Cryptanalysis of the TRMS Signature Scheme of PKC'05

Luk Bettale, Jean-Charles Faugère and Ludovic Perret

INRIA, Centre Paris-Rocquencourt, SALSA Project
UPMC, Univ Paris 06, LIP6
CNRS, UMR 7606, LIP6
104, avenue du Président Kennedy
75016 Paris, France
luk.bettale@free.fr, jean-charles.faugere@grobner.org,
ludovic.perret@lip6.fr

Abstract. In this paper, we investigate the security of the Tractable Rationale Maps Signature (TRMS) signature scheme [9] proposed at PKC'05. To do so, we present a hybrid approach for solving the algebraic systems naturally arising when mounting a signature-forgery attack. The basic idea is to compute Gröbner bases of several modified systems rather than a Gröbner basis of the initial system. We have been able to provide a precise bound on the (worst-case) complexity of this approach. For that, we have however assumed a technical condition on the systems arising in our attack; namely the systems are *semi-regular* [3, 5]. This claim is supported by experimental evidences. Finally, it turns out that our approach is efficient. We have obtained a complexity bounded from above by 2^{57} to forge a signature on the parameters proposed by the designers of TRMS [9]. This bound can be improved; assuming an access to 2^{16} processors (which is very reasonable), one can actually forge a signature in approximately 51 hours.

1 Introduction

Multivariate Cryptography is the set of all the cryptographic primitives using multivariate polynomials. The use of algebraic systems in cryptography dates back to the mid eighties [15, 29], and was initially motivated by the need for alternatives to number theoretic-based schemes. Indeed, although quite a few problems have been proposed to construct public-key primitives, those effectively used are essentially factorization (e.g. in RSA [33]) and discrete logarithm (e.g. in Diffie-Hellman key-exchange [16]). It has to be noted that multivariate systems enjoy low computational requirements; moreover, such schemes are not concerned with the quantum computer threat, whereas it is well known that number theoretic-based schemes like RSA, DH, or ECDH are [34].

Multivariate cryptography has become a dynamic research area, as reflected by the ever growing number of papers in the most famous cryptographic conferences.

This is mainly due to the fact that an European project (NESSIE[1]) has advised in 2003 to use such a scheme (namely, SFLASH [11]) in the smart-card context. Unfortunately, Dubois, Fouque, Shamir and Stern [14] discovered a sever flaw in the design of SFLASH, leading to an efficient cryptanalysis of this scheme. In this paper, we investigate the security of another multivariate signature scheme, the so-called Tractable Rationale Maps Signature (TRMS) [9].

1.1 Organization of the Paper. Main Results.

After this introduction, the paper is organized as follows. In Section 2, we introduce the main concern of this paper, namely the Tractable Rationale Maps Signature (TRMS) scheme presented at PKC'05 [9]. Note that the situation of this scheme is a bit fuzzy. A cryptanalysis of a preprint/previous version [36] of such scheme has been presented at PKC'05 [25]. However, no attack against the version presented at PKC'05 [9] has been reported so far. In [25], the authors remarked that one can – more or less – split the public-key of [36] in two independent algebraic systems which can be solved efficiently. We tried to mount this attack on the TRMS version of PKC'05 [9] without success. Thus, it makes sense to study the security of [9]. By the way, the authors of [25] also proposed an "improved" version of the XL algorithm, the so-called *linear method*. We will not much detail this point in this paper, but this linear method is actually very similar to the F_5 [19] algorithm in its matrix form [20]. We briefly come back to this point in Section 3. We will explain why the linear method cannot be more efficient than F_5.

In Section 3, we will introduce the necessary mathematical tools (ideals, varieties and Gröbner bases), as well as the algorithmic tools (F_4/F_5), allowing to address the problem of solving algebraic systems. We will give the definition of *semi-regular* sequences which will be useful to provide a precise complexity bound on our attack. The reader already familiar with these notions can skip this part. However, we would like to emphasize that the material contained in this section is important for understanding the behavior of the attack presented in Section 4. By the way, the notion presented in this section will permit to compare F_5 [19] with the linear method of [25].

