are satisfied by the components $y_i(x_1,\ldots,x_n)$ of the ciphers of the TTM schemes. We further demonstrate that, for the case of the most recent two implementation schemes in two versions of the paper [CM], where the inventor of TTM used them to refute a claim in [CG], if we do a linear substitution with the linear equations derived from the linearization equations for a given ciphertext, we can find the plaintext easily by an iteration of the procedure of first search for linear equations by linear combinations and then linear substitution. The computation complexity of the attack on these two schemes is less than 2^{35} over a finite field of size 2^8.

Keywords: open-key, multivariable, quadratic polynomials, linearization

1. Introduction

Recently new methods were invented to construct multivariable cryptosystems, namely cryptosystems based on multivariable functions instead of single variable functions. The security of such systems in general relies on how difficult it is to solve polynomial equations with many variables, a proven NP-hard problem in general.

Matsumoto and Imai suggested one of the first constructions of such cryptosystems [MI], which unfortunately has been defeated [P]. Another interesting one is the TTM cryptosystem [M]. Despite the inventor's claim that TTM systems are very secure from all standard attacks, in [CG], the authors claimed that they completely defeated TTM schemes using the Minirank method and demonstrated it by defeating one of the challenges set by the inventor of TTM, however the inventor of TTM refuted the claim with [CM], where they gave a new implementation scheme to support their claim. In [DH], another method was found to defeat the first TTM implementation scheme in [M]. Though this new method can also be applied to other TTM implementation schemes [CGC], it can not be directly applied to all existing implementation schemes, such as the new ones in two versions of [CM]. In this article, we will show that actually all existing implementation schemes by now for the TTM cryptosystem have a common defect that could make them insecure. For the case of the most recent two TTM implementation schemes in two different versions of the paper [CM], we use this defect to defeat the schemes.

The key idea comes from an observation that we can also extend the linearization method by Patarin[P] to attack all the TTM implementation schemes.

For all the TTM implementation schemes, a cipher F is made of m degree two polynomials of n variables on a finite field K of characteristic 2, namely,

$$F(x_1,\ldots,x_n)=(y_1(x_1,\ldots,x_n),\ldots,y_m(x_1,\ldots,x_n)), \tag{1.1}$$

where $m>n$. This cipher, the m polynomials y_i, are made public. The cipher F is given as a map from K^n to K^m and it is derived as

$$F=\phi_4\circ\phi_3\circ\phi_2\circ\Phi_1, \tag{1.2}$$

where $\circ$ denotes a composition of maps, ϕ_4 and Φ_1 are affine linear maps, ϕ_4 is an invertible map from K^m to K^m, Φ_1 is an injective map from K^n to K^m and ϕ_3 and ϕ_2 are nonlinear maps of the de Jonquiere type on K^m. Given an element $X = (z_1, \ldots, z_m)$ in K^m, a de Jonquiere map $J(X)$ is defined as a map from K^m to K^m:

$$J(X) = (z_1 + g_1(z_2, \ldots, z_m), z_2 + g_2(z_3, \ldots, z_m), \ldots, z_{m-1} + g_{m-1}(z_m), z_m), \tag{1.3}$$

where g_i are polynomial functions.

A linearization equation is an equation in the form of

$$\sum_{i=1,j=1}^{n,m} a_{ij} x_i y_j(x_1, \ldots, x_n) + \sum_{i=1}^{n} b_i x_i + \sum_{j=1}^{m} c_j y_j(x_1, \ldots, x_n) + d = 0, \tag{1.4}$$

which is satisfied by the set of polynomials y_i of the cipher F and its variables x_i. This equation was first used by Patarin successfully to attack the Matsumoto-Imai cryptosystems, which we call the linearization method.

From the construction of the TTM implementation schemes, we found that, for all the existing TTM implementation schemes, there exist a lot of linearization equations that are satisfied by the quadratic polynomials y_i of the TTM cipher F. For example, for the most recently proposed implementation scheme [CM](the revised version on IACR e-Archive, the former version has a different implementation scheme), where $m = n+52$, we found all the linearization equations and computed that the dimension of V is actually 347, where V is the linear space of all the linearization equations satisfied by the quadratic polynomials y_i.

This is the source of the common defect among all the TTM implementation schemes, because the existence of the linearization equations means that given a ciphertext $(y'_1, \ldots, y'_m)$, we can immediately produce some linear equations satisfied by the plaintext $(x'_1, \ldots, x'_n)$, which is something a secure open key cryptosystem should not have. For the case of the revised implementation scheme [CM], we found that, with the probability $1 - \frac{C_{17}^5}{2^{12\times 8}} > 1 - 2^{-82}$, the linearization equations will produce 17 linearly independent linear equations satisfied by x_i.

For this case, we can move one step further by performing a substitution of this 17 linear equations into y_i, which makes y_i quadratic polynomials with 17 fewer variables, which we denote by $(x_{v_1}, \ldots, x_{v_{31}})$. Now F becomes a new map $\hat{F}$ from K^{n-17} to K^m, which in the composition form can be equivalently rewritten as:

$$\hat{F} = \hat{\phi}_4 \circ \phi_2 \circ \hat{\Phi}_1, \tag{1.5}$$

where $\hat{\phi}_4$, which is invertible, and $\hat{\Phi}_1$, which is injective, are some affine linear maps. The procedure of the substitution of the 17 linear equations eliminates one of the composition factors of the de Jonquiere type. Then solving the equations

$$\hat{F} = (y'_1, \ldots, y'_m)$$

for the given ciphertext becomes straightforward because of the triangular form of the de Jonquiere type of maps and it is accomplished by an iteration of the procedure of first search for linear equations by linear combinations and then linear substitution. Finally the plaintext can be derived by substituting the solution of the values of $(x_{v_1}, \ldots, x_{v_{31}})$ into the original 17 linear equations. For the practical example $m = 100$ proposed in [CM], we can show that it takes about 2^{32} computations on a finite field of size 2^8 to defeat the scheme and we performed a computation example on a PC (450MHz) and defeated it in a few hours. Similarly, we can defeat the TTM scheme in the original version of [CM].

We arrange the paper in the following way. In Section 2, we will first discuss the basic idea of TTM. Then, we will present the details of our attacks on two different implementation schemes of the TTM: the first one is the one in the revised version (July 2002)of [CM], the second one is the one suggested in the first version of (August 2001) [CM]; and we will discuss the other cases including the first TTM implementation scheme [M]. In Section 3, we will present the conclusion.

2. The common defect of the TTM schemes.

2.1. **Basic technical idea of the TTM schemes.** Let $\bar{F}(x_1,\ldots,x_m)$ be a map on the space K^m. $\bar{F}(x_1,\ldots,x_m)$ is a composition of several maps G_i on K^m, i=1,...,k :

$$\bar{F} = G_1 \circ G_2 \ldots . \circ G_k,$$

which has the following properties:

(I)$\bar{F}(x_1,\ldots,x_m)$ is easy and fast to compute if we are given specific value of all x_i.

(II) The factorization of F in terms of composition of G_i is very difficult to compute if we only know expanded version of $\bar{F}(x_1,\ldots,x_m)$, $\bar{F}^{-1}$ is very difficult to compute without such a decomposition, and G_i are very easy to invert.

With such a $\bar{F}(x_1,\ldots,x_m)$ and if the equations $\bar{F}(x_1,,,,x_m) = (a_1,\ldots,a_m)$ is impossible to solve directly, we can use $\bar{F}$ to build an open-key public cryptosystem. The Matsumoto-Imai construction [MI] is an attempt of such a type of construction.

For the TTM construction, one uses only the following two types of maps.

1) The Linear Type.

Given the space K^m, we can apply all invertible affine linear maps to the m variables: $f(X) = aX + b$, where a is a $m \times m$ invertible matrix, and X and b are in K^m.

2) The de Jonquiere Type

These maps give isomorphisms of the corresponding polynomial rings, which are called the tamed transformation in algebraic geometry, and they can be easily inverted. TTM stands for the Tamed Transformation Method.

However due to the consideration of the size of public key and the complexity of public computations, any practical and efficient system requires to have the polynomial components of the cipher to be of degree 2, which seems to be very difficult to accomplish.

