EMBEDDED SURFACE ATTACK ON MULTIVARIATE PUBLIC KEY CRYPTOSYSTEMS FROM DIOPHANTINE EQUATIONS

JINTAI DING, ADAMA DIENE, AI REN, CHENG DONG TAO

ABSTRACT. In 2011, Gao and Heindl presented a new multivariate public key cryptosystems from Diophantine equations. However, by observing the decryption process, Ding et al. found recently that some special embedded surfaces could be used to decrypt the message efficiently. They showed that the three systems proposed by Gao and Heindl could be broken at the complexity of $2^{52}, 2^{61}$ and 2^{52} respectively. But, the attack by Ding et al. can not implemented on a usual PC. It was mainly theoretical and worked only for finite fields of characteristic 2. In this paper, We present a practical attack on this family of multivariate public key cryptosystems. Our attack is implemented on a PC and it works for all finite fields. By doing computer experiments, we find many new lower-degree embedded surfaces which help us to break the cryptosystem. Let $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_m)$ be a pair of corresponding plaintext and ciphertext for a cryptosystem. We define a degree 3 embedded surface of this cryptosystem as any polynomial equation:

$$\sum \alpha_{i,j,t} x_i x_j y_t + \sum \beta_{i,j} x_i x_j + \sum \gamma_{i,t} x_i y_t + \sum \mu_k y_t + \sum \nu_i x_i + \lambda = 0$$

which is satisfied by all such pairs. Thus, for any given ciphertext, we can derive many new independent quadratic equations which can be used to break the cryptosystem directly. We could break the three systems proposed at the complexity of $2^{41}, 2^{48}$ and 2^{41} respectively.

1. INTRODUCTION

1.1. Background. Multivariate public key cryptosystems (MPKC) was first introduced in 1988 by Matsumoto and Imai[17] with their scheme called C* or MI. The public key of a MPKC is a system of multivariate polynomials, mostly quadratic, over a finite field. In general, MPKCs have the structure describe in the following paragraph. Let k be a finite field with q elements. A public key is a map $\bar{F} : k^n \rightarrow k^m$, constructed as: $\bar{F} = L_1 \circ F \circ L_2$, where L_1 and L_2 are two random invertible affine transformations over k^n and k^m respectively. The central map $F : k^n \rightarrow k^m$ is a nonlinear multivariate polynomial map that has the property of being easily invertible computation-wise. The key of building a good MPKC is to find a good polynomial system, F , that makes the cryptosystem secure.

The security of a MPKC is based on the knowledge that solving a set of multivariate polynomial equations over a finite field, in general, is proven to be an NP-hard problem. However, this does not guarantee that MPKCs are secure. Nevertheless, this property makes the family of MPKCs to be among the good candidates of the Post Quantum Cryptography (PQC) era if well designed. On the other hand, the

Date: January 28, 2014.

Key words and phrases. HFE, degree of regularity.

well-known number theoretic based cryptosystems such as RSA, ECC, and Diffie-Hellman key exchange scheme are proven to be insecure if a quantum computer is built due to Peter Shors algorithm [12].

These facts have inspired many researchers to be involved in this new research area and MPKCs had a very fast development in the late 1990s. Since then, there are many attempts in building MPKC for encryption. Researchers have successfully designed some very good, secure and efficient MPKC signature schemes such as the Rainbow scheme proposed by Ding and Schmidt[6]. However despite the several attempts, the construction of MPKCs encryption have not been very successful. In general the encryption schemes of MPKCs are much slower than the signature schemes and many encryption schemes of MPKCs are proven insecure. For instance C* was proven to be insecure by Jacques Patarin [13] few years later after it was proposed.

One of the few MPKCs for encryption that was very promising was proposed by Gao and Heindl [10] in 2009. A completely new construction was used for this novel idea. Namely one uses a special function solution to certain special Diophantine equations. Even though this construction uses a triangular and an Oil-Vinegar construction [14], the key component comes from certain special solutions of the Diophantine equation:

$$AB = CD + EF + GH + IJ + KL;$$

which is inspired by the construction in MFE[16], a generalization of HFE[12].

For this new family, the authors propose in [9] three concrete cases for practical applications. These cases have very strong security claims at the level of 2^{113} or higher. The decryption process is also very efficient. This family of schemes remained secure for the past 4 years. No weaknesses in this system have been found prior to this work. In this paper, we prove the insecurity of this family of cryptosystems.

Let $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_n)$ be a pair of corresponding plaintext and ciphertext for any cryptosystem. We define an embedded surface of this cryptosystem as any polynomial equation:

$$E(X, Y) = E(x_1, \dots, x_n, y_1, \dots, y_n) = 0;$$

which is satisfied by all such pairs. This name comes from the fact that this equation defines an algebraic surface in the space k^{m+n} , where all the plaintext and ciphertext pairs for this cryptosystems belong, or we can embed all the such pairs in such a surface. The first non-trivial embedded surface is used as the linearization equation by Patarin to defeat the Matsumoto-Imai MPKCs[10]:

If we know the value of a ciphertext, this equation will give us a linear equation satisfied by the plaintext. This is very useful in attacking the system and is the idea used to build algebraic attack on AES. Later, another embedded surface, a high order linearization equation is used in breaking the MFE MPKCs[4].

A natural question arises from this. Can more general embedded surfaces other than linearization type of equations be useful to attack cryptosystems?

1.2. The Contributions of this paper. We show in this paper that the answer to this question is positive. We first observe that in the case of the new MPKCs using the Diophantine equations, new type of embedded surfaces can be swiftly generated from the decryption process to decrypt the message efficiently. Namely,

we show in section 3 that the embedded surfaces obtained in this case have the form of

$$\begin{aligned} & \sum a_{i,j,s,t} y_i y_j y_s y_t + \sum b_{i,j,s} y_i y_j y_s + \sum c_{i,j} y_i y_j + \sum d_j y_j + \\ & \sum a'_{i,j,s,t} y_i y_j x_s^2 x_t^2 + \sum b'_{i,s,t} y_i x_s^2 x_t^2 + \sum c'_{i,j,s} y_i y_j x_s^2 + \sum d'_{i,s} y_i x_s^2 + \\ & \sum e'_{s,t} x_s^2 x_t^2 + \sum h'_s x_s^2 + e = 0, \end{aligned}$$

which is satisfied by all plaintext and ciphertext pairs.

As a result, it enables us to derive new quadratic equations due to the fact that the field is of characteristic 2. This allows us to break the system. However, the degree of these embedded surfaces are 4. In order to finding these embedded surfaces, one has to solve a large linear system which will consume more memories and computation. Therefore the attack in [3] could not be implemented in a usual PC. It was mainly theoretical and worked only for finite fields of characteristic 2. These results can be also found in [3]

Second, we also show in section 3 that by doing substantial computer experiment, we can find many interesting equations associated to the polynomials of the central map. These equations imply that we can define an embedded surface of the new MPKCs from Diophantine equations as any polynomial equation:

$$\sum \alpha_{i,j,t} x_i x_j y_t + \sum \beta_{i,j} x_i x_j + \sum \gamma_{i,t} x_i y_t + \sum \mu_k y_t + \sum \nu_i x_i + \lambda = 0$$

which is satisfied by all plaintext and ciphertext pairs. As a result, many independent quadratic equations which help to break the cryptosystem can be obtained for any given ciphertext y_i .

Magma will be used to implement the attack for the first suggested set of parameters. We run all the experiments on a Sun X4440, with four Quad-Core AMD Opteron™ Processor 8356 CPUs and 128 GB of main memory. Each CPU is running at 2.3 GHz. The experiment results obtained are showed in the follow table:

Parameters	Claimed security level	Attack complexity in [3]	Our attack complexity
$q = 2^{16}, d = 1$ $n = 56, m = 74$	2^{113}	2^{52}	2^{41} Time(s): 4229.75
$q = 2^{16}, d = 2$ $n = 112, m = 148$	2^{221}	2^{61}	2^{48}
$q = 2^{32}, d = 1$ $n = 56, m = 74$	2^{114}	2^{52}	2^{41}

Table 1. Our attack complexity and time required of the first suggested set of parameters.

Furthermore, it can be seen easily that our embedded surface implies that the degree of regularity [1] of the systems is actually 4, therefore we can attack the new MPKCs from Diophantine equations by using direct attack with Gröebner basis algorithms. The experiment results obtained are showed in the follow table:

Parameters	Time(s)	Memory(MB)	Degree of regularity
$q = 2^{16}, d = 1, n = 56, m = 74$	33338.938	14723.5	4

Table 2. The time required of direct attack.

This paper is organized in the following order. Section 2 gives a description of the new MPKCs from Diophantine equations. Our main contributions are then stated in section 3 and 4. Finally, Section 5 concludes the paper.

2. MULTIVARIATE PUBLIC KEY CRYPTOSYSTEMS FROM DIOPHANTINE EQUATIONS

In this section, we will present the MPKCs from Diophantine equations, and we will follow the notations in [9].

Let k be a finite field with q elements, and let $\mathcal{F}$ be an extension of k with degree d . In an MFE type ("medium field") construction, we fix a basis $\alpha_1, \dots, \alpha_d$ of $\mathcal{F}$ over k , which identifies $\mathcal{F}$ with k^d via the natural map $p : \mathcal{F} \rightarrow k^d$:

$$(1) \quad p(a_1\alpha_1 + \dots + a_d\alpha_d) = (a_1, \dots, a_d).$$

Then we view a polynomial $f \in \mathcal{F}[X_1, \dots, X_n]$ component-wise over k by writing

$$X_i = x_{i1}\alpha_1 + \dots + x_{id}\alpha_d,$$

and then

$$f = f_1\alpha_1 + \dots + f_d\alpha_d,$$

with $f_i \in k[x_{11}, \dots, x_{nd}]$.

