TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 04/10 pp 671–679
DOI: 10.26599 / TST.2018.9010028
Volume 23, Number 6, December 2018

Practical Cryptanalysis of a Public Key Cryptosystem Based on the Morphism of Polynomials Problem

Jaihui Chen*, Chik How Tan, and Xiaoyu Li

Abstract: Multivariate Public Key Cryptography (MPKC) has intensively and rapidly developed during the past three decades. MPKC is a promising candidate for post-quantum cryptography. However, designing it is universally regarded as a difficult task to design a secure MPKC foundation scheme, such as an encryption scheme and key exchange scheme. In this work, we investigate the security of a new public key cryptosystem that is based on the Morphism of Polynomials (MP). The public key cryptosystem proposed by Wang et al. (Wuhan University, China) comprises a key exchange scheme and encryption scheme. Its security can be provably reduced to the hardness of solving a new difficult problem, namely, the Decisional Multivariate Diffie Hellman (DMDH) problem. This problem is a variant of the MP problem, which is difficult to solve by random systems. We present a proposition that reduces the DMDH problem to an easy example of the MP problem. Then, we propose an efficient algorithm for the Key Recover Attack (KRA) on the schemes of the public key cryptosystem. In practice, we are able to entirely break the cryptosystem's claimed parameter of 96 security levels in less than 17.252 s. Furthermore, we show that finding parameters that yield a secure and practical scheme is impossible.

Key words: cryptanalysis; post-quantum cryptography; multivariate public key cryptosystems; morphism of polynomials problem

1 Introduction

Since the invention of Shor's algorithm[1], current cryptography algorithms based on number theory can be broken in polynomial time by a quantum computer. Therefore an alternative to these algorithms must be identified. The alternative, which must be resistant to the attack of quantum computers, is called post-quantum cryptography. Besides code-based cryptography, i.e., schemes based on non-quasi cyclic Codes[2], Multivariate Public Key Cryptography (MPKC) is also considered as one of the most promising candidates for post quantum cryptography. Its security is based on solving a set of random Multivariate Quadratic (MQ) equations on a finite field that has been proven to be Non-deterministic Polynomial (NP) hard. However, no evidence has shown that quantum computers can solve this kind of problem efficiently. Furthermore, MPKC schemes are considerably more efficient than current cryptography algorithms, such as Rivest, Shamir, and Adleman (RSA) scheme, in computing and are friendlier to resource-restricted environment, i.e., wireless sensor networks[3].

Since the Matsumoto-Imai scheme introduced the first MPKC encryption scheme[4], a number of MPKC schemes have been proposed. However, the construction of these MPKC schemes relies not only on the MQ

- Jiahui Chen is with the School of Computer, Guangdong University of Technology, Guangzhou 510006, China. E-mail: csjhchen@gmail.com
- Chik How Tan is with Temasek Laboratories, National University of Singapore, Singapore 117411, Singapore. E-mail: tsltch@nus.edu.sg.
- Xiaoyu Li is with the School of Computer, Zhengzhou University of Aeronautics, Zhengzhou 450046, China. E-mail: Xiaoyu1987@163.com.
- * To whom correspondence should be addressed.
Manuscript received: 2017-05-26; revised: 2017-11-07; accepted: 2017-11-14

problem but also on the Isomorphism of Polynomials (IP) problem. Given the uncertainty of the IP problem, most MPKC schemes have been broken, i.e., the Matsumoto-Imai scheme and, balanced Oil and Vinegar (OV) scheme[5]. Rare exceptions, such as the simple matrix encryption (ABC)[6] and double hidden field encryption (ZHFE)[7], are considered as probable MPKC encryption schemes. Nonetheless, these schemes exhibit different disadvantages. For example, ABC may experience decryption failure, and ZHFE has limited key generation space[8]. Thus, existing promising MPKC schemes are all signature schemes, such as Unbalanced Oil and Vinegar (UOV) scheme[9], Rainbow[10], Quartz[11], Gui[12], MQ-based Digital Signature Scheme (MQDSS)[13], and HS-Sign[14].

In addition, a number of attempts have been undertaken to tackle the provable security problem of MPKC. For example, Courtois[15] studied provable security against the key-only attack on Quartz, but failed to clarify security against the chosen-message attack. Bulygin et al.[16] presented a concept for reducing the public key size of the UOV signature scheme and provided a provable security against direct attacks. Then, Sakumoto et al.[17] provided a provable security proof of UOV against the chosen-message attack by using the idea given by Bellare and Rogaway[18] wherein a random seed r is concatenated with the signed message M to render the basic trapdoor one-way function into full domain hash. In Crypto2011, Sakumoto et al.[19] proposed provably secure identification/signature schemes based on the MQ problem. These schemes have greatly improved the security of MPKCs. However, whether their techniques can be translated into a secure encryption scheme remains unknown.

Recently, together with their work on cryptanalysis of cryptosystems based on non-abelian factorization problems[20], Wang et al.[21] (WZM) proposed a novel public key cryptosystem that is based on the Morphism of Polynomials (MP) problem. In contrast to other schemes, their schemes can be provably reduced to the hardness of solving a new difficult problem, namely the Decisional Multivariate Diffie Hellman (DMDH) problem. This problem is a variant of the MP problem, which is known to be NP-hard for random systems.