In [M], a quadratic construction is obtained by instead using the map

$$F(x_1,\ldots,x_n) = \bar{F}(x_1,\ldots,x_n,0,0,\ldots,0),$$

where

$$\bar{F}(x_1,\ldots,x_m) = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1,x_2,\ldots,x_m),$$

ϕ_1 and ϕ_4 are of invertible linear type, ϕ_3, ϕ_2 are of the de Jonquiere type, ϕ_2 is of degree 2 and ϕ_3 is of a high degree (8). This map F, which can be viewed as a map from K^n to K^m, is an "invertible" map in the sense that it is injective, and given any element in the image of F, we can use $\bar{F}^{-1}$ to recover its preimage easily.

The key component of the construction of the TTM systems is based on a special multivariable polynomial $Q_8(z_1,\ldots,z_l)$ and a special set of quadratic polynomials $q_i(z_1,\ldots,z_k)$, $i = 1,\ldots,l$, such that $Q_8(q_1,\ldots,q_l)$ is still quadratic in z_i. Though the constructions of the TTM schemes are very interesting from a both theoretical and practical point of view, in particular from the point view of algebraic geometry, no principle was given about how Q_8 and q_i are constructed. Our attack starts from an observation of a special property of the polynomials Q_8 and q_i.

2.2. **The case of implementation scheme in the revised** [CM].

2.2.1. *The scheme.* In this subsection, we will use essentially the notation in the revised version of [CM].

First the finite field K is of size 2^8, and $m = n + 52$. The map $\bar{F}$ is made of ϕ_1, ϕ_2, ϕ_3, ϕ_4; $\bar{F} = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1, x_2, \dots, x_{n+52})$, which are maps from the (n+52) dimensional space to itself defined in [CM]. $\phi_1 = (\phi_{1,1}, \dots, \phi_{1,n+52})$, $\phi_4 = (\phi_{4,1} \dots, \phi_{4,n+52})$ are invertible affine linear maps, and $\phi_{1i} = x_i$, for $i > n$; ϕ_2 and ϕ_3 are nonlinear maps of the de Jonquiere type.

The map

$$F(x_1, \dots, x_n) = (y_1, \dots, y_{n+52}) =$$
$$\phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1, x_2, \dots, x_n, 0, \dots, 0) = \phi_4 \circ \phi_3 \circ \phi_2 \circ \Phi_1(x_1, \dots, x_n)$$

is the cipher, which is public, but ϕ_1, ϕ_4 are private. $\Phi_1(x_1, \dots, x_n) = \phi_1(x_1, x_2, \dots, x_n, 0, \dots, 0)$ is an injective map from K^n to K^{n+52}. In the expansion formula, the components y_i of the map F are degree two polynomials of variables $(x_1, \dots, x_n)$.

To attack this cryptosystem is to solve the set of equations

$$y_i(x_1, \dots, x_n) = y'_i$$

for $i = 1, \dots, m$, with the variables x_j, $j = 1, \dots, n$ and an element in K^m: $(y'_1, \dots, y'_{n+52})$. Here $(y'_1, \dots, y'_{n+52})$ can be viewed as the ciphertext, and the solution $(x'_1, \dots, x'_n) \in K^n$ the plaintext.

In [CM], it is claimed that, if $n = 48 (m = 100)$, no practical methods can work efficiently to attack such a system, in particular, the Minirank method in [CG], and the complexity of the attack by Minirank method is far bigger that 2^{84}.

In this scheme, $\phi_2(x_1, \dots, x_n) = (\phi_{2,1}, \dots, \phi_{2,100})$ is given as:

$$\phi_{2,i} = x_i, \mathrm{i} = 1;$$
$$\phi_{2,i} = x_i + f_i(x_1, \dots, x_{i-1}), \mathrm{i} = 2, 3, \dots, 41;$$
$$\phi_{2,i} = q_{i-41}(x_{38}, \dots, x_{48}), \mathrm{i} = 42, \dots, 48;$$
$$\phi_{2,i} = x_i + q_{i-41}(x_{38}, \dots, x_{48}), \mathrm{i} = 49, \dots, 76;$$
$$\phi_{2,i} = x_i + q_{i-72}(x_{36}, x_{39}, x_{40}, \dots, x_{45}, x_{37}, x_{47}, x_{48}) \mathrm{i} = 77, \dots, 84;$$
$$\phi_{2,i} = x_i + q_{i-80}(x_{34}, x_{39}, x_{40}, \dots, x_{45}, x_{35}, x_{47}, x_{48}) \mathrm{i} = 85, \dots, 92;$$
$$\phi_{2,i} = x_i + q_{i-88}(x_{32}, x_{39}, x_{40}, \dots, x_{45}, x_{33}, x_{47}, x_{48}) \mathrm{i} = 93, \dots, 100;$$

where a_1 are a_3 can be any nonzero number in the field K,

$$Q_8(z_1, \dots, z_{35}) = (z_5 z_{13} + z_8 z_{14})(z_{19} z_{32} + z_2(z_{18} + z_{24}))^2 (z_{20} z_{19} + z_{23} z_{18}) +$$
$$(z_{32} z_3 + (z_{18} + z_{24}) z_{21})^2 (z_{22} z_{19} + z_{23} z_{24})(z_9 z_{13} + z_8 z_{15}) +$$
$$a_1^8((z_{25} z_{26} + z_{27} z_{28})(z_6 z_{29} + z_7 z_{16}) + (z_{10} z_{30} + z_{11} z_{31})(z_{17} z_1 + z_{18} z_4)) + a_1^{12}(z_6 z_{33} + z_{34} z_7 + z_5 z_{35} + z_{14} z_{12}),$$

$$(q_1(z_1, \dots, z_{11}), q_2(z_1, \dots, z_{11}), \dots, q_{35}(z_1, \dots, z_{11})) =$$
$$(z_4 z_2 + a_1 z_5, z_3 z_4 + a_1 z_6, z_2 z_5 + a_1 z_7, z_4 z_7 + a_1 z_8, z_1 z_5 + a_1 z_9,$$
$$z_1 z_2 + a_1 z_{10}, z_2 z_9 + a_1 z_{11}, z_3 z_9 + a_1 z_1, z_1 z_3, z_1 z_7 + a_1 z_9,$$
$$z_4 z_9 + a_1 z_1, z_7 z_9 + a_1 z_1, z_3 z_{11} + a_1 z_{10}, z_5 z_{10} + a_1 z_{11}, z_3 z_{10},$$
$$z_2 z_{10}, z_7 z_8 + a_1 z_7, z_5 z_7 + a_1 z_2, z_2 z_3 + a_1 z_7, z_5 z_8 + a_1 z_5,$$
$$z_4 z_5 + a_1 z_6, z_3 z_8, z_3 z_5 + a_1 z_8, z_3 z_7, z_6 z_8 + a_3 z_5,$$
$$z_2 z_6, z_5 z_6, z_6 z_7 + a_3 z_2, z_2 z_{11}, z_4 z_{11} + a_1 z_{10},$$
$$z_7 z_{10} + a_1 z_{11}, z_3 z_6 + z_5 z_6 + a_1 z_4, z_8 z_{11}, z_8 z_{10}, z_7 z_{11} + a_1 z_{10}),$$

$f_i(x_1, \dots, x_{i-1})$ are randomly chosen quadratic functions.

And $\phi_3(x_1, \dots, x_n) = (\phi_{3,1}, \dots, \phi_{3,100})$ is given as:

$$\phi_{3,i} = x_i, \mathrm{i} = 5, \dots, 100;$$

$$\phi_{3,4} = x_4 + R_i(x_1, \ldots, x_{100}), \mathrm{i} = 1, 2, 3, 4;$$

where

$$R_i(x_1, \ldots, x_{100}) = \sum_1^4 \beta_{ij} P_j,$$

which are linearly independent and P_i, for $i = 1, 2, 3$, is given as

$$P_i = Q_8(x_{42}, \ldots, x_{45}, x_{101-8i}, \ldots, x_{108-8i}, x_{54}, \ldots, x_{76});$$

and

$$P_4 = Q_8(x_{42}, \ldots, x_{76}).$$

Remark In the new version of [CM], the polynomials Q_8 and q_i actually have three free parameters a_1, a_2 and a_3. We checked the formulas and found out that in order to make the cipher F to be of degree 2, one must make a_1 equal to a_2. We impose this condition on this implementation scheme.