Throughout this paper, we assume that the finite field $\mathcal{F}$ has characteristic two. If the field is not $GF(2)$, the embedded surfaces of (1) and (2) should be modified slightly, since there are not square terms.

2.1. The origin of the Diophantine equations. In the MFE MPKC, the key idea comes the fact that

$$\det(M_1 M_2) = \det(M_1) \times \det(M_2),$$

for two 2×2 matrices:

$$M_1 = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}, \quad M_2 = \begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix}.$$

This gives a quadratic polynomial solution to the the Diophantine equation over a polynomial ring $\mathcal{F}[X_1, \dots, X_8]$:

$$AB = CD + EF,$$

namely

$$(X_1 X_4 + X_2 X_3)(X_5 X_8 + X_6 X_7) = (X_1 X_5 + X_2 X_7)(X_3 X_6 + X_4 X_8) + (X_1 X_6 + X_2 X_8)(X_3 X_5 + X_4 X_7).$$

To build new type of MPKCs, Gao and Heindl were able to find solutions to the following new Diophantine equation:

$$AB = CD + EF + GH + IJ + KL,$$

over the ring $\mathcal{F}[X_1, \dots, X_8, Y_1, \dots, Y_8]$. In the context of their work, they rewrite this equation as

$$(2) \quad \psi_1 \psi_2 = f_1 f_2 + \dots + f_9 f_{10},$$

where each polynomial is quadratic and

- $\psi_1 \in \mathcal{F}[X_1, \dots, X_8], \psi_2 \in \mathcal{F}[Y_1, \dots, Y_n]$;
- $f_i \in \mathcal{F}[X_1, \dots, X_8, Y_1, \dots, Y_8]$, for $0 < i < 9$, are oil-vinegar polynomials;
- $f_i \in \mathcal{F}[X_1, \dots, X_8, Y_1, \dots, Y_8]$, $i = 9, 10$.

An oil-vinegar polynomial is a quadratic polynomial, where we divide the variables into two sets: the oil variables and the vinegar variables, and an oil-vinegar polynomial has not any quadratic terms with only oil variables:

$$\sum a_{ij}x_ix'_j + \sum b_{ij}x'_ix'_j + \sum c_jx_i + \sum d_jx'_j + e = 0,$$

where x_i are oil variables and x'_j are the vinegar variables.

The design starts with the polynomial ring

$$R = \mathcal{F}[x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4, z_1, z_2, z_3, z_4, w_1, w_2, w_3, w_4].$$

Let

$$p_{xy}^{ij} = x_iy_j + x_jy_i,$$

for $1 = i < j = 4$;

$$p^{ij}(x, y, z, w) = p_{xz}^{ij} + p_{yz}^{ij} + p_{yw}^{ij},$$

for $1 = i < j = 4$.

In algebraic geometry terms, the p_{xy}^{ij} are simply Plücker coordinates, which are known to satisfy the quadratic relations:

$$\begin{aligned} 0 = & (p_{xy}^{12} + p_{zw}^{12})p^{34}(x, y, z, w) + (p_{xy}^{13} + p_{zw}^{13})p^{24}(x, y, z, w) + \\ & (p_{xy}^{14} + p_{zw}^{14})p^{23}(x, y, z, w) + (p_{xy}^{23} + p_{zw}^{23})p^{14}(x, y, z, w) + \\ (3) \quad & (p_{xy}^{24} + p_{zw}^{24})p^{13}(x, y, z, w) + (p_{xy}^{34} + p_{zw}^{34})p^{12}(x, y, z, w). \end{aligned}$$

2.2. The central map. Let ρ be a ring homomorphism from R to $\mathcal{F}[X_1, \dots, X_8, Y_1, \dots, Y_8]$ induced by the map:

$$\begin{aligned} & (x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4, z_1, z_2, z_3, z_4, w_1, w_2, w_3, w_4) \rightarrow \\ (X_1, X_3, Y_1 + Y_5, Y_3 + Y_7, X_4, X_2, Y_5, Y_7, X_5, X_7, Y_4 + Y_8, Y_2 + Y_6, X_8, X_6, Y_8, Y_6). \end{aligned}$$

Let

$$\begin{aligned} p_1 &= \rho(p_{xy}^{12} + p_{zw}^{12}) = X_1X_2 + X_3X_4 + X_5X_6 + X_7X_8 \\ p_2 &= \rho(p^{34}(x, y, z, w)) = Y_1Y_2 + Y_3Y_4 + Y_5Y_6 + Y_7Y_8 \\ f_1 &= \rho(p_{xy}^{13} + p_{zw}^{13}) = X_4Y_1 + X_8Y_4 + (X_1 + X_4)Y_5 + X_5Y_8 \\ f_2 &= \rho(p^{24}(x, y, z, w)) = (X_2 + X_3)Y_2 + X_7Y_3 + X_2Y_6 + X_6Y_7 \\ f_3 &= \rho(p_{xy}^{14} + p_{zw}^{14}) = X_8Y_2 + X_4Y_3 + X_5Y_6 + (X_1 + X_4)Y_7 \\ f_4 &= \rho(p^{23}(x, y, z, w)) = X_7Y_1 + (X_2 + X_3)Y_4 + X_6Y_5 + X_2Y_8 \\ f_5 &= \rho(p_{xy}^{23} + p_{zw}^{23}) = X_2Y_1 + X_6Y_4 + (X_2 + X_3)Y_5 + X_7Y_8 \\ f_6 &= \rho(p^{14}(x, y, z, w)) = (X_1 + X_4)Y_2 + X_5Y_3 + X_4Y_6 + X_8Y_7 \\ f_7 &= \rho(p_{xy}^{24} + p_{zw}^{24}) = X_6Y_2 + X_2Y_3 + X_7Y_6 + (X_2 + X_3)Y_7 \\ f_8 &= \rho(p^{13}(x, y, z, w)) = X_5Y_1 + (X_1 + X_4)Y_4 + X_8Y_5 + X_4Y_8 \\ f_9 &= \rho(p_{xy}^{34} + p_{zw}^{34}) = Y_1Y_7 + Y_2Y_8 + Y_3Y_5 + Y_4Y_6 \\ f_{10} &= \rho(p^{12}(x, y, z, w)) = X_1X_7 + X_2(X_5 + X_8) + X_3X_5 + X_4(X_6 + X_7) \end{aligned}$$

Note that $f_1, \dots, f_8$ are Oil-vinegar polynomials, where we can take either $X = (X_1, \dots, X_8)$ or $Y = (Y_1, \dots, Y_8)$ be the vinegar variables. This implies that, if either X or Y is known, we can use the polynomial equations coming from knowing the value of these polynomials to find the value of the other one by plugging in the values of variables given.

In terms of the original notation, we rename each ψ_j and f_j above as $\psi_{1,j}$ and $f_{1,j}$ respectively, and define $\psi_{i,1} = \psi_{1,1}$, and $f_{i,j} = f_{1,j}$, for $i = 2, 3, 4$, $j = 1, 3, 5, 7, 9$.

Again, in terms of the original notation, interchanging z with w in (3), we define

$$(4) \quad \begin{aligned} \psi_{2,2} &= \rho(p^{34}(x, y, w, z)), \\ f_{2,2} &= \rho(p^{24}(x, y, w, z)), \\ f_{2,4} &= \rho(p^{23}(x, y, w, z)), \\ f_{2,6} &= \rho(p^{14}(x, y, w, z)), \\ f_{2,8} &= \rho(p^{13}(x, y, w, z)), \\ f_{2,10} &= \rho(p^{12}(x, y, w, z)). \end{aligned}$$

Similarly, by interchanging x with y in (3), we define $\psi_{3,2}$ and $f_{3,j}$, for $j = 2, 4, 6, 8, 10$; by interchanging x with y , and z with w in (3), we define $\psi_{4,2}$ and $f_{4,j}$, for $j = 2, 4, 6, 8, 10$. Then we have four identities:

$$(5) \quad \psi_{i,1}\psi_{i,2} = f_{i,1}f_{i,2} + \dots + f_{i,9}f_{i,10},$$

for $0 < i < 5$.