Our work mainly focuses on the WZM cryptosystem. We provide two major contributions according to the WZM cryptosystem. On the theoretical side, we explore the minimal polynomial property of a matrix and present a proposition that reduces the DMDH problem to an easy instance of the MP problem. In addition, we supply an efficient algorithm to break the WZM schemes. On the practical side, we implement a plenty of attack experiments to attack the WZM scheme. We are able to completely break their claimed parameter at 96 security levels in less than 17.252 s. Accordingly, we conclude that finding parameters that yield a secure and practical scheme is impossible.

This paper is structured as follows: In Section 2, we describe the hard problems underlying MQ schemes and Key Recovery Attacks (KRAs). Then, we overview WZM schemes in Section 3. In Section 4, we illustrate that WZM schemes with their recommended parameter of 96 security levels are entirely broken within a few seconds. Concluding remarks is provided in Section 5.

2 Preliminaries

The basic objective of multivariate cryptography is to utilize a system of multivariate quadratic polynomials over a finite field $\mathbb{F}_q$. Thus, the security of multivariate cryptosystems is based on the MQ problem which is defined as follows.

Definition 1 Given m quadratic polynomials $p_1,...,p_m$ in n variables over a finite field $\mathbb{F}_q$, find a vector $x=(x_1,...,x_n)\in\mathbb{F}_q^n$ such that $p_1(x)=\cdots=p_m(x)=0$.

This problem is NP hard even for quadratic systems over the field of two elements[22].

However, for most existing MPKCs, the coefficients of the public system P (a collection of m quadratic polynomials $p_1,...,p_m$ in n variables) are not chosen randomly. Instead, one begins with an easily invertible quadratic map F (called central map) and combines it with two invertible affine maps S and T to obtain a public key of the form $P=S\circ F\circ T$. Therefore, the security of the scheme is based not only on the MQ problem, but also on the IP problem, which is defined as follows.

Definition 2 The IP problem is the problem of finding an isomorphism (S,T) from P to F, where P and F are the two public sets of u quadratic equations, and S and T are isomorphic.

The scarcity of existing knowledge on the hardness of the IP problem mainly prevents researchers from providing security proofs for their MPKCs.

However, when the above affine transformations S and T are not bijective, the problem is called MP problem that has been proven to be NP hard for any finite field[23].

3 New Public Key Cryptosystem Based on the MP Problem

In this section we review the new WZM public key

cryptosystem[21] and the novel hard problem underlying WZM schemes. WZM proposed two schemes in their public key cryptosystem: a key exchange scheme and a public key encryption scheme. In addition, given that their public key encryption scheme uses the same trapdoor and is based on the same problem as their key exchange scheme, we only describe the key exchange scheme in this paper. Below, we provide an overview of their key exchange scheme.

Let $\mathrm{MQ}(n,m,\mathbb{F}_q)$ be a family of multivariate quadratic functions as follows:

$$F(x)=(f_1(x),...,f_m(x)),$$
$$f_i(x)=\sum_{j\leqslant k}^{n}\alpha_{i,j,k}x_jx_k+\sum_{j}^{n}\beta_{i,j}x_j+\gamma_i,$$

where $\alpha_{i,j,k}$, $\beta_{i,j}$, and $\gamma_i\in\mathbb{F}_q$ for $i=1,...,m$.

Let T be an $m\times m$ matrix, define $f(T)=\alpha_mT^m+\alpha_{m-1}T^{m-1}+\cdots+\alpha_1T^1+\alpha_0I$ and $\mathcal{K}_T=\{f(T)|\forall T\in\mathcal{M}_m(\mathbb{F}_q)\}$, where $\mathcal{M}_m(\mathbb{F}_q)$ is a set of $m\times m$ matrices over $\mathbb{F}_q$, $\alpha_i\in\mathbb{F}_q, 0\leqslant i\leqslant m$.

Similarly, let U be an $n\times n$ matrix, define $f(U)=\beta_nU^n+\beta_{n-1}U^{n-1}+\cdots+\beta_1U^1+\beta_0I$ and $\mathcal{K}_U=\{f(U)|\forall U\in\mathcal{M}_n(\mathbb{F}_q)\},\beta_i\in\mathbb{F}_q, 0\leqslant i\leqslant n$, where $\mathcal{M}_n(\mathbb{F}_q)$ is a set of $n\times n$ matrices over $\mathbb{F}_q$, $\beta_i\in\mathbb{F}_q, 0\leqslant i\leqslant n$.

The two elements T_a and T_b in $\mathcal{K}_T$ satisfy the multiplication commutative law, that is, $T_aT_b=T_bT_a$. In addition, the two elements U_a and U_b in $\mathcal{K}_U$ satisfy the multiplication commutative law, that is $U_aU_b=U_bU_a$.

In reference to the above situation, we describe the key exchange scheme proposed by Wang et al.[21] as follows:

(1) Let Alice and Bob be two parties that agree on publicly available system parameters $(\mathbb{F}_q,F,T,U)$, where $F\in\mathrm{MQ}(n,m,\mathbb{F}_q)$, and two singular matrix $T\in_R\mathcal{M}_m$ and $U\in_R\mathcal{M}_n$, of which the degrees of minimal polynomials are attained in m and n.

