From: (b) (6)
To: Perlner, Ray A. (Fed)
Subject: Zhang Tan Paper
Date: Thursday, June 23, 2016 10:17:43 AM
Attachments: 0025.pdf

Hi, Ray,

I'm working with a few guys on a practical attack on ZHFE, and I think that it may work on ZHFE-, too. The idea is a KS attack with minors modelling. It's basically the same attack as on multi-HFE, but there are a couple of major differences. First, there is only one variable over the extension field instead of multiple (this makes the attack easier) and then there is another step after the minrank attack.

We have had some weird results in experiments, though it seems like the attack works and is fairly efficient for most instances. The problem is that there are these two papers from Zhang and Tan. The first is attached and was their submission to PQCRYPTO 2016 that got rejected. Here they seem to be stating some of the steps of our attack. The second has recently been published and there is an eprint version, http://eprint.iacr.org/2016/637.pdf and has what seems to be the exact same attack and a different complexity estimate. In the first paper, they say the attack is feasible and should break ZHFE (although they are missing several steps in the attack), and then in the second paper, they say the complexity should be over 2^120.

I'm wondering what you think about this. It's clear that they aren't doing any experiments, but their more recent paper claims to prove security for ZHFE though they state that they break ZHFE in the previous paper.

Cheers,
Daniel

On the Security of ZHFE and Its Enhancement

Wenbin Zhang and Chik How Tan

Temasek Laboratories
National University of Singapore
tslzw@nus.edu.sg and tsltch@nus.edu.sg

Abstract. At PQCrypto'14 Porras, Baena and Ding proposed a new interesting encryption scheme called ZHFE. They used two high degree polynomials as the core map which are related by a secret low degree polynomial. ZHFE is relatively efficient on decryption but is inefficient on generation of the private key. They argued that ZHFE is secure against the main attacks that have threatened the security of HFE. In this paper, we investigate the security of ZHFE by computing its Q-Rank explicitly. We compute a small upper bound for the Q-Rank of ZHFE which is determined by the low degree of the secret polynomial but independent of the number of variables. We then formulate an attack to ZHFE using the Kipnis-Shamir MinRank attack to recover the private key of ZHFE, and use this upper bound to give an estimation on the security of ZHFE. Especially we find that the proposed parameters ZHFE(7,55,105), claimed to be of security greater than 2^{80}, has Q-Rank 4 only and its security is $2^{57.8}$. Hence ZHFE is far less secure than as claimed. To resolve the security issue of ZHFE, we then propose a solution to enhance ZHFE to have high Q-Rank and high enough security level against the Kipnis-Shamir MinRank attack. In addition, our enhanced ZHFE improves significantly the generation of the private key. We also propose a few practical parameter sets for implementation with security around 2^{100} or higher.

Keywords: post-quantum cryptography, multivariate public key cryptography, HFE

1 Introduction

Multivariate public key cryptography (MPKC) [DGS06] is a candidate of post-quantum cryptographyp to resist future quantum computers. MPKC uses multivariate polynomials to represent its public key and its security is backed by the fact that solving a random multivariate quadratic polynomial system is NP-hard [GJ79].

1.1 Hidden Field Equations (HFE)

There have been numerous schemes in MPKC since 1980's, but most of them have been broken. One of the most important schemes is Patarin's Hidden Field Equations (HFE) encryption schemes [Pat96]. Though Patarin's original HFE has been broken thoroughly [KS99, GJS06, BFP13], it has been developed into a big family. Some of its variants remain unbroken until now, for example HFEv for encryption and HFEv- for signature.

In 2014 Porras, Baena and Ding [PBD14] gave a very interesting new construction of HFE trapdoor for encryption. Namely, the core map is designed to satisfy

certain relation which is kept secret, and it is infeasible to directly solve the central map, but it is efficient to solve it with the help of the secret relation. They achieve this by choosing two high degree partially random polynomials F_1, F_2 which are related nonlinearly by a third low degree polynomial Ψ, and then setting the pair F_1, F_2 as the central map and keeping Ψ secret. Such an encryption scheme is called ZHFE in [PBD14]. They showed that ZHFE is relatively efficient on decryption, but is inefficient to generate the private key.

ZHFE belongs to the HFE family. So attacks applicable to it are direct algebraic attacks [FJ03] and the Kipnis-Shamir MinRank attack, KS attack for short, [KS99, BFP13]. They showed that ZHFE can resist direct algebraic attack by theoretical and experimental results. They then further provided some experimental results on applying KS attack to ZHFE and claimed that ZHFE has growing Q-Rank as the number of variables grows to resist the KS attack. They also recommended a practical parameter set ZHFE(7,55,105) and claimed that its security level is greater than 2^{80}.

1.2 Contribution of this Paper

The construction of ZHFE is novel and at first glance it seems promising. However we find that it is not as secure as claimed. In this paper, we give an detailed investigation on the security of ZHFE by computing its Q-Rank explicitly. We find that it indeed has a small upper bound dependent of the low degree of the secret third polynomial and is independent on the number of variables. We then formulate a key-recovery attack by applying recent results of KS attack to HFE [BFP13], and estimate the security level of ZHFE as $O(n^{2(r+1)})$ where n is the number of variables and r is the Q-Rank. For example, we find that the proposed parameter set ZHFE(7,55,105) has Q-Rank at most 4 only and its security level is $2^{57.8}$.