The central map:

$$(Z_1, \dots, Z_{74}) = F(X_1, \dots, X_{24}, Y_1, \dots, Y_{32}),$$

is defined as

$$\begin{aligned} Z_1 &= X_1 + \psi_{1,1}(X_1, \dots, X_8) + \phi_1(X_1) \\ Z_2 &= X_2 + \psi_{1,2}(Y_1, \dots, Y_8) + \phi_2(X_1, X_2) \\ Z_3 &= X_3 + \psi_{2,2}(Y_9, \dots, Y_{16}) + \phi_3(X_1, X_2, X_3) \\ Z_4 &= X_4 + \psi_{3,2}(Y_{17}, \dots, Y_{24}) + \phi_4(X_1, X_2, X_3, X_4) \\ Z_5 &= X_5 + \psi_{2,1}(X_9, \dots, X_{16}) + \phi_5(X_1, X_2, X_3, X_4, X_5) \\ Z_6 &= X_6 + \psi_{3,1}(X_{17}, \dots, X_{24}) + \phi_6(X_1, X_2, X_3, X_4, X_5, X_6) \\ Z_7 &= X_7 + \psi_{4,2}(Y_{25}, \dots, Y_{32}) + \phi_7(X_1, X_2, X_3, X_4, X_5, X_6, X_7) \\ Z_{7+i} &= f_{1,i}(X_1, \dots, X_8, Y_1, \dots, Y_8) \quad i = 1, \dots, 10 \\ Z_{17+i} &= f_{2,i}(X_1, \dots, X_8, Y_9, \dots, Y_{16}) \quad i = 1, \dots, 10 \\ Z_{27+i} &= f_{2,i}(Y_1, \dots, Y_8, Y_9, \dots, Y_{16}) \quad i = 1, \dots, 8 \\ Z_{36} &= f_{2,10}(Y_1, \dots, Y_8, Y_9, \dots, Y_{16}) \\ Z_{36+i} &= f_{3,i}(X_1, \dots, X_8, Y_{17}, \dots, Y_{24}) \quad i = 1, \dots, 10 \\ Z_{46+i} &= f_{2,i}(X_9, \dots, X_{16}, Y_9, \dots, Y_{16}) \quad i = 1, \dots, 8 \\ Z_{55} &= f_{2,10}(X_9, \dots, X_{16}, Y_9, \dots, Y_{16}) \\ Z_{55+i} &= f_{3,i}(X_{17}, \dots, X_{24}, Y_{17}, \dots, Y_{24}) \quad i = 1, \dots, 8 \\ Z_{64} &= f_{3,10}(X_{17}, \dots, X_{24}, Y_{17}, \dots, Y_{24}) \\ Z_{64+i} &= f_{4,i}(X_9, \dots, X_{16}, Y_{25}, \dots, Y_{32}) \quad i = 1, \dots, 10 \end{aligned}$$

Here ϕ_i are randomly chosen quadratic functions.

Note that $f_{2,9}(Y_1, \dots, Y_8, Y_9, \dots, Y_{16})$ is not explicitly in the central map due to the redundancy that

$$(6) \quad f_{2,9}(Y_1, \dots, Y_8, Y_9, \dots, Y_{16}) = f_{2,9}(X_1, \dots, X_8, Y_9, \dots, Y_{16}) = Z_{26}.$$

Similar situation is also true for:

$$f_{2,9}(X_9, \dots, X_{16}, Y_9, \dots, Y_{16}) \text{ and } f_{3,9}(X_{17}, \dots, X_{24}, Y_{17}, \dots, Y_{24}).$$

There are 7 oil-vinegar polynomial systems inside the central map:

- (1) $Z_8, \dots, Z_{15}$, where $Y_1, \dots, Y_8$ are viewed as oil variable and $X_1, \dots, X_8$ are viewed as vinegar variables;
- (2) $Z_{18}, \dots, Z_{25}$, where $Y_9, \dots, Y_{16}$ are viewed as oil variable and $X_1, \dots, X_8$ are viewed as vinegar variables;
- (3) $Z_{28}, \dots, Z_{35}$, where $Y_1, \dots, Y_8$ can be viewed either as oil or vinegar variables and $Y_9, \dots, Y_{16}$ are viewed as the opposite variables;
- (4) $Z_{37}, \dots, Z_{44}$, where $Y_{17}, \dots, Y_{24}$ are viewed as oil variable and $X_1, \dots, X_8$ are viewed as vinegar variables;
- (5) $Z_{47}, \dots, Z_{54}$, where $X_9, \dots, X_{16}$ are viewed as oil variable and $Y_9, \dots, Y_{16}$ are viewed as vinegar variables;
- (6) $Z_{56}, \dots, Z_{63}$, where $X_{17}, \dots, X_{24}$ are viewed as oil variable and $Y_{17}, \dots, Y_{24}$ are viewed as vinegar variables;
- (7) $Z_{65}, \dots, Z_{72}$, where $Y_{25}, \dots, Y_{32}$ are viewed as oil variable and $X_9, \dots, X_{16}$ are viewed as vinegar variables.

The public key is constructed as a map from $k^{56 \times d}$ to $k^{74 \times d}$:

$$Y = (y_1, \dots, y_{74 \times d}) = \bar{F}(x_1, \dots, x_{56 \times d}) = L_1 \circ \Phi \circ F \circ \Phi' \circ L_2(x_1, \dots, x_{56 \times d}),$$

where Φ' is the map from k^{56d} to $\mathcal{F}^{56}$ induced from the map P in (3), and Φ is the map from $\mathcal{F}^{74}$ to k^{74d} induced from P^{-1} .

2.3. The decryption process. The decryption process requires to invert the maps and the key is how to invert the central map.

The key step is first to unmask the triangular system $(Z_1, \dots, Z_7)$, and to derive the value of $X_1, \dots, X_7$ and then X_8 . The rest is just to solve the oil-vinegar systems. We start by focusing on the first three equations of the central map system.

We will first write some of the Diophantine equations satisfied by components of the central map. Let

$$\begin{aligned}
 g_1 &= Z_8 Z_9 + Z_{10} Z_{11} + Z_{12} Z_{13} + Z_{14} Z_{15} + Z_{16} Z_{17} \\
 &= \psi_{1,1}(X_1 \dots X_8) \psi_{1,2}(Y_1 \dots Y_8), \\
 g_2 &= Z_{18} Z_{19} + Z_{20} Z_{21} + Z_{22} Z_{23} + Z_{24} Z_{25} + Z_{26} Z_{27} \\
 &= \psi_{2,1}(X_1 \dots X_8) \psi_{2,2}(Y_9 \dots Y_{16}), \\
 g_3 &= Z_{28} Z_{29} + Z_{30} Z_{31} + Z_{32} Z_{33} + Z_{34} Z_{35} + Z_{26} Z_{36} \\
 &= \psi_{2,1}(Y_1 \dots Y_8) \psi_{2,2}(Y_9 \dots Y_{16}).
 \end{aligned}
 \tag{7}$$

Note that Z_{26} appears in both g_2 and g_3 because of (5), (6).

Since

$$\psi_{2,1}(X_1, \dots, X_8) = \psi_{1,1}(X_1, \dots, X_8), \quad \psi_{2,1}(Y_1, \dots, Y_8) = \psi_{1,2}(Y_1, \dots, Y_8),$$

we have

$$\begin{aligned}
 (8) \quad h_1 &= (g_1 g_2 g_3^{-1})^{1/2} = \psi_{1,1}(X_1, \dots, X_8) \\
 (9) \quad h_2 &= g_1 h_1^{-1} = \psi_{1,2}(Y_1, \dots, Y_8) \\
 (10) \quad h_3 &= g_2 h_1^{-1} = \psi_{2,2}(Y_9, \dots, Y_{16}).
 \end{aligned}$$

This allows us to compute the value of $\psi_{1,1}(X_1, \dots, X_8)$, $\psi_{1,2}(Y_1, \dots, Y_8)$ and $\psi_{2,2}(Y_9, \dots, Y_{16})$.

Since we also have the Diophantine equations:

$$(11) \quad \begin{aligned} g_4 &= Z_{37}Z_{38} + Z_{39}Z_{40} + Z_{41}Z_{42} + Z_{43}Z_{44} + Z_{45}Z_{46} \\ &= \psi_{3,1}(X_1, \dots, X_8)\psi_{3,2}(Y_{17}, \dots, Y_{24}), \end{aligned}$$

$$(12) \quad \begin{aligned} g_5 &= Z_{47}Z_{48} + Z_{49}Z_{50} + Z_{51}Z_{52} + Z_{53}Z_{54} + Z_{26}Z_{55} \\ &= \psi_{2,1}(X_9, \dots, X_{16})\psi_{2,2}(Y_9, \dots, Y_{16}), \end{aligned}$$

$$(13) \quad \begin{aligned} g_6 &= Z_{56}Z_{57} + Z_{58}Z_{59} + Z_{60}Z_{61} + Z_{62}Z_{63} + Z_{45}Z_{64} \\ &= \psi_{3,1}(X_{17}, \dots, X_{24})\psi_{3,2}(Y_{17}, \dots, Y_{24}), \end{aligned}$$

$$(14) \quad \begin{aligned} g_7 &= Z_{65}Z_{66} + Z_{67}Z_{68} + Z_{69}Z_{70} + Z_{71}Z_{72} + Z_{73}Z_{74} \\ &= \psi_{4,1}(X_9, \dots, X_{16})\psi_{4,2}(Y_{25}, \dots, Y_{32}). \end{aligned}$$

Then, since

$$\psi_{3,1}(X_1, \dots, X_8) = \psi_{1,1}(X_1, \dots, X_8),$$

and

$$\psi_{4,1}(X_9, \dots, X_{16}) = \psi_{2,1}(X_9, \dots, X_{16}),$$

we have

$$(15) \quad h_4 = g_4 h_1^{-1} = \psi_{3,2}(Y_{17}, \dots, Y_{24}),$$

$$(16) \quad h_5 = g_5 h_3^{-1} = \psi_{2,1}(X_9, \dots, X_{16}),$$

$$(17) \quad h_6 = g_6 h_4^{-1} = \psi_{3,1}(X_{17}, \dots, X_{24}),$$

$$(18) \quad h_7 = g_7 h_5^{-1} = \psi_{4,2}(Y_{25}, \dots, Y_{32}).$$

Using the value of $h_1, \dots, h_7$, we can restore the triangular structure of $Z_1, \dots, Z_7$, and recover the values of $X_1, \dots, X_7$.

Then we recover X_8 by using the value of $h_1 = \psi_{1,1}(X_1, \dots, X_8)$, as long as X_7 is nonzero, or we can use Z_{17} to recover X_8 as long as X_2 is not zero.

