New candidates for multivariate trapdoor functions

Jaiberth Porras[1], John B. Baena[1], Jintai Ding[2,✉]

[1]Universidad Nacional de Colombia, Medellín, Colombia

[2]University of Cincinnati, Cincinnati, OH, USA

Abstract. We present a new method for building pairs of HFE polynomials of high degree, such that the map constructed with such a pair is easy to invert. The inversion is accomplished using a low degree polynomial of Hamming weight three, which is derived from a special reduction via Hamming weight three polynomials produced by these two HFE polynomials. This allows us to build new candidates for multivariate trapdoor functions in which we use the pair of HFE polynomials to fabricate the core map. We performed the security analysis for the case where the base field is $GF(2)$ and showed that these new trapdoor functions have high degrees of regularity, and therefore they are secure against the direct algebraic attack. We also give theoretical arguments to show that these new trapdoor functions over $GF(2)$ are secure against the MinRank attack as well.

Key words and phrases. Multivariate cryptography, HFE polynomials, HFE cryptosystem, trapdoor functions, Zhuang-zi algorithm.

1. Introduction

The public key cryptosystems currently used in practice are based on the difficulty of factoring large integers or solving the Discrete Logarithm Problem. In 1996 P. Shor published an algorithm to solve both problems in polynomial time on a quantum computer [17]. Some experts argue that it is possible to build in the coming years a quantum computer, which is a threat to our modern communication system. This leads to the recent fast development of Post-Quantum Cryptography . Post-Quantum Cryptography refers to the study of cryptosystems that have the potential to resist the possible future quantum computer attacks [1].

Multivariate Public Key Cryptography (MPKC) is part of the Post-Quantum Cryptography. In MPKC, the public key consists of a set of multivariate quadratic polynomials over a finite field. One of the main cryptosystems in MPKC is named Hidden Field Equations (HFE), proposed by Patarin in 1996 [16]. The public key in HFE is formed by "hiding" a core polynomial F by two invertible affine transformations, and using the vector space structure of a field extension of the base field.

A crucial part in HFE is the choice of the degree D of the core polynomial F. The degree D cannot be too big, since otherwise the decryption process would not be efficient. The main attacks against HFE, direct algebraic attack ([10, 5, 6, 14, 15]) and the Kipnis-Shamir MinRank attack (KS attack [13]), exploit this fact. For characteristic 2, HFE is vulnerable to the direct algebraic attack [10]. Recently, some authors improved the KS attack and were able to break certain HFE systems, over both odd and even characteristic [2].

We propose a special reduction method to construct new candidates for trapdoor functions using HFE polynomials of high degree. The use of these high degree polynomials prevents the known attacks against HFE. This opens the possibility to build a secure variant of the HFE cryptosystem.

The idea of the construction is inspired by the first steps of the Zhuang-Zi algorithm [7]. Given a finite field k of size q and a field extension K of degree n, we consider two high degree HFE polynomials over K of the form $F(X) = \sum a_{ij} X^{q^i+q^j} + \sum b_i X^{q^i} + c$ and $\tilde{F}(X) = \sum \tilde{a}_{ij} X^{q^i+q^j} + \sum \tilde{b}_i X^{q^i} + \tilde{c}$, where the coefficients $a_{ij}, b_i, c, \tilde{a}_{ij}, \tilde{b}_i, \tilde{c} \in K$ are to be determined. The idea behind the method is to construct a low degree polynomial Ψ of Hamming weight three of the form

$$\begin{aligned}\Psi =& X\left(\alpha_1 F_0 + \cdots + \alpha_n F_{n-1} + \beta_1 \tilde{F}_0 + \cdots + \beta_n \tilde{F}_{n-1}\right) + \\ & X^q\left(\alpha_{n+1} F_0 + \cdots + \alpha_{2n} F_{n-1} + \beta_{n+1} \tilde{F}_0 + \cdots + \beta_{2n} \tilde{F}_{n-1}\right),\end{aligned}$$

where $F_0, F_1, \cdots, F_{n-1}$ are the Frobenius powers of F, and $\tilde{F}_0, \tilde{F}_1, \cdots, \tilde{F}_{n-1}$ are the Frobenius powers of $\tilde{F}$.

To obtain such a polynomial Ψ we need to determine the coefficients of F and $\tilde{F}$, also the scalars α_i and β_i, such that the degree of Ψ is less than or equal to a fixed positive integer D_0 (the integer D_0 is such that we can easily invert Ψ using Berlekamp's algorithm). To achieve this, we derive a system of equations from the vanishing coefficients of the terms in Ψ of degree higher than D_0. After randomly choosing in this system the scalars α_i and β_i, we get a linear system with more variables than equations, and thus we can guarantee nontrivial solutions for it. This linear system has about n^3 variables and therefore we have to deal with huge matrices to reach large values of n. On the plus side we have that these matrices are sparse, which is an advantage in terms of efficiency.

The new multivariate trapdoor function is built in a similar way to the HFE scheme (composition with invertible affine transformations), except that now the core map is replaced by the map $G = (F, \tilde{F})$. The main part of the inversion of the trapdoor function is to invert the map G, which is achieved using the low degree Hamming weight three polynomial Ψ and the scalars α_i, β_i.

To construct our new candidate for trapdoor function we use these high degree HFE polynomials for the core map with the expectation that this trapdoor

function has high degree of regularity, very different from what was observed by Faugère and Joux [10] for a system of quadratic equations derived from a single HFE polynomial with low degree. For the case of $q = 2$, our extensive experiments confirmed that the new trapdoor function has high degree of regularity (it increases as n increases). This high degree of regularity shows that our new candidates for multivariate trapdoor functions are secure against the direct algebraic attack.

