Attacks to polynomial cryptosystems

Pasqualina Fragneto (pasqualina.fragneto@st.com)
STMicroelectronics Advanced System Technology
Agrate - Milano, Italy

Ilaria Simonetti (simonet@mat.unimi.it)
Department of Mathemathics - University of Milan, Italy

Abstract
In this paper we review some possible attacks to cryptosystems based on the problem of multivariate quadratic equations (MQ). After introducing the MQ problem and sketching schemes to generate private and public keys for these cryptosystems, we present well-known attacks based attacks based on the Gröbner bases computation to solve multivariate systems.

The MQ problem

The MQ problem consists in solving multivariate quadratic equations over finite fields $\mathbb{F}_q$. Let

$$f_1, \ldots, f_m \in \mathbb{F}_q[x_1, \ldots, x_n]$$

be m multivariate quadratic polynomials, that is

$$f_i(x_1, \ldots, x_n) = \sum_{1 \leq j \leq k \leq n} \gamma_{i,j,k} x_j x_k + \sum_{1 \leq j \leq n} \beta_{i,j} x_j + \alpha_i \,,$$

with $i = 1, \ldots, m$ and $\alpha_i, \beta_{i,j}, \gamma_{i,j,k} \in \mathbb{F}_q$, the problem is to find $(a_1, \ldots, a_n) \in (\mathbb{F}_q)^n$ which solves the system

$$\mathcal{A} : \left\{ \begin{array}{rcl} f_1(x_1, \ldots, x_n) & = & 0 \\ \vdots & & \\ f_m(x_1, \ldots, x_n) & = & 0 \end{array} \right.$$

In the generic case, this problem is NP-complete.

MQ cryptosystems

Cryptosystems based on the MQ problem have the same structure in the generation of the keys.

Generation of the key

Private key. Its a triple

$$(S, \mathcal{A}', T)\,,$$

where S and T are two affine transformations over $(\mathbb{F}_q)^n$ and $(\mathbb{F}_q)^m$ respectively, and $\mathcal{A}'$ is a multivariate quadratic system generated by $(f'_1, \ldots, f'_m)$.

Public key. Its computed as composition function of the affine transformations

$$S : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^n \text{ and } T : (\mathbb{F}_q)^m \to (\mathbb{F}_q)^m\,,$$

and the central equation

$$\mathcal{A} : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^m\,,$$

i.e., we have

$$\mathcal{A} = T \circ \mathcal{A}' \circ S\,,$$

which is a multivariate quadratic system. The main difference between MQ schemes lies in their special construction of the central equations $\mathcal{A}'$. The security of MQ cryptosystems lies in the fact that the system obtained by the public key has to be apparently indistinguishable from a random system.

Encryption

Let $X = (x_1, \ldots, x_n) \in (\mathbb{F}_q)^n$ be the plaintext. The encryption step is the same for all MQ schemes: we evaluate the polynomials of the public key $\mathcal{A} = (f_1, \ldots, f_m)$ in X .

Decryption

Let $Y \in (\mathbb{F}_q)^m$ the ciphertext, the receiver can obtain the original message $X \in (\mathbb{F}_q)^n$ because he knows the private key $(S, \mathcal{A}', T)$. Indeed S and T are two bijections, which can be written as matrices. So the receiver is able to compute $S^{?1}$ and $T^{?1}$ inverting the matrices which represent them. Hence, the difficulty lies in inverting $\mathcal{A}'$, and the method to do this depends on the used scheme.

Some MQ cryptosystems

- The Imai-Matsumoto scheme C* [MI88];
- Two variants of C* , introduced by Patarin, C^{*--} [PCG98] and HFE [Pat96]. On these schemes are based two digital signature schemes: Sflash [PCG01a] and Quartz [PCG01b];
- The Moh scheme TTM [Moh99].

Attacks

We analyse attacks to MQ cryptosistems based on algorithms to solve the multivariate quadratic system generated by $\mathcal{A}$.