Besides computing the Q-Rank and giving an estimation of the security of ZHFE, we also provide a solution to enhance ZHFE. Our enhanced ZHFE outperforms ZHFE on security. It does not only increase the Q-Rank of ZHFE and thus its security, but also simplified its generation of private key significantly. We thus provide a secure alternative to ZHFE and propose several parameter sets for our enhanced ZHFE to be of security level around 2^{100} or higher.

This paper is organized as follows. In Section 2, we review the ZHFE scheme, and then give a cryptanalysis of ZHFE in next section. In Section 4, we present an enhancement of ZHFE, compare it with ZHFE and discuss its security. Finally Section 5 concludes this paper.

2 The Encryption Scheme ZHFE

In this section, we shall recall Porras, Baena and Ding's novel encryption scheme ZHFE [PBD14]. Let $\mathbb{K}$ be a degree n extension of $\mathbb{F}_q$ and $\phi : \mathbb{K} \to \mathbb{F}_q^n$ the canonical isomorphism of vector spaces over $\mathbb{F}_q$.

2.1 Design of the Core Map

Firstly we describe how to generate the core map of the scheme which consists of two high degree polynomials over $\mathbb{K}$,

$$F_1(X) = \sum a_{ij} X^{q^i+q^j} + \sum b_i X^{q^i} + c,$$

$$F_2(X) = \sum a'_{ij} X^{q^i+q^j} + \sum b'_i X^{q^i} + c'$$

whose coefficients will be determined by a linear system later. Let

$$\Psi(X, F_1, F_2) = X(u_1F_1 + u_2F_1^q + \cdots + u_nF_1^{q^{n-1}} + v_1F_2 + v_2F_2^q \cdots + v_nF_2^{q^{n-1}})$$
$$+X^q(u_{n+1}F_1 + u_{n+2}F_1^q + \cdots + u_{2n}F_1^{q^{n-1}} + v_{n+1}F_2 + v_{n+2}F_2^q + \cdots + v_{2n}F_2^{q^{n-1}})$$

be a polynomial with coefficients chosen randomly from $\mathbb{K}$. Choose a positive integer D. The coefficients of F_1, F_2 are required to satisfy the following condition

$$\deg \Psi(X, F_1(X), F_2(X)) \leq D. \tag{1}$$

From this condition, we can get a very large linear system for the coefficients of F_1, F_2. Any nonzero solution to this system gives a pair of F_1, F_2. We then compute the expression

$$\begin{aligned} \Psi_D(X) &= \Psi(X, F_1(X), F_2(X)) \\ &= \sum_{0\leq i\leq 1} \sum_{q^i+q^j+q^k\leq D} a''_{ijk} X^{q^i+q^j+q^k} \sum_{q^i+q^j\leq D} b''_{ij} X^{q^i+q^j} + \sum_{q^i\leq D} c''_i X^{q^i}. \end{aligned} \tag{2}$$

Secondly, we describe how to invert the central map, i.e., the pair of polynomials (F_1, F_2). Given any $Y_1, Y_2 \in \mathbb{K}$, since F_1, F_2 are of high degree, it is expected infeasible generally to solve

$$\begin{cases} F_1(X) = Y_1 \\ F_2(X) = Y_2 \end{cases} \tag{3}$$

directly. It is shown in [PBD14] that equations (3) can be solved with the help of

$$\Psi_D(X) = \Psi(X, Y_1, Y_2);$$

namely, solutions to Equation (3) are also solutions to

$$\Psi_D(X) - \Psi(X, Y_1, Y_2) = 0. \tag{4}$$

In other words, we can solve equation (4) and then check which solution of it is also a solution to the original equations (3). Since degree of equation (4) is bounded by D, we can be choose a relatively small D so that it can be solved efficiently using Berlekamp's algorithm.

2.2 ZHFE

We can now describe the encryption scheme ZHFE. Its public map is

$$P = T \circ (\phi \times \phi) \circ (F_1, F_2) \circ \phi^{-1} \circ S : \mathbb{F}_q^n \to \mathbb{F}_q^{2n}$$

where $S : \mathbb{F}_q^n \to \mathbb{F}_q^n$ and $T : \mathbb{F}_q^{2n} \to \mathbb{F}_q^{2n}$ are two randomly chosen invertible affine transformations.

Public Key The public key includes $\mathbb{F}_q$ and the polynomial map $P(x_1, \ldots, x_n)$.
Private Key The private key includes Ψ, D, Ψ_D and S, T.
Encryption The ciphertext of a plaintext $(x_1, \ldots, x_n) \in \mathbb{F}_q^n$ is obtained by computing $(y_1, \ldots, y_n) = P(x_1, \ldots, x_n)$.
Decryption A given ciphertext $\mathbf{y}$ is decrypted as follows:

1. Compute $(w_1, \ldots, w_{2n}) = T^{-1}(\mathbf{y})$.
2. Compute $(Y_1, Y_2) = (\phi^{-1}(w_1, \ldots, w_n), \phi^{-1}(w_{n+1}, \ldots, w_{2n}))$;
3. Substitute (Y_1, Y_2) into equation (4), solve it by Berlekamp's algorithm, and let $\mathcal{Z}$ be the set of solutions.
4. For each $X \in \mathcal{Z}$, compute $S^{-1}(\phi(X))$ and check whether it is a solution to $P(\mathbf{x}) = \mathbf{y}$. Each solution is a candidate for the plaintext — additional redundant information must be added to determine which candidate is the correct plaintext.