Furthermore, for the case of $q = 2$, we can give a theoretical argument to show why the MinRank attack does not work against the new trapdoor functions, based on some results about the degree of regularity obtained by Ding and Hodges [8]. From those results, we see that, for $q = 2$, it suffices to make sure that our trapdoor function has a high degree of regularity to conclude that both the direct algebraic and KS MinRank attacks do not work against these new trapdoor functions. We show that this is indeed the case and that our new trapdoor functions are secure against these two attacks. For larger values of q, this argument cannot be used and the MinRank attack must be directly performed against these new trapdoor functions.

The method described here was not our first attempt to reduce high degree HFE polynomials along the same line. Among failed attempts, we considered using a single polynomial F, but the linear systems we needed to solve had more equations than variables and then we could not guarantee nontrivial solutions for them. This lead us to use two HFE polynomials instead of one in order to get a linear system with more variables than equations and this gives the construction in this paper.

This paper is organized as follows. First, we present some background material about HFE cryptosystems. Secondly, we describe the method for building the new candidates for multivariate trapdoor functions. Next, we present a toy example to explain step by step our method, and two big examples. Then, we carry out a security analysis and discuss future work. In the appendix we show some data about the generation of the new trapdoor function.

2. Background

The cryptosystem Hidden Field Equations (HFE) was proposed by Patarin in 1996 [16]. The public key is formed by "hiding" a core polynomial F via two invertible affine transformations, and taking advantage of the vector space structure of an extension of the base field.

Let k be a finite field of size q. Fix $n \in \mathbb{N}$ and take an irreducible polynomial g over k of degree n. Consider the field extension $K = k[y]/(g(y))$. Then $K \cong k^n$, via the isomorphism $\varphi\colon K \to k^n$ defined by

$$\varphi\left(u_1 + u_2 y + \ldots + u_n y^{n-1}\right) = (u_1, u_2, \ldots, u_n)\,.$$

Notice that $\left\{1, y, \cdots, y^{n-1}\right\}$ is a basis for K over k.

We say that a polynomial has *Hamming weight* W if the maximum of the q-Hamming weights of all its exponents is W. The q-Hamming weight of a non-negative integer is the sum of the q-digits of its q-nary expansion. Let $F : K \to K$ be a Hamming weight two polynomial of the form

$$F(X) = \sum_{0 \leq j \leq i}^{n-1} a_{ij} X^{q^i + q^j} + \sum_{i=0}^{n-1} b_i X^{q^i} + c,$$

where the coefficients a_{ij}, b_i, c are chosen randomly in K. Such a polynomial F is called an *HFE polynomial*. If in addition, we require that $\deg(F) \leq D$, where D is a fixed positive integer, we say that F is an *HFE polynomial with bound* D.

For a fixed D, an HFE cryptosystem is built as follows. First, we randomly choose an HFE polynomial with bound D, say $F \colon K \to K$. Then, we randomly choose two invertible affine transformations S and T over k^n. The public key P is the composition of F with the transformations S and T, together with the isomorphism φ, i.e.,

$$P = T \circ \varphi \circ F \circ \varphi^{-1} \circ S.$$

Notice that P is an $n-$tuple of the form

$$P = \left(P_1(x_1, \ldots, x_n), \cdots, P_n(x_1, \ldots, x_n)\right),$$

where each P_i is a multivariate quadratic polynomial. The private key consists of the core map F together with the transformations S and T.

When constructing an HFE cryptosystem we need to be very careful with the choice of the bound D. This bound cannot be too high, since this would affect the decryption process, making it inefficient. Also, D cannot be too small, because this would make the system vulnerable to the algebraic and KS attacks.

Many attempts have been made to build safe HFE variants for both digital signatures and encryption schemes [12, 9, 4, 11]. However, most of them have not been successful. One of the latest, Multi-HFE [11], proposes to use as core map a system of multivariate polynomials over K, instead of a single HFE polynomial. This cryptosystem was broken by means of a generalization of the Kipnis-Shamir MinRank attack [2].

In the next section we present a procedure to generate a low degree polynomial of Hamming weight three, which can be used to invert a map constructed with two high degree HFE polynomials. This idea enables us to build candidates for trapdoor functions using high degree HFE core polynomials, preventing the attacks that we mentioned earlier.

3. Construction of new candidates for multivariate trapdoor functions

The weakness of the HFE cryptosystem lies on the use of a low degree core polynomial F. This polynomial is used for both encryption and decryption.

The process of decryption involves inverting the map F (search of pre-images). Therefore, if we take a polynomial of high degree the decryption could be impossible, and if otherwise we take a polynomial of low degree the attacks mentioned above would work.

To overcome this weakness, we developed a method for building pairs of HFE polynomials of very high degree, and such that the map constructed with such a pair is easy to invert, using a low degree polynomial derived from a special reduction via Hamming weight three polynomials. This low degree polynomial is easy to invert by means of Berlekamp's algorithm. In this way, we are able to use two HFE polynomials of high degree to construct a new candidate for a trapdoor function, and a polynomial of small degree as the trapdoor used to invert such trapdoor function.

3.1. The Reduction method

Let $F : K \to K$ and $\tilde{F} : K \to K$ be two high degree HFE polynomials of the form

$$F(X) = \sum_{0\leq j\leq i}^{n-1} a_{ij} X^{q^i+q^j} + \sum_{i=0}^{n-1} b_i X^{q^i} + c,$$

$$\tilde{F}(X) = \sum_{0\leq j\leq i}^{n-1} \tilde{a}_{ij} X^{q^i+q^j} + \sum_{i=0}^{n-1} \tilde{b}_i X^{q^i} + \tilde{c},$$

where the coefficients $a_{ij}, b_i, c, \tilde{a}_{ij}, \tilde{b}_i, \tilde{c} \in K$ are to be determined. Next, let $F_0, F_1, \cdots, F_{n-1}$ be the Frobenius powers of F and let $\tilde{F}_0, \tilde{F}_1, \cdots, \tilde{F}_{n-1}$ be the Frobenius powers of $\tilde{F}$, i.e.,

$$F_i(X) = [F(X)]^{q^i} \text{ and } \tilde{F}_i(X) = \left[\tilde{F}(X)\right]^{q^i}, \text{ for } i = 0, 1, \cdots, n-1.$$

Let D_0 be an upper bound for the degree of a univariate polynomial equation that can be solved efficiently using Berlekamp's algorithm.