XL algorithm

The XL algorithm, introduced by Courtois [CKPS00] as an alternative to Gröbner bases computation to solve multivariate quadratic systems, is based on the idea of relinearization of the problem.
Let $\mathcal{A}$ be the system generated by quadratic polynomials f_i over the field $\mathbb{F}_q$, the idea is to multiply the polynomials for all monomials, until we obtain polynomials of degree d_{max} at most, where d_{max} is a positive integer to give as input, to build the Macaulays matrix of degree d_{max} and to perform Gaussian elimination on this matrix.
Let t_k be a monomial of degree k. We denote

$$t_k f = \{ {}_t k f_i \,|\, i = 1, \ldots, m \} .$$

```
Algorithm 1: XL
  Input: Let $\mathcal{A}$ be the system generated by quadratic polynomials
         $f_1, \dots, f_m$, let $>$ be a monomial ordering and let $d_{max}$ be a
         positive integer
  Output: Solution of the system $\mathcal{A}$
    1. For $i = 1, \dots, m$ compute all products $t_k f_i \in I_{d_{max}}$ with
       $0 \leq k \leq d - 2$.
    2. Store polynomial coefficients in a matrix, in which the columns
       represent monomial.
    3. Perform Gaussian elimination on the matrix obtained at the
       step 2.
    4. Solve the univariate equation, corresponding to the last row of the
       matrix obtained at the step 3.
    5. Simplify the equations and repeat the process to find the values
       of the other variables.
```

The monomial ordering which we use has to be such that in the Gaussian elimination, a variable x_i is the last eliminated. Its possible extend XL algorithm to systems of any degree: its sufficient to replace $d_{max} - 2$ with $d_{max} - 1$.

F$_4$ algorithm

Its an algorithm to compute Gröbner basis, due to Faugère [Fau99], which is based on the choice of a suitable subset of the polynomials to reduce in the Buchberger algorithm [Buc65].

Let T be the monomial set. A critical pair of two polynomials $f_i, f_j \in S_n = \mathbb{F}_q[x_1, \dots, x_n]$ is an element of $T^2 \times S_n \times T \times S_n$,

$$pair(f_i, f_j) := (t_{i,j}, t_i, f_i, t_j, f_j)$$

such that $t_{i,j}$ is the least common multiple between leading terms of f_i and f_j . Therefore

$$t_{i,j} = \mathrm{LT}(t_i f_i) = \mathrm{LT}(t_j f_j)\,.$$

The degree of the critical pair $p_{i,j} = pair(f_i, f_j)$ is the degree of $t_{i,j}$.
We define the two projections:

$$\mathrm{Left}(p_{i,j}) = (t_i, f_i) \quad \mathrm{Right}(p_{i,j}) = (t_j, f_j)\,.$$

We give a basic version of the algorithm, which uses the normal strategy: it chooses, in the set of all critical pairs, the subset formed by minimal degree critical pairs.

Algorithm 2: F_4

```
Input: Let $F = \{f_1, \dots, f_m\}$ be a set of generators for the ideal
       $I \subset S_n$ and let $>$ be a monomial ordering
Output: A Gröbner basis $G = \{g_1, \dots, g_t\}$ for $I$ with respect the
        ordering $>$, with $F \subset G$
$G := F$
$\tilde{F}_0 := F$
$d := 0$
$P := \{Pair(f_i, f_j) \,|\, i, j = 1, \dots, m, i \neq j\}$
while $P \neq \emptyset$ do
    $d := d + 1$
    $P_d := \{p_{i,j} \in P \,|\, \deg(p_{i,j}) = \min(\deg(p_{r,s}) \,|\, p_{r,s} \in P)\}$
    $P := P - P_d$
    $L_d := Left(P_d) \cup Right(P_d)$
    $E := \{t \cdot f \,|\, (t, f) \in L_d\}$
    Store the polynomials of $E$ in the matrix $\mathcal{M}$.
    Perform Gaussian elimination to obtain $\tilde{\mathcal{M}}$.
    $\tilde{F}_d :=$ {polynomials corresponding to the rows of $\tilde{\mathcal{M}}$ such that
        the leading terms are different from those of $\mathcal{M}$}
    for $h \in \tilde{F}_d$ do
        $P := P \cup \{Pair(h, f) \,|\, f \in G\}$
        $G := G \cup \{h\}$
return $G$
```