Because the specific form of ϕ_1, we can write:

$$\phi_1(x_1, x_2, \ldots, x_{48}, 0, \ldots, 0) = \Phi_1(x_1, \ldots, x_{48}) = \pi \circ \hat{\phi}_1(x_1, \ldots, x_{48}),$$

where π is the standard embedding that maps K^{48} into K^{100}:

$$\pi(x_1, \ldots, x_{48}) = (x_1, \ldots, x_{48}, 0, 0, \ldots, 0),$$

and $\hat{\phi}_1(x_1, \ldots, x_{48}) = (\hat{\phi}_{1,1}(x_1, \ldots, x_{48}), \ldots, \hat{\phi}_{1,48}(x_1, \ldots, x_{48}))$ is an invertible affine linear transformation from K^{48} to itself.

Let $\phi_3 \circ \phi_2 \circ \pi = \bar{\phi}_{32}$, then

$$F(x_1, \ldots, x_{48}) = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1, x_2, \ldots, x_{48}, 0, \ldots, 0) =$$
$$\phi_4 \circ \phi_3 \circ \phi_2 \circ \pi \circ \hat{\phi}_1(x_1, \ldots, x_{48}) = \phi_4 \circ \bar{\phi}_{32} \circ \hat{\phi}_1(x_1, \ldots, x_{48}).$$

Let $\bar{\phi}_{32}(x_1, \ldots, x_{48}) = (\bar{\phi}_{32,1}, \ldots, \bar{\phi}_{32,100})$, then

$$\bar{\phi}_{32,i} = x_i + a_1^{14}\beta_{i4}(x_{38}x_{48} + x_{47}x_{46}) + a_1^{14} \sum_1^3 \beta_{ij}(x_{38-2j}x_{48} + x_{39-2j}x_{47}), \tag{2.1}$$

for i =1;

$$\bar{\phi}_{32,i} = x_i + f_i(x_1, .., x_{i-1}) + a_1^{14}\beta_{i4}(x_{38}x_{48} + x_{37}x_{46}) + a_1^{14} \sum_1^3 \beta_{ij}(x_{38-2j}x_{48} + x_{39-2j}x_{47}), \tag{2.2}$$

for $i = 2, 3, 4$;

$$\bar{\phi}_{32,i} = x_i + f_i(x_1, .., x_{i-1}), \mathrm{i} = 5, 6, \ldots, 41;$$
$$\bar{\phi}_{32,i} = q_{i-31}(x_{38}, \ldots, x_{48}), \mathrm{i} = 42, \ldots, 48;$$
$$\bar{\phi}_{32,i} = q_{i-31}(x_{38}, \ldots, x_{48}), \mathrm{i} = 49, \ldots, 76;$$
$$\bar{\phi}_{32,i} = q_{i-72}(x_{36}, x_{39}, x_{40}, \ldots, x_{45}, x_{37}, x_{47}, x_{48})\mathrm{i} = 77, \ldots, 84;$$
$$\bar{\phi}_{32,i} = q_{i-85}(x_{34}, x_{39}, x_{40} \ldots, x_{45}, x_{35}, x_{47}, x_{48})\mathrm{i} = 85, \ldots, 92;$$
$$\bar{\phi}_{32,i} = q_{i-93}(x_{32}, x_{39}, x_{40}, \ldots, x_{45}, x_{33}, x_{47}, x_{48})\mathrm{i} = 93, \ldots, 100.$$

The formula above is due to the fact that

$$Q_8(q_1, \ldots, q_{35}) = a_1^{14}(z_9 z_{10} + z_1 z_{11}),$$

which is the reason F is of degree 2.

2.2.2. *The basic idea of the cryptanalysis.* Our attack starts from the observation that all the q_i are very simple quadratic polynomials, which have only one quadratic term. In this case, Q_8 has 35 variables and q_i has 11 variables, and we have

$$q_9 = z_1 z_3, \quad q_{15} = z_3 z_{10}.$$

This implies that

$$z_{10} q_9 - z_1 q_{15} = 0. \tag{2.3}$$

In this implementation scheme, the map $\bar{\phi}_{32}$ has actually 4 sets of q_i as its components (with intersections). Because, F is derived from $\bar{\phi}_{32}$ by composing from both the left side and the right side by an invertible linear map, the equation (2.3) above implies that we must have linearization equations for the y_i, the components of F. This means there is a possibility to actually use such linearization equations to attack this scheme, which is the only method used by Patarin to defeat the Matsumoto-Imai scheme.

Let V denote the linear space of the linearization equations

$$\sum_{i=1,j=1}^{n,m} a_{ij} x_i y_j(x_1, \ldots, x_n) + \sum_{i=1}^{n} b_i x_i + \sum_{j=1}^{m} c_j y_j(x_1, \ldots, x_n) + d = 0,$$

satisfied by y_i of F and let D be its dimension.

Let $\bar{V}$ denote the linear space of the linearization equations satisfied by $\bar{\phi}_{32,i}(x_1, \ldots, x_{48})$ of $\bar{\phi}_{32}$:

$$\sum_{i=1,j=1}^{n,m} \bar{a}_{ij} x_i \bar{\phi}_{32,j}(x_1, \ldots, x_{48}) + \sum_{i=1}^{n} \bar{b}_i x_i + \sum_{j=1}^{m} \bar{c}_j \bar{\phi}_{32,j}(x_1, \ldots, x_{48}) + \bar{d} = 0,$$

and let $\bar{D}$ be the dimension of $\bar{V}$.

Let $\hat{\phi}_{32}(x_1, \ldots, x_{48}) = (\hat{\phi}_{32,1}, \ldots, \hat{\phi}_{32,100}) = \bar{\phi}_{32} \circ \hat{\phi}_1(x_1, \ldots, x_{48})$.

Let $\hat{V}$ denote the linear space of the linearization equations satisfies by $\hat{\phi}_{32,i}(x_1, \ldots, x_{48})$ of $\hat{\phi}_{32}$:

$$\sum_{i=1,j=1}^{n,m} \hat{a}_{ij} x_i \hat{\phi}_{32,j}(x_1, \ldots, x_{48}) + \sum_{i=1}^{n} \hat{b}_i x_i + \sum_{j=1}^{m} \hat{c}_j \hat{\phi}_{32,j}(x_1, \ldots, x_{48}) + \hat{d} = 0,$$

and let $\hat{D}$ be the dimension of $\hat{V}$.

Let $\phi_{4,i}$ denote the components of ϕ_4 and $\hat{\phi}_{1,i}$ denote the components of $\hat{\phi}_1$. Let $(\phi^{-1})_{4,i}$ denote the components of ϕ_4^{-1} and $(\hat{\phi}^{-1})_{1,i}$ denote the components of $\hat{\phi}_1^{-1}$.

Let M be the map from $\hat{V}$ to V given by:

$$M : (\Sigma \hat{a}_{ij} x_i \hat{\phi}_{32,j}(x_1, \ldots, x_{48}) + \Sigma \hat{b}_i x_i + \Sigma \hat{c}_j \hat{\phi}_{32,j}(x_1, \ldots, x_{48}) + \hat{d} = 0) \rightarrow$$

$$(\Sigma \hat{a}_{ij} x_i (\phi)_{4,j}^{-1}(y_1((x_1, \ldots, x_{48}), .., y_{100}(x_1, \ldots, x_{48})) + \Sigma \hat{b}_i x_i +$$

$$\Sigma \hat{c}_j (\phi)_{4,j}^{-1}(y_1((x_1, .., x_{48}), .., y_{100}(x_1, \ldots, x_{48})) + \hat{d} = 0).$$

Let $\hat{M}$ be the map from $\bar{V}$ to $\hat{V}$ given by:

$$\hat{M} : (\Sigma \bar{a}_{ij} x_i \bar{\phi}_{32,j}(x_1, \ldots, x_{48}) + \Sigma \bar{b}_i x_i + \Sigma \bar{c}_j \bar{\phi}_{32,j}(x_1, \ldots, x_{48}) + \bar{d} = 0) \rightarrow$$

$$(\Sigma \bar{a}_{ij} \hat{\phi}_{1,i}(x_1, \ldots, x_{48}) \hat{\phi}_{32,j}(x_1, \ldots, x_{48}) + \Sigma \bar{b}_i \hat{\phi}_{1,i}(x_1, \ldots, x_{48}) + \Sigma \bar{c}_j \hat{\phi}_{32,j}(x_1, \ldots, x_{48}) + \bar{d} = 0).$$

Theorem 1. *M and $\hat{M}$ are invertible linear maps and $D = \bar{D} = \hat{D}$.*

The proof follows from the fact that both ϕ_4 are $\bar{\phi}_1$ are invertible affine linear maps. Essentially the map $\hat{M}$ is a change of basis of x_i and the map M is an affine linear transformation of the substitution of $\hat{\phi}_{32,i}$ by y_i.