One finishes the inversion of the central map by using the 1,2,4,5,6,7 oil-vinegar systems described in the section above to derive the remaining variables $X_9, \dots, X_{24}$ and $Y_1, \dots, Y_{32}$.

Note that the decryption process succeeds with a high probability but not 1, and our attack only deals with ciphertexts, whose decryption can be performed successfully as above.

2.4. Practical parameters and security claims. The authors [9] suggested the following practical parameters:

- (1) $q = 2^{16}$, $d=1$, the number of variables 56, the number of public key polynomials 74, and the security level claim is 2^{113} ;
- (2) $q = 2^{16}$, $d=2$, the number of variables $56 \times 2 = 112$, the number of public key polynomials $74 \times 2 = 148$, and the security level claim is 2^{221} ;
- (3) $q = 2^{32}$, $d=1$, the number of variables 56, the number of public key polynomials 74, and the security level claim is 2^{114} .

3. THE IDEAL GENERATED BY EMBEDDED SURFACES

In this section, we show the existence of a new type of embedded surfaces for the new cryptosystem from Diophantine equations described in the previous section different from the linearization equations types ones. We use the same notation as in the previous section.

3.1. Embedded surfaces from the decryption process. By analysing the decryption process, one can derive and prove the following theorem for the new cryptosystem from Diophantine equations.

Theorem 1. *Let $q = 2^{16}$, $d=1$, $\mathcal{F} = k = GF(2^{16})$, Assume that the public key of the new cryptosystem from Diophantine equations is $\bar{F} : k^{56} \rightarrow k^{74}$, with*

$$Y = (y_1, \dots, y_{74}) = \bar{F}(x_1, \dots, x_{56}) = L_1 \circ \Phi \circ F \circ \Phi' \circ L_2(x_1, \dots, x_{56}),$$

where Φ' is the map from k^{56d} to $\mathcal{F}^{56}$ induced from the map P in (3), and Φ is the map from $\mathcal{F}^{74}$ to k^{74d} induced from P^{-1} . Then for a pair of ciphertext and plaintext $(x_1, \dots, x_{56})$ and $(y_1, \dots, y_{74})$, there exist relations in the form:

$$(19) \quad \begin{aligned} & \sum a_{i,j,s,t} y_i y_j y_s y_t + \sum b_{i,j,s} y_i y_j y_s + \sum c_{i,j} y_i y_j + \sum d_j y_j + \\ & \sum a'_{i,j,s,t} y_i y_j x_s^2 x_t^2 + \sum b'_{i,s,t} y_i x_s^2 x_t^2 + \sum c'_{i,j,s} y_i y_j x_s^2 + \sum d'_{i,s} y_i x_s^2 + \\ & \sum e'_{s,t} x_s^2 x_t^2 + \sum h'_s x_s^2 + e = 0 \end{aligned}$$

Proof. Let g_1, g_2 , and g_3 be as defined in the relations (7) obtained during the decryption process. (19)

Assume that $g_1 = h_1 \times h_2$,

$g_2 = h_1 \times h_3$,

$g_3 = h_2 \times h_3$. Then

$$(20) \quad \begin{aligned} g_1 g_2 &= g_3 \times h_1^2, \\ g_2 g_3 &= g_1 \times h_3^2, \\ g_1 g_3 &= g_2 \times h_2^2. \end{aligned}$$

By rewriting g_1, g_2 , and g_3 using the Z_i variables, the first equation yields

$$(21) \quad \begin{aligned} & (Z_8 Z_9 + Z_{10} Z_{11} + Z_{12} Z_{13} + Z_{14} Z_{15} + Z_{16} Z_{17}) \times \\ & (Z_{18} Z_{19} + Z_{20} Z_{21} + Z_{22} Z_{23} + Z_{24} Z_{25} + Z_{26} Z_{27}) = \\ & (Z_{28} Z_{29} + Z_{30} Z_{31} + Z_{32} Z_{33} + Z_{34} Z_{35} + Z_{26} Z_{36}) \times (\psi_{1,1}(X_1, \dots, X_8))^2. \end{aligned}$$

i.e.

$$(22) \quad \sum a_{i,j,s,t} Z_i Z_j Z_s Z_t = \sum a'_{i,j,s,t} Z_i Z_j X_s^2 X_t^2.$$

or

$$(23) \quad \sum a_{i,j,s,t} Z_i Z_j Z_s Z_t - \sum a'_{i,j,s,t} Z_i Z_j X_s^2 X_t^2 = 0.$$

for some constants $a_{i,j,s,t}$ and $a'_{i,j,s,t}$

Since $\mathcal{F} = k$ is of characteristics 2, we have

$$(a + b)^2 = a^2 + b^2.$$

for all $a, b \in \mathcal{F} = k$

Therefore, since the decryption requires

$$g_1 \times g_2 \times g_3 \neq 0,$$

given the values of Z_i , we can derive the values of

$$\begin{aligned} h_1^2 &= (\psi_{1,1}(X_1, \dots, X_8))^2, \\ h_2^2 &= (\psi_{1,2}(Y_1, \dots, Y_8))^2, \\ h_3^2 &= (\psi_{2,2}(Y_9, \dots, Y_{16}))^2. \end{aligned}$$

Hence, we can find $h_1 = \psi_{1,1}(X_1, \dots, X_8)$, $h_2 = \psi_{1,2}(Y_1, \dots, Y_8)$ and $h_3 = \psi_{2,2}(Y_9, \dots, Y_{16})$, by taking square-root in k . Now by applying the 2 affine transformations L_1 and L_2 , we have:

$$\begin{aligned} &\sum a_{i,j,s,t} y_i y_j y_s y_t + \sum b_{i,j,s} y_i y_j y_s + \sum c_{i,j} y_i y_j + \sum d_j y_j + \\ &\sum a'_{i,j,s,t} y_i y_j x_s^2 x_t^2 + \sum b'_{i,s,t} y_i x_s^2 x_t^2 + \sum c'_{i,j,s} y_i y_j x_s^2 + \sum d'_{i,s} y_i x_s^2 + \\ (24) \quad &\sum e'_{s,t} x_s^2 x_t^2 + \sum h'_s x_s^2 + e = 0, \end{aligned}$$

for any pair of ciphertext and plaintext $(x_1, \dots, x_{56})$ and $(y_1, \dots, y_{74})$ □

Notice that if the values of all the ciphertext y_i are given, the relation obtained in the conclusion of Theorem 1 becomes

$$(25) \quad \sum a''_{s,t} x_s^2 x_t^2 + \sum b''_s x_s^2 + c'' = 0,$$

which implies that Theorem 1 implies the existence of a none-trivial embedded surface which allows us to find the value of polynomials corresponding to h_1 , h_2 and h_3 , which are not components in the central map. This is true, as long as the corresponding value of $g_1 \times g_2 \times g_3$ is not zero, an assumption required for decryption.

3.2. Embedded surface obtained from the central map. Instead of analysing the decryption process, we perform substantial computer experiments and find many interesting equations associated to the polynomials of the central map. Using the same set of parameters, the following equations can be deduced:

$$\begin{aligned}
& (X_2 + X_3)Z_8 + X_8Z_{11} + (X_1 + X_4)Z_{12} + X_6Z_{15} + Y_8Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_1 \\
& X_4Z_9 + X_7Z_{10} + X_2Z_{13} + X_5Z_{14} + Y_7Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_2 \\
& X_8Z_9 + (X_2 + X_3)Z_{10} + X_6Z_{13} + (X_1 + X_4)Z_{14} + Y_6Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_3 \\
& X_7Z_8 + X_4Z_{11} + X_5Z_{12} + X_2Z_{15} + Y_5Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_4 \\
& X_2Z_8 + X_5Z_{11} + X_4Z_{12} + X_7Z_{15} + Y_4Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_5 \\
& (X_1 + X_4)Z_9 + X_6Z_{10} + (X_2 + X_3)Z_{13} + X_8Z_{14} + Y_3Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_6 \\
& X_5Z_9 + X_2Z_{10} + X_7Z_{13} + X_4Z_{14} + Y_2Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_7 \\
& X_6Z_8 + (X_1 + X_4)Z_{11} + X_8Z_{12} + (X_2 + X_3)Z_{15} + Y_1Z_{17} &= \psi_{1,1}(X_1, \dots, X_8)Y_8 \\
& X_2Z_{18} + X_3Z_{18} + X_5Z_{21} + X_1Z_{22} + X_4Z_{22} + X_7Z_{25} + Y_{12}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_9 \\
& X_1Z_{19} + X_4Z_{19} + X_7Z_{20} + X_2Z_{23} + X_3Z_{23} + X_5Z_{24} + Y_{11}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{10} \\
& X_5Z_{19} + (X_2 + X_3)Z_{20} + X_7Z_{23} + (X_1 + X_4)Z_{24} + Y_{10}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{11} \\
& X_7Z_{18} + (X_1 + X_4)Z_{21} + X_5Z_{22} + (X_2 + X_3)Z_{25} + Y_9Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{12} \\
& X_2Z_{18} + X_8Z_{21} + X_4Z_{22} + X_6Z_{25} + Y_{16}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{13} \\
& X_4Z_{19} + X_6Z_{20} + X_2Z_{23} + X_8Z_{24} + Y_{15}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{14} \\
& X_8Z_{19} + X_2Z_{20} + X_6Z_{23} + X_4Z_{24} + Y_{14}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{15} \\
& X_6Z_{18} + X_4Z_{21} + X_8Z_{22} + X_2Z_{25} + Y_{13}Z_{27} &= \psi_{1,1}(X_1, \dots, X_8)Y_{16} \\
& (X_2 + X_3)Z_{37} + X_8Z_{40} + (X_1 + X_4)Z_{41} + X_6Z_{44} + Y_{24}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)Y_{17} \\
& X_4Z_{38} + (X_6 + X_7)Z_{39} + X_2Z_{42} + (X_5 + X_8)Z_{43} + Y_{23}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)(Y_{18} + Y_{22}) \\
& X_5Z_{38} + X_3Z_{39} + X_7Z_{42} + X_1Z_{43} + Y_{18}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)(Y_{19} + Y_{23}) \\
& X_8Z_{38} + (X_2 + X_3)Z_{39} + X_6Z_{42} + (X_1 + X_4)Z_{43} + Y_{22}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)Y_{19} \\
& (X_6 + X_7)Z_{37} + X_4Z_{40} + (X_5 + X_8)Z_{41} + X_2Z_{44} + Y_{21}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)(Y_{20} + Y_{24}) \\
& X_3Z_{37} + X_5Z_{40} + X_1Z_{41} + X_7Z_{44} + Y_{20}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)(Y_{17} + Y_{21}) \\
& (X_1 + X_4)Z_{38} + X_6Z_{39} + (X_2 + X_3)Z_{42} + X_8Z_{43} + Y_{19}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)Y_{22} \\
& X_6Z_{37} + (X_1 + X_4)Z_{40} + X_8Z_{41} + X_2Z_{44} + X_3Z_{44} + Y_{17}Z_{46} &= \psi_{1,1}(X_1, \dots, X_8)Y_{24} \\
& Y_8Z_9 + Y_6Z_{11} + Y_3Z_{12} + Y_1Z_{14} + (X_2 + X_3)Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)X_6 \\
& Y_5Z_9 + Y_7Z_{11} + Y_2Z_{12} + Y_4Z_{14} + X_7Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)X_2 \\
& Y_4Z_9 + Y_2Z_{11} + Y_7Z_{12} + Y_5Z_{14} + X_2Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)X_7 \\
& Y_1Z_9 + Y_3Z_{11} + Y_6Z_{12} + Y_8Z_{14} + X_6Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)(X_2 + X_3) \\
& Y_7Z_8 + Y_5Z_{10} + Y_4Z_{13} + Y_2Z_{15} + X_4Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)X_5 \\
& Y_6Z_8 + Y_8Z_{10} + Y_1Z_{13} + Y_3Z_{15} + X_8Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)(X_1 + X_4) \\
& Y_3Z_8 + Y_1Z_{10} + Y_8Z_{13} + Y_6Z_{15} + (X_1 + X_4)Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)X_8 \\
& Y_2Z_8 + Y_4Z_{10} + Y_5Z_{13} + Y_7Z_{15} + X_5Z_{16} &= \psi_{1,2}(Y_1, \dots, Y_8)X_4 \\
& Y_7Z_{28} + (Y_1 + Y_4)Z_{31} + Y_5Z_{32} + (Y_2 + Y_3)Z_{35} + Y_9Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{12} \\
& Y_5Z_{29} + (Y_2 + Y_3)Z_{30} + Y_7Z_{33} + (Y_1 + Y_4)Z_{34} + Y_{10}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{11} \\
& (Y_1 + Y_4)Z_{29} + Y_7Z_{30} + (Y_2 + Y_3)Z_{33} + Y_5Z_{34} + Y_{11}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{10} \\
& (Y_2 + Y_3)Z_{28} + Y_5Z_{31} + (Y_1 + Y_4)Z_{32} + Y_7Z_{35} + Y_{12}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_9 \\
& Y_6Z_{28} + Y_4Z_{31} + Y_8Z_{32} + Y_2Z_{35} + Y_{13}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{16} \\
& Y_8Z_{29} + Y_2Z_{30} + Y_6Z_{33} + Y_4Z_{34} + Y_{14}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{15} \\
& Y_4Z_{29} + Y_6Z_{30} + Y_2Z_{33} + Y_8Z_{34} + Y_{15}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{14} \\
& Y_2Z_{28} + Y_8Z_{31} + Y_4Z_{32} + Y_6Z_{35} + Y_{16}Z_{36} &= \psi_{1,2}(Y_1, \dots, Y_8)Y_{13} \\
& Y_7Z_{26} + Y_9Z_{29} + Y_{11}Z_{31} + Y_{10}Z_{32} + Y_{12}Z_{34} &= \psi_{2,2}(Y_9, \dots, Y_{16})(Y_2 + Y_3) \\
& Y_2Z_{26} + Y_{16}Z_{29} + Y_{14}Z_{31} + Y_{15}Z_{32} + Y_{13}Z_{34} &= \psi_{2,2}(Y_9, \dots, Y_{16})Y_6 \\
& Y_{11}Z_{18} + Y_9Z_{20} + Y_{12}Z_{23} + Y_{10}Z_{25} + (X_1 + X_4)Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_5 \\
& (Y_1 + Y_4)Z_{26} + Y_{11}Z_{28} + Y_9Z_{30} + Y_{12}Z_{33} + Y_{10}Z_{35} &= \psi_{2,2}(Y_9, \dots, Y_{16})Y_5 \\
& Y_4Z_{26} + Y_{15}Z_{28} + Y_{13}Z_{30} + Y_{16}Z_{33} + Y_{14}Z_{35} &= \psi_{2,2}(Y_9, \dots, Y_{16})Y_8 \\
& Y_{13}Z_{19} + Y_{15}Z_{21} + Y_{14}Z_{22} + Y_{16}Z_{24} + X_6Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_2 \\
& Y_{15}Z_{18} + Y_{13}Z_{20} + Y_{16}Z_{23} + Y_{14}Z_{25} + X_4Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_8 \\
& Y_{12}Z_{19} + Y_{10}Z_{21} + Y_{11}Z_{22} + Y_9Z_{24} + (X_2 + X_3)Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_7 \\
& Y_6Z_{26} + Y_{13}Z_{29} + Y_{15}Z_{31} + Y_{14}Z_{32} + Y_{16}Z_{34} &= \psi_{2,2}(Y_9, \dots, Y_{16})Y_2 \\
& Y_5Z_{26} + Y_{10}Z_{28} + Y_{12}Z_{30} + Y_9Z_{33} + Y_{11}Z_{35} &= \psi_{2,2}(Y_9, \dots, Y_{16})(Y_1 + Y_4) \\
& Y_{14}Z_{18} + Y_{16}Z_{20} + Y_{13}Z_{23} + Y_{15}Z_{25} + X_8Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_4 \\
& Y_{16}Z_{19} + Y_{14}Z_{21} + Y_{15}Z_{22} + Y_{13}Z_{24} + X_2Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_6 \\
& Y_8Z_{26} + Y_{14}Z_{28} + Y_{16}Z_{30} + Y_{13}Z_{33} + Y_{15}Z_{35} &= \psi_{2,2}(Y_9, \dots, Y_{16})Y_4 \\
& Y_9Z_{19} + Y_{11}Z_{21} + Y_{10}Z_{22} + Y_{12}Z_{24} + X_7Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})(X_2 + X_3) \\
& Y_{10}Z_{18} + Y_{12}Z_{20} + Y_9Z_{23} + Y_{11}Z_{25} + X_5Z_{26} &= \psi_{2,2}(Y_9, \dots, Y_{16})(X_1 + X_4) \\
& (Y_2 + Y_3)Z_{26} + Y_{12}Z_{29} + Y_{10}Z_{31} + Y_{11}Z_{32} + Y_9Z_{34} &= \psi_{2,2}(Y_9, \dots, Y_{16})Y_7 \\
& (X_{10} + X_{11})Z_{26} + Y_{12}Z_{48} + Y_{10}Z_{50} + Y_{11}Z_{51} + Y_9Z_{53} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_{15} \\
& X_{15}Z_{26} + Y_9Z_{48} + Y_{11}Z_{50} + Y_{10}Z_{51} + Y_{12}Z_{53} &= \psi_{2,2}(Y_9, \dots, Y_{16})(X_{10} + X_{11}) \\
& X_{10}Z_{26} + Y_{16}Z_{48} + Y_{14}Z_{50} + Y_{15}Z_{51} + Y_{13}Z_{53} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_{14} \\
& X_{14}Z_{26} + Y_{13}Z_{48} + Y_{15}Z_{50} + Y_{14}Z_{51} + Y_{16}Z_{53} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_{10} \\
& (X_9 + X_{12})Z_{26} + Y_{11}Z_{47} + Y_9Z_{49} + Y_{12}Z_{52} + Y_{10}Z_{54} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_{13} \\
& X_{13}Z_{26} + Y_{10}Z_{47} + Y_{12}Z_{49} + Y_9Z_{52} + Y_{11}Z_{54} &= \psi_{2,2}(Y_9, \dots, Y_{16})(X_9 + X_{12}) \\
& X_{12}Z_{26} + Y_{15}Z_{47} + Y_{13}Z_{49} + Y_{16}Z_{52} + Y_{14}Z_{54} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_{16} \\
& X_{16}Z_{26} + Y_{14}Z_{47} + Y_{16}Z_{49} + Y_{13}Z_{52} + Y_{15}Z_{54} &= \psi_{2,2}(Y_9, \dots, Y_{16})X_{12} \\
& Y_{20}Z_{38} + Y_{18}Z_{40} + (Y_{19} + Y_{23})Z_{41} + (Y_{17} + Y_{21})Z_{43} + X_3Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_7 \\
& Y_{17}Z_{38} + Y_{19}Z_{40} + Y_{22}Z_{41} + Y_{24}Z_{43} + X_6Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})(X_2 + X_3) \\
& Y_{24}Z_{38} + Y_{22}Z_{40} + Y_{19}Z_{41} + Y_{17}Z_{43} + (X_2 + X_3)Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_6 \\
& Y_{21}Z_{38} + Y_{23}Z_{40} + (Y_{18} + Y_{22})Z_{41} + (Y_{20} + Y_{24})Z_{43} + (X_6 + X_7)Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_2 \\
& (Y_{19} + Y_{23})Z_{37} + (Y_{17} + Y_{21})Z_{39} + Y_{20}Z_{42} + Y_{18}Z_{44} + X_1Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_5 \\
& Y_{22}Z_{37} + Y_{24}Z_{39} + Y_{17}Z_{42} + Y_{19}Z_{44} + X_8Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})(X_1 + X_4) \\
& Y_{19}Z_{37} + Y_{17}Z_{39} + Y_{24}Z_{42} + Y_{22}Z_{44} + (X_1 + X_4)Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_8 \\
& (Y_{18} + Y_{22})Z_{37} + (Y_{20} + Y_{24})Z_{39} + Y_{21}Z_{42} + Y_{23}Z_{44} + (X_5 + X_8)Z_{45} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_4 \\
& X_{19}Z_{45} + Y_{20}Z_{57} + Y_{18}Z_{59} + (Y_{19} + Y_{23})Z_{60} + (Y_{17} + Y_{21})Z_{62} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_{23} \\
& X_{22}Z_{45} + Y_{17}Z_{57} + Y_{19}Z_{59} + Y_{22}Z_{60} + Y_{24}Z_{62} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})(X_{18} + X_{19}) \\
& (X_{18} + X_{19})Z_{45} + Y_{24}Z_{57} + Y_{22}Z_{59} + Y_{19}Z_{60} + Y_{17}Z_{62} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_{22} \\
& (X_{22} + X_{23})Z_{45} + Y_{21}Z_{57} + Y_{23}Z_{59} + (Y_{18} + Y_{22})Z_{60} + (Y_{20} + Y_{24})Z_{62} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_{18} \\
& X_{17}Z_{45} + (Y_{19} + Y_{23})Z_{56} + (Y_{17} + Y_{21})Z_{58} + Y_{20}Z_{61} + Y_{18}Z_{63} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_{21} \\
& X_{24}Z_{45} + Y_{22}Z_{56} + Y_{24}Z_{58} + Y_{17}Z_{61} + Y_{19}Z_{63} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})(X_{17} + X_{20}) \\
& (X_{17} + X_{20})Z_{45} + Y_{19}Z_{56} + Y_{17}Z_{58} + Y_{24}Z_{61} + Y_{22}Z_{63} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_{24} \\
& (X_{21} + X_{24})Z_{45} + (Y_{18} + Y_{22})Z_{56} + (Y_{20} + Y_{24})Z_{58} + Y_{21}Z_{61} + Y_{23}Z_{63} &= \psi_{3,2}(Y_{17}, \dots, Y_{24})X_{20} \\
& X_{15}Z_{47} + (X_9 + X_{12})Z_{50} + X_{13}Z_{51} + (X_{10} + X_{11})Z_{54} + Y_9Z_{55} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{12} \\
& X_{13}Z_{48} + (X_{10} + X_{11})Z_{49} + X_{15}Z_{52} + (X_9 + X_{12})Z_{53} + Y_{10}Z_{55} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{11} \\
& (X_9 + X_{12})Z_{48} + X_{15}Z_{49} + (X_{10} + X_{11})Z_{52} + X_{13}Z_{53} + Y_{11}Z_{55} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{10}
\end{aligned}$$

