International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:12, 2007

Generalisation of Kipnis and Shamir Cryptanalysis of the HFE public key cryptosystem

Omessaad Hamdi, Ammar Bouallegue, and Sami Harari

***Abstract* —** In [4], Kipnis and Shamir have cryptanalised a version of HFE of degree 2. In this paper, we describe the generalization of this attack of HFE of degree more than 2. We are based on Fourier Transformation to acheive partially this attack.

***Keywords* —** Public, cryptosystem, cryptanalisis, HFE.

I. Introduction

Public key cryptography depends on a handful of algebraic problems which try to achieve security. The original RSA problem requires large blocs sizes. Other altenatives with small size have been proposed: Elliptic curves and recently the family of quadratic multivariate schemes such as HFE (Hidden Field Equations)[3][1].

The security of this system is based on the difficulty of solving large systems of quadratic multivariate polynomial equations[2]. This problem is NP-hard over any field. The most efficient attack is the one of Kipnis and Shamir that consist in determining the secret key from the public key.

This attack is based on a non standard representation of the HFE. In this paper, we generalise the idea of Kipnis and Shamir to attack partially the HFE cryptosystem of degree 3.

II. HFE scheme

The encoder takes a finite field $K = GF(q)$ of a cardinal q and a characteristic p $(q = p^m)$, L_n is an extension field of degree n. L_n is also a $GF(q)$vector space of dimension n over K or nm over $GF(p)$. Next, he choose a generic polynomial of degree d.

$$p : F_q^n \longrightarrow F_q^n$$

$$x \mapsto \sum_{i,j=0}^{n-1} a_{ij} x^{q^i+q^j} + \sum_{i=0}^{n-1} b_i x^{q^i} + \gamma_0 \qquad a_{ijs},\, b_i \text{ and } \gamma_0 \in L_n.$$

In addition, he choose two secret affine transformations; ie; two invertible matrix $S = \{s_{ij}\}$ and $t = \{t_{ij}\}$ with entries in $GF(q)$ and two constant vectors $s = (s_1, s_2, ..., s_n)$ and $t = (t_1, t_2, ..., t_n)$ and sets

$$t(x) = T.x + t \quad \text{and} \quad s(x) = S.x + s.$$

The attack of Kipnis and Shamir is an attack that consists to guess the secret key from the public key. In this attack, the original HFE scheme is simplified, in particularly, they consider only homogenous polynomial p and linear mappings S and T.

A. Keys

- Secret items: T, p, S
- Public entities: $T(p(S(x))) = (g_1(x_1, \ldots, x_n), \ldots, g_n(x_1, \ldots, x_n)) = g(x)$.

B. Encryption

Let $m = (m_1, \ldots, m_n)$ be the clear to encode bourred by bits of redundancy from a hash function or a linear code.

The ciphering consists to evaluate the message m by the public equations. We obtain therefore $(y_1, \ldots, y_n)$ with $y_i = g_i(m_1, \ldots, m_n)$; $i = 1, \ldots, n$.

The decrypted message is $y = (y_1, y_2, ..., y_n)$.

C. Decryption

The decoder receives the encoded message $y = (y_1, \ldots, y_n) = T(f(S(m)))$. It decrypts:

1) $T^{-1}(y) = f(S(m))$.
2) solves
$$f(z) - a = 0 \qquad (1)$$
$a \in F_{q^n}, a = T^{-1}(y)$
3) Apply S^{-1} to the gotten solution.

The equation (1) can admit more than one solution. The redundancy permits to determine the good solution.

The attacker who hasn't S, T and p can not use ths procedure.

Kipnis and Shamir introduce a new technique to decode:

1) Transform S and T from matrix representations to polynomial representations.
2) Convert the n quadratic polynomials in a matrix representation.
3) Solve the fundamental equation.
4) Use the condition of the rank of the polynomial p to determine T, p, S.