The key part of this method is to construct a polynomial Ψ of the form

$$\begin{aligned}\Psi =& X\left(\alpha_1 F_0 + \cdots + \alpha_n F_{n-1} + \beta_1 \tilde{F}_0 + \cdots + \beta_n \tilde{F}_{n-1}\right) + \\ & X^q\left(\alpha_{n+1} F_0 + \cdots + \alpha_{2n} F_{n-1} + \beta_{n+1} \tilde{F}_0 + \cdots + \beta_{2n} \tilde{F}_{n-1}\right),\end{aligned}$$

such that $\deg(\Psi) \leq D_0$. Notice that Ψ is a Hamming weight three polynomial.

To accomplish this, we need to determine the coefficients of F and $\tilde{F}$, also the scalars α_i and β_i, such that the coefficients of the terms in Ψ of degree greater than D_0 are equal to zero. We derive a system of equations from these

vanishing coefficients in Ψ of degree higher than D_0. This yields a system of equations of the form

$$g_1(z_1, z_2, \cdots, z_N) = 0, \cdots, g_t(z_1, z_2, \cdots, z_N) = 0,$$

where the variables $z_1, z_2, \cdots, z_N$ are the coefficients of F and $\tilde{F}$, together with the scalars α_i and β_i.

The number t of equations of this system depends on how small we want the degree bound D_0 to be. More precisely, t is the number of different terms in Ψ with degree higher than D_0. To invert the trapdoor function, which we will describe in Section 3.3, via the use of the polynomial Ψ, we require that the polynomial Ψ has degree smaller than D_0.

If we write each variable z_j in terms of the basis $\left\{1, y, \cdots, y^{n-1}\right\}$, we obtain a system of quadratic equations. More precisely, each variable z_j in this system can be written in the form

$$z_j = u_{1j} + u_{2j}y + \cdots + u_{nj}y^{n-1}, \tag{1}$$

where $u_{1j}, \cdots, u_{nj}$ are n new variables. Next, by the linearity of the Frobenius powers, we get

$$z_j^{q^i} = u_{1j} + u_{2j}y^{q^i} + \cdots + u_{nj}y^{(n-1)q^i}. \tag{2}$$

After we write each power y^m as a linear combination of the elements of the basis $1, y, \cdots, y^{n-1}$ with coefficients in k, and group like terms, we get that

$$z_j^{q^i} = h_{1j}\left(u_{1j}, \cdots, u_{nj}\right) + h_{2j}\left(u_{1j}, \cdots, u_{nj}\right)y^2 + \cdots + h_{nj}\left(u_{1j}, \cdots, u_{nj}\right)y^{n-1}, \tag{3}$$

where each h_{ij} is a linear function with coefficients in k.

We now write each variable of the system $g_1 = 0, \cdots, g_t = 0$ in the form (1), and proceed like in (2) and (3). By comparing the coefficients of the elements of the basis $\left\{1, y, y^2, \cdots, y^{n-1}\right\}$ we obtain a system of nt quadratic equations in $n\left[n(n+1) + 6n + 2\right] = n^3 + 7n^2 + 2n$ variables over k. These equations are in fact bilinear, i.e., each term of these equations has the product of a variable that comes from the coefficients and a variable that comes from the scalars. Thus, if we randomly fix the variables associated to the scalars we obtain a sparse linear system coming from the coefficients of F and $\tilde{F}$. Due to its construction, this linear system has more variables than equations, i.e., $nt < n^3 + 7n^2 + 2n$, and hence we can always get nontrivial solutions. We then randomly choose one of those solutions to build the high degree polynomials F and $\tilde{F}$ and the reduced polynomial Ψ of degree less than or equal to D_0, as explained above.

One could be tempted to randomly choose the variables coming from the coefficients of F and $\tilde{F}$, and then try to solve the linear system for the variables coming from the scalars, with the purpose of having generic core polynomials F and $\tilde{F}$. However, this approach leads to a linear system with more equations than variables, and thus, in general, this system has no nontrivial solutions.

3.2. Complexity of the reduction method and dimension of the solution space

The described method leads to a sparse linear system over the small field k with more variables than equations. This system has about n^3 variables and thus the complexity of the reduction method is polynomial: $O\left(\left(n^3\right)^{\omega}\right)$, where ω is a constant that depends on the elimination algorithm used to solve the sparse linear system.

On the other hand, after we choose the $4n$ scalars α_i and β_i in the system $\{g_i(z_1, z_2, \cdots, z_N) = 0 : i = 1, \cdots, t\}$, we get a new system over the big field K with t equations and $N - 4n$ variables (the coefficients of F and $\tilde{F}$). Therefore, in this new system the number of variables exceeds the number of equations by $(N - 4n) - t$. Hence the final linear system over the small field k has at least $n\left((N - 4n) - t\right)$ free variables. Then we have at least $q^{n((N-4n)-t)} > q^n$ possible choices for the coefficients of the polynomials F and $\tilde{F}$. Thus, if we choose large parameters q and n, and if we randomly choose a solution from the solution space, it is impossible for anyone to guess correctly the polynomials we will use. The large dimension of the solution space also ensures that there are sufficiently many choices for the core map.

3.3. How to build and invert the trapdoor function

For building a new candidate for multivariate trapdoor function, we make use of a map of the form $G = (F, \tilde{F}) : K \to K \times K$, in which F and $\tilde{F}$ have been constructed by the method described in Section 3.1. We select two invertible affine transformations $S : k^n \to k^n$ and $T : k^{2n} \to k^{2n}$. Similar to HFE, the multivariate trapdoor function will be the composition from k^n to k^{2n} given by $P = T \circ (\varphi \times \varphi) \circ G \circ \varphi^{-1} \circ S$ (see Figure 1).

