Nonlinear Phenomena in Complex Systems, vol. **17**, no. 3 (2014), pp. 278 - 283

Analysis of Hidden Field Equations Cryptosystem over Odd-Characteristic Fields

N. G. Kuzmina and E. B. Makhovenko
Saint-Petersburg State Polytechnic University,
29 Politechnicheskaya Str, 195251 Saint-Petersburg, RUSSIA
(Received 15 May, 2014)

Some results on cryptanalysis of Hidden Field Equations (HFE) cryptosystem over odd-characteristic fields are presented. Using of odd-char HFE schemes reduces key generation, encryption and decryption time. Possible attacks are analyzed. HFE parameters, for which cryptosystem is resistant to a given set of attacks, are revealed. Recommendations for HFE parameters choice are provided.

AMS Subject Classification: 11T71, 14G50, 81P68
Keywords: postquantum cryptography, Hidden Field Equations, odd-characteristic fields

1. Introduction

At present public key cryptography is represented mainly by cryptosystems based on factorization and discrete logarithm problems. These problems are vulnerable to a quantum computer: provided such a computer is developed, the problems can be solved with polynomial complexity.

Hash-based, code-based, lattice-based cryptosystems, cryptosystems based on isogenous groups of elliptic curves, as well as cryptosystems using multivariate quadratic polynomials are presumably resistant to a quantum computer.

The latest approach is based on the following observation: the problem of solving multivariate system over a finite field is NP-complete:

Proposition 1 *Let $f_1, \dots, f_m \in K[x_1, \dots, x_n]$ are random quadratic polynomials with the coefficients from a finite field K. Given vector $(y_1, \dots, y_m)$ the problem of solving simultaneous equations $f_1(x_1, \dots, x_n) = y_1, \dots, f_m(x_1, \dots, x_n) = y_m$ is NP-complete [1].*

One of the earliest signature schemes based on such multivariate simultaneous equations is that by Schnorr and Shamir, cracked by Pollard and Schnorr soon after it was published. Henceforth some new schemes were published, but they turned out to be insecure.

Some methods of hiding a secret key structure within the public key (in a set of polynomials) were provided by Patarin. The simplest construction was provided in Oil and Vinegar signature scheme (cracked by Kipnis and Shamir). Some other encryption and signature schemes (Dragon, Little Dragon) provided by Patarin are less secure in comparison with HFE cryptosystems [2].

Up to date, the main focus of researchers was on HFE cryptosystems over the fields of characteristic 2, since computations in these fields are quick in both software and hardware implementation.

In HFE cryptosystem a finite field $\mathbb{F}$ of q elements is used (the recommended parameter is $q = 2$) and the extension $\mathbb{E}$ of degree n of the field $\mathbb{F}$ where n is sufficiently large (the recommended degree is $n = 128$, so that the field $\mathbb{E}$ is of 2^{128} elements). The field extension is defined by irreducible over $\mathbb{F}$ polynomial of degree n. The number of elements of the fields $\mathbb{F}^n$ and $\mathbb{E}$ is the same, so it is possible to give a bijection between them. The map $\mathbb{E} \leftrightarrow \mathbb{F}^n$ can be defined by the basis of n elements $w_1, \dots, w_n$ of the field $\mathbb{E}$: $\sum_{i=1}^{n} t_i w_i \leftrightarrow (x_1, \dots, x_n)$.

HFE cryptosystem secret key is given by affine transformations $S, T : \mathbb{F}^n \to \mathbb{F}^n$ and polynomial $P(x) \in \mathbb{E}[x]$. Formally, the secret key is given as $(S, P, T) \in \mathbb{A}_n(\mathbb{F}) \times \mathbb{E}[x] \times \mathbb{A}_n(\mathbb{F})$.