Omessad Hamdi is with SYSCOM Laboratory, ENIT, TUNISIA, and with SIS Laboratory, UTV, FRANCE(e-mail:omessaad.hamdi@laposte.net)
Ammar Bouallegue is with SYSCOM Laboratory, ENIT, TUNISIA(e-mail:ammar.bouallegue@enit.rnu.tn)
Sami Harari is with SIS Laboratory, UTV, FRANCE(e-mail:harari@univ-tln.fr)

III. Polynomial representation of S and T

Lemma 1: Let $A : F_q^n \longrightarrow F_q^n$
$f(x_1,\ldots,x_n) \mapsto \ (y_1,\ldots y_n)$
A is a linear application only if $\exists(a_1,\ldots a_n)$ so that $y = \sum_{i=1}^{n} a_i x^{q^i}$ with $x = \sum_{i=1}^{n} x_i w_i$, $y = \sum_{j=1}^{n} y_j w_j$; $(w_1,\ldots,w_n)$ is a basis of F_q^n.
Proof:
x^{q^i} is linear over F_q^n therefore $\sum_{i=1}^{n} a_i x^{q^i}$ is linear $\forall a_i \in F_q$.
On the other hand, F_q^n is an extension field of the field F_q. $\exists$ an element $\beta \in F_{q^n}$ so that $(\beta, \beta^q, \ldots, \beta^{q^{n-1}})$ is a basis of F_q^n. So, all elements of F_{q^n} can be decomposed in this basis. Thus, $x = \sum_{i=1}^{n} \beta^{q^i} x_i$ and $y = \sum_{i=1}^{n} \beta^{q^i} y_i$; x_i, y_i are elements of F_q. By hypothesis, y is linear in x,

$$y_i = \sum_{j=0}^{n-1} t_{ij} x_j, t_{ij} \in F_q$$
$$y = \sum_{j=0}^{n-1}(\sum_{i=0}^{n-1} t_{ij} x_i)\beta^{q^j} = \sum_{i=0}^{n-1} x_i(\sum_{j=0}^{n-1} t_{ij}\beta^{q^j})$$
$$y = \sum_{i=0}^{n-1} x_i P_i(\beta) \qquad (2)$$

In other hand, $y = \sum_{j=0}^{n-1} a_j x^{q^j} = \sum_{j=0}^{n-1} a_j(\sum_{i=0}^{n-1} x_i \beta^{q^i})^{q^j}$ $= \sum_{i=0}^{n-1} x_i(\sum_{j=0}^{n-1} a_j \beta^{q^{i+j}})$

From (2), it is sufficient to show that $\forall P_i(\beta); i = 0 \cdots n-1, \exists\ a_j\ , j = 0 \cdots n-1$ so that $P_i(\beta) = \sum_{j=10}^{n-1} a_j \beta^{q^{i+j}}$.

The matrix $M = (m_{ij})$, $m_{ij} = trace(\beta^{q^i}\beta^{q^j})$ is regular.
$\forall\ P = (P_0(\beta),\ldots,P_{n-1}(\beta)), \exists R = (R_0(\beta),\ldots,R_{n-1}(\beta)))$ so that

$$RM = P$$

$RM = \sum_{i=0}^{n-1}(R_i m_{ij}) = \sum_{i=0}^{n-1} R_i(\sum_{k=0}^{n-1} \beta^{q^{i+k}}\beta^{q^{j+k}}) = \sum_{k=1}^{n}(\sum_{i=1}^{n} R_i \beta^{q^{i+k}})\beta^{q^{j+k}} = P_j(\beta)$

We choose $a_k = \sum_{i=0}^{n-1} R_i \beta^{i+k}$.

From this lemma, we can represent the two standard applications of the cryptosystem HFE by the polynomial representations.