$$\begin{aligned}
& X_{15}Z_{65} + (X_9 + X_{12})Z_{68} + X_{13}Z_{69} + (X_{10} + X_{11})Z_{72} + Y_{25}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{28} \\
& X_{13}Z_{66} + (X_{10} + X_{11})Z_{67} + X_{15}Z_{70} + (X_9 + X_{12})Z_{71} + Y_{26}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{27} \\
& (X_9 + X_{12})Z_{66} + X_{15}Z_{67} + (X_{10} + X_{11})Z_{70} + X_{13}Z_{71} + Y_{27}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{26} \\
& (X_{10} + X_{11})Z_{65} + X_{13}Z_{68} + (X_9 + X_{12})Z_{69} + X_{15}Z_{72} + Y_{28}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})Y_{25} \\
& (X_{14} + X_{15})Z_{65} + X_{12}Z_{68} + (X_{13} + X_{16})Z_{69} + X_{10}Z_{72} + Y_{29}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})(Y_{28} + Y_{32}) \\
& X_{16}Z_{66} + X_{11}Z_{67} + X_{14}Z_{70} + X_9Z_{71} + Y_{30}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})(Y_{27} + Y_{31}) \\
& X_{12}Z_{66} + (X_{14} + X_{15})Z_{67} + X_{10}Z_{70} + (X_{13} + X_{16})Z_{71} + Y_{31}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})(Y_{26} + Y_{30}) \\
& X_{11}Z_{65} + X_{16}Z_{68} + X_9Z_{69} + X_{14}Z_{72} + Y_{32}Z_{74} &= \psi_{2,1}(X_9, \dots, X_{16})(Y_{25} + Y_{29}) \\
& X_{22}Z_{56} + (X_{17} + X_{20})Z_{59} + X_{24}Z_{60} + (X_{18} + X_{19})Z_{63} + Y_{17}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})Y_{24} \\
& X_{21}Z_{57} + X_{19}Z_{58} + X_{23}Z_{61} + X_{17}Z_{62} + Y_{18}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})(Y_{19} + Y_{23}) \\
& (X_{17} + X_{20})Z_{57} + X_{22}Z_{58} + (X_{18} + X_{19})Z_{61} + X_{24}Z_{62} + Y_{19}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})Y_{22} \\
& X_{19}Z_{56} + X_{21}Z_{59} + X_{17}Z_{60} + X_{23}Z_{63} + Y_{20}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})(Y_{17} + Y_{21}) \\
& (X_{22} + X_{23})Z_{56} + X_{20}Z_{59} + (X_{21} + X_{24})Z_{60} + X_{18}Z_{63} + Y_{21}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})(Y_{20} + Y_{24}) \\
& X_{24}Z_{57} + (X_{18} + X_{19})Z_{58} + X_{22}Z_{61} + (X_{17} + X_{20})Z_{62} + Y_{22}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})Y_{19} \\
& X_{20}Z_{57} + (X_{22} + X_{23})Z_{58} + X_{18}Z_{61} + (X_{21} + X_{24})Z_{62} + Y_{23}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})(Y_{18} + Y_{22}) \\
& (X_{18} + X_{19})Z_{56} + X_{24}Z_{59} + (X_{17} + X_{20})Z_{60} + X_{22}Z_{63} + Y_{24}Z_{64} &= \psi_{3,1}(X_{17}, \dots, X_{24})Y_{17} \\
& Y_{28}Z_{66} + Y_{26}Z_{68} + Y_{27}Z_{69} + Y_{25}Z_{71} + (X_{10} + X_{11})Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})X_{15} \\
& Y_{25}Z_{66} + Y_{27}Z_{68} + Y_{26}Z_{69} + Y_{28}Z_{71} + X_{15}Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})(X_{10} + X_{11}) \\
& Y_{32}Z_{66} + Y_{30}Z_{68} + (Y_{27} + Y_{31})Z_{69} + (Y_{25} + Y_{29})Z_{71} + X_{11}Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})X_{14} \\
& Y_{29}Z_{66} + Y_{31}Z_{68} + (Y_{26} + Y_{30})Z_{69} + (Y_{28} + Y_{32})Z_{71} + (X_{14} + X_{15})Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})X_{10} \\
& Y_{27}Z_{65} + Y_{25}Z_{67} + Y_{28}Z_{70} + Y_{26}Z_{72} + (X_9 + X_{12})Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})X_{13} \\
& Y_{26}Z_{65} + Y_{28}Z_{67} + Y_{25}Z_{70} + Y_{27}Z_{72} + X_{13}Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})(X_9 + X_{12}) \\
& (Y_{27} + Y_{31})Z_{65} + (Y_{25} + Y_{29})Z_{67} + Y_{32}Z_{70} + Y_{30}Z_{72} + X_9Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})X_{16} \\
& (Y_{26} + Y_{30})Z_{65} + (Y_{28} + Y_{32})Z_{67} + Y_{29}Z_{70} + Y_{31}Z_{72} + (X_{13} + X_{16})Z_{73} &= \psi_{4,2}(Y_{25}, \dots, Y_{32})X_{12}
\end{aligned}$$

From the first two equations, we have

$$\frac{(X_2 + X_3)Z_8 + X_8Z_{11} + (X_1 + X_4)Z_{12} + X_6Z_{15} + Y_8Z_{17}}{X_4Z_9 + X_7Z_{10} + X_2Z_{13} + X_5Z_{14} + Y_7Z_{17}} = \frac{Y_1}{Y_2},$$

i.e.,

$$\begin{aligned}
& Y_2((X_2 + X_3)Z_8 + X_8Z_{11} + (X_1 + X_4)Z_{12} + X_6Z_{15} + Y_8Z_{17}) \\
& - (X_4Z_9 + X_7Z_{10} + X_2Z_{13} + X_5Z_{14} + Y_7Z_{17})Y_1 = 0.
\end{aligned}$$

Let $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_m)$ be a pair of corresponding plaintext and ciphertext for a cryptosystem. Since

$$\begin{aligned}
& (X_1, \dots, X_{24}, Y_1, \dots, Y_{32}) = L_2(x_1, \dots, x_n), \\
& (Z_1, \dots, Z_m) = L_1^{-1}(y_1, \dots, y_m),
\end{aligned}$$

which yield

$$(26) \quad \sum \alpha_{i,j,t} x_i x_j y_t + \sum \beta_{i,j} x_i x_j + \sum \gamma_{i,t} x_i y_t + \sum \mu_t y_t + \sum \nu_i x_i + \lambda = 0$$

for any plaintext-ciphertext pair. Clearly, given the values of all the ciphertext y_i , we obtain

$$(27) \quad \sum a''_{s,t} x_s^2 x_t^2 + \sum b''_s x_s^2 + c'' = 0,$$

That is, we obtain new embadded surface and the following theorem is proved.