We can describe XL algorithm as F$_4$ algorithm, where the choice of critical pairs that we consider is trivial: we choose all critical pairs. Differently from the original description of XL, the idea is to begin with $d_{max} = 1$ and to iterate the XL algorithm until we obtain the solution, increasing d_{max} of one at every iteration. At each step, the system $\mathcal{A}$ is replaced by the system obtained at the previous iterate.

Matrix-F$_5$ algorithm

The idea of the F$_5$ algorithm, introduced by Faugère [Fau02], is to construct a submatrix $M_{d,m}$ of the Macaulays matrix $\mathcal{M}_{d,m}$ incrementally in d, eliminating rows which are reduced to zero by the relations $f_i f_j = f_j f_i$. Rows are labelled and ordered, therefore that the row (t, f_j) is linear combination of the previous rows if t is the leading term of an element in $\langle f_1, \dots, f_m \rangle$. To apply this criteria, we build the matrix $\tilde{M}_{d,m}$, performing Gaussian elimination on $M_{d,m}$, and the only operation we can do on the i-th row is:

$$\mathrm{row}_i \leftarrow c \times \mathrm{row}_i + c' \times \mathrm{row}_{i-j} \text{ with } 0 < j < i, c, c' \in \mathbb{F}_q \text{ and } c \neq 0\,.$$

Algorithm 3: Matrix-F$_5$

```
Input: Let $F = \{f_1, \dots, f_m\}$ be a set of homogeneous polynomials in
        $S_n$ with degree $d_1 \leq d_2 \dots \leq d_m \leq d_{max}$, with $d_{max}$ fixed.
Output: $G_{d_{max}}$ a $d_{max}$-Gröbner basis for $\langle f_1, \dots, f_m \rangle$.
for $d$ from $d_1$ to $d_{max}$ do
    $M_{d,0}$ := matrix with 0 rows
    for $i$ from 1 to $m$ do
        build $M_{d,i}$ adding to $M_{d,i-1}$ the rows:
        if $d_i = d$ then
            add the row $(1, f_i)$
        else
            for any row $(e, f_i)$ of $\tilde{M}_{d-1,i}$ such that $x_\lambda$ is the biggest
            variable of $e$, add the $n - \lambda + 1$ rows
            $(x_\lambda e, f_i), (x_{\lambda+1} e, f_i), \dots, (x_n e, f_i)$, except rows such that
            $x_{\lambda+k} e$ is a leading term in $\tilde{M}_{d-d_i,i-1}$
        perform Gaussian elimination (without pivots) on $M_{d,m}$ to
        obtain $\tilde{M}_{d,m}$
        Give $G_d$ =
        {polynomials corresponding to the rows of $\tilde{M}_{d,m}$, such that
        the leading coefficients are different from those of $M_{d,m}$}
return $G_{d_{max}} := \bigcup_{d \leq d_{max}} G_d$.
```

If $f_1, \dots, f_m$ is a semi-regular sequence with degree of regularity D_{reg} , there arent any reductions to zero during the F$_5$ algorithm until degree $d = D_{reg} - 1$ [Bar01].

The whole complexity is dominated by the cost of linear algebra on the matrix of biggest degree, that is the matrix of degree D_{reg} if the sequence is semi-regular. So its important to have an evaluation of D_{reg} to be able to estimate the computational cost of F$_5$.