IV. Univariate representation of a system of multivariate equations

Lemma 2:
let $(P_0,\cdots,P_{n-1})$ a multivariate polynomial system over F_q. $y_j = P_j(x_0,\cdots,x_{n-1})$; $j = 0\ldots n-1$ only if $\exists(a_0, a_2, \cdots, a_{q^{n-1}}) \in F_{q^n}$ so that $y = \sum_{i=0}^{n-1} a_i x^i$ with $x = \sum_{i=0}^{n-1} x_i w_i$, $y = \sum_{i=0}^{n-1} y_i w_i$, $(w_1,\ldots,w_n)$ is F_{q^n} basis.
Proof
$y_j = P_j(x_0,\cdots,x_{n-1})$, $j = 0\ldots n-1$.
$P_j(x_0,\cdots,x_{n-1}) = t_{j,0} + \sum_{i_0=0}^{n-1} t_{j,i_0} x_{i_0} +$
$\sum_{i_0,i_1=0}^{n-1} t_{j,i_0,i_1} x_{i_0} x_{i_1}$
$+ \ldots + \sum_{i_0,i_1,\cdots,i_m}^{n-1} t_{j,i_0,\cdots,i_m} x_{i_0} x_{i_1} \ldots x_{i_m}$
A term of degree m can be written as:
$\sum_{\vec{i}_m=\vec{0}}^{(\vec{n-1})_m} t_{j,\vec{i}_m} x_{\vec{i}_m}$ with $x_{\vec{i}_m} = x_{i_0} \ldots x_{i_{m-1}}$
and $\vec{i}_m = (i_0,\cdots,i_{m-1})$
so $y_j = P_j(x_0,\ldots,x_{n-1}) = \sum_m \sum_{\vec{i}_m=\vec{0}}^{(\vec{n-1})_m} t_{j,\vec{i}_m} x_{\vec{i}_m}$.

Or, $y = \sum_{j=0}^{n-1} y_j \beta^{q^j} = \sum_m \sum_{\vec{i}_m=\vec{0}}^{(\vec{n-1})_m} (\sum_{j=0}^{n-1} t_{j,\vec{i}_m} \beta^{q^j}) x_{\vec{i}_m}$

$= \sum_m \sum_{\vec{i}_m=\vec{0}}^{(\vec{n-1})_m} P_{\vec{i}_m}(\beta) x_{\vec{i}_m}$
Thus, a term of degree m has the following form:
$\sum_{i_0,\cdots,i_m} P_{i_0,\cdots,i_m}(\beta) x_{i_0} x_{i_1} \ldots x_{i_m}$

In other hand, $y = \sum_{l=0}^{q^n-1} a_l x^l$ with $x = \sum_{i=0}^{n-1} x_i \beta^{q^i}$.
$\forall l = 0\ldots n-1$; l can be written *in one way*: $l = \gamma_0 + \gamma_1 q + \ldots + \gamma_{n-1} q^{n-1}$ with $0 <= \gamma < q$
The set of definition of l is $E = \{0,\cdots,q^n-1\}$.
$\forall l \in E$, we associate the vector $\vec{\gamma} = (\gamma_0,\cdots,\gamma_{n-1})$. we divide the whole E equivalence classes:
l $\in E_m$ if $\gamma_0 + \ldots + \gamma_{n-1} = m$; $0 <= m < n(q-1)$.
So, $\forall l \in E, l \in E_m$; $E = E_1 \cup E_2 \cup ... \cup E_{n(q-1)}$.
These classes are disconnected since if $l \in E_i$; $l \notin E_j$,
$\forall l \in E_m; l = \gamma_0 + \gamma_1 q + \ldots + \gamma_{n-1} q^{n-1}$ with $\gamma_0 + \gamma_1 + \cdots + \gamma_{n-1} = m$.
Let's show that there are integers
$j_0,\cdots,j_{m-1}; j_k \in \{0,1,\cdots,n-1\}$ so that
$l = q^{j_0} + \ldots + q^{jm-1}$.
Indeed, $l = \underbrace{q^0 + \ldots + q^0}_{\gamma_0} + \underbrace{q^1 + \ldots + q^1}_{\gamma_1} + \ldots + \underbrace{q^{n-1} + \ldots + q^{n-1}}_{\gamma_{n-1}}$; $\gamma_0 + \ldots + \gamma_{n-1} = m$;
$m < n(q-1)$.
$\forall l \in E_m$; we can associate a vector
$\vec{j}_m = (j_0,\cdots,j_{m-1})$ so that $l = q^{j_0} + q^{j_1} + \ldots + q^{j_{m-1}}$; $j_k \in \{0,\cdots,n-1\}$