Public key k is given by polynomials

$(p_1, \ldots, p_n)$ of n variables, determined over the field $\mathbb{F}$. The public key is formed as follows: random polynomial $P(x)$ of one variable over the field $\mathbb{E}$ is generated:

$$P(x) = \sum_{\substack{0 \leqslant i < j < n, \\ q^i + q^j < d}} a_{ij} x^{q^i + q^j} + \sum_{\substack{0 \leqslant i < n, \\ q^i < d}} b_i x^{q^i} + c$$

where d is a small constant, which restricts the degrees of the polynomial $P(x)$ and is usually of the order of several hundreds. The degree of polynomial $P(x)$ is such that it could be efficiently inverted (e.g. with Berlekamp algorithm).

In order to encrypt a message m, one needs to convert it into a vector $(x_1, \ldots, x_n)$ over $\mathbb{F}^n$. Transformation S maps this vector into the vector $x' \in \mathbb{F}^n$. The value $y' = P(x')$ is the element of the field $\mathbb{E}$. Then y' is presented as a vector $(y'_1, \ldots, y'_n)$ and is transformed by T into a vector $(y_1, \ldots, y_n)$. Ciphertext for the message m is the value y with the redundancy value evaluated for it.

To solve the public simultaneous equations, the receiver applies transformation T^{-1} to ciphertext, solves his secret simultaneous equations, interpreting the transformed text as the element of the field $\mathbb{E}$. Then he applies transformation S^{-1} to the components of the solution. An intruder, not knowing secret transformations S and T, is unable to carry out such a procedure. The confusion transformations can be implicitly interpreted over the field $\mathbb{F}$ but not over $\mathbb{E}$. Here it is a priori unknown how n public polynomials can be described by a single variable polynomial over $\mathbb{E}$. Even if such a polynomial exists it can be of exponential number of coefficients and/or a larger degree that makes the inversion problem practically unsolvable.

Nonetheless, in order cryptosystem be sufficiently secure, the requirements to its parameters are such that decryption algorithm is inefficient, and it makes the practical usage of cryptosystem difficult.

2. Analysis of algorithm parameters

One of the main problems in applying HFE cryptosystems is the generation of algorithm parameters, that will provide the fastest encryption/decryption (verification/signature), guarantee cryptographical strength and give optimum values to other algorithm characteristics such as signature length, public key length, time of the key pair generation.

In today HFE cryptosystem implementations (Quartz, Flash, SFlash) the field $\mathbb{F}_2$ is used, that considerably reduces the space of choosing other algorithm parameters. Thus some implementations such as Quartz-513d are not practiced as for providing resistance to Gröbner basis attack, the degree of the secret polynomial must be at least 513, but with this value the expected time of decryption is proved to be excessively large.

The monomials in polynomial $P(x)$ have degrees only of the form $q^a + q^b$ where $a, b \in \mathbb{N}$ so initially there are no other restrictions imposed to degree d. Herewith decreasing of d increases the secret key operations, including evaluation the argument by the value, i.e. solution of the equation $P(x) = y$ towards x, which is the hardest decryption operation.

Currently, to provide cryptographic security, parameter d varies from 129 to 513 for different cryptosystem modifications. These values significantly increase the time of message encryption, so using HFE cryptosystem in practice with $d \geqslant 256$ is hindered, whatever how q and n are chosen (with regard to the restrictions of safety requirements).

Odd-characteristic fields allow using less-degree polynomials for providing the same cryptographic security level. For $\deg(P) = 2$ one can use standard formula of getting roots for second-degree equation for solving $P(x) = y$ equation. Herewith, if $q^n \equiv 3 \pmod 4$ then to find the quadratic root from the field element, one is to raise this element to $(q^n + 1)/4$ power, which is rapidly carried out by successive squaring and multiplying. As q is prime, $q \neq 2$ (q and

d can not be simultaneously equal to 2), the congruence $q^n \equiv 3 \pmod 4$ holds if and only if $q \equiv 3 \pmod 4$.