This means that we only need to find $\bar{D}$ to find D and we did so by computations.

First we choose the field K to be $K = \mathbf{Z}_2[x]/x^8 + x^5 + x + 1$. Because a_1 and a_3 can be any nonzero constants, we choose them both to be 1. Then we choose $f_i(x_1, \ldots, x_{i-1})$, i=2,...,41, randomly quadratic polynomials over K and β_{ij} randomly in K (but satisfying the condition R_i are linearly independent). We chose 10 different sets of $f_i(x_1, \ldots, x_{48})$ and β_{ij} for testing. For all these 10 choices, our computation showed that

a) $\bar{D} = 347$.

b)All the linearization equations are in the form of

$$\Sigma_{i>31}\Sigma_{j>41}\bar{a}_{ij}x_i\bar{\phi}_{32,j}(x_1, \ldots, x_{48}) + \Sigma_{i>31}\bar{b}_ix_i + \Sigma_{j>41}c_j\bar{\phi}_{32,j}(x_1, \ldots, x_{48}) = 0,$$

and the polynomials $\phi_{32,j}(x_1, \ldots, x_{48}), j > 41$ are polynomials of only 17 variables x_i, $i > 31$.

Though we have such a large number of linearization equations, but we are not sure that if we are given the ciphertext y'_i, how many linearly independent equations they will produce.

Let $(x'_1, \ldots, x'_{48})$ be an element in K^{48}. Let $y'_i = y_i(x'_1, \ldots, x'_{48})$, $\hat{\phi}'_{32,i} = \hat{\phi}_{32,i}(x'_1, \ldots, x'_{48})$.

Let U be the space of linear equations derived from substitution of y_i by the values y'_i in V.

Let $\hat{U}$ be the linear space of linear equations derived from substitution of $\hat{\phi}_{32,i}$ by the values $\hat{\phi}'_{32,i}$ in $\hat{V}$.

Let $\bar{U}$ be the linear space of linear equations derived from substitution of $\bar{\phi}_{32,i}$ by the values $\hat{\phi}'_{32,i}$ and x_i by $(\hat{\phi})^{-1}_{1,i}(x_1, \ldots, x_{48})$ in $\bar{V}$.

For a linear equation $\sum_1^{48} a_ix_i + b = 0$, we define $\tilde{M}$ to be the linear map:

$$\tilde{M}(\sum_1^{48} a_ix_i + b = 0) \to (\sum_1^{48} a_i(\hat{\phi})_{1,i}(x_1, \ldots, x_{48}) + b = 0).$$

Theorem 2. The dimension of U is equal to the dimension of $\bar{U}$, the dimension of $\hat{U}$ and the dimension of $\tilde{U}$. $\hat{U} = U = \tilde{M}(\bar{U})$.

This is proven easily by using the maps M and $\hat{M}$.

Because all the linearization relations in $\bar{V}$ are expressed in the last 59 components $\bar{\phi}_{32,j}(x_1, \ldots, x_{48})$, $j > 41$ and they are all expressed in terms of the quadratic polynomial q_i; and they involve only the last 17 variables x_i, $i = 32, \ldots, 48$, we did 200 samples of randomly chosen values $\bar{x}'_{32}$,$\bar{x}'_{33}$,..., $\bar{x}'_{48}$ for $x_{32}, \ldots, x_{48}$, computed the corresponding values of $\bar{\phi}_{32,j}, j > 41$ for these $\bar{x}'_{32}$,$\bar{x}'_{33}$,..., $\bar{x}'_{48}$ and then substituted the values of $\bar{\phi}_{32,j}, j > 41$ into the 347 linearization equations. We found out that these 347 linearization equations in $\bar{V}$ actually produce 17 linearly independent equations of x_i, $i > 31$, and by solving those equations we have $x_i = \bar{x}'_i$, $i = 32, \ldots, 48$.

Then we notice that if all the x_i are set to be zero, which means $\bar{\phi}_{32,i}(0, \ldots, 0) = 0$ for any i, the linearization equations in $\bar{V}$ will not produce 17 linearly independent equations at all. So instead of choosing randomly the values, we chose $(\bar{x}'_{32}, \ldots, \bar{x}'_{48})$ to be the ones with many zeros, and we found out (with 500 random samples) that as long as at least 5 of $x_{32}, \ldots, x_{48}$ are not zero, by substituting the corresponding values of $\bar{\phi}_{32,j}, j > 41$ into the 347 linearly independent linearization equations in $\bar{V}$, these 347 linearization equations will actually produce 17 linearly independent linear equations of $x_i, i > 31$ and by solving those equations we again recover the values of $\bar{x}'_i$ by the solution $x_i = \bar{x}'_i$, $i = 32, \ldots, 48$. Among all the possible values of x_i, $i = 32, \ldots, 48$, the probability that at most 5 of them among x_i, $i = 32, \ldots, 48$ to be non zero is $\frac{C^5_{17}2^{5\times 8}}{2^{17\times 8}} = \frac{C^5_{17}}{2^{12\times 8}} < 2^{-82}$ Therefore we have a probability $1 - \frac{C^5_{17}}{2^{12\times 8}} > 1 - 2^{-82}$ that the linearization equations will produce 17 linearly independent equations for a given set values of $\bar{\phi}_{32,i}$ and solving those equations will recover the value of $\bar{x}'_{32}, \ldots, \bar{x}'_{48}$ if we are given the corresponding values of $\bar{\phi}_{32,j}, j > 41$.

With Theorem 1 and Theorem 2, we conclude that with the probability $1 - \frac{C_{17}^5}{2^{12\times 8}} > 1 - 2^{-82}$, the linearization equations of y_i in V will produce 17 linearly independent equations satisfied by x_i for a given ciphertext $(y'_1, \ldots, y'_{100})$. This is the first step of our attack. Here we would like to emphasize that the statement about the probability to derive 17 linearly independent linear equations from a ciphertext is based on computational experiments not on any theoretical argument and it seems possible to actually prove it.

Let's assume that we now have 17 linearly independent equations in U derived from a ciphertext $(y'_1, \ldots, y'_{100})$ and its substitution in V. Let's $(x'_1, \ldots, x'_{48})$ be the corresponding plaintext. This set of linear equations surely is not enough to recover the original plaintext. However, we know that if we have seventeen linearly independent equations, we can use Gaussian elimination method to find two sets: $A = \{u_1, \ldots, u_{17}\}$, $B = \{v_1, \ldots, v_{31}\}$, $A \cap B = \emptyset$ and $A \cup B = \{1, \ldots, 48\}$, such that we can derive 17 linearly independent linear equations in the form $x_{u_j} = h_j(x_{v_1}, \ldots, x_{v_{31}})$.

Then we substitute these 17 equations into the y_i, which will become quadratic polynomials with only 31 variables. We will call this new set of polynomials $\hat{y}_i$. They can be viewed as components of a map from K^{31} to K^{100}, which will be denoted by $\hat{F}$.

Let ϕ_0 be the map from K^{31} to K^{48}, which is given by:

$$\phi_{0,i}(x_{v_1}, \ldots, x_{v_{31}}) = x_i$$

if $i \in B$, otherwise

$$\phi_{0,i}(x_{v_1}, \ldots, x_{v_{31}}) = h_i(x_{v_1}, \ldots, x_{v_{31}}).$$

Then

$$\hat{F} = \phi_4 \circ \phi_3 \circ \phi_2 \circ \pi \circ \hat{\phi}_1 \circ \phi_0.$$

From the point view of algebraic geometry, the substitution process is nothing but evaluation of the y_i on the variety defined by the 17 linearly independent linear equations $x_{u_i} = h_i(x_{v_1}, \ldots, v_{31})$ and the existing variables are nothing but the coordinates of this variety.