Table of comparison

We give some results of the computational cost of F$_5$, in the case the input is a semi-regular sequence of m polynomials with n variables over the field

$\mathbb{F}_2$ [M.B03]:

Number of equations $m(n)$	Degree of regularity D_{reg}	Size of matrix	Global complexity
$\sim Nn$ $N \geq 1/4$	$\sim D_0 n$ $0 < D_0 \leq 1/4$	$\sim 2^{(D_1/\log 2)^n}$ $0 < D_1 \leq \log \frac{2^3}{\sqrt[4]{3^3}}$	exponential
$n \ll m \ll n^2$	$\sim n^2/8m$	$\sim \frac{n}{8} \frac{\log m/n}{m/n}$	sub-exponential
$\sim Nn^2$	$\sim 1/8N$	$\sim n^{1/8N}$	polynomial

Acknowledgement

The second author would like to thank her two supervisors: M. Sala and the first author.

References

[Bar01] Magali Bardet, *Étude des systèmes algébriques surdéterminés. applications aux codes correcteurs et à la cryptographie*, Ph.D. thesis, University of Paris 6, Paris, France, 2001.

[Buc65] B. Buchberger, *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal*, Ph.D. thesis, Innsbruch, 1965.

[CKPS00] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir, *Efficient algorithms for solving overdefined systems of multivariate polynomial equations*, Advances in cryptology—EUROCRYPT 2000 (Bruges), Lecture Notes in Comput. Sci., vol. 1807, Springer, Berlin, 2000, pp. 392–407. MR MR1772028

[Fau99] Jean-Charles Faugére, *A new efficient algorithm for computing Gröbner bases* (F_4), J. Pure Appl. Algebra **139** (1999), no. 1-3, 61–88, Effective methods in algebraic geometry (Saint-Malo, 1998). MR MR1700538 (2000c:13038)

[Fau02] Jean-Charles Faugère, *A new efficient algorithm for computing Gröbner bases without reduction to zero* (F_5), Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation (New York), ACM, 2002, pp. 75–83 (electronic). MR MR2035234 (2005c:13033)

[M.B03] B.Salvy M.Bardet, J.Faugère, *Complexity of groebner basis computation for semi-regular overdetermined sequences over* $\mathbb{F}_2$ *with solutions in* $\mathbb{F}_2$, Inria Research Report RR-5049, INRIA, INRIA, France, 2003.

[MI88] Tsutomu Matsumoto and Hideki Imai, *Public quadratic polynomial-tuples for efficient signature-verification and message-encryption*, Advances in cryptology—EUROCRYPT '88 (Davos, 1988), Lecture Notes in Comput. Sci., vol. 330, Springer, Berlin, 1988, pp. 419–453. MR MR994679 (90d:94008)

[Moh99] T. Moh, *A public key system with signature and master key functions*, Comm. Algebra **27** (1999), no. 5, 2207–2222. MR MR1683861

[Pat96] Jacques Patarin, *Hidden Field Equation (HFE) and isomorphisms of polynomial (IP). Two new families of asymmetric algorithms*, Advances in cryptology—EUROCRYPT 1999 (Saragossa), Lecture Notes in Comput. Sci., vol. 1070, Springer, Berlin, 1996, pp. 33–48.

[PCG98] Jacques Patarin, Nicolas Courtois, and Louis Goubin, *C^*_{-+} and HM: Variations around two schemes of T. Matsumoto and H. Imai*, Advances in cryptology, Lecture Notes in Comput. Sci., vol. 1514, Springer-Verlag, London, 1998, pp. 35–49.

[PCG01a] ———, *FLASH, a fast multivariate signature algorithm*, Topics in cryptology—CT-RSA 2001 (San Francisco, CA), Lecture Notes in Comput. Sci., vol. 2020, Springer, Berlin, 2001, pp. 298–307. MR MR1907105

[PCG01b] ———, *QUARTZ, 128-bit long digital signatures*, Topics in cryptology—CT-RSA 2001 (San Francisco, CA), Lecture Notes in Comput. Sci., vol. 2020, Springer, Berlin, 2001, pp. 282–297. MR MR1907104