In Section 4, we present a hybrid approach for solving the algebraic systems arising when attacking TRMS. The basic idea is to compute Gröbner bases of several modified systems rather than one Gröbner basis of the (bigger) initial system. We have been able to provide a precise bound on the (worst-case) complexity of this approach. For that, we have assumed that the systems arising in our attack are semi-regular. This claim is supported by experimental evidences. This approach approach is efficient; we have obtained a complexity bounded from above by 2^{57} (fields operations) to forge a signature on the parameters proposed by the designers of TRMS [9]. This bound can be improved; assuming an access to 2^{16} processors (which is very reasonable), one can actually forge a signature in approximately 51 hours.

[1] https://www.cosic.esat.kuleuven.be/nessie/

2 Tractable Rationale Maps Signature Schemes

To the best of our knowledge, multivariate public-key cryptosystems are mainly constructed from two different one-way functions. The first one, that we only mention for the sake of completeness is as follows. Let $\mathcal{I} = \langle f_1, \ldots, f_u \rangle$ be an ideal of the polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$ ($\mathbb{K}$ is a finite field) then :

$$f_{\mathrm{PC}} : m \in \mathbb{K} \longmapsto e_{\mathcal{I}} + m \in \mathbb{K}[x_1, \ldots, x_n],$$

with $e_{\mathcal{I}}$ a random element of $\mathcal{I}$.

The one-way function f_{PC} gave rise to a family of public-key encryption schemes that are named *Polly Cracker cryptosystems* [23, 27]. The public-key of such systems is an ideal $\mathcal{I} = \langle f_1, \ldots, f_u \rangle \subset \mathbb{K}[x_1, \ldots, x_n]$, and the secret-key (or trapdoor) is a zero $\mathbf{z} \in \mathbb{K}^n$ of $\mathcal{I}$. Although the security study of Polly Cracker-type systems led to interesting mathematical and algorithmic problems, several evidences have been presented showing that those schemes are not suited for the design of secure cryptosystems (for a survey, we refer the reader to [28]). Moreover, such systems suffer from efficiency problems, namely a poor encryption rate and a large public-key size.

From a practical point of view, the most interesting type of one-way function used in multivariate cryptography is based on the evaluation of a set of algebraic polynomials $\mathbf{p} = \big(p_1(x_1, \ldots, x_n), \ldots, p_u(x_1, \ldots, x_n)\big) \in \mathbb{K}[x_1, \ldots, x_n]^u$, namely :

$$f_{\mathrm{MI}} : \mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{K}^n \longmapsto \mathbf{p}(\mathbf{m}) = \big(p_1(\mathbf{m}), \ldots, p_u(\mathbf{m})\big) \in \mathbb{K}^u.$$

Here, the mathematical hard problem associated to this one-way function is :
Polynomial System Solving (PoSSo)
INSTANCE : polynomials $p_1(x_1, \ldots, x_n), \ldots, p_u(x_1, \ldots, x_n)$ of $\mathbb{K}[x_1, \ldots, x_n]$.
QUESTION : Does there exists $(z_1, \ldots, z_n) \in \mathbb{K}^n$ s. t. :

$$p_1(z_1, \ldots, z_n) = 0, \ldots, p_u(z_1, \ldots, z_n) = 0.$$

It is well known that this problem is NP-COMPLETE [24]. Note that PoSSo remains NP-COMPLETE even if we suppose that the input polynomials are quadratics. This restriction is sometimes called MQ [10].

To introduce a trapdoor, we start from a carefully chosen algebraic system :

$$\mathbf{f}(\mathbf{x}) = \big(f_1(x_1, \ldots, x_n), \ldots, f_u(x_1, \ldots, x_n)\big) \in \mathbb{K}[x_1, \ldots, x_n]^u,$$

which is *easy* to solve. That is, for all $\mathbf{c} = (c_1, \ldots, c_u) \in \mathbb{K}^u$, we have an efficient method for describing/computing the zeroes of :

$$f_1(x_1, \ldots, x_n) = c_1, \ldots, f_u(x_1, \ldots, x_n) = c_u.$$

In order to hide the specific structure of $\mathbf{f}$, we usually choose two linear transformations – given by invertible matrices – $(S, U) \in GL_n(\mathbb{K}) \times GL_u(\mathbb{K})$ and set

$$\big(p_1(\mathbf{x}), \ldots, p_u(\mathbf{x})\big) = \big(f_1(\mathbf{x} \cdot S), \ldots, f_u(\mathbf{x} \cdot S)\big) \cdot U,$$

abbreviated by $\mathbf{p}(\mathbf{x}) = \mathbf{f}(\mathbf{x} \cdot S) \cdot U \in \mathbb{K}^u$ to shorten the notation.