(2) Upon obtaining the system parameters $(\mathbb{F}_q,F,T,U)$, Alice chooses $T_a\in\mathcal{K}_T$ and $U_a\in\mathcal{K}_U$ at random and computes $G_a=T_a\circ F\circ U_a$. Then Alice sends G_a to Bob.

(3) Bob receives G_a and obtains the system parameters $(\mathbb{F}_q,F,T,U)$. He chooses $T_b\in\mathcal{K}_T$ and $U_b\in\mathcal{K}_U$ at random and computes $G_b=T_b\circ F\circ U_b$. Bob sends G_b to Alice and computes the shared key $k_B=G_{ab}=T_b\circ G_a\circ U_b$.

(4) Alice receives G_b and computes the shared key $k_A=G_{ba}=T_a\circ G_b\circ U_a$.

Given that T_a and T_b in $\mathcal{K}_T$, U_a and U_b in $\mathcal{K}_U$ satisfy the multiplication commutative law, and Alice and Bob successfully establish a common session key $\mathrm{sk}=k_A=k_B$, which is a multivariate quadratic function in $\mathrm{MQ}(n,m,\mathbb{F}_q)$.

Now, we will revisit the fact that the hardness of this new problem is related to the difficulty of solving a DMDH problem which is defined below.

Definition 3 Computational Multivariate Diffie Hellman (CMDH) problem. Given a triple (F,G_x,G_y), the CMDH problem involves finding the MQ function G_{xy} such that $G_{xy}=T_x\circ T_y\circ F\circ U_x\circ U_y$, where $G_x=T_x\circ F\circ U_x$, $G_y=T_y\circ F\circ U_y$, $G_{xy}=T_x\circ T_y\circ F\circ U_x\circ U_y$ and $F\in\mathrm{MQ}(n,m,\mathbb{F}_q)$, $T_x,T_y\in\mathcal{K}_T$, $U_x,U_y\in\mathcal{K}_U$.

Definition 4 DMDH problem. Given a 4-tuple (F,G_x,G_y,G_z), the DMDH problem involves deciding whether $G_z=G_{xy}$, where $G_x=T_x\circ F\circ U_x$, $G_y=T_y\circ F\circ U_y$, $G_{xy}=T_x\circ T_y\circ F\circ U_x\circ U_y$, and $F\in\mathrm{MQ}(n,m,\mathbb{F}_q)$, $T_x,T_y\in\mathcal{K}_T$, $U_x,U_y\in\mathcal{K}_U$.

The relationship among these problems, including the MP problem, is similar to the relationship in the case of the computing Diffie Hellman problem, the decision Diffie Hellman problem, and the discrete-logarithm problem. That is, we do not know whether CMDH and DMDH belong to an NP-hard problem. A straightforward method to solve this problem is to identify the two non-bijective affine transformations T_x and U_x when given G_z and G_{xy}.

Thus, the basic hypothesis for parameter setting is to assume that solving the above problem is essentially not easier than solving the MP problem with random non-bijective affine transformations.

Assumption Hardness Hypothesis by Wang et al.[21] Solving the above DMDH problem is as hard as solving the MP problem.

However, this assumption is wrong. In the next section, we will show that this DMDH problem is an easy instance of the MP problem.

4 Cryptanalysis of the New Public Key Cryptosystem Based on the MP Problem

Now, we analyze the security of WZM schemes against KRA.

4.1 Revisitation of the first attack

The goal of breaking WZM schemes is to evaluate their basic problems: the MP problem and the proposed DMDH problem.

As stated by Ref. [21], the first KRA attempts to directly find the private key pair (T_a,U_a) from G_a or (T_b,U_b) from G_b. Let $T_a=\sum_{i=1}^{m}\alpha_iT^i+\alpha_0I$ and $U_a=\sum_{i=1}^{n}\beta_iU^i+\beta_0I$. For any $x\in\mathbb{F}_q^n$, we have

$$G_a(x)-T_a\circ F\circ U_a(x)\equiv 0.$$

Now we consider the first attack discussed by Wang et al.[21] The problem of finding (T_a,U_a) from G_a or (T_b,U_b) from G_b is viewed as a random instance of the reduced MP problem, and the number of unknown variables is reduced from m^2+n^2 to $m+n$. Given that $T_a=\sum_{i=1}^{m}\alpha_i T^i+\alpha_0 I$ and $U_a=\sum_{i=1}^{n}\beta_i U^i+\beta_0 I$, we can obtain the following q^n cubic equations with respect to $\alpha_1,..,\alpha_m,\beta_1,...,\beta_n$ by substituting all x into the equation $G_a(x)-T_a\circ F\circ U_a(x)\equiv 0$.

$$\sum_{j,k,t}\xi_{ijkt}^{(1)}\alpha_j\beta_k\beta_t+\sum_{j,k}\xi_{ijk}^{(2)}\alpha_j\beta_k+\sum_{j,k}\xi_{ijk}^{(3)}\beta_j\beta_k+\sum_{j}\xi_{ij}^{(4)}\alpha_j+\sum_{j}\xi_{ij}^{(5)}\alpha_j+\xi_i^{(6)}=0 \tag{1}$$

where $1\leqslant i\leqslant q^n$ and all coefficients are in $\mathbb{F}_q$.