In [PBD14], the parameter set $(q, n, D) = (7, 55, 105)$ is suggested for ZHFE and its security level is claimed to be greater than 2^{80}. Features of ZHFE(7,55,105) are listed in Table 1. A drawback of ZHFE is the low efficiency of generating the secret F_1, F_2 which is $O((n^3)^\omega)$ where $2 \leq \omega \leq 3$ depends on algorithm used.

Table 1. ZHFE(7,55,105)

Public Key	Private Key	Encryption Time	Decryption Time	Claimed Security
66 KB	11 KB	0.024 s	0.427 s	2^{80}

3 Cryptanalysis of ZHFE

ZHFE belongs to the big family of HFE schemes. So the authors of [PBD14] analyzed the security of ZHFE against current major attacks to the HFE family. There are two kinds of such attacks, direct algebraic attacks [FJ03] and Kipnis-Shamir MinRank Attack (KS attack) [KS99, BFP13].

Each of F_1, F_2 individually is kind of random, but as a pair they are not, and they are related by a third polynomial which is kept secret. Moreover, F_1, F_2 generally have very high degree, or even full degree. Hence it is expected that ZHFE would be just like random systems against direct algebraic attack. This expectation is

confirmed by the strong (theoretical and experimental) evidences given in [PBD14]. Therefore we shall consider only Kipnis-Shamir MinRank Attack (KS attack) in this section.

3.1 Kipnis-Shamir MinRank Attack (KS Attack)

KS attack was originally proposed by Kipnis and Shamir in 1999 [KS99] to recover the secret key of HFE. It relies on the fact that the core map of HFE has a small rank and thus can be converted into the MinRank Problem.

The MinRank Problem: Let $\mathbb{K}$ be a finite field and $M_1, \ldots, M_m$ $t \times t$ matrices over $\mathbb{K}$. Given a positive integer $r \leq t$, find scalars $\lambda_1, \ldots, \lambda_m$, not all zero, such that

$$\mathrm{Rank}(\lambda_1 M_1 + \cdots + \lambda_m M_m) \leq r.$$

It is well known that this is generally an NP-hard problem, but if r is small, then the MinRank problem is not too hard. Kipnis and Shamir proposed a method to solve the MinRank problem for small r and thus gave an attack to HFE [KS99]. In 2013, Bettale, Faugère and Perret improved KS attack significantly and break HFE and multi-HFE [BFP13]. Their attack used the notion of quadratic rank (Q-Rank for short) which is the minimal rank of all the linear combinations of the associated matrices of the quadratic polynomials of HFE. They showed that this rank is exactly the minimal quadratic rank of all linear combinations of the Frobenius powers of the core map of HFE. Using this Q-Rank, the complexity of Bettale, Faugère and Perret's, BFP's for short, KS attack is estimated as $O(n^{(r+1)\omega})$ for small q where r is the Q-Rank of the HFE scheme.

In [PBD14], Porras, Baena and Ding suggested ZHFE(7,55,105) and showed by experiments that its Q-Rank is greater than 3. However for the case that whether the Q-Rank is 4, their experiment did not stop but reached the set limit on time and memory. In [PBD14], they claimed that for ZHFE:

1. Q-rank grows as n grows;
2. Q-rank is independent of D;
3. In principle there seems no obvious way to recover Ψ.

In the rest of this section, we will compute a small upper bound for the Q-rank of ZHFE which is independent on n and find that the degree D of the secret Ψ determines this upper bound. Moreover, we will show how to apply BFP's KS attack [BFP13] to recover the secret Ψ. Therefore the above claims of [PBD14] are all wrong.

3.2 Small Upper Bound for the Q-Rank of ZHFE

Let

$$\bar{F}_1 = u_1 F_1 + u_2 F_1^q + \cdots + u_n F_1^{q^{n-1}} + v_1 F_2 + v_2 F_2^q \cdots + v_n F_2^{q^{n-1}}$$
$$\bar{F}_2 = u_{n+1} F_1 + u_{n+2} F_1^q + \cdots + u_{2n} F_1^{q^{n-1}} + v_{n+1} F_2 + v_{n+2} F_2^q + \cdots + v_{2n} F_2^{q^{n-1}}.$$

Then

$$\Psi(X, F_1, F_2) = X\bar{F}_1 + X^q\bar{F}_2.$$

We will compute the rank of $\bar{F}_1$ and $\bar{F}_2$.

First case $q > 2$. Write

$$\bar{F}_1(X) = \sum_{0\le i\le j\le n-1} \bar{a}_{ij}X^{q^i+q^j} + \sum_{0\le i\le n-1} \bar{b}_i X^{q^i} + \bar{c}$$
$$\bar{F}_2(X) = \sum_{0\le i\le j\le n-1} \bar{a}'_{ij}X^{q^i+q^j} + \sum_{0\le i\le n-1} \bar{b}'_i X^{q^i} + \bar{c}'$$