The public-key of such systems will be the polynomials of $\mathbf{p}$ and the secret-key is the two matrices $(S, U) \in GL_n(\mathbb{K}) \times GL_u(\mathbb{K})$ and the polynomials of $\mathbf{f}$.

To generate a signature $\mathbf{s} \in \mathbb{K}^n$ of a digest $\mathbf{m} \in \mathbb{K}^u$, we compute $\mathbf{s}' \in \mathbb{K}^n$ such that $\mathbf{f}(\mathbf{s}') = \mathbf{m} \cdot U^{-1}$. This can be done efficiently due to the particular choice of $\mathbf{f}$. Finally, the signature is $\mathbf{s} = \mathbf{s}' \cdot S^{-1}$ since :

$$\mathbf{p}(\mathbf{s}) = \mathbf{f}(\mathbf{s}' \cdot S^{-1} \cdot S) \cdot U = \mathbf{m} \cdot U^{-1} \cdot U = \mathbf{m}.$$

To verify the signature $\mathbf{s} \in \mathbb{K}^n$ of the digest $\mathbf{m} \in \mathbb{K}^u$, we check whether the equality :

$$\text{“}\mathbf{p}(\mathbf{s}) = \mathbf{m}\text{” holds.}$$

We would like to emphasize that most of the multivariate signature schemes proposed so far (e.g. [11, 26, 37]), including TRMS [9], follow this general principle.

The specificity of TRMS lies in the way of constructing the inner polynomials $\mathbf{f}(\mathbf{x}) = \big(f_1(x_1, \ldots, x_n), \ldots, f_u(x_1, \ldots, x_n)\big) \in \mathbb{K}[x_1, \ldots, x_n]^u$. The designers of TRMS propose to use so-called *tractable rational maps*, which are of the following form :

$$\begin{aligned}
f_1 &= r_1(x_1)\\
f_2 &= r_2(x_2) \cdot \frac{g_2(x_1)}{q_2(x_1)} + \frac{h_2(x_1)}{s_2(x_1)}\\
&\vdots\\
f_k &= r_k(x_k) \cdot \frac{g_k(x_1, \ldots, x_{k-1})}{q_k(x_1, \ldots, x_{k-1})} + \frac{h_k(x_1, \ldots, x_{k-1})}{s_k(x_1, \ldots, x_{k-1})}\\
&\vdots\\
f_n &= r_k(x_n) \cdot \frac{g_k(x_1, \ldots, x_{n-1})}{q_k(x_1, \ldots, x_{n-1})} + \frac{h_k(x_1, \ldots, x_{n-1})}{s_k(x_1, \ldots, x_{n-1})}
\end{aligned}$$

where for all $i, 2 \leq i \leq n, g_i, q_i, h_i, s_i$ are polynomials of $\mathbb{K}[x_1, \ldots, x_n]$, and $i, 2 \leq i \leq n, r_i$ is a permutation polynomial on $\mathbb{K}$. Remember that r_i is a univariate polynomial. As explained in [9], tractable rational maps can be explicitly inverted on a well chosen domain. We will not detail this point, as well as how the polynomials g_i, q_i, h_i, s_i and r_i are constructed. This is not relevant for the attack that we will present. We refer the reader to the initial paper [9]. We just mention that we finally obtain quadratic polynomials for the f_is. We quote below the set of parameters recommended by the authors :

- $\mathbb{K} = \mathbb{F}_{2^8}$
- $n = 28$ and $u = 20$

We will show that this set of parameters does not guaranty a sufficient level of security.