WZM suggested that in general, α_0 and β_0 are equal to 0 to ensure that T_a and T_b are singular, and the total number of combined unknowns in the above Eq. (1) is $N=mn^2+mn+n(n+1)/2+m+n$.[21] If we can select more than this number of linearly independent equations from Eq. (1), then solving the above system through linearization is easy. However, WZM claimed that because F and $(T_a\ ,T_b)$ are selected at random, the construction of such linearly independent equations search process and the complexity is approximately $\mathcal{O}(q^n)$.

However, the above attack can be constructed through another approach. Given that $G_a(x)-T_a\circ F\circ U_a(x)\equiv 0$ for all $x\in\mathbb{F}_q^n$, the coefficients of $G_a(x)$ and the coefficients of $T_a\circ F\circ U_a(x)$ are equal. (F,T,U) are publicly known system parameters. Thus, if we use symmetric matrices to represent the quadratic component of the public key F (omitting the linear part of F will not affect cryptanalysis), we can obtain the following $\frac{mn(n+1)}{2}$ cubic equations with $m+n$ variables.

$$\sum_{j,k,t}\lambda_{ijkt}^{(1)}\alpha_j\beta_k\beta_t+\sum_{j,k}\lambda_{ijk}^{(2)}\alpha_j\beta_k+\sum_{j,k}\lambda_{ijk}^{(3)}\beta_j\beta_k+\sum_{j}\lambda_{ij}^{(4)}\alpha_j+\sum_{j}\lambda_{ij}^{(5)}\alpha_j+\lambda_i^{(6)}=0 \tag{2}$$

where $1\leqslant i\leqslant\frac{mn(n+1)}{2}$ and all variables are in $\mathbb{F}_q$.

4.2 KRA with direct attacks

Solving such a system of $\frac{mn(n+1)}{2}$ cubic equations with $m+n$ variables requires the use of the direct attacks technique. The direct attacks use equation solvers, including extended linearization (XL)[24] and Gröbner Basis algorithms, such as F_4[25] and F_5[26]. We briefly describe these equation solvers below.

Bettale et al.[27] asserted that for a semiregular system, the computational complexity of F_4 is bounded by $\mathcal{O}\left(\left(t\binom{n+d_{\mathrm{reg}}-1}{d_{\mathrm{reg}}}\right)^{\omega}\right)$, where n is the number of variables, t is the number of equations, and ω is a linear algebraic constant and $2\leqslant\omega\leqslant 3$. In general, we set $\omega=2$ for lower bound complexity and $\omega=3$ for upper bound complexity. d_{reg} is the degree of regularity of the system, which is the index of the first non-positive coefficient in the Hilbert series $S_{m,n}$ with

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

where d_i is the degree of the i-th equation and z is the variable of the Hilbert series.

The HybridF_5 (HF_5) algorithm[27] is currently the fastest algorithm among all the direct attacks algorithms. Its underlying principle is the deduction of some variables required to create overdetermined systems before the F_5 algorithm is applied. Thus, one has to run the F_5 algorithm several times (depending on how many variables he/she guesses) to find a solution for the original system. When guessing u variables over $\mathbb{F}_q$, this number is given by q^u. The complexity of solving a semiregular system of t multivariate equations in n variables over $\mathbb{F}_q$ by the HF_5 algorithm can be estimated as $q^u\mathcal{O}\left(\left(t\binom{n-u+d_{\mathrm{reg}}-1}{d_{\mathrm{reg}}}\right)^{\omega}\right)$.

Generally speaking, the best direct attack algorithm for solving multivariate polynomial equations over medium fields is the HF_5 algorithm, and that for solving multivariate polynomial equations over large fields is the F_4 (or F_5) Gröbner Basis algorithms.

When the underground field is small and under the assumption that $\frac{mn(n+1)}{2}=\varepsilon(m+n)^2$, $\varepsilon>0$, as is discussed in Ref. [24], the best algorithm is the XL algorithm, and the complexity is $\mathcal{O}((m+n)^{\omega D}/D!)$, where $D\approx\lceil 1/\sqrt{\varepsilon}\rceil$.

However, the above result cannot be used directly in the WZM scheme because it is based on the DMDH problem which seems as hard as the MP problem. Next we discuss why direct attack algorithms cannot be efficiently used to solve the MP problem.

Recall that for a system based on the MP problem, we assume that G_a and F are two public sets of m quadratic equations, and T_a and U_a are two non-bijective morphism transformations, where $G_a(x)-T_a\circ F\circ U_a(x)\equiv 0$ for all $x\in\mathbb{F}_q^n$. The target of solving MP problem is to recover T_a

and U_a. Similarly, we can obtain a system of $\frac{mn(n+1)}{2}$ cubic equations with m^2+n^2 variables, but because T_a and U_a are singular, the resultant $\frac{mn(n+1)}{2}$ cubic equations are not independent and can be reduced to a small number of cubic equations.