Theorem 2. Let $q = 2^{16}$, $d=1$, $\mathcal{F} = k = GF(2^{16})$, Assume that the public key of the new cryptosystem from Diophantine equations is $F : k^{56} \rightarrow k^{74}$ is the central map of the new cryptosystem from Diophantine equations. Then by performing substantial computer experiments, we can derive relations in the form:

$$(28) \quad \sum \alpha_{i,j,t} x_i x_j y_t + \sum \beta_{i,j} x_i x_j + \sum \gamma_{i,t} x_i y_t + \sum \mu_t y_t + \sum \nu_i x_i + \lambda = 0$$

for any given pair of ciphertext and plaintext $(x_1, \dots, x_{56})$ and $(y_1, \dots, y_{74})$.

It is important to emphasize that this comes as a surprise. Without computer experiments, we would not be able to derive these equations. We did not want to include in this paper the complete set of equations of the central map but it is available.

4. EMBEDDED SURFACE ATTACKS

In this section, we show that the new embedded surface obtained in the previous section can be used to recover the message.

4.1. Embedded surface attack from the decryption process and its complexity. By considering the linear space spanned by all the embedded surfaces in the form of (24), we will be able to derive the corresponding values of h_1 and h_2 and h_3 for any valid (decryption possible) of the ciphertext $(y'_1, \dots, y'_{74})$. If we amend those equations to the original system, in the context of original polynomial system derived from the known ciphertext, we will be able to derive the values corresponding to X_1, X_2, X_3 .

On the other hand, looking at the relations that are used to derive $h_3, \dots, h_7$ in the decryption process we obtain:

$$(29) \quad \begin{aligned} h_4 h_1 &= \psi_{3,2}(Y_{17}, \dots, Y_{24}) h_1 = g_4 = \\ &Z_{37} Z_{38} + Z_{39} Z_{40} + Z_{41} Z_{42} + Z_{43} Z_{44} + Z_{45} Z_{46}, \end{aligned}$$

$$(30) \quad \begin{aligned} h_5 h_3 &= \psi_{2,1}(X_9, \dots, X_{16}) h_1 = g_5 = \\ &Z_{47} Z_{48} + Z_{49} Z_{50} + Z_{51} Z_{52} + Z_{53} Z_{54} + Z_{26} Z_{55}. \end{aligned}$$

and

$$(31) \quad \begin{aligned} h_6 h_4 &= \psi_{3,1}(X_{17}, \dots, X_{24}) h_1 = g_6 = \\ &Z_{56} Z_{57} + Z_{58} Z_{59} + Z_{60} Z_{61} + Z_{62} Z_{63} + Z_{45} Z_{64}, \end{aligned}$$

$$(32) \quad \begin{aligned} h_7 h_5 &= \psi_{4,2}(Y_{25}, \dots, Y_{32}) h_5 = g_7 = \\ &Z_{65} Z_{66} + Z_{67} Z_{68} + Z_{69} Z_{70} + Z_{71} Z_{72} + Z_{73} Z_{74}. \end{aligned}$$

This means, in the central map system, if we amend the h_1 and h_2 and h_3 to the map, and if we apply either Groebner basis algorithm like F_4 or F_5 of Faugere or the mutant XL family of algorithms [11],[2], in the first computation round when the algorithm reaches degree 4, we will derive the values of h_4 and h_5 as mutants, and in the next computation round, which is still at degree 4, we will derive h_6 and h_7 as mutants. This implies that we can derive the values of $X_4, \dots, X_7$ and therefore the value of X_8 as in the decryption process. Since the rest are just Oil-Vinegar type of systems, this further implies that we can solve the system at degree 4 using polynomial solving algorithms once the values h_1, h_2, h_3 are derived using the embedded surfaces in the form (24).

The above enables us to develop the following complete algorithm to attack the system.

4.1.1. *Attacking Algorithm.*(1) **Step 1. Find all the embedded surfaces**

Randomly pick

$$\binom{74+4}{4} + \binom{56+2}{2} \times \binom{74+2}{2} = 1426425 + 4711050 = 6137475$$

ciphertext and plaintext pairs derived from the public key, and substitute them into the equation in the form of (24), where the coefficients of the system are treated as variables. This will give us a set of linear equations with 6137475 variables over $GF(2^{16})$ and the same number of equations.

Find the solutions for this set of linear equations. The solution space should be of dimension:

$$(76 \times 75)/2 \times 76 + 76 \times 76 + 76 + 3 = 222455,$$

where 222455 of them come from equations derived from the trivial relations from terms like $Z_i Z_j Z_k^2$, $Z_j^2 Z_i$ and Z_i^2 , and only 3 of them is what we really need, namely the ones coming from (20).

(2) Step 2. Derive new equations from the embedded surfaces

With the 222455 embedded surfaces, once given any valid ciphertext, we substitute it into the embedded surfaces, we will derive $74 + 3 = 77$ linearly independent degree 4 equations in the form:

$$\sum A_{s,t} x_s^2 x_t^2 \sum B_s x_s^2 + C = 0.$$

Then, we take the square root of these equations due to the fact that it is over a field of characteristic 2, namely

$$\sum A_{s,t}^{1/2} x_s x_t \sum B_s^{1/2} x_s + C^{1/2} = 0,$$

since

$$(\sum A_{s,t}^{1/2} x_s x_t \sum B_s^{1/2} x_s + C^{1/2})^2 = \sum A_{s,t} x_s^2 x_t^2 \sum B_s x_s^2 + C,$$

and computing the square root over a field of characteristic 2 is easy to do. This gives us a set of 77 linearly independent quadratic equations. All the public equations derived from the public key and the known ciphertext are already included in the span of this set of equations.

(3) Step 3. Reduce three variables

Perform Gaussian elimination on this set of equations to look for an equation in the form of

$$\sum a_i x_i^2 + \sum b_i x_i + c = 0,$$

where

$$a_i = b_i^2 \times \alpha,$$

for a fixed constant α . Solving this quadratic equation will give us a linear in the form

$$\sum a_i x_i + b = 0.$$

This corresponds to deriving the value of X_1 , which comes from Z_1 by eliminating $\psi_{1,1}(X_1, \dots, X_8)$ due to the known value of $h_1 = \psi_{1,1}(X_1, \dots, X_8)$.

We will then substitute this linear equation into the system, and perform Gaussian elimination, which gives us again an equation in the form of

$$\sum a_i x_i^2 + \sum b_i x_i + c = 0,$$

where again

$$a_i = b_i^2 \times \alpha',$$

for a some fixed constant α' . Solving this quadratic equation will give us a linear equation in the form $\sum a_i x_i + b = 0$.

This corresponds to deriving the value of X_2 , which comes from Z_2 and the known value of h_2 .

We will repeat the process to derive a new linear equation, which corresponds to deriving the value of X_3 coming from Z_3 and the known value of h_3 .

(4) **Step 4. Solve the system**

We will feed the new system including the 3 new linear equations into a Groebner solver like F_4 or the mutant XL algorithm. We will solve the systems at degree 4. This will give us the value of the whole plaintext.

In principle, we can merge Step 3 and Step 4 by using directly the algebraic solver like F_4 or the mutant XL algorithm, which will yield the same results.

4.1.2. *The complexity of the attack.* From the attacking steps, it is clear that the complexity of the attacks concentrates on **Step 1** and **Step 4**.

In Step 1, the key part is to solve a system of linear equations with $N = 6137475$ variables and the same number of equations. If we use usual Gaussian elimination, the complexity will be roughly

$$2N^3/3 = 154126724635276031250 \approx 2^{68},$$

over the field $GF(2^{16})$. Assume that we use the best optimized linear solver, we should have the complexity $N^{2.3} \approx 2^{52}$ theoretically.

In Step 4, we will need to solve a linear system roughly with $\binom{53+4}{4} = 395010$ variables and the same number of equations. Clearly this system is much smaller than the system above, whose complexity is much smaller.

Therefore, we conclude that if we use the optimized Gaussian elimination, the complexity will be roughly 2^{52} theoretically, and the complexity will be 2^{68} with usual Gaussian elimination. This complexity is based on operations over $GF(2^{16})$. The original security claim for the system is 2^{113} with the assumption of using the optimized Gaussian elimination.

4.1.3. *The complexity for attacking the other two systems.* For the case, where $q = 2^{32}$ and $d=1$, the attack complexity will be precisely the same except that everything will be on a field over $GF(2^{32})$. The original security assumption is 2^{114} with the assumption of using the optimized Gaussian elimination.

As for the case, where $q = 2^{16}$, $d=2$, it is clear the complexity will be determined by solving a set of linear equations with the number of variables and equations as

$$N = \binom{2 * 74 + 4}{4} + \binom{2 * 56 + 2}{2} * \binom{2 * 74 + 2}{2} = 93352225.$$

If we use normal Gaussian elimination, the complexity will be roughly

$$2N^3/3 \approx 2^{79},$$

over the field $GF(2^{16})$. Assume that we use the best optimized linear solver, we should have the complexity $N^{2.3} \approx 2^{61}$ theoretically. The original security claim is 2^{221} with assumption of using the optimized Gaussian elimination.