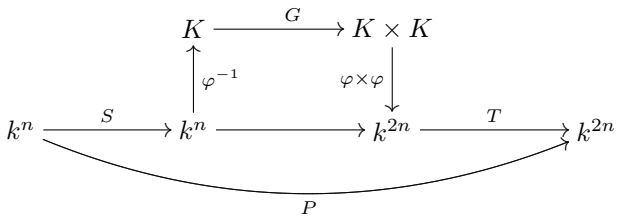

FIGURE 1. New candidate for multivariate trapdoor function.

The crucial part to invert the trapdoor function is to invert the core map $G = (F, \tilde{F})$, since the transformations S and T, and the isomorphism φ, are easy to invert. In what follows we explain how to invert G. Consider an element $X_0 \in K$ and let $(Y_1, Y_2) = G(X_0) = (F(X_0), \tilde{F}(X_0))$. We show how to recover X_0 from (Y_1, Y_2). Let $F_0, \cdots, F_{n-1}$ be the Frobenius powers of F and $\tilde{F}_0, \cdots, \tilde{F}_{n-1}$

be the Frobenius powers of $\tilde{F}$. By the construction of F and $\tilde{F}$, there exist scalars $\alpha_1, \cdots, \alpha_{2n}, \beta_1, \cdots, \beta_{2n}$ such that the polynomial

$$\Psi = \sum_{j=1}^{2} X^{q^{j-1}} \sum_{i=1}^{n} \alpha_{i+n(j-1)} F_{i-1} + \beta_{i+n(j-1)} \tilde{F}_{i-1}$$

has degree less than or equal to D_0.

We define $F' = F - Y_1$ and $\tilde{F}' = \tilde{F} - Y_2$, and

$$\Psi' = \sum_{j=1}^{2} X^{q^{j-1}} \sum_{i=1}^{n} \alpha_{i+n(j-1)} F'_{i-1} + \beta_{i+n(j-1)} \tilde{F}'_{i-1}.$$

Clearly, $F'(X_0) = 0$ and $\tilde{F}'(X_0) = 0$, and therefore $\Psi'(X_0) = 0$. Given that $F'_i = (F - Y_1)^{q^i} = F^{q^i} - Y_1^{q^i} = F_i - Y_1^{q^i}$ and $\tilde{F}'_i = (\tilde{F} - Y_2)^{q^i} = \tilde{F}^{q^i} - Y_2^{q^i} = \tilde{F}_i - Y_2^{q^i}$, the polynomial Ψ', just like Ψ, has degree less than or equal to D_0, when we choose $D_0 \geq q$. Thus, we can find the roots of Ψ' by means of Berlekamp's algorithm and therefore we can recover the common root X_0 of the polynomials F' and $\tilde{F}'$.

We now discuss the complexity of the trapdoor function inversion. The isomorphism φ and its inverse φ^{-1} can be represented in matrix form [2]. Thus, except for the inversion of the core map G, the computational cost of each step of the algorithm to invert the trapdoor function is the cost of a matrix multiplication. The degrees of the polynomials F and $\tilde{F}$, which are the components of the map G, are extremely high (usually close to q^{n-1}), which makes impossible to invert G directly for practical values of n. However, as noted above, the inversion of the map G can be reduced to finding the roots of the low degree polynomial Ψ', which can be done efficiently using Berlekamp's algorithm. For the particular case of $q = 2$, in all the computations that we performed we were able to obtain a function Ψ' of degree less than 500, whose roots can be found very quickly using Berlekamp's algorithm. Therefore, inverting the trapdoor function is a very efficient process.

4. Examples

We now show some examples built by the method described in Section 3.1. We begin by presenting a toy example in which we explain step by step the procedure. We next show two large scale cases.

Example 1. Let $q = 2$ and $n = 2$, and consider the field with two elements $k = GF(2)$. We select the irreducible polynomial $g(y) = y^2 + y + 1 \in k[y]$. A degree n extension field of k is $K = k[y]/(g(y))$. We choose a generator $b \in K$ of the multiplicative group of K such that $g(b) = 0$. Two HFE polynomials in

the ring $K[X]/\left(X^{q^n}-X\right)$ are of the form

$$F(X) = a_{01}X^3 + b_0X + b_1X^2 + c,$$
$$\tilde{F}(X) = \tilde{a}_{01}X^3 + \tilde{b}_0X + \tilde{b}_1X^2 + \tilde{c},$$

where $a_{01}, b_0, b_1, c, \tilde{a}_{01}, \tilde{b}_0, \tilde{b}_1, \tilde{c} \in K$.

The Frobenius powers of F and $\tilde{F}$, in that order, are:

$$F_0 = a_{01}X^3 + b_0X + b_1X^2 + c,$$
$$F_1 = a_{01}^2X^3 + b_0^2X^2 + b_1^2X + c^2,$$

$$\tilde{F}_0 = \tilde{a}_{01}X^3 + \tilde{b}_0X + \tilde{b}_1X^2 + \tilde{c},$$
$$\tilde{F}_1 = \tilde{a}_{01}^2X^3 + \tilde{b}_0^2X^2 + \tilde{b}_1^2X + \tilde{c}^2.$$

We now multiply the Frobenius powers by X and X^q and we obtain

$$XF_0 = a_{01}X + b_0X^2 + b_1X^3 + cX,$$
$$XF_1 = a_{01}^2X + b_0^2X^3 + b_1^2X^2 + c^2X,$$
$$X^2F_0 = a_{01}X^2 + b_0X^3 + b_1X + cX^2,$$
$$X^2F_1 = a_{01}^2X^2 + b_0^2X + b_1^2X^3 + c^2X^2,$$

$$X\tilde{F}_0 = \tilde{a}_{01}X + \tilde{b}_0X^2 + \tilde{b}_1X^3 + \tilde{c}X,$$
$$X\tilde{F}_1 = \tilde{a}_{01}^2X + \tilde{b}_0^2X^3 + \tilde{b}_1^2X^2 + \tilde{c}^2X,$$
$$X^2\tilde{F}_0 = \tilde{a}_{01}X^2 + \tilde{b}_0X^3 + \tilde{b}_1X + \tilde{c}X^2,$$
$$X^2\tilde{F}_1 = \tilde{a}_{01}^2X^2 + \tilde{b}_0^2X + \tilde{b}_1^2X^3 + \tilde{c}^2X^2.$$