The number of reduced equations depends on the rank of T_a and U_a, and the final number of reduced cubic equations is attained by $\frac{\mathrm{Rank}(U_a)\cdot\mathrm{Rank}(T_a)\cdot(\mathrm{Rank}(T_a)+1)}{2}$. Taking parameters $q=2^{16}, n=12$, and $m=10$ as an example, if we treat all polynomials of F as homogeneous in general, we will obtain a system of 780 cubic equations with 244 variables. However, assuming that the rank of T_a is 9 and that of U_a is 11, the resultant cubic equations can be reduced to 594 cubic equations with 244 variables. When the rank of T_a and U_a is small, the resultant cubic equations can be reduced to a small number of equations. Given that we do not know the number of independent equations, direct attack algorithms cannot be efficiently used to solve this MP problem. The same situation requires the use of direct attack algorithms to attack WZM schemes that provide a system of $\frac{mn(n+1)}{2}$ cubic equations with $m+n$ variables for reducing cubic equations. Similarly, taking the parameters $q=2^{16}, n=12$, and $m=10$ recommended by Wang et al.[21] as an example, we will obtain a system of 780 cubic equations with 22 variables. However, assuming that the rank of T_a is 8 and that of U_a is 10, the resultant cubic equations can be reduced to 360 cubic equations with 22 variables. These equations will be difficult to solve with direct attack algorithms. Thus, the equation seems to be solvable only by an exhaustive search algorithm, as claimed by Wang et al.[21], who stated that their schemes are based on the reduced MP problem wherein the number of variables are reduced from m^2+n^2 to $m+n$. In addition, this problem remains intractable because it can only be dealt with by an exhaustive search attack.

Through the above discussion, we find that the use of a direct attack to solve the MP problem is mainly hindered by the following: The resultant cubic equations are not independent and reducible. Thus, if we can find some method to convert the resultant cubic equations into nonreduced equations, we can use direct attack algorithms to solve this problem. Luckily, we find a property of the minimum polynomial to break the intractability of DMDH problem. Below, we present a proposition to illustrate that the DMDH problem can be reduced to an easy instance of the MP problem.

Proposition The DMDH problem of WZM schemes can be reduced to an easy instance of the MP problem. This reduction shows that recovering the two transformations is always equal to solving a system of $\frac{mn(n+1)}{2}$ independent cubic equations with $m+n$ variables.

Let $f_{\min T}(x)$ and $f_{\min U}(x)$ be the corresponding minimum polynomials of matrices T and U, respectively. In the construction of WZM schemes, the degree of the minimum polynomial of matrices T and U is equal to m and n, respectively. Let $f_{\min T}(x)=x^m+\sum_{i=0}^{m-1}\alpha_i' x^i$ and $f_{\min U}(x)=x^n+\sum_{i=0}^{n-1}\beta_i' x^i$, we have

$$T^m=-\sum_{i=0}^{m-1}\alpha_i' T^i$$

and

$$U^n=-\sum_{i=0}^{n-1}\beta_i' T^i.$$

Recall that $T_a=\alpha_m T^m+\alpha_{m-1}T^{m-1}+\cdots+\alpha_1 T$ and $U_a=\beta_n U^n+\beta_{n-1}U^{n-1}+\cdots+\beta_1 U$ in the construction of WZM schemes, we have

$T_a = \alpha_m(-\sum_{i=0}^{m-1}\alpha_i' T^i)+\alpha_{m-1}T^{m-1}+\cdots+\alpha_1 T^1 = \sum_{i=1}^{m-1}(\alpha_i-\alpha_m\alpha_i')T^i-\alpha_m\alpha_0' T^0$

and

$U_a = \beta_n(-\sum_{i=0}^{n-1}\beta_i' U^i)+\beta_{n-1}U^{n-1}+\cdots+\beta_1 U^1 = \sum_{i=1}^{n-1}(\beta_i-\beta_n\beta_i')U^i-\beta_n\beta_0' U^0.$

If we let $\gamma_0=-\alpha_m\alpha_0'$, $\gamma_i=\alpha_i-\alpha_m\alpha_i'$ and $\eta_0=-\beta_n\beta_0'$, $\eta_j=\beta_j-\beta_n\beta_j'$, where $1\leqslant i\leqslant(m-1), 1\leqslant j\leqslant(n-1)$, and all elements are in $\mathbb{F}_q$.

Finally we will get

$$T_a=\sum_{i=0}^{m-1}\gamma_i T^i$$

and

$$U_b=\sum_{i=0}^{n-1}\eta_i U^i.$$

The above equations show that because of the restraint of the degree of minimum polynomial of matrices T and U must be equal to m and n, respectively, and T_a and U_a are always bound by the coefficients of the minimum polynomial of matrices T and U, respectively. T_a and U_a are not randomly generated and are always equal to new ${T'}_a$ and ${U'}_b$ as shown by the following form.

$${T'}_a=\sum_{i=0}^{m-1}\gamma_i' T^i$$

and

$${U'}_b=\sum_{i=0}^{n-1}\eta_i' U^i.$$

Recall that solving the DMDH problem involves finding the two nonbijective affine transformations T_a and

U_a, when given F and G_a.

In addition, because T and U are known, knowing $f_{\min T}(x)$ and $f_{\min U}(x)$ is easy. Consequently, if we can solve the above γ'_i, $0 \leqslant i \leqslant (m-1)$ and η'_j, $0 \leqslant j \leqslant (n-1)$, we can recover the underlying T_a and U_a. Then, we have $G_a(x) - T'_a \circ F \circ U'_a(x) \equiv 0$ for all $x \in \mathbb{F}_q^n$. Thus we can obtain a system of $\frac{mn(n+1)}{2}$ cubic equations with $m+n$ variables γ'_i, $0 \leqslant i \leqslant (m-1)$ and η'_j, $0 \leqslant j \leqslant (n-1)$. However, because

$$T'_a = \sum_{i=0}^{m-1} \gamma'_i T^i$$

and

$$U'_b = \sum_{i=0}^{n-1} \eta'_i U^i,$$

the resultant equations are always independent equations.