Then by calculation, we have

$$\begin{aligned}\Psi(X, F_1, F_2) = &\sum_{1\le i\le j\le n-1} \bar{a}'_{ij}X^{q+q^i+q^j} + \sum_{2\le i\le j\le n-1} \bar{a}_{ij}X^{1+q^i+q^j} + \sum_{1\le i\le n-1} \bar{b}'_i X^{q+q^i}\\ &+ \sum_{1\le j\le n-1} (\bar{a}_{1j} + \bar{a}'_{0j})X^{1+q+q^j} + \sum_{1\le j\le n-1} \bar{a}_{0j}X^{2+q^j} + \sum_{2\le i\le n-1} \bar{b}_i X^{1+q^i}\\ &+ \bar{a}'_{00}X^{2+q} + (\bar{b}_1 + \bar{b}'_0)X^{1+q} + \bar{c}'X^q + \bar{a}_{00}X^3 + \bar{b}_0X^2 + \bar{c}X.\end{aligned}$$

Separate D according to $q + 2q^{s-1} < D \le q + 2q^s$. Then $\deg\Psi(X, F_1, F_2) \le D$ if and only if

$$\begin{aligned}&\bar{a}'_{ij} = 0, \quad 1 \le i < j, j > s\\ &\bar{a}_{ij} = 0, \quad 2 \le i < j, j > s\\ &\bar{a}_{0j} = 0, \bar{a}_{1j} + \bar{a}'_{0j} = 0, \quad j > s\\ &\bar{b}_i = 0, \bar{b}'_i = 0, \quad i > s\end{aligned}$$

$$\bar{F}_1 = \sum_{1\le j\le n-1} \bar{a}_{1j}X^{q+q^j} + \sum_{0\le i\le j\le s, i\ne 1} \bar{a}_{ij}X^{q^i+q^j} + \sum_{0\le i\le s} \bar{b}_i X^{q^i} + \bar{c}$$
$$\bar{F}_2 = \sum_{0\le j\le n-1} \bar{a}'_{0j}X^{1+q^j} + \sum_{1\le i\le j\le s} \bar{a}'_{ij}X^{q^i+q^j} + \sum_{0\le i\le s} \bar{b}'_i X^{q^i} + \bar{c}'$$

and their relation is

$$\bar{a}_{1j} = -\bar{a}'_{0j}, \quad j > s$$

So $\bar{F}_1, \bar{F}_2$ can have maximum degree $q + q^{n-1}$ and $1 + q^{n-1}$ respectively. Although their maximum degree can be this high, their ranks are very small. If q odd, their representing matrices are

$$\begin{pmatrix} \bar{a}_{00} & \bar{a}_{01}/2 & \cdots & \bar{a}_{0s}/2 & & \\ \bar{a}_{01}/2 & \bar{a}_{11} & \cdots & \bar{a}_{1s}/2 & \cdots & \bar{a}_{1,n-1}/2 \\ \vdots & \vdots & & \vdots & & \\ \bar{a}_{0s}/2 & \bar{a}_{1s}/2 & \cdots & \bar{a}_{ss} & & \\ & \vdots & & & & \\ & \bar{a}_{1,n-1}/2 & & & & \end{pmatrix}, \quad \begin{pmatrix} \bar{a}'_{00} & \cdots & \bar{a}'_{0s}/2 & \cdots & \bar{a}'_{0,n-1}/2 \\ \vdots & & \vdots & & \\ \bar{a}'_{0s}/2 & \cdots & \bar{a}'_{ss} & & \\ \vdots & & & & \\ \bar{a}'_{0,n-1}/2 & & & & \end{pmatrix}$$

respectively. If $q > 2$ even, their representing matrices are

$$\begin{pmatrix} 0 & \bar{a}_{01} & \cdots & \bar{a}_{0s} & & \\ \bar{a}_{01} & 0 & \cdots & \bar{a}_{1s} & \cdots & \bar{a}_{1,n-1} \\ \vdots & \vdots & & \vdots & & \\ \bar{a}_{0s} & \bar{a}_{1s} & \cdots & 0 & & \\ & \vdots & & & & \\ & \bar{a}_{1,n-1} & & & & \end{pmatrix}, \quad \begin{pmatrix} 0 & \cdots & \bar{a}'_{0s} & \cdots & \bar{a}'_{0,n-1} \\ \vdots & & \vdots & & \\ \bar{a}'_{0s} & \cdots & 0 & & \\ \vdots & & & & \\ \bar{a}'_{0,n-1} & & & & \end{pmatrix}$$

respectively, with zero diagonal. It is then obvious that their ranks are both $\leq s+2$.

Second case $q = 2$. Write

$$\begin{aligned} \bar{F}_1(X) &= \sum_{0 \leq i < j \leq n-1} \bar{a}_{ij} X^{2^i+2^j} + \sum_{0 \leq i \leq n-1} \bar{b}_i X^{2^i} + \bar{c} \\ \bar{F}_2(X) &= \sum_{0 \leq i < j \leq n-1} \bar{a}'_{ij} X^{2^i+2^j} + \sum_{0 \leq i \leq n-1} \bar{b}'_i X^{2^i} + \bar{c}' \end{aligned}$$

Notice that there is no $\bar{a}_{ii}, \bar{a}'_{ii}$ (or they can be regarded as 0) since $X^{2^i+2^i} = X^{2^{i+1}}$.