Then we form the polynomial $\Psi = X(\alpha_1F_0+\alpha_2F_1+\beta_1\tilde{F}_0+\beta_2\tilde{F}_1)+X^2(\alpha_3F_0+\alpha_4F_1+\beta_3\tilde{F}_0+\beta_4\tilde{F}_1)$. In this example we want to determine the coefficients $a_{ij}, b_i, c, \tilde{a}_{ij}, \tilde{b}_i, \tilde{c}$ and the scalars α_i, β_i such that the terms of degree ≥ 2 in Ψ vanish. In order to do that, we have to solve the following two equations

$$\alpha_1b_0 + \alpha_2b_1^2 + \alpha_3(a_{01}+c) + \alpha_4\left(a_{01}^2+c^2\right) + \beta_1\tilde{b}_0 + \beta_2\tilde{b}_1^2 + \beta_3(\tilde{a}_{01}+\tilde{c}) + \beta_4\left(\tilde{a}_{01}^2+\tilde{c}^2\right) = 0,$$
$$\alpha_1b_1 + \alpha_2b_0^2 + \alpha_3b_0 + \alpha_4b_1^2 + \beta_1\tilde{b}_1 + \beta_2\tilde{b}_0^2 + \beta_3\tilde{b}_0 + \beta_4\tilde{b}_1^2 = 0.$$

We randomly choose the scalars $(\alpha_1, \cdots, \alpha_4) = (0, 0, b, 1)$ and $(\beta_1, \cdots, \beta_4) = (0, b, b^2, b^2)$. Then we write the variables $a_{01}, b_0, b_1, c, \tilde{a}_{01}, \tilde{b}_0, \tilde{b}_1, \tilde{c}$ in terms of the basis $1, y, \cdots, y^{n-1}$, as follows:

$$a_{01} = u_1 + u_2y, \cdots, \tilde{c} = u_{15} + u_{16}y.$$

Proceeding as explained in Section 3.1, we get the linear equations

$$\begin{aligned} u_1 + u_7 + u_{10} + u_{14} + u_{16} &= 0, \\ u_1 + u_7 + u_{10} + u_{13} + u_{16} &= 0, \\ u_4 + u_5 + u_6 + u_{11} + u_{13} &= 0, \\ u_3 + u_4 + u_6 + u_{13} + u_{14} &= 0. \end{aligned}$$

One of the solutions of this system is

$$(u_1, \cdots, u_{16}) = (0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1)\,.$$

This solution leads to the coefficients

$$(a_{01}, b_0, b_1, c, \tilde{a}_{01}, \tilde{b}_0, \tilde{b}_1, \tilde{c}) = \left(0, b^2, 1, b^2, b, b^2, b^2, b\right).$$

With these coefficients we get the polynomials $F = X^2 + b^2X + b^2$ and $\tilde{F} = bX^3 + b^2X^2 + b^2X + b$. Then, we use the scalars α_i and β_i to form the reduced polynomial $\Psi = b^2X$.

Example 2. In this example, for convenience in the presentation, we consider the coefficients and the scalars in the small field $k = GF(2)$ so we can nicely present the polynomials here. Of course, a realistic example would require the coefficients to be taken in the big field K and then the coefficients would not only be ones and zeros. Let $q = 2$ and $n = 17$, and consider the field with two elements $k = GF\,(2)$. We select the irreducible polynomial $y^{17} + y^3 + 1 \in k[y]$. A degree n extension field of k is $K = k[y]/\,(g(y))$. We choose a generator $b \in K$ of the multiplicative group of K such that $g(b) = 0$. We require the terms of degree ≥ 40 in Ψ to vanish. Following the same procedure explained

in the previous example, we obtain the polynomials

$$\begin{aligned}F =&X^{98304}+X^{81920}+X^{67584}+X^{66560}+X^{66048}+X^{65664}+X^{65568}+\\&X^{65552}+X^{65540}+X^{49152}+X^{40960}+X^{33792}+X^{32896}+X^{32832}+\\&X^{32800}+X^{32772}+X^{32770}+X^{32769}+X^{18432}+X^{16896}+X^{16640}+\\&X^{16416}+X^{16400}+X^{16392}+X^{16388}+X^{16386}+X^{10240}+X^{8208}+\\&X^{8192}+X^{6144}+X^{4608}+X^{4352}+X^{4112}+X^{4104}+X^{3072}+X^{2080}+\\&X^{2064}+X^{2049}+X^{1536}+X^{1152}+X^{1056}+X^{1040}+X^{1028}+X^{1025}+\\&X^{768}+X^{640}+X^{520}+X^{384}+X^{320}+X^{288}+X^{272}+X^{258}+X^{257}+\\&X^{192}+X^{136}+X^{132}+X^{130}+X^{96}+X^{68}+X^{66}+X^{36}+X^{34}+\\&X^{18}+X^{12}+X^{10}+X^{8}+X^{5}+X^{3},\end{aligned}$$