Finally, the DMDH problem of WZM schemes is reduced to an easy instance of the MP problem. Its solution is equal to solving a system of $\frac{mn(n+1)}{2}$ independent cubic equations with $m+n$ variables.

4.3 Our attack algorithm

Now we will describe our modified attack algorithm, which uses direct attacks.

The above discussion shows that we can use other algorithms rather than the exhaustive search algorithm for KRA on WZM schemes. Our modified attack algorithm on KRA for WZM schemes over a finite field is described in Algorithm 1.

In our attack algorithm, despite $T_a = \alpha_m T^m + \alpha_{m-1} T^{m-1} + \cdots + \alpha_1 T$ and $U_a = \beta_n U^n + \beta_{n-1} U^{n-1} + \cdots + \beta_1 U$, where $\alpha_i, 1 \leqslant i \leqslant m$ and $\beta_i, 1 \leqslant i \leqslant n$ are the constructed unknowns, we do not directly recover these unknowns but instead construct other unknowns $\gamma'_i, 0 \leqslant i \leqslant (m-1)$ and $\eta'_j, 0 \leqslant j \leqslant (n-1)$ which can ensure direct attack algorithms run efficiently. According to our above analysis in **Proposition**, by using $\gamma'_i, 0 \leqslant i \leqslant (m-1)$ and $\eta'_j, 0 \leqslant j \leqslant (n-1)$, we can also recover T_a and U_a.

4.4 Practical attack under the recommended parameters

Wang et al.[21] proposed concrete parameters for their scheme at the commendable security level of 96. The parameters are chosen as follows: $q = 2^{16}$, $n = 12$, and $m = 10$. In this section, we show that security is essentially incompatible for WZM schemes.

Algorithm 1 Our algorithm for KRA on WZM schemes (G_a, F, T, U)

Require:
G_a: the key Alice sends to Bob;
F, T, U: the public system parameters;
Ensure:
T_a, U_a: the private key;
1: Construct two new transformation T'_a and U'_a in the form:

$$T'_a = \sum_{i=0}^{m-1} \gamma'_i T^i$$

and

$$U'_a = \sum_{i=0}^{n-1} \eta'_i U^i;$$

2: Let $G_a(x) - T'_a \circ F \circ U'_a(x) \equiv 0$ for all $x \in \mathbb{F}_q^n$ and generate a system of $\frac{mn(n+1)}{2}$ cubic equations with $m+n$ variables $\gamma'_i, 0 \leqslant i \leqslant (m-1)$ and $\eta'_j, 0 \leqslant j \leqslant (n-1)$;
3: Solve the system using the appropriate direct attack algorithms in accordance with the finite field and obtain all solutions (defined as sol);
4: For every $sol[i]$ do
5: Compute

$$T'_a = \sum_{j=0}^{m-1} sol[i]_j T^i$$

and

$$U'_a = \sum_{j=0}^{n-1} sol[i]_{j+m} U^i;$$

6: Check if $T'_a \circ F \circ U'_a = G_a$ and mark the position as pos when it is correct;
7: End for;
8: **return** $T_a = \sum_{j=0}^{m-1} sol[pos]_j T^i$, $U_a = \sum_{j=0}^{n-1} sol[pos]_{j+m} U^i$;

The result of our cryptoanalysis to the lower/upper bound complexities against KRA with direct attack by F_4 which are $2^{22}/2^{30}$ for WZM $(\mathbb{F}_{2^{16}}, 10, 12)$ not 2^{96} claimed security in Ref. [21]. Table 1 shows the improvements of lower bound ($\omega = 2$) and upper bound ($\omega = 2.8$) in the complexity of solving such a system by using different attack methods for WZM $(\mathbb{F}_{2^{16}}, 10, 12)$.

Table 1 shows that the rank of T and U is assumed to be 9 and 11, respectively. Table 1 also shows that the number of reduced equations in KRA with direct attack is smaller than that with our modified attack. Thus, even in the simplest non-bijective form (the rank of T and U is $n-1$ and $m-1$, respectively), the complexity remains as high as 2^{52} and memory will be insufficient when performing such an attack in practice. However, when using our modified attack algorithm, all equations are independent equations, and the attack complexity is only 2^{22}.

Furthermore, we are able to attack WZM schemes. We

Table 1 Lower and upper bounds of the complexity of different KRA algorithms against WZM schemes with $(\mathbb{F}_q, n, m) = (\mathbb{F}_{2^{16}}, 12, 10)$.

Attack algorithm	Number of equations (reduced)	Number of variables	Complexity (lower/upper)
KRA with exhaustive search	2^{192} (linear)	1660	$2^{96}/2^{192}$
KRA with direct attacks (directly)	594 (cubic)	22	$2^{54}/2^{62}$ (Out of Memory)
KRA with our attacks	780 (cubic)	22	$2^{22}/2^{30}$

completely break WZM schemes under their recommended parameters $(\mathbb{F}_{2^{16}}, 12, 10)$. We program and run our modified attack algorithm using MAGMA[28] v2.20-5. All experiments are run on a workstation with a Dual XEON Quad Core 2.27 GHz processor, 24 GB of main random access memory, and operation system of Scientific Linux 5.11 (Boron). In addition to the recommended parameters, we also attack other parameters $(\mathbb{F}_{2^{12}}, 14, 12)$ and $(\mathbb{F}_{2^6}, 28, 20)$ under reasonably selected conditions in accordance with the security analysis provided by Wang et al.[21] to evaluate the security of different fields under our attack. We perform each attack 1000 times during each test. The results are listed in Table 2. We break their recommended parameters at 96 security levels in less than 17.252 s. The results of the practical attacks summarized in Table 2 shows that we can efficiently break the schemes in different fields at their claimed security within several seconds.