$$\begin{aligned} \Psi(X, F_1, F_2) = &\sum_{1 \leq i < j \leq n-1} \bar{a}'_{ij} X^{2+2^i+2^j} + \sum_{2 \leq i < j \leq n-1} \bar{a}_{ij} X^{1+2^i+2^j} \\ &+ \sum_{2 \leq i \leq n-1} (\bar{a}_{1i} + \bar{a}'_{0i}) X^{1+2+2^i} + \sum_{2 \leq i \leq n-1} (\bar{a}_{0i} + \bar{b}'_i) X^{2+2^i} \\ &+ \sum_{3 \leq i \leq n-1} \bar{b}_i X^{1+2^i} + (\bar{a}'_{01} + \bar{b}_2) X^5 + (\bar{a}_{01} + \bar{b}'_1) X^4 \\ &+ (\bar{b}_1 + \bar{b}'_0) X^3 + (\bar{b}_0 + \bar{c}') X^2 + \bar{c} X \end{aligned}$$

For $2 + 2^{s-2} + 2^{s-1} < D \leq 2 + 2^{s-1} + 2^s$, $\deg \Psi(X, F_1, F_2) \leq D$ if and only if

$$\begin{aligned} &\bar{a}'_{ij} = 0, \quad 1 \leq i < j, j > s \\ &\bar{a}_{ij} = 0, \quad 2 \leq i < j, j > s \\ &\bar{a}_{1i} + \bar{a}'_{0i} = 0, \bar{a}_{0i} + \bar{b}'_i = 0, \bar{b}_i = 0, \quad i > s \end{aligned}$$

$$\begin{aligned} \bar{F}_1 &= \sum_{1 < j \leq n-1} \bar{a}_{1j} X^{2+2^j} + \sum_{0 < j \leq n-1} \bar{a}_{0j} X^{1+2^j} + \sum_{2 \leq i < j \leq s} \bar{a}_{ij} X^{2^i+2^j} + \sum_{0 \leq i \leq s} \bar{b}_i X^{2^i} + \bar{c} \\ \bar{F}_2 &= \sum_{0 \leq i \leq n-1} \bar{a}'_{0i} X^{1+2^i} + \sum_{0 \leq i \leq n-1} \bar{b}'_i X^{2^i} + \sum_{1 \leq i < j \leq s} \bar{a}'_{ij} X^{2^i+2^j} + \bar{c}' \end{aligned}$$

and their relation is

$$\bar{a}_{1j} + \bar{a}'_{0j} = 0, \bar{a}_{0j} + \bar{b}'_j = 0, \quad j > s$$

So $\bar{F}_1, \bar{F}_2$ can have maximum degree $2+2^{n-1}$ and $1+2^{n-1}$ respectively. Although their maximum degree can be this high, their ranks are very small. Their matrices are

$$\begin{pmatrix} 0 & \bar{a}_{01} & \cdots \bar{a}_{0s} \cdots & \bar{a}_{0,n-1} \\ \bar{a}_{01} & 0 & \cdots \bar{a}_{1s} \cdots & \bar{a}_{1,n-1} \\ \vdots & \vdots & \vdots & \\ \bar{a}_{0s} & \bar{a}_{1s} & \cdots 0 & \\ \vdots & \vdots & & \\ \bar{a}_{0,n-1} & \bar{a}_{1,n-1} & & \end{pmatrix}, \quad \begin{pmatrix} 0 & \cdots \bar{a}'_{0s} \cdots & \bar{a}'_{0,n-1} \\ \vdots & \vdots & \\ \bar{a}'_{0s} & \cdots 0 & \\ \vdots & & \\ \bar{a}'_{0,n-1} & & \end{pmatrix}$$

respectively. It is then obvious that their ranks are $\leq s+3$ and ≤ 2 respectively.

Theorem 1. *Let r be the Q-Rank of ZHFE(q,n,D).*

1. *If $q>2$, $D \leq q+2q^s$, then $r \leq s+2$.*
2. *If $q=2$, $2+2^{s-2}+2^{s-1} < D \leq 2+2^{s-1}+2^s$, then $r \leq s+3$.*

As an example, for ZHFE(7,55,105), $105 = q+2q^s$ where $q=7$ and $s=2$, so its Q-Rank $r \leq 4$.

3.3 Recovering Private Key by KS Attack

ZHFE is can be viewed as a multi-HFE with two brunches. BFP's KS attack [BFP13] is applicable to recover $\bar{F}_1, \bar{F}_2$. The first brunch has high degree $q+q^{n-1}$ but small Q-Rank $s+2$ (if $q>2$ and $s+3$ if $q>3$), and the second brunch also has high degree $1+q^{n-1}$ but small Q-Rank $s+2$. Since $\bar{F}_1, \bar{F}_2$ have high degree, it is infeasible to invert any one of them directly. Namely it is insufficient to only recover $\bar{F}_1, \bar{F}_2$, but need to recover Ψ as well. Nevertheless, recovering Ψ can be done easily due the simple relation $\Psi(X, F_1, F_2)$.

More explicitly, we can first apply BFP's KS attack to find two linearly independent G_1, G_2 with rank $s+2$ (if $q>2$ and $s+3$ if $q>3$) and $s+2$. And then recover Ψ by finding the coefficients of Ψ

$$\Psi(X, G_1, G_2) = X(u_1 G_1 + u_2 G_1^q + \cdots + u_n G_1^{q^{n-1}} + v_1 G_2 + v_2 G_2^q \cdots + v_n G_2^{q^{n-1}})$$
$$+X^q(u_{n+1} G_1 + u_{n+2} G_1^q + \cdots + u_{2n} G_1^{q^{n-1}} + v_{n+1} G_2 + v_{n+2} G_2^q + \cdots + v_{2n} G_2^{q^{n-1}})$$

satisfying the degree condition deg $\leq D$. This second step is simply a small linear problem. Hence the complexity of BFP's KS attack is

$$O(n^{2(r+1)}) = O(n^{2s+6}) \quad \text{if } q>2,$$

and

$$O(n^{2(r+1)}) = O(n^{2s+8}) \quad \text{if } q=2$$

for small q.