Because for the case of $\bar{\phi}_{32}$, if the dimension of $\bar{U}$ is 17, the variety is defined by $x_i = \bar{x}'_i$ for $i = 32, \ldots, 48$ and $\bar{x}'_i \in K$, with Theorem 1 and Theorem 2, we know that the variety defined by linear equation in U is the same variety defined by $\hat{\phi}_{1,i}(x_1, \ldots, x_{48}) = \hat{\phi}_{1,i}(x'_1, \ldots, x'_{48})$, for $i > 31$ and we denote this variety by W. The linear equations in U is nothing but linear combinations of this set of linear equations.

Let

$$\hat{\phi}_{32} = \bar{\phi}_{32} \circ \hat{\phi}_1 \circ \phi_0(x_{v_1}, \ldots, x_{v_{31}}) = (\bar{\phi}_{32,1}, \ldots, \bar{\phi}_{32,100}),$$
$$\phi_{10}(x_{v_1}, \ldots, x_{v_{31}}) = \hat{\phi}_1 \circ \phi_0(x_{v_1}, \ldots, x_{v_{31}}) = (\phi_{10,1}, \ldots, \bar{\phi}_{10,100}).$$

Then using the expansion formula of $\bar{\phi}_{32}$, in particular the formulas (2.1) and (2.2), we have:

$$\hat{\phi}_{32,i} = \phi_{10,i} + a_1^{14}\beta_{i4}(\phi_{10,38}\phi_{10,48} + \phi_{10,47}\phi_{10,46}) + a_1^{14}\sum_1^3 \beta_{ij}(\phi_{10,38-2j}\phi_{10,48} + \phi_{10,39-2j}\phi_{10,47}) = \phi_{10,i} + R'_i, \tag{2.4}$$

for i =1;

$$\hat{\phi}_{32,i} = \phi_{10,i} + f_i(\phi_{10,1}, .., \phi_{10,i-1}) + a_1^{14}\beta_{i4}(\phi_{10,38}\phi_{10,48} + \phi_{10,37}\phi_{10,46}) + a_1^{14}\sum_1^3 \beta_{ij}(\phi_{10,38-2j}\phi_{10,48} + \phi_{10,39-2j}\phi_{10,47}) = \phi_{10,i} + R'_i, \tag{2.5}$$

for $i = 2, 3, 4$;

$$\hat{\phi}_{32,i} = \phi_{10,i} + f_i(\phi_{10,1}, \dots, \phi_{10,i-1}), \mathrm{i} = 5, 6, \dots, 41;$$

$$\hat{\phi}_{32,i} = q_{i-31}(\phi_{10,38}, \dots, \phi_{10,48}), \mathrm{i} = 42, \dots, 48;$$

$$\hat{\phi}_{32,i} = q_{i-31}(\phi_{10,38}, \dots, \phi_{10,48}), \mathrm{i} = 49, \dots, 76;$$

$$\hat{\phi}_{32,i} = q_{i-72}(\phi_{10,36}, \phi_{10,39}, \phi_{10,40}, \dots, \phi_{10,45}, \phi_{10,37}, \phi_{10,47}, \phi_{10,48})\mathrm{i} = 77, \dots, 84;$$

$$\hat{\phi}_{32,i} = q_{i-85}(\phi_{10,34}, \phi_{10,39}, \phi_{10,40} \dots, \phi_{10,45}, \phi_{10,35}, \phi_{10,47}, \phi_{10,48})\mathrm{i} = 85, \dots, 92;$$

$$\hat{\phi}_{32,i} = q_{i-93}(\phi_{10,32}, \phi_{10,39}, \phi_{10,40}, \dots, \phi_{10,45}, \phi_{10,33}, \phi_{10,47}, \phi_{10,48})\mathrm{i} = 93, \dots, 100,$$

where $R'_i = \sum_1^4 \beta_{ij} P'_i$, and P'_i, for $i = 1, 2, 3$, is given as

$$P'_i = \hat{\phi}_{1,31+i+1}(x'_1, \dots, x'_{48})\hat{\phi}_{1,48}(x'_1, \dots, x'_{48}) + \hat{\phi}_{1,31+i}(x'_1, \dots, x'_{48})\hat{\phi}_{1,47}(x'_1, \dots, x'_{48}).$$

$$P'_4 = \hat{\phi}_{1,42}(x'_1, \dots, x'_{48})\hat{\phi}_{1,48}(x'_1, \dots, x'_{48}) + \hat{\phi}_{1,46}(x'_1, \dots, x'_{48})\hat{\phi}_{1,47}(x'_1, \dots, x'_{48}),$$

which are constants. Namely $R_i(\bar{\phi}_{32}(x_1, \dots, x_{48}))$ are constants on the variety W. Therefore

$$\hat{F}(x_{v_1}, \dots, x_{v_{31}})) = (\hat{y}_1, \dots \hat{y}_{100}) = \phi_4 \circ \hat{\phi}_3 \circ \phi_2 \circ \pi \circ \hat{\phi}_1 \circ \phi_0(x_{v_1}, \dots, x_{v_{31}}).$$

where $\bar{\phi}_3 = (\bar{\phi}_{3,1}, \dots, \bar{\phi}_{3,100})$ is given by

$$\hat{\phi}_{3,i} = x_i, \mathrm{i} = 5, \dots, 100; \quad \hat{\phi}_{3,4} = x_4 + R'_i, \mathrm{i} = 1, 2, 3, 4.$$

Therefore ϕ_3 on the variety W is equivalent to $\hat{\phi}_3$, which is linear and is just a translation.

Then

$$\hat{F}(x_{v_1}, \dots, x_{v_{31}}) = (\phi_4 \circ \hat{\phi}_3) \circ \phi_2 \circ (\pi \circ \hat{\phi}_1 \circ \phi_0) = \hat{\phi}_4 \circ \phi_2 \circ \hat{\Phi}_1,$$

where $\hat{\phi}_4 = \phi_4 \circ \hat{\phi}_3$, $\hat{\Phi}_1 = \pi \circ \hat{\phi}_1 \circ \phi_0$ and both $\hat{\phi}_4$, which is invertible, and $\hat{\Phi}_1$, which is injective, are linear type.

Then $\hat{F}(x_{v_1}, \dots, x_{v_{31}}) = (y'_1, \dots, y'_{100})$ can be easily solved because of the triangular form of ϕ_2, namely the equation above is equivalent to the equations:

$$\hat{\phi}_4^{-1}(\hat{y}_1, \dots, \hat{y}_{100}) = \phi_2 \circ \hat{\Phi}_1(x_{v_1}, \dots, x_{v_{31}}) = \hat{\phi}_4^{-1}(y'_1, \dots, y_{100'}),$$

whose first nontrivial equation is always a linear equation.

This shows that the equations can be solved by iteration of the procedure of first searching for linear equations by linear combinations of quadratic equations, and then substituting the linear equations into the quadratic equations. Each time of iteration, we reduce the variable by 1. This eventually will require 31 iterations to find the 31 linearly independent linear equations in the triangular form, whose solution gives the values of the 31 variables x_{v_i}. Then we can substitute the values of x_{v_i}, $i = 1, \dots, 31$ back into the first 17 substitution equations $x_{u_i} = h_i(x_{v_1}, \dots, v_{31})$, $j = 1, \dots, 17$, which recovers the complete set of (x'_i), the plaintext.

Overall, our general method is first to search all the linearization equations. Then, for a given ciphertext $(y'_1, \dots, y'_m)$ corresponding to a plaintext $(x'_1, \dots, x'_n)$, we use the linearization equations to produce enough (17) linearly independent linear equations satisfied by x_i. Then we do a substitution using these linear equations, which essentially makes ϕ_3 linear on the variety defined by the 17 linear equations. The rest becomes straightforward.

2.2.3. *The practical attack procedure and its complexity.* We have three steps to derive the plaintext $(x'_1, \dots, x'_{48})$ from a ciphertext $(y'_1, \dots, y'_{100})$, and the first step is a common step for any given ciphertext.

Step 1 of the attack

We first look for a basis for the space V, namely the basis of solutions of a_{ij}, b_i, c_j and d for the equations:

$$\sum_{i=1,j=1}^{n,m} a_{ij}x_iy_j(x_1,\dots,x_n) + \sum_{i=1}^{n} b_ix_i + \sum_{j=1}^{m} c_jy_j(x_1,\dots,x_n) + d = 0.$$

For this set of equations, we have $4949 = 4800 + 48 + 100 + 1$ variables and $19697 = 1 + 48 + (24 \times 47 + 48) + (48 + 24 \times 47 + (8 \times 47 \times 46))$ equations. We know that the dimension of the solutions is 347.