$$\sum_{l=0}^{q^n-1} a_l x^l = \sum_{l\in E_0} a_l x^l + \sum_{l\in E_1} a_l x^l + \cdots + \sum_{l\in E_m} a_l x^l.$$

Or, $\sum\limits_{l\in E_m} a_l x^l = \sum\limits_{\vec{j}_m\ =\vec{0}}^{\vec{(}n-1)_m} a_{\vec{j}_m} x^{q^{j_0}}.x^{q^{j_1}} \cdots x^{q^{j_{m-1}}}$.

$x^{q^{j_k}} = (\sum\limits_{i=0}^{n-1} x_i \beta^{q^i})^{q^{j_k}}$
$=\sum_{i=0}^{n-1} x_i \beta^{q^{i+j_k}}$.
$\Rightarrow \sum\limits_{l\in E_l} a_l x^l =$
$\sum\limits_{\vec{j}_m=\vec{0}}^{\vec{(}n-1)} a_{\vec{j}_m}(\sum\limits_{i_0}^{n-1} x_{i_0}\beta^{q^{j_0+i_0}})..(\sum\limits_{i_{m-1}}^{n-1} x_{i_{m-1}}\beta^{q^{j_{m-1}+i_{m-1}}})$
$= \sum\limits_{\vec{i}_m=\vec{0}}^{\vec{(}n-1)} (\sum\limits_{\vec{j}_m=\vec{0}}^{\vec{(}n-1)} a_{\vec{j}_m}\beta^{q^{\vec{i}_m+\vec{j}_m}}) x_{\vec{i}_m}$.
It is necessary to show that $\forall\ P_{\vec{i}_m}(\beta); \vec{i}_m = (i_0, i_1, \cdots, i_{m-1})$, $\exists\ a_{\vec{j}_m}$; $\vec{j}_m = (j_0, j_1, \cdots, j_{m-1})$ so that

$$\sum_{\vec{j}_m} a_{\vec{j}_m}\beta^{q^{\vec{j}_m+\vec{i}_m}} = P_{\vec{i}_m}$$

Lets the matrix $M = m_{ij}$;
$m_{ij} = trace(\beta^{q^i}\beta^{q^j}) = \sum_{k=0}^{n-1}\beta^{q^{i+k}}\beta^{q^{j+k}}$.
The matrix $N = M \otimes M \otimes \ldots \otimes M$ is regular, so, there is a vector $R = \{R_{\vec{k}_m}\}$ so that $R.N = P$.
$\Rightarrow \sum\limits_{\vec{k}_m} R_{\vec{k}_m}.m_{\vec{k}_m,\vec{j}_m} = P_{\vec{j}_m}$
$m_{\vec{k}_m,\vec{j}_m} = (\sum\limits_{j_0}\beta^{q^{j_0+k_0}}.\beta^{q^{i_0+j_0}})\ldots$
$(\sum\limits_{j_m}\beta^{q^{j_m+k_m}}.\beta^{q^{i_m+j_m}}). = \sum\limits_{\vec{j}_m}\beta^{q^{\vec{j}_m+\vec{k}_m}}.\beta^{q^{\vec{j}_m+\vec{i}_m}}$
$\sum\limits_{\vec{k}_m} R_{\vec{k}_m} m_{\vec{k}_m,\vec{j}_m}$ = $\sum_{\vec{k}_m} R_{\vec{k}_m}\sum_{\vec{j}_m}\beta^{q^{\vec{j}_m+\vec{k}_m}}\beta^{q^{\vec{j}_m+\vec{i}_m}}$
we puts $a_{\vec{j}_m} = \sum\limits_{\vec{k}_m} R_{\vec{k}_m}\beta^{q^{\vec{j}_m+\vec{k}_m}}$.