Increasing the degree n of extension results in increasing the number of elements number of the field $\mathbb{E}$, that, in turn, is followed by the rising field operations time, growing the public key size, and rising key pair generation time. Among other things (in spite of the fact that attacks to HFE algorithm using the subfields of the field $\mathbb{E}$ are unknown yet), it is recommended to choose n being prime, as, on the one hand, it does not impose essential restrictions to the implementation, and, on the other, makes possible to hold the system security, provided such attacks be implemented. Furthermore, in consequence of n rising, the length of the plaintext block being encrypted grows: given q and n it is equal to q^n. For today, parameter n is varied from 80 to 256 for different cryptosystem modifications.

From the theoretical point of view, it is possible to implement HFE cryptosystem for any given finite field $\mathbb{F}$. Nonetheless, the characteristic q of the field $\mathbb{F}$ should not be large for two reasons. Firstly, operations in the finite non-large characteristic field are accomplished easier than in the large ones. The second reason is contained in the way of secret key P is formed: parameter d depends exponentially on the characteristic of the field $\mathbb{F}$, since d depends on $q^a + q^b$ where $a, b \in \mathbb{N}$.

When the value d is fixed ($d = 2$), the field characteristic does not affect the polynomial degree, so this restriction is not considered further.

At the same time, when q decreases, the number of coefficients of public polynomials increases, hence, the size of public key grows as well.

With the restrictions indicated, it may be summed up that while choosing characteristic of the finite field $\mathbb{F}$, the balance should be found between the rate of encryption, on the one hand, and the permissible maximum length of the public key, on the another. A particular task can be of higher priority for different applications. In this case the increase in the encryption rate may be reached at the cost of that in the public key length, and vice versa.

3. Security analysis

Let (x, y) be a pair plaintext/ciphertext, k be the public key of HFE cryptosystem. Possible attacks on a cryptosystem are:

1. Inversion: given y and k find x.
2. Reveal of the inner structure: given public key k compute the secret key (S, P, T).

We will call such attacks *inversion attack* and *structural attack* respectively.

3.1. Inversion attacks

HFE cryptosystem public key k is an algebraic simultaneous equations of not above the second degree. Given ciphertext y and public key $k = (p_1, \ldots, p_n)$, it is possible to form simultaneous equations

$$\begin{cases} p_1(x_1, \ldots, x_n) - y_1 = 0, \\ \ldots \\ p_n(x_1, \ldots, x_n) - y_n = 0, \end{cases}$$

the solutions of which in the field $\mathbb{F}_q$ represent the plaintext. Attacks based on solving of such a system, without knowing the inner cipher structure, are called *algebraic attacks*. They can be classified as follows:

- attacks based on the algorithms of Gröbner basis computation (Buchberger algorithm, F_4 and F_5 algorithms);
- attacks based on linearization (relinearization) technique;
- attacks specified for the secret key of a special form.

The best published implementation of Buchberger algorithm allows solving efficiently only not large systems of equations, the solutions of which are of the size of approximately 20 bits.

The most efficient inversion attacks are those based on F_4 and F_5 algorithms. It is shown in [3] that when solving simultaneous equations over the field $\mathbb{F}_2$ with F_4 and F_5 algorithms, to provide security of at least 2^{128} operations, it is necessary for extension degree to be of at least 134. For such the parameters encryption, decryption and key generation in HFE algorithm will be carried out excessively slow in practice.

Let consider secure parameters for HFE cryptosystem over odd-characteristic fields.

Proposition 2 *For a semiregular system the number of arithmetic operations implemented by F_5 algorithm in the field $\mathbb{F}_q$ is at most*

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

where $\omega < 2.39$ shows the complexity of matrix multiplication; the regularity degree d_{reg} is the number of the smallest nonpositive member of the Gilbert series $S_{m,n} = \frac{\prod_{i=1}^{m}(1-z^{d_i})}{(1-z)^n}$. Here $m = 2$ is the number of equations, n is the number of variables (coincides with the degree of the field extension), d_i is the degree of i-th equation, which for $1 \leqslant i \leqslant n$ is equal to 2, for $n+1 \leqslant i \leqslant 2n$ is equal to q [4]. For semiregular systems F_5 algorithm does not carry out the reduction of monomials, whose degree is less than d_{reg}.