Though we have 19697 equations, we have only 4949 variables, we do not need to use all those equations to find the solutions. We can actually randomly choose 6000 equations, the probability that we will not find the complete solution is essentially zero. To solve these linear equations, is to do row operations on a 6000×4949 matrix. However, because we are working on a finite field with only 2^8 elements, the row operations corresponding the elimination procedure on each column requires at most $2^8 - 1$ multiplication of a given row. To eliminate each variable, on average, it takes $(2^8 - 1) \times 6000/2$ multiplications. Therefore to solve these equations, it requires at most $4600 \times (2^8 - 1) \times 6000/2 \doteq 2^{32}$ computations on the finite field K. This step is also the common step for any attack.

However, because we are working over the fixed field K, we can perform the computation of multiplication on K by finding first a generator g of the multiplicative group of K, and storing the table of elements γ in K as g^k, then computing the multiplication by two searches and one addition. This will improve the speed by at least a factor of 2. Therefore, this step takes at most 2^{31} computations.

Step 2 of the attack

For a given ciphertext $(y'_1, \dots, y'_{100})$, we substitute the polynomials of y_i by y'_i into the 347 linearly independent solutions of the linearization equations in V and derive 17 linearly independent linear equations of x_i by the Gaussian elimination method in the form of $x_{u_j} = h_j(x_{v_1}, \dots, x_{v_{31}})$, where h_j is a linear function, $A = \{u_1, \dots, u_{17}\}$, $B = \{v_1, \dots, v_{31}\}$, $A \cap B = \emptyset$ and $A \cup B = \{1, \dots, 48\}$. We, then, substitute them into y_i to make it into polynomials with only 31 variables $\{v_1, \dots, v_{31}\}$.

First with a probability 2^{-82}, we might fail to get 17 linearly independent equations, which surely can be neglected.

In this step, for first part, we need to do $347 \times (4800 + 100) \doteq 2^{21}$ computations when we substitute y_i by y'_i. Then, to reduce 347 equations to 17 equations for substitution, it takes $(2^8 - 1) \times 48 \times 347/2 \doteq 2^{21}$ computations . Then we perform the substitution of the 17 equations into y_i and it takes $100 \times (2 \times 17^2 + 17 \times 31 + 17) \doteq 2^{17}$ computations.

For the new 100 polynomials with 31 variables, which we denote by $\hat{y}_i$, we will write down first the 100 equations $\hat{y}_i - y'_i = \tilde{y}_{1i}(x_{v_1}, \dots, x_{v_{31}}) = 0$, and they are linearly dependent and the dimension is only actually 41.

Step 3 of the attack

For the equations $\tilde{y}_{1i}(x_{v_1}, \dots, x_{v_{31}}) = 0$ $i = 1, \dots, 100$, we will use Gaussian elimination method, first on the quadratic terms, to derive $\hat{m}$ ($\hat{m} = 41$) linearly independent equations $\hat{y}_{1i}(x_{v_1}, \dots, x_{v_{31}}) = 0$, $i = 1, \dots, m$, and the last one is linear. Then we take the linear equation out and substitute it back into the leftover $\hat{m}-1$ quadratic equations (the linear one is taken out) $\hat{y}_{1i}(x_{v_1}, \dots, x_{v_{31}}) = 0$, $i = 1, \dots, m-1$. We denote the new equations $\tilde{y}_{2i}(x_{v_1}, \dots, x_{v_i}, x_{v_{i+2}}, \dots, x_{v_{31}}) = 0$. Then we repeat the same process on these new equations, and later again and again for total of 31 times. We then collect all the 31 linear equations derived in this process, a set of 31 linearly independent equations in the triangular *form. The solution gives us all the values of x_{v_i}, then we plug them back into 17 linear equations $x_{u_j} = h_j(x_{v_1}, \ldots, x_{v_{31}})$ in Step 2, which will give us x_{u_i}. We recover the plaintext.*

For the first part of this step, we need at most $100 \times (31 \times 16 + 31) \times 100/2 \doteq 2^{22}$ computations to perform the Gaussian elimination and then the substitution takes at most $42 \times 31^2 \doteq 2^{15}$ computations. Then we need at most to perform 31 times these two procedures and therefore it, at most, takes 2^{25} computations to solve the equations for the 31 variables and then need to do 2^9 computations to find the values for the other 17 variables.

If we add all the three steps together, it takes at most 2^{32} computations. We did simulation of an example of this scheme and it took us a few hours to find the plaintext for any given ciphertext. The best way to test our method is definitely to attack the challenges set by T, Moh. However, at this moment the web site(www.usdsi.com), where the challenge is posted, does not allow public access to the challenge's data and we plan to do so once it is available.

2.3. **Cryptanalysis for the scheme in the first version of** [CM]**.** Our work was first done for the scheme in the original version of [CM]. When we just finished the work on this scheme in July last year, the new version appeared. In this section, we will present our work on the implementation scheme in the original version of [CM].

The construction of the scheme is similar to the revised case above.

We again work on the field K of size 2^8. A map $\bar{F}$ is made of ϕ_1, ϕ_2, ϕ_3, ϕ_4, which are maps from the (n+68) dimensional space to itself. ϕ_1, ϕ_4 are invertible affine maps, and ϕ_2 and ϕ_3 are nonlinear of de Jonquiere type but different from that of the section above.

Again a map

$$F(x_1, \ldots., x_n) = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1, x_2, \ldots, x_n, 0, \ldots 0) = (y_1, \ldots, y_{n+68})$$

is the cipher, which is public, but ϕ_1, ϕ_4 are private. To make sure the system is of degree 2, another set of polynomials $Q_8(z_1, \ldots, z_{48})$ and $q_i(z_1, \ldots, z_{14})$ are used. The detail of ϕ_3, ϕ_2 and Q_8 and q_i are given in appendix 1.1.

Through computations and similar argument as in the section above, we have:

1) The dimension of V of the space of linearization equations for the components y_i of F is 286;

2)For a given ciphertext $(y'_1, \ldots, y'_{68+n})$, with a probability of $(1 - (\frac{C_{14}^4}{2^{10}})^2 \doteq 1 - 2 \times 2^{-60}$, the linearization will produce 28 linearly independent linear equations of x_i.

3)For the case of 28 linearly independent equations, we can again do a substitution using these 28 linear equations into y_i to derive a new operator $\hat{F}$ which is a map from K^{n-28} to K^{n+68} and $\hat{F} = \hat{\phi}_4 \circ \phi_2 \circ \hat{\Phi}_1$, for some linear maps $\hat{\phi}_4$ invertible, and $\hat{\Phi}_1$ injective.

This allows us to use exactly the same attack steps as in the previous section to defeat this scheme.

Here if we choose $n = 52$, $m = 120$, we estimate that it takes about 2^{35} computations on K to defeat the scheme. We performed a computation example on such a scheme on a PC and it took a few days to find the plaintext from a given ciphertext.

2.4. **Cases for other implementation schemes.** By now, there are several implementation schemes that have been suggested by the inventor. We notice that for all cases, due to the fact that the q_i components are all very simple and they never have more than 2 quadratic monomials, it is easy to see that for all the schemes, the dimension of linear space of all the linearization equations for the components y_i of F is not small. This is the common defect for the implementation schemes, which is not in any way desirable for a secure open key cryptosystem. However even with those linearization equations, it does not necessarily mean that finding the plaintext from a given ciphertext is easy. We will use the example of the first implementation [M] to demonstrate the situation.

The construction of the implementation scheme in [M] is the first implementation scheme for the TTM systems, and is also set on the field K of size 2^8. A map $\bar{F}$ is again made of ϕ_1, ϕ_2, ϕ_3, ϕ_4, which are maps from the (n+36) dimensional space to itself, and $n = 64$. ϕ_1, ϕ_4 are invertible affine maps, and ϕ_2 and ϕ_3 are nonlinear of de Jonquiere type but different from that of the sections above.