Finally, we attempt to estimate the secure and optimal parameters that will enable WZM schemes to resist our attacks. We select secure and optimal parameters for WZM schemes at different security levels in accordance with the theoretical complexity of our attacks. The selected parameters are summarized in Table 3, which shows that more than 10 million multiplication operations are required to encrypt just one plaintext to achieve 80-bit security. This requirement is impractical because encrypting just one plaintext requires an estimated running time of more than 10 min. Thus, finding parameters that yield a practical scheme for the WZM cryptosystem is impossible.

Table 2 Results of theoretical complexities and practical KRA using our modified attack algorithm on WZM schemes.

Parameter $(\mathbb{F}_q, n, m)$	Claimed security	Theoretical complexity (lower/upper)	Attack time (s)
$(\mathbb{F}_{2^{16}}, 12, 10)$ (recommend)	2^{96}	$2^{22}/2^{30}$	17.252
$(\mathbb{F}_{2^{12}}, 14, 12)$ (select)	2^{84}	$2^{32}/2^{43}$	91.87
$(\mathbb{F}_{2^6}, 28, 20)$ (select)	2^{84}	$2^{32}/2^{43}$	330.86

Table 3 Estimated parameters for WZM schemes at given security levels.

Parameter $(\mathbb{F}_q, n, m)$	Security level	Encryption (Muls)	Decryption (Muls)
$(\mathbb{F}_{2^{16}}, 28, 26)$	2^{80}	9 956 926	8 885 240
$(\mathbb{F}_{2^{12}}, 32, 36)$	2^{80}	31 876 164	29 147 904
$(\mathbb{F}_{2^6}, 64, 70)$	2^{80}	841 995 030	801 207 680

5 Conclusion

We used modified attacks to investigate the security of the WZM schemes, a new public key cryptosystem that is based on the MP problem, against the KRA. We showed that our attack algorithm completely broke the recommended parameters for WZM schemes at 96 security levels in less than 17.252 s. We attempted to estimate secure and optimal parameters at 80-bit security levels to resist our attacks. However, WZM schemes with these parameters are considerably slower than expected. Therefore, finding parameters that yield a secure and practical scheme is impossible.

References

[1] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput*., vol. 26, no. 5, pp. 1484–1509, 1997.

[2] B. H. Lin, Y. K. Pei, L. G. Yin, and J. H. Lu, Design and efficient hardware implementation schemes for Non-Quasi-Cyclic LDPC codes, *Tsinghua Sci. Technol*., vol. 22, no. 1, pp. 92–103, 2017.

[3] J. H. Chen, S. H. Tang, D. J. He, and Y. Tan, Online/offline signature based on uov in wireless sensor networks, *Wirel. Networks*, vol. 23, no. 6, pp. 1719–1730, 2017.

[4] T. Matsumoto and H. Imai, Public quadratic polynomial-tuples for efficient signature-verification and message-encryption, in *Advances in Cryptology-EUROCRYPT 98*, D. Barstow, W. Brauer, P. B. Hansen, D. Gries, D. Luckham, C. Moler, A. Pnueli, G. Seegmüller, J. Stoer, N. Wirth, et al., eds. Springer, vol. 330, pp. 419–453, 1988.

[5] J. Patarin, The oil and vinegar signature scheme, in *the Dagstuhl Workshop on Cryptography*, 1997.

[6] C. Tao, A. Diene, S. Tang, and J. Ding, Simple matrix scheme for encryption, in *Proc. 5^{th} Int. Post-Quantum Cryptography*, Waterloo, Canada, 2013, pp. 231–242.

[7] J. Porras, J. Baena, and J. Ding, Zhfe, a new multivariate public key encryption scheme, in *Proc. 6^{th} Int. Post-Quantum Cryptography*, Waterloo, Canada, 2014, pp. 229–245.

[8] W. B. Zhang and C. H. Tan, On the security and key generation of the ZHFE encryption scheme, in *Proc. 11^{th} Int. Workshop on Security*, Tokyo, Japan, 2016, pp. 289–304.

[9] A. Kipnis, J. Patarin, and L. Goubin, Unbalanced oil and vinegar signature schemes, in *Advances in Cryptology-EUROCRYPT 99*, J. Stern, ed. Springer, pp. 206–222, 1999.

[10] J. Ding and D. Schmidt, Rainbow—A new multivariable

polynomial signature scheme, in *Applied Cryptography and Network Security*, J. Ioannidis, A. Keromytis, and M. Yung, eds. Springer, 2005, pp. 164–175.

[11] J. Patarin, N. Courtois, and L. Goubin, Quartz, 128-bit long digital signatures, in *Proc. Cryptographer's Track at RSA Conf.* San Francisco, CA, USA, 2001, pp. 282–297.