$$\begin{aligned}\tilde{F} =&X^{131072}+X^{98304}+X^{81920}+X^{66560}+X^{65664}+X^{65568}+X^{65552}+\\&X^{65544}+X^{65540}+X^{65538}+X^{33280}+X^{32832}+X^{32784}+X^{32776}+\\&X^{32770}+X^{17408}+X^{16896}+X^{16640}+X^{16512}+X^{16400}+X^{16392}+\\&X^{16384}+X^{12288}+X^{10240}+X^{9216}+X^{8704}+X^{8256}+X^{8208}+X^{8200}+\\&X^{8194}+X^{8193}+X^{8192}+X^{4224}+X^{4160}+X^{4112}+X^{4100}+X^{4097}+\\&X^{2304}+X^{2176}+X^{2080}+X^{2052}+X^{2049}+X^{1280}+X^{1088}+X^{1056}+\\&X^{1028}+X^{1025}+X^{768}+X^{576}+X^{544}+X^{528}+X^{520}+X^{514}+X^{512}+\\&X^{384}+X^{258}+X^{257}+X^{256}+X^{136}+X^{132}+X^{130}+X^{128}+X^{96}+\\&X^{68}+X^{66}+X^{48}+X^{40}+X^{36}+X^{34}+X^{32}+X^{20}+X^{18}+X^{10}+\\&X^{6}+X^{3}+X^{2}.\end{aligned}$$

The scalars $(\alpha_1, \cdots, \alpha_{34})$ are

$$(1,1,0,0,0,1,1,0,1,1,0,1,1,0,0,1,1,1,1,1,0,0,1,0,0,1,0,0,0,0,0,1,0,1,1)\,,$$

and the scalars $(\beta_1, \cdots, \beta_{34})$ are

$$(1,0,1,0,0,0,0,1,0,1,1,0,0,0,1,1,0,0,0,1,1,1,0,1,0,1,0,0,0,1,1,1,1,0)\,.$$

These values, together with the polynomials F and $\tilde{F}$, lead to the reduced polynomial

$$\Psi = X^{36}+X^{35}+X^{33}+X^{26}+X^{25}+X^{22}+X^{19}+X^{12}+X^{11}+X^{8}.$$

The high degrees of F and $\tilde{F}$ prevent us to invert $G = (F, \tilde{F})$ directly, but we can invert G using Ψ as explained in Section 3.3. To show how to invert G, we randomly choose $X_0 = b^{51298} \in K$. Then we calculate $(Y_1, Y_2) =$

$(F(X_0), \tilde{F}(X_0)) = (b^{114562}, b^{126611})$. We now show how to recover X_0 from (Y_1, Y_2). First we put $F' = F - Y_1$ and $\tilde{F}' = \tilde{F} - Y_2$ and use the scalars α_i and β_i to form the low degree polynomial

$$\Psi' = X^{36} + X^{35} + X^{33} + X^{26} + X^{25} + X^{22} + X^{19} + X^{12} + X^{11} + X^8 + X^5 + b^{117898}X^2 + b^{101296}X.$$

The set of roots of the polynomial Ψ', found very quickly by Berlekamp's algorithm, is $\{0, b^{51298}\}$. Notice that X_0 is one of these roots.

Example 3. With $q = 2$ and $n = 60$, and taking the coefficients and the scalars in the big field K, we found a pair of polynomials F and $\tilde{F}$ with the same degree D and $\lfloor \log_q D \rfloor = 59$. The reduced polynomial Ψ that we found has degree 386. Hence, we can invert easily the map $G = (F, \tilde{F})$ via the low degree polynomial Ψ using Berlekamp's algorithm. These polynomials are too big to be displayed here. In this example we needed to deal with a sparse matrix of size 208080×226800 and the time and memory used were 4.6 days and 52.7 GB, respectively.

5. Security analysis

As we noted before, there are two attacks that have threatened the security of HFE schemes: direct algebraic attack and Kipnis-Shamir MinRank attack. Here we analyze these attacks against the new candidates for multivariate trapdoor functions in the special case $q = 2$.

We would like to recall the recent result on the degree of regularity of Ding and Hodges [8]. We know that for an HFE system P the degree of regularity is bounded by

$$\frac{(q-1)\,\text{Q-Rank}(P)}{2} + 2,$$

where Q-Rank(P) is the quadratic rank for the quadratic operator P. So for $q = 2$ we have that this degree of regularity is bounded by

$$\frac{\text{Q-Rank}(P)}{2} + 2.$$

Since the corresponding quadratic rank used in the Kipnis-Shamir MinRank attack is given by Q-Rank(P), we see that if an HFE system has a high degree of regularity, this HFE system must have a high quadratic rank for the Kipnis-Shamir attack. From this we conclude that it suffices to show that our new trapdoor functions have high degree of regularity, in order to demonstrate that the MinRank attack will not work against these new trapdoor functions. We dedicate the rest of this section to discuss the algebraic attack against the new trapdoor functions.

Suppose that someone, who does not know the private trapdoor information, wants to invert the multivariate trapdoor function $P\colon k^n \to k^{2n}$, $P =$

n	Average time [s]	Minimum time [s]	Maximum time [s]	$\lfloor \log_q D \rfloor$
18	0.100	0.100	0.100	17
20	0.205	0.200	0.210	19
22	0.434	0.420	0.440	21
24	0.849	0.840	0.860	23
26	7.981	7.950	8.020	25
28	32.046	31.550	32.690	27
30	90.770	76.430	110.250	29
32	225.557	221.310	230.720	31

TABLE 1. Algebraic attack against the new trapdoor function for $q = 2$.

$(P_1, \ldots, P_{2n})$, i.e., she wants to find a pre-image of an element $(y_1, \ldots, y_{2n}) \in \operatorname{Im} P \subseteq k^{2n}$. This person only has access to the trapdoor function P. Attempting to solve directly the system of equations

$$\begin{aligned} P_1(x_1, \ldots, x_n) - y_1 &= 0 \\ P_2(x_1, \ldots, x_n) - y_2 &= 0 \\ &\vdots \\ P_{2n}(x_1, \ldots, x_n) - y_{2n} &= 0, \end{aligned} \tag{4}$$

is what we call the direct algebraic attack. One way to do this is with the help of a Gröbner basis. We ran extensive experiments using the F_4 algorithm of MAGMA, [3], to perform the direct algebraic attack for $q = 2$ and several values of n, on a Sun X4440 server, with four Quad-Core AMD OpteronTM Processor 8356 CPUs and 128 GB of main memory (each CPU is running at 2.3 GHz). For the trapdoor functions used in these experiments we utilized $D_0 = 500$. The results of our experiments are shown in Table 1 and 2, and Figure 2, 3 and 4. The F_4 function of MAGMA is the most efficient implementation of the Gröbner F_4 algorithm that is currently available.