For example, the security level of ZHFE(7,55,105) is $2^{57.8}$ which is not high enough but also not small. This explains why Porras, Baena and Ding' experiment [PBD14] did not terminate in 10 day when solving the case that $r=4$.

4 Enhancement of ZHFE

In this section, we present an enhancement of ZHFE, abbreviated EZHFE, to repair the problem of low Q-Rank. Our enhancement preserves the advantages of ZHFE and can increase the Q-Rank so that there are practical parameters to have high enough security level. Moreover, it indeed also simplifies the construction of ZHFE and makes its generation of private key very efficient.

4.1 Construction of EZHFE

It is very interesting to design the trapdoor using to high degree HFE polynomials which are related by and inverted by inverting a secret low degree polynomial $\Psi(X, F_1, F_2)$. In the secret $\Psi(X, F_1, F_2)$ of ZHFE, two linear combinations of the Frobenius powers $F_1^{q^i}, F_2^{q^i}$ are multiplied with X and X^q respectively and then added together. This design was intended to make the secret $\Psi(X, F_1, F_2)$ secure. However we find that this design has two drawbacks making it fail. Firstly the two linear combinations of the Frobenius powers are indeed unnecessary in the sense that they are just linear transformations of the core map. It only makes the design complex and distracts the efficiency of generating the private keys. Secondly the two multipliers X and X^q are too simple.

In the following we shall show how to enhance ZHFE so that the original construction of ZHFE is not only simplified but also extended. For convenience, we consider only the case that $q \geq 3$. The case that $q = 2$ is similar. Let

$$F_1(X) = \sum_{0 \leq i \leq j \leq n-1} a_{ij} X^{q^i + q^j}, \quad F_2(X) = \sum_{0 \leq i \leq j \leq n-1} a'_{ij} X^{q^i + q^j}$$

whose coefficients will be determined later. Choose randomly two linearly independent polynomials $L_1(X) = \sum b_i X^{q^i}$ and $L_2(X) = \sum b'_i X^{q^i}$. Let

$$\Phi(X, F_1, F_2) = L_1(X)F_1(X) + L_2(X)F_2(X).$$

Notice that the form of this Φ is much simpler than the Ψ of ZHFE. We want L_1, L_2 and Φ to be of low degree on X,

$$\deg L_1, \deg L_2, \deg \Phi(X, F_1, F_2) \leq D, \tag{5}$$

Let $\Phi_D(X) = \Phi(X, F_1(X), F_2(X))$ be the polynomial satisfying the above degree condition. Since D is relatively small, Berlekamp's algorithm can be used to solve the following equation efficiently

$$\Phi_D(X) = L_1(X)Y_1 + L_2(X)Y_2, \quad Y_1, Y_2 \in \mathbb{K}.$$

These secret F_1, F_2, Ψ will be generated by solving a linear system, which is much simpler than the one of ZHFE, from the condition (5) on the degree.

Before discussing the generation of the secret key, we shall first give our enhancement of ZHFE, EZHFE for short, below.

Public Key $P = T \circ ((\phi \times \phi) \circ (F_1, F_2) \circ \phi^{-1}) \circ S$.
Private Key $S, T, F_1, F_2, \Phi, D, \Phi_D, L_1, L_2$.
Encryption A plaintext $\mathbf{x} \in \mathbb{F}_q^n$ is encrypted by computing $\mathbf{y} = P(\mathbf{x})$.
Decryption A given ciphertext $\mathbf{y}$ is decrypted in the following procedure:

1. Compute $(w_1, \ldots, w_{2n}) = T^{-1}(\mathbf{y})$.
2. Compute $(Y_1, Y_2) = (\phi^{-1}(w_1, \ldots, w_n), \phi^{-1}(w_{n+1}, \ldots, w_{2n}))$.
3. Solve $\Phi_D(X) = L_1(X)Y_1 + L_2(X)Y_2$ by Berlekamp's algorithm.
4. If there are more than one solutions, then check which one satisfies $F_1(X) = Y_1$ and $F_2(X) = Y_2$.
5. For each X satisfying $F_1(X) = Y_1$ and $F_2(X) = Y_2$, compute $\mathbf{x} = S^{-1}(\phi(X))$ which is then a candidate for the plaintext.

Like ZHFE, there may be more than one candidates for the plaintext, so redundant information should be added to help determine the correct plaintext.

It should be remarked that here we consider only homogeneous F_1, F_2 and linearized polynomials L_1, L_2 just for simplicity. One can of course consider inhomogeneous F_1, F_2 and can also replace L_1, L_2 by other polynomials like $\sum X^{q^i+q^j}$ etc. Moreover, the ZHFE scheme is indeed just the simplest case of EZHFE, i.e., the case that $L_1(X) = X$ and $L_2(X) = X^q$.

The major difference between EZHFE and ZHFE is only on the secret third polynomials Φ of EZHFE and Ψ of ZHFE. So EZHFE shares many advantages of ZHFE, and with the same parameters, they have the same public key size, private key size, same encryption and decryption time. We next show how our design of the secret third polynomial Φ can significantly simplify the generation of private key and improve the security level.