From this last expression, the transformation is feasible for all degrees of the hidden polynomial p provided that it is homogeneous.

A. Example

We consider the quadratic equation system

$$\begin{aligned} y_0 =&\ \ x_1x_2 + x_0x_1 \\ y_1 =&\ \ x_1x_2 + x_0x_2 \\ y_2 =&\ \ x_0x_2 \end{aligned}$$

$$y = \textstyle\sum_i y_i\beta^{q^i} = $$
$$x_0x_1\underbrace{\beta^{q^0}}_{p_{10}} + x_1x_2\underbrace{(\beta^{q^0}+\beta^{q^1})}_{p_{12}} + x_0x_2\underbrace{(\beta^{q^1}+\beta^{q^2})}_{p_{02}}$$

If we choose the normal basis $(\beta^{q^0}, \beta^{q^1}, ..., \beta^{q^n n-1})$ so that: $trace(\beta^{q^i}\beta^{q^j}) = \gamma_{ij}$

The matrix M is I_3 and the matrix N is I_9.
$RN = P \Leftrightarrow R = P$

The transformation in a matrix representation consists in determining the $a_{ij}; i, j = 0, 1, 2$ which verify: $a_{ij} = \sum\limits_{k,l} p_{kl}\beta^{q^{i+j+k+l}}$ which represent the coefficients of the matrix.

V. SHAMIR AND KIPNIS ATTACK

The principle of the crypting consists in applying the transformation g to the ciphered

$$g(x) = T(p(S(x)) \tag{3}$$

From [3],

$$p(S(x)) = xWPW^tx^t; W = w_{ij} = s_{j-i}^{q^i}; p = p_{ij} \tag{4}$$

$$T^{-1}(g(x)) = xG'x^t; \tag{5}$$

$$G' = \sum_{k=0}^{n-1} t_k G^{*k}; G^{*k} = (g_{i-k,j-k}^{q^k}) \tag{6}$$

(5) et (6)$\Rightarrow$

$$G' = WPW^t \tag{7}$$

(7) is the fundamental equation.
The first stage of the attack consists in determining T by resolving (7), than S and finally p. Their hypotheses are:

- $rang(p) = r << n$
- $rang(W) = n$

VI. HFE OF DEGREE 3

We have $T(f(S(x))) = G(x)$;
T is invertible so T^{-1} exists and it has the same form as T.

$$\Rightarrow T^{-1}(G(x)) = P(S(x)) \tag{8}$$

$P(x) = \sum\limits_{ijk=0}^{n-1} x^{q^i+q^j+q^k}$, $S(x) = \sum_k^{n-1} s_k x^{q^k}$.
$\Rightarrow P(S(x)) = \sum\limits_{ijk=0}^{n-1} p_{ijk}(S(x))^{q^i+q^j+q^k}$
$= \sum\limits_{w,k=0}^{n-1}(xWP_kW^tx^t)s_{w-k}^{q^k}x^{q^w}$

In other hand,
$t^-1(g(x)) = = \sum\limits_k (xG'_kx^t)x^{q^k}$
with $G'_k = \sum_{l=0}^{n-1} t_l G_k^{*l}$ and $G_k^{*l} = g_{i-l,j-l,k-l}^{q^l}$
so (9) becomes

$$\sum_{w,k=0}^{n-1}(xWP_kW^tx^t)s_{w-k}^{q^k}x^{q^w} = \sum_k (xG'_kx^t)x^{q^k} \tag{9}$$

There is no obvious matrix representation ?