In Table 1 and Figure 2 we can observe that the time needed to solve the equations by F_4 has an exponential growth in n. We can also see this behaviour with the memory used by the F_4 algorithm. This situation is different from the one observed by Faugere and Joux in [10]. The difference lies on the fact that in [10] the quadratic equations are produced using a polynomial of fixed low degree as core map in the HFE cryptosystem, and in our new trapdoor function the quadratic equations are generated via two high degree polynomials. In our experiments, in general, these two high degree polynomials have the same degree D and this degree increases as n increases (see Table 1). This is the fundamental security improvement of our new method.

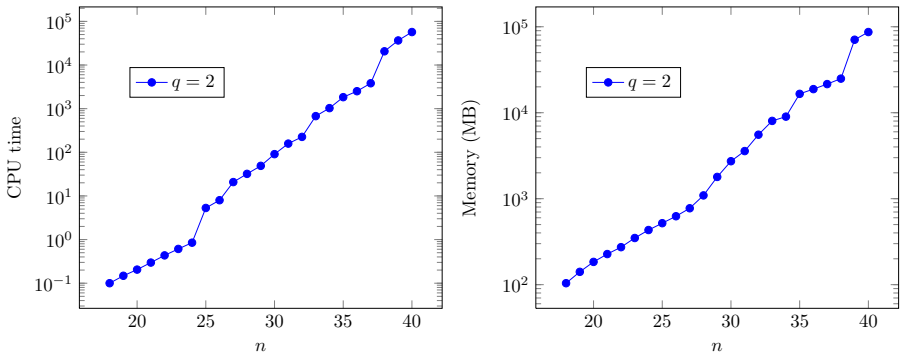

FIGURE 2. Algebraic attack against the new trapdoor function for $q = 2$.

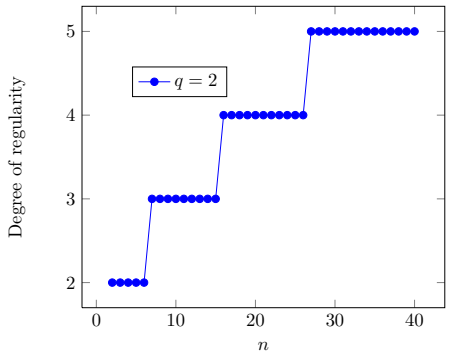

FIGURE 3. Degree of regularity for the algebraic attack against the new trapdoor function.

Another evidence that the complexity of the algebraic attack against the new trapdoor functions is exponential, is that the degree of regularity of the trapdoor function increases as n increases. This behaviour can be observed in Figure 3. As we mentioned earlier, the fact that this degree of regularity increases as n increases, not only says that the direct algebraic does not work against the new trapdoor functions, but also that the Kipnis-Shamir MinRank attack is not successful against these new trapdoor functions.

We also chose random quadratic equations of the same dimensions ($k^n \to k^{2n}$) and found that the time needed to solve such equations using Gröbner bases is essentially the same that is needed to solve the quadratic equations from the new trapdoor function. Table 2 and Figure 4 show this comparison between the new trapdoor functions and random equations for different values of n.

(a) New trapdoor function

n	Average time [s]	Memory [MB]	Degree of regularity
14	0.019	3	3
16	0.142	5	4
18	0.100	8	4
20	0.205	13	4
22	0.434	20	4
24	0.849	33	4
26	7.981	118	4
28	32.046	1121	5
30	90.770	2769	5
32	225.557	5610	5

(b) Random equations

n	Average time [s]	Memory [MB]	Degree of regularity
14	0.040	12	3
16	0.060	13	4
18	0.100	16	4
20	0.200	21	4
22	0.440	31	4
24	0.830	46	4
26	7.800	105	4
28	34.700	1087	5
30	87.810	2725	5
32	239.260	5549	5

TABLE 2. Algebraic attack comparison between the new trapdoor function and random equations for $q = 2$.

In Table 2 we can notice that the degree of regularity of the new trapdoor function is the same as the degree of regularity of the set of random equations. In both cases we observe that the degree of regularity increases as n increases. From all the information we collected with our experiments, it seems that the F_4 algorithm is no more efficient in solving the equations from the new trapdoor function than a set of random equations of the same dimensions. In other words, with respect to the direct algebraic attack, the new trapdoor function behaves as if it were a system of random quadratic equations.

6. Conclusions and future work

We have created a procedure to build candidates for multivariate trapdoor functions using pairs of HFE polynomials of high degree. The way to invert these trapdoor functions is through a low degree polynomial of Hamming weight three. We have shown how the main attacks against HFE do not work for these new trapdoor functions for the particular case of $q = 2$.

The next step in our research is to use the ideas of this paper to construct candidates for multivariate trapdoor functions for larger values of q. The benefit of doing this is that the larger the value of q is, the smaller the value of n is, and so the smaller the sizes of the matrices needed to construct the trapdoor functions are. This would significantly reduce the time and memory needed to construct the multivariate trapdoor functions. It would also be important to study the effect that the matrix sparsity has on the complexity of the algorithm used to construct the new trapdoor functions.