Due to the memory constraints in our own equipments, we could not perform the experiments to attack the system in practice. But we did perform some small

scale toy experiments, where we set some of the variables in the central map to be 0, to confirm that our attack works indeed.

4.2. Embedded surface attack from the central map and its complexity.

4.2.1. *The Attack and its complexity.* The observation above helps us develop an efficient algorithm to attack the system. The attack is performed in two steps to derive the plaintext $(\bar{x}_1, \dots, \bar{x}_n)$ from the ciphertext $(\bar{y}_1, \dots, \bar{y}_m)$.

Step 1: Find a basis for the space V of solutions of the equations

$$\sum \alpha_{i,j,t} x_i x_j y_t + \sum \beta_{i,j} x_i x_j + \sum \gamma_{i,t} x_i y_t + \sum \mu_k y_t + \sum \nu_i x_i + \lambda = 0$$

This step is the same for any given ciphertext. Plugging a given plaintext-ciphertext pair into the equation above, a new equation is obtained in the $\binom{n+2}{2}(m+1)$ unknowns. $\alpha_{ijt}, \beta_{ij}, \gamma_{it}, \mu_t, \nu_i, \lambda \in k$.

Therefore, choosing roughly $N = \binom{n+2}{2}(m+1)$ plaintext-ciphertext pairs and plugging into the equation yield a system of N linear equations in N unknowns. In the context of the first set of proposed parameters, we have $N = \binom{n+2}{2}(m+1) = 123975$ i.e. 123975 equations in 123975 unknowns are obtained. The complexity of solving such system using Gaussian elimination is roughly

$$2N^3/3 = 1270314021656250 \approx 2^{51},$$

over the field $GF(2^{16})$, and is theoretically $N^{2.3766} \approx 2^{41}$ using the best optimized linear solver. Notice here that the dimension of the solution space of the system above is at least $2 \times \binom{24}{2} + 3 \times \binom{16}{2} + 2 \times \binom{8}{2} = 968$, that is, more than 968 independent quadratic equations are obtained.

Step 2: Use Mutant XL algorithm or F4 algorithm to solve these independent quadratic equations in N variables to recover the plaintext.

Mutant XL algorithm[2][11] is an algorithm for solving systems of polynomial equations such as $p_i(x_1, x_2, \dots, x_n) = 0, 1 \leq i \leq m$. It is a variant of the XL algorithm. Both are based on finding new elements that are univariate or linear polynomials in the ideal generated by the polynomials $p_i, 1 \leq i \leq m$. Let F be a finite field with q elements and

$$R = F[x_1, x_2, \dots, x_n] / (x_1^q - x_1, x_2^q - x_2, \dots, x_n^q - x_n)$$

be a ring of polynomial functions over F in the n variables $x_1, x_2, \dots, x_n$. Denote by P the element $(p_1, p_2, \dots, p_m) \in R^m$. The polynomial coordinates p_i of P will be changing throughout the operations of the algorithm. Let D be the maximum degree of the polynomials p_i . The Mutant XL algorithm extends the original system of polynomials equations by multiplying all the polynomials p_i by all monomials up to degree $D - \deg(p_i)$. Hence, a new linearized system is obtained by considering the monomials as new variables. We then apply Gaussian elimination on this resulting linear system. In [1], Barget et al. determined the computation complexity for F4 and XL algorithms over the finite field k to be

$$O\left(\binom{n + d_{reg}}{d_{reg}}^\omega\right),$$

where $\omega = 3$ in the usual Gaussian elimination algorithm and $\omega = 2.3766$ in improved algorithm, d_{reg} is the degree of regularity of the ideal formed by the polynomials in the system; it is given by the degree of the first term with negative coefficient in the expansion of

$$\frac{\prod_{i=1}^{m_1} (1 - z^{d_i})}{(1 - z)^n},$$

with d_i being the total degree of the i -th equations. For the given parameters, $m_1 = 968, n = 56$, the degree of regularity is 2. Therefore, using the usual Gaussian elimination, the complexity will be roughly 2^{33} over the field $GF(2^{16})$. We implemented the attack and the result is showed in Table 1.

4.2.2. Complexity of the attack for the other two sets of proposed parameters. For the case, where $q = 2^{32}$ and $d = 1$, the attack complexity will be mainly the same except that everything will be on the finite field $GF(2^{32})$. The original security assumption in this case is 2^{114} with the assumption of using the optimized Gaussian elimination.

As for the case, where $q = 2^{16}$, $d = 2$, a system of $N = \binom{112+2}{2} \times (148 + 1) = 959709$ linear equations with N variables is obtained. The complexity for using the usual Gaussian elimination is roughly

$$2N^3/3 \approx 2^{59},$$

over the field $GF(2^{16})$ and is theoretically $N^{2.3766} \approx 2^{48}$ for the best optimized linear solver. The original security claim is 2^{221} with assumption of using the optimized Gaussian elimination.

4.3. Direct algebraic attack. One may think to use F_4 or the Mutant XL directly to attack the new MPKCs. In this case, it can be seen easily that our embedded surface implies that the degree of regularity [1] of the systems is actually 4, since each y_i is of degree 2 and our surface is of degree 2 in y_i . Therefore the attack complexity using F_4 or the mutant XL directly is

$$O\left(\binom{n+4}{4}^\omega\right).$$

For the case, where $q = 2^{16}, d = 1, n = 56, m = 74$, F4 algorithm which is implemented in Magma is used directly to attack the new MPKCs. The experiment results obtained are showed in Table 2.

It is clear that our new attack is much more efficient despite the fact that our linear algebra implementation is much less optimized than the one in Magma.

5. CONCLUSION

This paper presented a new attack on the new MPKCs from Diophantine equations proposed by Gao and Hendl. This attack uses lower-degree embedded surfaces associated with the new MPKCs, which was not used before. It showed that this new attack can totally break the system efficiently and one may speculate that such an approach is a very useful and can be applied on other types of systems including symmetric cryptosystems.

REFERENCES

- [1] Bardet M, Faugère J C, Salvy B. On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations[C]. Proceedings of the International Conference on Polynomial System Solving, 2004: 71-74.
- [2] Mohamed M S E, Mohamed W S A E, Ding J, et al. MXL2: Solving polynomial equations over GF(2) using an improved mutant strategy[M]. Post-Quantum Cryptography. Springer Berlin Heidelberg, 2008: 203-215.
- [3] Ding J, Ren A, Tao C. Embedded Surface Attack on Multivariate Public Key Cryptosystems from Diophantine Equations[C]. Information Security and Cryptology. Springer Berlin Heidelberg, 2013: 122-136.
- [4] Ding J, Hu L, Nie X, et al. High order linearization equation (HOLE) attack on multivariate public key cryptosystems[M]. Public Key Cryptography-PKC 2007. Springer Berlin Heidelberg, 2007: 233-248.
- [5] Ding J, Schmidt D, Werner F. Algebraic attack on HFE revisited[M]. Information Security. Springer Berlin Heidelberg, 2008: 215-227.
- [6] Ding J, Schmidt D. Rainbow, a new multivariable polynomial signature scheme[C]. Applied Cryptography and Network Security. Springer Berlin Heidelberg, 2005: 164-175.
- [7] Dubois V, Gama N. The degree of regularity of HFE systems[M]. Advances in Cryptology-ASIACRYPT 2010. Springer Berlin Heidelberg, 2010: 557-576.
- [8] Gary M R, Johnson D S. Computers and Intractability: A Guide to the Theory of NP-completeness[J]. 1979.
- [9] Gao S, Heindl R. Multivariate public key cryptosystems from diophantine equations[J]. Designs, Codes and Cryptography, 2013, 67(1): 1-18.
- [10] Raymond A. Heindl, New directions in multivariate public key cryptography , Ph.D. Thesis, Clemson University. Mathematical Science - 2009
- [11] Mohamed M S E, Cabarcas D, Ding J, et al. MXL3: An efficient algorithm for computing Gröbner bases of zero-dimensional ideals[M]. Information, Security and Cryptology-ICISC 2009. Springer Berlin Heidelberg, 2010: 87-100.
- [12] Patarin J. Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms[C]. Advances in Cryptology-Eurocrypt'96. Springer Berlin Heidelberg, 1996: 33-48.
- [13] Patarin J. Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt88[M]. Advances in Cryptology-CRYPTO95. Springer Berlin Heidelberg, 1995: 248-261.
- [14] Patarin J.: The oil and vinegar signature scheme. Presented at the Dagstuhl Workshop on Cryptography (1997).
- [15] Shor P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer[J]. SIAM journal on computing, 1997, 26(5): 1484-1509.
- [16] Wang L C, Yang B Y, Hu Y H, et al. A medium-field multivariate public-key encryption scheme[M]. Topics in Cryptology-CT-RSA 2006. Springer Berlin Heidelberg, 2006: 132-149.
- [17] Matsumoto T, Imai H. Public quadratic polynomial-tuples for efficient signature-verification and message-encryption[C]. Advances in Cryptology-EUROCRYPT88. Springer Berlin Heidelberg, 1988: 419-453.

DEPARTMENT OF MATHEMATICAL SCIENCES, UNITED ARAB EMIRATES UNIVERSITY

E-mail address: jintai.ding@gmail.com, renai@mail.uc.edu, chengdongtao2010@gmail.com, adiene@uaeu.ac.ae