[12] A. Petzoldt, M. S. Chen, B. Y. Yang, C. D. Tao, and J. T. Ding, Design principles for HFEv-based multivariate signature schemes, in *Proc. 21st Int. Conf. on the Theory and Application of Cryptology and Information Security*, Auckland, New Zealand, 2015, pp. 311–334.

[13] M. S. Chen, A. Hülsing, J. Rijneveld, S. Samardjiska, and P. Schwabe, From 5-pass MQ-based identification to MQ-based signatures, in *Proc. 22nd Int. Conf. on the Theory and Application of Cryptology and Information Security*, Hanoi, Vietnam, 2016, pp. 135–165.

[14] J. H. Chen, S. H. Tang, and X. L. Zhang, HS-sign: A security enhanced UOV signature scheme based on hyper-sphere, *KSII Trans. Int. Inf. Syst.*, vol. 11, no. 6, pp. 3166–3187, 2017.

[15] N. T. Courtois, Generic attacks and the security of Quartz, in *Public Key Cryptography 2003*, Y. G. Desmedt, ed. Springer, 2002, pp. 351–364.

[16] S. Bulygin, A. Petzoldt, and J. Buchmann, Towards provable security of the unbalanced oil and vinegar signature scheme under direct attacks, in *Progress in Cryptology–INDOCRYPT 2010*, G. Gong and K. C. Gupta, eds. Springer, 2010, pp. 17–32.

[17] K. Sakumoto, T. Shirai, and H. Hiwatari, On provable security of UOV and HFE signature schemes against chosen-message attack, in *Post-Quantum Cryptography*, B. Y. Yang, ed. Springer, 2011, pp. 68–82.

[18] M. Bellare and P. Rogaway, The exact security of digital signatures—How to sign with RSA and Rabin, in *Advances in Cryptology-EUROCRYPT'96*, U. Maurer, ed. Springer, pp. 399–416, 1996.

[19] K. Sakumoto, T. Shirai, and H. Hiwatari, Public-key identification schemes based on multivariate quadratic polynomials, in *Advances in Cryptology–CRYPTO 2011*, P. Rogaway, ed. Springer, 2011, pp. 706–723.

[20] J. H. Liu, A. W. Fan, J. W. Jia, H. G. Wang, H. Z. Zhang, and S. W. Mao, Cryptanalysis of public key cryptosystems based on non-abelian factorization problems, *Tsinghua Sci. Technol.*, vol. 21, no. 3, pp. 345–351, 2016.

[21] H. Z. Wang, H. G. Zhang, S. W. Mao, W. Q. Wu, and L. Q. Zhang, New public-key cryptosystem based on the morphism of polynomials problem, *Tsinghua Sci. Technol.*, vol. 21, no. 3, pp. 302–311, 2016.

[22] J. Patarin and L. Goubin, Trapdoor one-way permutations and multivariate polynomials, in *Information and Communications Security*, J. Patarin and L. Coubin, eds. Springer, 1997, pp. 356–368.

[23] J. Patarin, L. Goubin, and N. Courtois, Improved algorithms for isomorphisms of polynomials, in *Advances in Cryptology–EUROCRYPT 1998*, Springer, 1998, pp. 184–200.

[24] N. Courtois, A. Klimov, J. Patarin, and A. Shamir, Efficient algorithms for solving overdefined systems of multivariate polynomial equations, in *Advances in Cryptology–EUROCRYPT 2000*, B. Preneel, ed. Springer, pp. 392–407, 2000.

[25] J. C. Faugére, A new efficient algorithm for computing Gröbner bases (F4), *J. Pure Appl. Algebra*, vol. 139, nos. 1–3, pp. 61–88, 1999.

[26] J. C. Faugère, A new efficient algorithm for computing Gröbner bases without reduction to zero (F5), in *Proc. Int. Symp. on Symbolic and Algebraic Computation*, New York, NY, USA, pp. 75–83, 2002.

[27] L. Bettale, J. C. Faugère, and L. Perret, Hybrid approach for solving multivariate systems over finite fields, *J. Math. Cryptol.*, vol. 3, no. 3, pp. 177–197, 2009.

[28] W. Bosma, J. Cannon, and C. Playoust, The Magma algebra system I: The user language, *J. Symbol. Computat.*, vol. 24, nos. 3 & 4, pp. 235–265, 1997.

Jiahui Chen received the BS degree from South China Normal University, China, in 2009, and MS and PhD degrees from South China University of Technology, China, in 2012 and 2016, respectively. He joined National University of Singapore as a research scientist between March 2017 and May 2018. He is currently a senior lecturer in the School of Computer at Guangdong University of Technology. His research interests mainly focus on multivariate public key cryptography, quantum cryptography, and information security.

Chik How Tan received the BS degree from National University of Singapore in 1984, and MS and PhD degrees in mathematics from University of Wisconsin-Madison, USA, in 1990 and 1992, respectively. He is currently a principal research scientist with Temasek Laboratories at National University of Singapore, Singapore. His research interests include cryptography, discrete mathematics, and information security.

Xiaoyu Li received the PhD degree from South China University of Technology, China, in 2016. He is currently an assistant professor in the School of Computer at Zhengzhou University of Aeronautics, China. His research interests include applied cryptography, cloud computing security, and privacy protection.