VII. Fourier Transformation

We tried to use fourier transformations to attack HFE of degree 3.
From (8), we have the following equality:

$$\sum_{uvw}(\sum_{h} t_h g^{q^h}_{u-h,v-h,w-h})x^{q^u+q^v+q^w}$$
$$= \sum_{ijk=0}^{n-1} p_{ijk}(\sum_{u,v,w} s^{q^i}_{u-i}s^{q^j}_{v-j}s^{q^k}_{w-k}x^{q^u+q^v+q^w})$$
$$\forall x \ ; \ \ (\sum_{h} t_h g^{q^h}_{u-h,v-h,w-h}) = \sum_{ijk=0}^{n-1} p_{ijk}(s^{q^i}_{u-i}s^{q^j}_{v-j}s^{q^k}_{w-k})$$

If we permute $i --> j --> k --> i$
and $u --> v --> w --> u$:

$$\Longrightarrow \ (\sum_{h} t_h g^{q^h}_{i-h,j-h,k-h}) = \sum_{ijk=0}^{n-1} p_{ijk}(s^{q^i}_{u-i}s^{q^j}_{v-j}s^{q^k}_{w-k})$$
$$= \sum_{ijk=0}^{n-1} p_{ijk}(s^{q^j}_{v-j}s^{q^k}_{w-k}s^{q^i}_{u-i}) = \sum_{ijk=0}^{n-1} p_{ijk}(s^{q^k}_{w-k}s^{q^i}_{u-i}s^{q^j}_{v-j})$$
$$= H$$

We are in F_{2^n} so, $H + H + H = H$

$$\Longrightarrow \sum_{jki=0}^{n-1}(p_{ijk} + p_{jki} + p_{kij})(\sum_{v,w,u} s^{q^j}_{v-j}s^{q^k}_{w-k}s^{q^i}_{u-i}) = \sum_{ijk}\sum_{h} t_h g^{q^h}_{i-h,j-h,k-h}$$

If we apply fourier transformation :

$$\sum_{ijk=0}^{n-1} \beta_{ijk}(\sum_{v,w,u} s^{q^j}_{v-j}s^{q^k}_{w-k}s^{q^i}_{u-i}X^uY^vZ^w)$$
$$= \sum_{h} t_h(\sum_{ijk} g^{q^h}_{i-h,j-h,k-h}X^iY^jZ^k)$$
$$\Rightarrow \textstyle\sum_{ijk}\beta_{ijk}R_i(X)R_j(Y)R_k(Z) = \sum_h E_h(X,Y,Z)t_h$$
$$\Rightarrow \textstyle\sum_{ijk}\beta_{ijk}R_i(X)R_j(Y)R_k(Z)$$
$$= \textstyle\sum_{abc} X^aY^bZ^cP_{abc}(t_0, t_1, ..., t_{n-1})$$

However,

$$P_{abc}(t_0, t_1, ..., t_{n-1}) = P_{bca}(t_0, t_1, ..., t_{n-1})$$
$$= P_{cab}(t_0, t_1, ..., t_{n-1})$$

Thus, we get $2n$ equations with n variables. the resolution in t_i becomes very simple.

VIII. Conclusion

In this paper we are interressted to the generalization of the attack of Kipnis and Shamir for HFE of degree more than 2. We have introduceda new technique to finish this attack which permits to determine the transformation T. This technique is based on Fourier transformation .

References

[1] Nicolas Courtois, Louis Goubin, Jacques Patarin: Quartz, 128-bit long digital signatures: in cryptographers' Track Rsa Confrence 2001, LNCS 2020, pp 282-297, Springer-Verlag.

[2] Nicolas Courtois:The security of Hidden Field Equations (HFE), Cryptographers Track Rsa Conference 2001,LNCS 2020, pp. 266-281, Springer-Verlag.

[3] PATARIN Jacques:"Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): two new families of Asymmetric Algorithms"; Eurocrypt'96, Springer Verlag, pp. 33-48.

[4] SHAMIR Adi, KIPNIS Aviad:"Cryptanalysis of the HFE public key cryptosystem"; Crypto'99. www.minrank.org hfe.