3 Gröbner Basics

In order to mount a signature-forgery attack against TRMS, we have to address the problem of solving an algebraic system of equations. To date, Gröbner bases [6, 7] provide the most efficient algorithmic solution for tackling this problem. We introduce here these bases and some of their useful properties (allowing in particular to find the zeroes of an algebraic system). We also describe efficient algorithms permitting to compute Gröbner bases. We will touch here only a restricted aspect of this theory. For a more thorough introduction, we refer the reader to [1, 12].

3.1 Definition – Property

We start by defining two mathematical objects naturally associated to Gröbner bases : *ideals* and *varieties*. We shall call *ideal generated* by $p_1, \ldots, p_u \in \mathbb{K}[x_1, \ldots, x_n]$ the set :

$$\mathcal{I} = \langle p_1, \ldots, p_u \rangle = \left\{ \sum_{k=1}^{u} p_k \cdot h_k : h_1, \ldots, h_k \in \mathbb{K}[x_1, \ldots, x_n] \right\} \subseteq \mathbb{K}[x_1, \ldots, x_n].$$

We will denote by :

$$V_{\mathbb{K}}(\mathcal{I}) = \big\{ \mathbf{z} \in \mathbb{K}^n : p_i(\mathbf{z}) = 0, \text{ for all } i, 1 \leq i \leq u \big\},$$

the *variety associated* to $\mathcal{I}$, i.e. the common zeros – over $\mathbb{K}$ – of $p_1, \ldots, p_u$.

Gröbner bases offer an explicit method for describing varieties. Informally, a Gröbner basis of an ideal $\mathcal{I}$ is a generating set of $\mathcal{I}$ with "good" algorithmic properties. These bases are defined with respect to *monomial ordering*. For instance, the *Lexicographical* (Lex) and *Degree Reverse Lexicographical* (DRL) orderings – which are widely used in practice – are defined as follows :

Definition 1. *Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$. Then:*
– $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \succ_{\mathrm{Lex}} x_1^{\beta_1} \cdots x_n^{\beta_n}$ if the left-most nonzero entry of $\alpha - \beta$ is positive.
– $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \succ_{\mathrm{DRL}} x_1^{\beta_1} \cdots x_n^{\beta_n}$ if $\sum_{i=1}^{n} \alpha_i > \sum_{i=1}^{n} \beta_i$, or $\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i$ and the right-most nonzero entry of $\alpha - \beta$ is negative.

Once a (total) monomial ordering is fixed, we define :

Definition 2. *We shall denote by* $\mathrm{M}(n)$ *the set of all monomials in n variables, and* $\mathrm{M}_d(n)$ *the set of all monomials in n variables of degree $d \geq 0$. We shall call* **total degree** *of a monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ the sum $\sum_{i=1}^{n} \alpha_i$. The* **leading monomial** *of $p \in \mathbb{K}[x_1, \ldots, x_n]$ is the largest monomial (w.r.t. some monomial ordering $\prec$) among the monomials of p. This leading monomial will be denoted by* $\mathrm{LM}(p, \prec)$*. The* **degree** *of p, denoted* $\deg(p)$*, is the total degree of* $\mathrm{LM}(p, \prec)$*.*

We are now in a position to define more precisely Gröbner bases.

Definition 3. *A set of polynomials $G \subset \mathbb{K}[x_1, \ldots, x_n]$ is a* **Gröbner basis** *– w.r.t. a monomial ordering $\prec$ – of an ideal $\mathcal{I} \subseteq \mathbb{K}[x_1, \ldots, x_n]$ if, for all $p \in \mathcal{I}$, there exists $g \in G$ such that* $\mathrm{LM}(g, \prec)$ *divides* $\mathrm{LM}(p, \prec)$.

Gröbner bases computed for a lexicographical ordering (Lex-Gröbner bases) permit to easily describe varieties. A Lex-Gröbner basis of a *zero-dimensional system* (i.e. with a finite number of zeroes over the algebraic closure) is always as follows

$$\{f_1(x_1) = 0, f_2(x_1, x_2) = 0, \ldots, f_{k_2}(x_1, x_2) = 0, \ldots, f_{k_n}(x_1, \ldots, x_n)\}$$

To compute the variety, we simply have to successively eliminate variables by computing zeroes of univariate polynomials and back-substituting the results.