4.2 Generation of Private Key and Conclusion on Security

Now we shall deduce the linear systems for generating the private key. Write $A_{ijk} = a_{ij}b_k + a'_{ij}b'_k$. Recall that we assume $q > 2$ (for simplicity). By calculation,

$$\begin{aligned}\Phi(X, F_1, F_2) &= \sum_{0\le k\le n-1} \sum_{0\le i\le j\le n-1} A_{ijk} X^{q^i+q^j+q^k} \\ &= \sum_{0\le i\le n-1} A_{iii} X^{3q^i} + \sum_{0\le i<j<k\le n-1} (A_{ijk} + A_{ikj} + A_{jki}) X^{q^i+q^j+q^k} \\ &+ \sum_{0\le i<j\le n-1} (A_{iij} + A_{iji}) X^{2q^i+q^j} + \sum_{0\le i<j\le n-1} (A_{ijj} + A_{jji}) X^{q^i+2q^j}\end{aligned}$$

For $3q^{s-1} < D \le 3q^s$, $\deg \Psi(X, F_1, F_2) \le D$ if and only if

$$\begin{cases} A_{iii} = 0, & i > s \\ A_{iij} + A_{iji} = 0, & 0 \le i < j, j > s \\ A_{ijj} + A_{jji} = 0, & 0 \le i < j, j > s \\ A_{ijk} + A_{ikj} + A_{jki} = 0, & 0 \le i < j < k, k > s \end{cases} \tag{6}$$

First case: $q > 3$. Since $\deg L_1, \deg L_2 \leq D$, we have $b_k = b'_k = 0$ for $k > s$. Thus

$$A_{ijk} = 0 \text{ for } k > s \text{ if } q > 3.$$

Equation (6) is then simplified as

$$\begin{cases} A_{iji} = a_{ij}b_i + a'_{ij}b'_i = 0, & 0 \leq i \leq s < j \\ A_{ijk} = a_{ij}b_k + a'_{ij}b'_k = 0, & 0 \leq k \leq s < i \leq j \\ A_{ijk} + A_{kji} = a_{ij}b_k + a'_{ij}b'_k + a_{kj}b_i + a'_{kj}b'_i = 0, & 0 \leq i < k \leq s < j \end{cases}$$

Recall that L_1, L_2, i.e., $(b_0, \ldots, b_s)$ and $(b'_0, \ldots, b'_s)$, are linearly independent. So from the above second identity, we have

$$a_{ij} = a'_{ij} = 0 \quad \text{for } s < i \leq j.$$

The rest of a_{ij}, a'_{ij}, $0 \leq i \leq s$ satisfy

$$\begin{cases} a_{ij}b_i + a'_{ij}b'_i = 0, & 0 \leq i \leq s < j \\ a_{ij}b_k + a'_{ij}b'_k + a_{kj}b_i + a'_{kj}b'_i = 0, & 0 \leq i < k \leq s < j \end{cases} \tag{7}$$

So to compute F_1, F_2, we need only solve the simple linear system (7). Moreover, their representing matrices are of shape $\begin{pmatrix} M_1 & M_2 \\ M_2^T & 0 \end{pmatrix}$ where M_1 is $(s+1) \times (s+1)$. If all b_i, b'_i, $0 \leq i \leq s$ are all nonzero, then M_1, M_2 are generally of full rank $s+1$. Hence $r = 2(s+1)$ is a sharp upper bound of the rank of F_1, F_2, as well as the Q-Rank of EZHFE. Therefore the security level of EZHFE against BFP' KS attack is

$$O(n^{2(r+1)}) = O(n^{2(2(s+1)+1)}) = O(n^{4s+6}).$$

Second case $q = 3$. We have $b_k = b'_k = 0$ for $k > s+1$. Then similarly we can deduce a simple linear system for the coefficients of F_1, F_2, and can prove that $r = 2(s+2)$ is a sharp upper bound of the rank of F_1, F_2, as well as the Q-Rank of EZHFE. It should be noted that this upper bound is bigger than the one in the first case. Therefore the security level of EZHFE against BFP' KS attack is

$$O(n^{2(r+1)}) = O(n^{2(2(s+2)+1)}) = O(n^{4s+10}).$$

Theorem 2. *Let $q \geq 3$ and r the Q-Rank of EZHFE(q, n, D).*

1. *If $q > 3$, $3q^{s-1} < D \leq 3q^s$, then $r \leq 2(s+1)$ and the security level of EZHFE against BFP' KS attack is $O(n^{2(r+1)}) = O(n^{4s+6})$.*
2. *If $q = 3$, $q^s < D \leq q^{s+1}$, then $r \leq 2(s+2)$ and the security level of EZHFE against BFP' KS attack is $O(n^{2(r+1)}) = O(n^{4s+10})$.*

Table 2. Comparison of ZHFE and EZHFE

Scheme(q, n, D)	Public Key	Q-Rank	KS Attack	q^n
ZHFE(7,55,105)	66 KB	4	$2^{57.8}$	2^{140}
EZHFE(7,55,105)	66 KB	6	$2^{80.9}$	2^{140}
ZHFE(4,63,132)	66 KB	5	$2^{71.7}$	2^{126}
EZHFE(4,63,132)	66 KB	8	$2^{107.5}$	2^{126}
ZHFE(3,63,81)	66 KB	6	$2^{83.6}$	$2^{99.8}$
EZHFE(3,63,81)	66 KB	10	$2^{131.5}$	$2^{99.8}$