Thus, the complexity of the attack depends on the number of variables and the degree of the system regularity. The proposition above allows generating a large set of parameter pairs n and q for which the complexity of the attack based on F_4 and F_5 algorithm exceeds 2^{80} (e.g. see Table 1).

Table 1. How the degree of regularity of simultaneous equations over the field $\mathbb{F}_q$ depends on the number of variables for F_4 and F_5 algorithms.

n	q	d_{reg}	Attack complexity
11	37	38	2^{80}
13	41	42	2^{93}
23	17	21	2^{95}
11	47	48	2^{88}

Other kinds of inversion attacks are those based on relinearization. Relinearization technique uses equations of the form $(x_a x_b)(x_c x_d) = (x_a x_c)(x_b x_d) = (x_a x_d)(x_b x_c)$, the so-called fourth-degree relinearization, and solves m quadratic equations of n variables for $m \geqslant 0.1n^2$. Examples of the most efficient algorithms of such a kind are those of XL and FXL.

The time of algorithm implementation is approximated by $O\left(\left(n^D/D!\right)^{\omega}\right)$, as its most cumbersome step is the Gaussian elimination of about $n^D/D!$ variables where D is a selectable algorithm parameter. For the base variant of the Gaussian elimination it is $\omega = 3$, for optimized variant it is $\omega = 2.3766$.

Proposition 3 *For overdetermined simultaneous equations, congruence $d_{reg} = \sqrt[4]{q} + \frac{m}{2} - h_{f-1,1}\sqrt{\frac{m}{4}} + O(1)$ is true where $h_{f,1}$ is the supreme null of kth Hermite polynomial and $f = m - n$ [5].*

As in the case of the Gröbner basis, the degree of regularity corresponds to the largest degree of a monomial from those computed during relinearization algorithm. In accordance with this proposition, let us build up the table with the degrees of regularity, as well as complexity of XL algorithm for the different values of n and q (Table 2).

Table 2. How the degree of regularity of simultaneous equations over the field $\mathbb{F}_q$ depends on the number of variables for XL algorithm.

n	q	d_{reg}	Attack complexity
37	53	22	2^{98}
41	53	22	2^{98}
43	53	22	2^{98}
47	53	22	2^{98}
53	53	22	2^{98}
59	53	22	2^{115}
61	53	22	2^{120}
67	53	22	2^{120}
71	53	22	2^{120}

Table 3. How the attack complexity depends on q and n, the complexity is at least 2^{80} operations.

q	n	Attack complexity	q	n	Attack complexity
37	11	80	17	19	81
41	11	84	19	19	83
43	11	85	23	19	91
47	11	88	11	23	84
53	11	92	13	23	87
59	11	95	17	23	95
29	13	81	5	29	84
31	13	84	7	29	981
37	13	90	5	31	86
41	13	93	3	37	84
43	13	95	3	41	88
23	17	86	3	43	94

Table 4. How the attack complexity depends on q and n, the complexity is at least 2^{128} operations.

q	n	Attack complexity	q	n	Attack complexity
59	17	128	23	31	128
61	17	130	29	31	135
67	17	134	31	31	138
71	17	137	37	31	150
73	17	138	13	37	129
79	17	142	17	37	137
83	17	145	19	37	140
89	17	148	23	37	146
53	19	132	11	41	135
59	19	138	13	41	141
61	19	139	17	41	147
67	19	144	11	43	141
71	19	148	13	43	144
73	19	149	7	47	136
41	23	133	11	47	150
43	23	136	5	53	135
47	23	141	5	59	150