Again we have a map: $F(x_1, \ldots, x_n) = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1, x_2, \ldots, x_n, 0, \ldots, 0) = (y_1, \ldots, y_{100})$. To make sure the system is of degree 2, again a set of polynomials $Q_8(z_1, \ldots, z_{48})$ and $q_i(z_1, \ldots, z_{14})$ are used. The detail of ϕ_3, ϕ_2 and Q_8 and q_i are given in appendix 1.2.

Through computation, we have:

1) The dimension of V of the space of linearization equations for the components y_i of F is 68;

2)For a given ciphertext $(y'_1, \ldots, y'_{68+n})$, the linearization will **NOT** produce enough number of linearly independent linear equations of x_i (often small) such that if we do a substitution using these linear equations into y_i to derive a new operator $\hat{F}$ that can be expressed as: $\hat{F} = \hat{\phi}_4 \circ \hat{\phi}_2 \circ \hat{\Phi}_1$, for some linear maps $\hat{\phi}_4$ and $\hat{\Phi}_1$.

This means that what we achieved is just a reduction of number of variables, surely it is useful, in general, but it is still unclear how we can obtain the plaintext fast. We actually performed a search for linearization on the components of $\hat{y}_i$ and we had some successes, but failures as well. Further work needs to be done to see if we can pursue further in this direction. We also know that the implementation scheme [M] is defeated by another completely different method [DH].

For the case of the most TTM schemes in the original version of [CM], we actually also observed that for the components y_i of the cipher F, a higher order type of linearization equation:

$$\Sigma a_{ij} y_i y_j + \Sigma b_{ij} y_i x_j + \Sigma c_k y_k + \Sigma d_l x_l + e = 0$$

is also satisfied. To find all the solutions for this case surely takes more time but not impossible. It could also produce non trivial linear equation of x_i if we are given a ciphertext y'_i. Definitely there is a possibility that these linearization equations will produce more linear equations to help us to defeat the scheme.

3. Conclusion

Due to the fact that the set of quadratic polynomials q_i, a key component of the TTM schemes, are very simple and often just one degree 2 monomial, through computation, we show that for all the TTM implementation schemes, the polynomial components y_i of the public cipher F satisfies linearization equations and for a given cipher text y'_i, we can obtain linear equations satisfied by the plaintext x'_i. This is something that a secure open key cryptosystem should not have. Though this defect does not necessary lead us to defeat the scheme easily for all the implementation schemes, for the cases of the two most recent implementation schemes suggested in two different versions of [CM], we show that with a very small probability of failure, this defect actually allows us to defeat the schemes easily. For the suggested practical example in revised [CM], we show that it takes 2^{32} computations on the finite field K to defeat the scheme and we confirmed this by a computation example; for the case of the scheme in the original version of [CM], it takes 2^{35} computations to defeat it, and it is confirmed by a computer experiment as well.

Also for the existing TTM implementation schemes, we can even find higher order type of linearization equations, which a secure scheme should avoid as well.

Considering all the works on attacking the TTM schemes, [CG], [DH] and this paper, we conclude that all the existing TTM schemes are insecure. But we do not, in anyway, suggest that our results imply that there does not exist good TTM schemes. We do conclude that to avoid the defect we found in this paper, more sophisticated construction of Q_8 and q_i is needed. We think this is a very interesting direction to pursue, but it needs some deep insight from algebraic geometry.

References

[CGC] Chou, G., Guan, J. Chen J. *A systematic Construction of a Q_{2^k}-model in TTM.*

[CM] Chen, J., Moh, T., *On the Goubin-Courtois Attack on TTM* , Cryptology ePrint Archive (2001/72).

[CG] Goubin, L., Courtois, N., *Cryptanalysis of the TTM cryptosystem*, Asiacrypt2000, LNCS 1976, 44-57.

[Dm] Dickerson, Matthew, *The inverse of an automorphism in polynomial time.* J. Symbolic Comput. 13 (1992), no. 2, 209–220.

[DH] Ding, Jintai, Timonthy Hodges, *Cryptanalysis of an implementation scheme of TTM.* Department of Mathematical Sciences, University of Cincinnati, Preprint.

[MI] Matsumoto, T., Imai, H., *Public quadratic polynomial-tuples for efficient signature-verification and message-encryption*, Advances in cryptology—EUROCRYPT '88 (Davos, 1988), 419–453, Lecture Notes in Comput. Sci., 330, Springer, Berlin, 1988.

[M] Moh, T. T., *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

[P] Patarin, J., *Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt'88.*, Des. Codes Cryptogr. 20 (2000), no. 2, 175–209.

[P1] Patarin, J., *Hidden filed equations and isomorphism of polynomials*, Eurocrypto'96, 1996.

Appendix 1.1 Here we give some details for the original implementation scheme in the original version of [CM]. $m = n + 68$.

We will not give the exact detail of ϕ_3 and ϕ_2 (we refer it to the original version of [CM]), rather we will give the detail of $\bar{\phi}_{32} = \phi_3 \circ \phi_2 \circ \pi$, where $\pi(x_1, \ldots, x_n) = (x_1, \ldots, x_n, 0, \ldots, 0)$ a map from K^n to K^{n+68}:

$$
\begin{gathered}
\bar{\phi}_{32,1} = y_1 = x_1 + P(y_3, \ldots, y_m) = x_1 + P(x_3 + f_3(x_1, x_2), \ldots, f_m(x_1, \ldots, x_m) = \\
x_1 + Q_8(y_{n-7}, \ldots, y_{n+34} + Q_8(y_{n-21}, y_{n-14}, y_{n+35} \cdots, y_{n+68}) = \\
x_1 + x_{n-5}x_{n-6} + x_{n-8}x_n + x_{n-19}x_{n-20} + x_{n-22}x_{n-14}, \\
\bar{\phi}_{32,2} = y_2 = x_2 + Q(y_3, \ldots . y_m) = x_2 + x_1^2 + Q_8(y_{n-21}, y_{n-14}, y_{n+35} \cdots, y_{n+68}) = \\
x_1 + x_2^2 + x_{n-19}x_{n-20} + x_{n-22}x_{n-14}, \\
\bar{\phi}_{32,3} = y_3 = x_3 + f_3(x_1, x_2), \\
\cdots \\
\bar{\phi}_{32,n-22} = y_{n-22} = x_{n-22} + f_{n-22}(x_1, \ldots, x_{n-23}, \\
\bar{\phi}_{32,n-23} = y_{n-23} = q_1(x_{n-27}, \ldots, x_{n-14}), \\
\cdots \\
\bar{\phi}_{32,n-14} = y_{n-14} = q_8(x_{n-27}, \ldots, x_{n-14}), \\
\bar{\phi}_{32,n-13} = y_{n-13} = x_{n-13} + f_{n-13}(x_1, \ldots, x_{n-14}), \\
\cdots \\
\bar{\phi}_{32,n-8} = y_{n-8} = x_{n-8} + f_{n-22}(x_1, \ldots, x_{n-9}), \\
\bar{\phi}_{32,n-7} = y_{n-7} = q_1(x_{n-13}, \ldots, x_n), \\
\cdots \\
\bar{\phi}_{32,n+34} = y_{n+34} = q_{42}(x_{n-13}, \ldots, x_n), \\
\bar{\phi}_{32,n+35} = y_{n+35} = q_9(x_{n-27}, \ldots, x_{n-14}), \\
\cdots \\
\bar{\phi}_{32,n+68} = y_{n+68} = q_{42}(x_{n-27}, \ldots, x_{n-14}),
\end{gathered}
$$

where

$$
\begin{gathered}
Q_8(q_1, \ldots, q_{42}) = [q_{14}q_{23} + q_{17}q_{24}][q_{10}q_9 + q_6q_4]^2[q_{11}q_{30} + q_1q_{31}] + \\
[q_{33}q_{34} + q_{35}q_{36}][q_{15}q_{37} + q_{16}q_{26}] + [q_{19}q_8 + q_{20}q_{38}][q_{13}q_7 + q_{12}q_5] + \\
[q_{21}q_{39} + q_{40}q_2 + q_{22}q_{41} + q_{42}q_3];
\end{gathered}
$$

q_i are functions of 14 variables $z_i, z_2, \ldots, z_{14}$:

$$
\begin{gathered}
q_1 = z_7 + z_2z_5, q_2 = z_8 + z_6z_7, \\
q_3 = z_9 + z_6z_5, q_4 = z_2z_4 + z_{10}, \\
q_5 = z_3z_5 + z_{11}, q_6 = z_1z_3 + z_{12}, \\
q_7 = z_3z_7 + z_{13}, q_8 = z_{14} + z_8z_3, \\
q_9 = z_3 + z_2z_{12}, q_{10} = z_4 + z_{10}z_1, \\
q_{11} = z_{13} + z_{11}z_2, q_{12} = z_4z_{13} + z_7, \\
q_{13} = z_4z_{11} + z_5, q_{14} = z_6 + z_9z_2, \\
q_{15} = z_{14} + z_9z_{10}, q_{16} = z_8 + z_{10}z_6, \\
q_{17} = z_9 + z_1z_6, q_{18} = z_9z_1, \\
q_{19} = z_9z_4 + z_6, q_{20} = z_6z_3 + z_9,
\end{gathered}
$$

$$q_{21} = z_7 z_9 + z_{14}, q_{22} = z_9 z_{13} + z_6,$$
$$q_{23} = z_1 z_8 + z_{14}, q_{24} = z_2 z_{14} + z_8,$$
$$q_{25} = z_{14} z_1, q_{26} = z_{10} z_{14},$$
$$q_{27} = z_2 z_{10}, q_{28} = z_2 z_3,$$
$$q_{29} = z_1 z_{11}, q_{30} = z_1 z_7 + z_5,$$
$$q_{31} = z_1 z_{13} + z_{11}, q_{32} = z_1 z_5,$$
$$q_{33} = z_{12} z_{11} + z_{13}, q_{34} = z_{12} z_7,$$
$$q_{35} = z_{12} z_{13}, q_{36} = z_{12} z_5 + z_7,$$
$$q_{37} = z_{10} z_8, q_{38} = z_4 z_{14} + z_8,$$
$$q_{39} = z_8 z_{11}, q_{40} = z_{14} z_{11},$$
$$q_{41} = z_8 z_5 + z_{14}, q_{42} = z_{14} z_{13} + z_8,$$

and

$$Q_8(q_1, \ldots, q_{42}) = z_8 z_9 + z_6 z_{14}.$$

Appendix 1.2

We give the details for the first TTM implementation scheme. The notation comes form that in [M].

Let $n = 64$, $r = 36$ and $m = n + r = 100$. ϕ_1, ϕ_2, ϕ_3, ϕ_4 are invertible maps from the 100 dimensional space K^{100} to itself. ϕ_1, ϕ_4 are invertible affine linear maps, and ϕ_2 and ϕ_3 are nonlinear maps of the de Jonquiere Type.

Let $[j] = j \bmod 8$, and $0 < [j] < 9$.

Let $\phi_2 = (\phi_{2,1}, \ldots, \phi_{2,100})$ and $\phi_3 = (\phi_{3,1}, \ldots, \phi_{3,100})$. We have:

$$\phi_{2,i} = x_i, \mathrm{i} = 1, 2;$$
$$\phi_{2,i} = x_i + x_{i-1} x_{i-2}, \mathrm{i} = 3, \ldots, 9;$$
$$\phi_{2,i} = x_i + x_{[i-1]}^2 + x_{[i]} x_{[i-5]} + x_{[i+1]} x_{i+6}, \mathrm{i} = 10, \ldots, 17;$$
$$\phi_{2,i} = x_i + x_{[i-1]} x_{[i+1]} + x_{[i]} x_{[i+4]}, \mathrm{i} = 18, \ldots, 25;$$
$$\phi_{2,i} = x_i + x_{[i-1]} x_{[i+1]} + x_{[i+2]} x_{[i+5]}, \mathrm{i} = 26, \ldots, 30;$$
$$\phi_{2,i} = x_i + x_{i-10}^2, \mathrm{i} = 31, \ldots, 60;$$
$$\phi_{2,61} = x_{61} + x_9^2,$$
$$\phi_{2,62} = x_{62} + x_{61}^2,$$
$$\phi_{2,63} = x_{63} + x_{10}^2,$$
$$\phi_{2,64} = x_{64} + x_{63}^2,$$
$$\phi_{2,i} = x_i + q_{i-64}(x_9, x_{11}, \ldots, x_{16}, x_{51}, x_{52} \ldots, x_{62}), \mathrm{i} = 65, \ldots, 92;$$
$$\phi_{2,i} = x_i + q_{i-92}(x_{10}, x_{17}, \ldots, x_{20}, x_{15}, x_{16}, x_{51}, \ldots, x_{60}, x_{63}, x_{64}), \mathrm{i} = 93, \ldots, 100;$$

where $q_i(z_1, \ldots, z_{19})$, i=1,...,30, are degree two polynomials of 19 variables $z_1, \ldots, z_{19}$:

$$q_1 = z_1 + z_2 z_6, q_2 = z_2^2 + z_3 z_7,$$
$$q_3 = z_3^2 + z_4 z_{10}, q_4 = z_3 z_5,$$
$$q_5 = z_3 z_{11}, q_6 = z_4 z_7,$$
$$q_7 = z_4 z_5, q_8 = z_7^2 + z_5 z_{11},$$
$$q_9 = z_6^2 + z_8 z_9, q_{10} = z_8^2 + z_{12} z_{13},$$
$$q_{11} = z_9^2 + z_{14} z_{15}, q_{12} = z_7 z_{10},$$

$$q_{13} = z_{10}z_{11}, q_{14} = z_{12}^2 + z_7z_8,$$
$$q_{15} = z_{13}^2 + z_{11}z_{16}, q_{16} = z_{14}^2 + z_{10}z_{12},$$
$$q_{17} = z_{15}^2 + z_{11}z_{17}, q_{18} = z_{12}z_{16},$$
$$q_{19} = z_{11}z_{12}, q_{20} = z_8z_{13},$$
$$q_{21} = z_7z_{13}, q_{22} = z_8z_{16},$$
$$q_{23} = z_{14}z_{17}, q_{24} = z_7z_{11},$$
$$q_{25} = z_{12}z_{15}, q_{26} = z_{10}z_{15},$$
$$q_{27} = z_{12}z_{17}, q_{28} = z_{11}z_{14},$$
$$q_{29} = z_{18} + z_1^2, q_{30} = z_{19} + z_{18}^2;$$

and the map ϕ_3 is defined as:

$$\phi_{3,i} = x_i, \mathrm{i} = 3, .., 100;$$
$$\phi_{3,2} = x_2 + Q_8(x_{93}, \dots, x_{100}, x_{73}, \dots, x_{92}, x_{63}, , x_{64}),$$
$$\phi_{3,1} = x_i + Q_8(x_{65}, x_{66}, \dots, x_{92}, x_{61}, x_{62}),$$

where Q_8 as a polynomial of degree 8 with 30 variables $z_1, \dots, z_{30}$ is defined as

$$Q_8(z_1, \dots, z_{30}) = z_1^8 + [z_2^4 + z_3^4 + z_3^2z_8^2 + z_4^2z_5^2 + z_6^2z_{12}^2 + z_7^2z_{13}^2]\times$$
$$[z_9^4 + (z_{10}^2 + z_{14}z_{15} + z_{18}z_{19} + z_{20}z_{21} + z_{22}z_{24})(z_{11}^2 + z_{16}z_{17} + z_{23}z_{28} + z_{25}z_{26} + z_{13}z_{27})] + z_{29}^4 + z_{30}^2.$$

$$F(x_1, \dots, x_{64}) = \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(x_1, x_2, \dots, x_{64}, 0, \dots, 0) = (y_1, \dots, y_{100})$$

is the cipher. In the expansion formula, the components of the map are degree two polynomials of $(x_1, \dots, x_{64})$, because

$$Q_8(q_1(z_1, .., z_{19}), \dots, q_{30}(z_1, \dots, z_{19})) =$$
$$q_1^8 + [q_2^4 + q_3^4 + q_3^2q_8^2 + q_4^2q_5^2 + q_6^2q_{12}^2 + q_7^2q_{13}^2]\times$$
$$[q_9^4 + (q_{10}^2 + q_{14}q_{15} + q_{18}q_{19} + q_{20}q_{21} + q_{22}q_{24})(q_{11}^2 + q_{16}q_{17} + q_{23}q_{28} + q_{25}q_{26} + q_{13}q_{27})] + q_{29}^4 + q_{30}^2$$
$$= z_{19}^2$$