In Section 5 we saw that for $q = 2$ it suffices to show that the new trapdoor functions have high degree of regularity to conclude that the Kipnis-Shamir

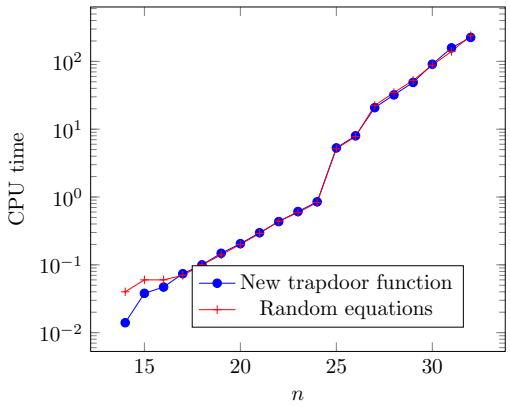

FIGURE 4. Algebraic attack comparison between the new trapdoor function and random equations.

MinRank attack does not work against these trapdoor functions. However, for larger values of q this argument cannot be used and the MinRank attack must be directly performed against these new trapdoor functions. We are currently starting to work these cases and we will publish the results in an upcoming paper.

We believe that these ideas have great potential to construct new variants of the HFE cryptosystem that are resistant against the direct algebraic and MinRank attacks.

References

[1] Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, *Post quantum cryptography*, Springer, 2009.

[2] Luk Bettale, Jean-Charles Faugère, and Ludovic Perret, *Cryptanalysis of hfe, multi-hfe and variants for odd and even characteristic*, Designs, Codes and Cryptography **69** (2013), no. 1, 1–52.

[3] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265, Computational algebra and number theory (London, 1993). MR MR1484478

[4] Crystal Clough, John Baena, Jintai Ding, Bo-Yin Yang, and Ming shing Chen, *Square, a new multivariate encryption scheme*, CT-RSA 2009, Springer, Berlin, 2009.

[5] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir, *Efficient algorithms for solving overdefined systems of multivariate polynomial equations*, Advances in Cryptology—EUROCRYPT 2000, Lecture Notes in Computer Science, vol. 1807, Springer Berlin Heidelberg, 2000, pp. 392–407.

[6] J. Ding, J. Buchmann, M. Mohamed, W. Mohamed, and R. Weinmann, *Mutant xl*, First International Conference on Symbolic Computation and Cryptography (SCC 2008), Beijing, 2008.

[7] Jintai Ding, Jason E. Gower, and Dieter S. Schmidt, *Multivariate public key cryptosystems*, Advances in Information Security, vol. 25, Springer, New York, 2006. MR MR2244659 (2007i:94049)

[8] Jintai Ding and Timothy J. Hodges, *Inverting hfe systems is quasi-polynomial for all fields*, Advances in Cryptology—CRYPTO 2011 (Phillip Rogaway, ed.), Lecture Notes in Computer Science, vol. 6841, Springer Berlin Heidelberg, 2011, pp. 724–742.

[9] Jintai Ding and Dieter Schmidt, *Cryptanalysis of HFEv and internal perturbation of HFE*, Public key cryptography—PKC 2005, Lecture Notes in Comput. Sci., vol. 3386, Springer, Berlin, 2005, pp. 288–301. MR MR2174048 (2006j:94061)

[10] Jean-Charles Faugère and Antoine Joux, *Algebraic cryptanalysis of hidden field equation (HFE) cryptosystems using Gröbner bases*, Advances in cryptology—CRYPTO 2003, Lecture Notes in Comput. Sci., vol. 2729, Springer, Berlin, 2003, pp. 44–60. MR MR2093185 (2005e:94140)

[11] Chia hsin Owen Chen, Ming shing Chen, Jintai Ding, Fabian Werner, and Bo yin Yang, *B.y.: Odd-char multivariate hidden field equations. cryptology eprint archive*, (2008).

[12] Aviad Kipnis, Jacques Patarin, and Louis Goubin, *Unbalanced oil and vinegar signature schemes*, Advances in cryptology—EUROCRYPT '99 (Prague), Lecture Notes in Comput. Sci., vol. 1592, Springer, Berlin, 1999, pp. 206–222. MR MR1717470

[13] Aviad Kipnis and Adi Shamir, *Cryptanalysis of the HFE public key cryptosystem by relinearization*, Advances in cryptology—CRYPTO '99 (Santa Barbara, CA), Lecture Notes in Comput. Sci., vol. 1666, Springer, Berlin, 1999, pp. 19–30. MR MR1729291 (2000i:94052)

[14] MohamedSaiedEmam Mohamed, Daniel Cabarcas, Jintai Ding, Johannes Buchmann, and Stanislav Bulygin, *Mxl3: An efficient algorithm for computing Gröbner bases of zero-dimensional ideals*, Information, Security and Cryptology—ICISC 2009, Lecture Notes in Computer Science, vol. 5984, Springer Berlin Heidelberg, 2010, pp. 87–100.

[15] MohamedSaiedEmam Mohamed, WaelSaidAbdElmageed Mohamed, Jintai Ding, and Johannes Buchmann, *Mxl2: Solving polynomial equations over gf(2) using an improved mutant strategy*, Post-Quantum Cryptography, Lecture Notes in Computer Science, vol. 5299, Springer Berlin Heidelberg, 2008, pp. 203–215.

[16] Jacques Patarin, *Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): Two new families of asymmetric algorithms*, Advances in Cryptology—EUROCRYPT 96 (Ueli Maurer, ed.), Lecture Notes in Computer Science, vol. 1070, Springer-Verlag, 1996, pp. 33–48.

[17] Peter W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. on Computing (1997), 1484–1509.

Departamento de Matemáticas
Universidad Nacional de Colombia
Facultad de Ciencias
Calle 59A No 63-20 - Núcleo El Volador
Medellín, Colombia
e-mail: jporras@unal.edu.co

Departamento de Matemáticas
Universidad Nacional de Colombia
Facultad de Ciencias
Calle 59A No 63-20 - Núcleo El Volador
Medellín, Colombia
e-mail: jbbaena@unal.edu.co

Department of Mathematical Sciences
University of Cincinnati
4199 French Hall West
Cincinnati, OH 45221-0025, USA
e-mail: jintai.ding@uc.edu