3.2. Structural attacks

Given public key, structural attacks find the secret key. In fact, these attacks use peculiarities of a cryptosystem, in contrast to inversion attacks aimed at solving complex mathematical problem (that of solving simultaneous quadratic equations over the finite field). Nowadays the only known attack on HFE cryptosystem is Kipnis and Shamir's one [6]. The attack of Kipnis and Shamir is based on sequential computing secret key parts, namely S and T transformations and polynomial P. Cryptanalysis gives either the real secret key or an equivalent key (i.e. secret key corresponding to the same public key). This attack usually appears to be exponential and inefficient even for the fields of characteristic 2. For odd-characteristic fields the complexity of this attack grows. In comparison with inversion attacks, structural attacks are less efficient for cryptosystems over odd-characteristic fields.

Table 5. How the attack complexity depends on q and n, the complexity is at least 2^{256} operations.

q	n	Attack complexity	q	n	Attack complexity
151	29	256	101	43	284
157	29	260	71	47	257
163	29	263	73	47	260
167	29	265	79	47	270
173	29	268	83	47	276
179	29	271	89	47	284
181	29	272	67	53	267
191	29	277	71	53	274
193	29	278	73	53	278
197	29	280	43	67	270
199	29	281	47	67	276
137	31	259	53	67	283
139	31	261	29	71	261
149	31	267	31	71	264
151	31	268	37	71	275
157	31	272	41	71	281
163	31	275	43	71	284
167	31	278	29	73	268
97	41	272	31	73	270
101	41	276	37	73	282
103	41	278	17	79	258
107	41	283	19	79	265
109	41	285	23	79	274
79	43	256	13	83	257
83	43	262	17	83	271
89	43	270	19	83	277
97	43	280	11	89	265

4. Choosing HFE cryptosystem parameters

Thus in order HFE cryptosystem over odd-characteristic fields be secure over inversion attacks (F_4, F_5, XL, FXL), its parameters should meet the following conditions:

- the field characteristic should be 3 modulo 4: $q \equiv 3 \pmod 4$;
- secret polynomial $P(x)$ should be of the form $ax^2 + bx + c$ where a, b, c are random elements of the field F_{q^n};
- the extension degree n should be prime.

Amongst others, to provide resistance to F_4, F_5, XL, FXL attacks, as well as to Kipnis and Shamir's one, it is recommended to choose the set of parameters in correspondence with tables 3–5 (we consider the complexity of the most effective attack).

References

[1] M.R. Garay, D.S. Johnson. Computers and intractability: a guide to the theory of NP-completeness. Available from: http://books.google.ru/books?id= b4KKQgAACAAJ&dq=%22computers+and+intractability%22&source=gbs_book_other_versions_r&cad=2.

[2] J. Patarin. Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): two new families of asymmetric algorithms. Available from: http://www.cryptosystem.net/hfe.pdf.

[3] M. Bardet, J.-C. Faugère, B. Salvy, B.-Y. Yang. Asymptotic expansion of the degree of regularity for semi-regular systems of equations. Available from: http://www-calfor.lip6.fr/~jcf/Papers/BFS05.pdf.

[4] M. Bardet. Étude des systèmes algébriques surdéterminés. Applications aux codes correcteurs et à la cryptographie (PhD thesis). Available from: http://tel.archives-ouvertes.fr/docs/00/44/96/09/PDF/these_Bardet.pdf.

[5] B.-Y. Yang, J.-M. Chen, N. T. Courtois. On Asymptotic Security Estimates in XL and. Gröbner Bases-Related Algebraic Cryptanalysis. Available from: http://by.iis.sinica.edu.tw/by-publ/recent/xxl3.pdf.

[6] A. Kipnis, A. Shamir. Cryptanalysis of the HFE public key cryptosystem. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.30.3115&rep=rep1&type=pdf.