Table 3. Comparison of ZHFE and EZHFE

Scheme(q, n, D)	Public Key	Q-Rank	KS Attack	q^n
ZHFE(7,60,105)	113 KB	4	2^{59}	2^{168}
EZHFE(7,60,105)	113 KB	6	$2^{82.6}$	2^{168}
ZHFE(4,77,132)	119 KB	5	$2^{75.2}$	2^{154}
EZHFE(4,77,132)	119 KB	8	$2^{112.8}$	2^{154}
ZHFE(3,77,81)	119 KB	6	$2^{87.7}$	2^{122}
EZHFE(3,77,81)	119 KB	10	$2^{137.8}$	2^{122}

We compare in Tables 2 and 3 the original ZHFE and EZHFE with various parameters (q, n, D) on their public key size, Q-Rank and security against KS attack. We do not compare other features, such as private key and decryption efficiency, because they are simply the same for the same parameters. Here security and complexity is expressed in terms of the number of the corresponding $\mathbb{F}_q$ operations needed.

In each table, we choose parameters such that their public key size is almost the same and compare their Q-Rank and security level. Compared to Table 2, the parameter n is slightly increased in Table 3 to see how the public key size, Q-Rank and security level change accordingly.

From the two tables, we have several interesting findings:

- ZHFE can still remain secure with bigger n, but with the same parameters, EZHFE has higher Q-Rank and much higher security level than ZHFE.
- If the degree D of different parameter sets are close, then the one with smaller q but bigger n has higher Q-Rank and higher security level than the one with bigger q but smaller n.
- As n increases, the public key size increases fast but the security level increases slowly.
- If the Q-Rank is relatively high, then KS attack seems no better than brute force.

Therefore we can conclude that EZHFE outperforms ZHFE on security with the same parameters, and thus recommend to use EZHFE instead of ZHFE. Moreover we suggest to use small q and moderate n. The parameters we recommend is EZHFE(3,63,81) which has public key size 66 KB and security $2^{99.8}$ ($\mathbb{F}_3$ operations) from Table 2.

4.3 Discussion

Although EZHFE can reach security level higher than 2^{128} with practical parameters, we see it also has restrictions. Namely its Q-Rank is independent on n but dependent on D, and its security is exponential on its Q-Rank but polynomial on n. This means that its security level is not very scalable on n. If very high level of security is desired, then one has to suffer large public key size and slow decryption. These restrictions are due to the structure of ZHFE and EZHFE.

Nevertheless, we find that there is solution to resolve this issue. Notice that adding vinegar variables to HFE can increase its rank and make BFP's attack becomes much less efficient. So we can simply add vinegar variables to F_1, F_2 to increase the Q-Rank of the scheme. Namely we can use high degree HFEv polynomials instead of HFE polynomials in $\Phi(X, F_1, F_2)$. We may call such a scheme EZHFEv which can have much higher security. As the technique of adding vinegar variables is well known in MPKC, we will not expand the resulted scheme EZHFEv in this paper.

5 Conclusion

In this paper, we investigate the security of ZHFE by calculating its Q-Rank explicitly. We then apply the Kipnis-Shamir MinRank attack to formulate a key-recovery attack to ZHFE and use our upper bound of the Q-Rank to give an estimation of the security of ZHFE. We find that ZHFE has a small upper bound for its Q-Rank and so it is not as secure as the authors claimed. Especially we find that the Q-Rank of ZHFE(7,55,105) which was claimed to have security 2^{80}, is at most 4 and its security is $2^{57.8}$. To repair the problem of ZHFE, we propose an enhanced version of ZHFE to have higher Q-Rank and much higher security. Moreover our enhancement is very efficient on generating the private key, unlike the slow generation of private key of ZHFE. We also propose a few practical parameter sets with security 2^{100} or even higher.

References

[BFP13] L. Bettale, J. C. Faugère, and L. Perret. Cryptanalysis of HFE, Multi-HFE and Variants for Odd and Even Characteristic. *Des. Codes Cryptography*, 69(1):1–52, 2013.

[DGS06] Jintai Ding, Jason E. Gower, and Dieter S. Schmidt. *Multivariate public key cryptosystems*, volume 25 of *Advances in Information Security*. Springer, 2006.

[FJ03] J.-C. Faugère and A. Joux. Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems Using Gröbner Bases. In D. Boneh, editor, *CRYPTO 2003*, volume 2729 of *LNCS*, pages 44–60. Springer, 2003.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and intractability: A guide to the theory of NP-completeness*. W. H. Freeman, 1979.

[GJS06] L. Granboulan, A. Joux, and J. Stern. Inverting HFE is Quasipolynomial. In C. Dwork, editor, *CRYPTO 2006*, volume 4117 of *LNCS*, pages 345–356. Springer, 2006.

[KS99] Aviad Kipnis and Adi Shamir. Cryptanalysis of the HFE public key cryptosystem by relinearization. In M. Wiener, editor, *CRYPTO'99*, volume 1666 of *LNCS*, pages 19–30. Springer, 1999.

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

[PBD14] Jaiberth Porras, John Baena, and Jintai Ding. ZHFE, a New Multivariate Public Key Encryption Scheme. In M. Mosca, editor, *PQCrypto 2014*, volume 8772 of *LNCS*, pages 229–245. Springer, 2014.