From a practical point of view, computing (directly) a Lex-Gröbner basis is much slower that computing a Gröbner basis w.r.t. another monomial ordering. On the other hand, it is well known that computing degree reverse lexicographical Gröbner bases (DRL-Gröbner bases) is much faster in practice. The FLGM algorithm [17] permits – in the zero-dimensional case – to efficiently solve this issue. This algorithm use the knowledge of a Gröbner basis computed for a given order to construct a Gröbner for another order. The complexity of this algorithm is polynomial in the number of solutions of the ideal considered. This leads to the following strategy for computing the solutions of a zero-dimensional system

$$p_1 = 0, \ldots, p_u = 0.$$

1. Compute a DRL-Gröbner basis G_{DRL} of $\langle p_1, \ldots, p_u \rangle$.
2. Compute a Lex-Gröbner basis of $\langle p_1, \ldots, p_u \rangle$ from G_{DRL} using FGLM.

This approach is sometimes called *zero-dim solving* and is widely used in practice. For instance, this is the default strategy used in the computer algebra system Magma[2] when calling the function **Variety**. In our context, the varieties will usually have only one solution. Thus, the cost of the zero-dim solving is dominated by the cost of computing a DRL-Gröbner basis. We now describe efficient algorithms for performing this task.

3.2 The $\mathbf{F_4/F_5}$ Algorithms

The historical method for computing Gröbner bases is Buchberger's algorithm [6, 7]. Recently, more efficient algorithms have been proposed, namely the F_4 and F_5 algorithms [18, 19]. These algorithms are based on the intensive use of linear algebra techniques. Precisely, F_4 can be viewed as the "gentle" meeting of Buchberger's algorithm and Macaulay's ideas [30]. In short, the arbitrary choices – limiting the practical efficiency of Buchberger's algorithm – are replaced in F_4 by computational strategies related to classical linear algebra problems (mainly the computation of a row echelon form).

[2] http://magma.maths.usyd.edu.au/magma/

In [19], a new criterion (the so-called F_5 criterion) for detecting useless computations has been proposed. It is worth pointing out that Buchberger's algorithm spends 90% of its time to perform these useless computations. Under some regularity conditions, it has been proved that all useless computations can be detected and avoided. A new algorithm, called F_5, has then been developed using this criterion and linear algebra methods. Briefly, F_5 (in its matrix form) constructs incrementally the following matrices in degree d :

$$A_d = \begin{array}{c|cccc} & m_1 \succ & m_2 \succ & m_3 & \ldots \\ \hline t_1 \cdot p_1 & \ldots & \ldots & \ldots & \ldots \\ t_2 \cdot p_2 & \ldots & \ldots & \ldots & \ldots \\ t_3 \cdot p_3 & \ldots & \ldots & \ldots & \ldots \\ \vdots & \ldots & \ldots & \ldots & \ldots \end{array}$$

where the indices of the columns are monomials sorted w.r.t. $\prec$ and the rows are products of some polynomials f_i by some monomials t_j such that $\deg(t_j f_i) \leq d$. In a second step, row echelon forms of theses matrices are computed, i.e.

$$A'_d = \begin{array}{c|cccc} & m_1 & m_2 & m_3 & \ldots \\ \hline t_1 \cdot p_1 & 1 & 0 & 0 & \ldots \\ t_2 \cdot p_2 & 0 & 1 & 0 & \ldots \\ t_3 \cdot p_3 & 0 & 0 & 1 & \ldots \\ \vdots & 0 & 0 & 0 & \ldots \end{array}$$

For d sufficiently large, A'_d contains a Gröbner basis.

In [25], the authors proposed an "improved" version of the XL algorithm [10], the so-called *linear method*. This method is very similar to F_5 [19]. It can be proved [2] that the matrices constructed by F_5, with Lex, are sub-matrices of the matrices generated by the linear method. One can argue that the goal of F_5 and the linear method is not the same. Namely, F_5 computes Gröbner bases whereas the linear method computes varieties. Again, using the same arguments of [2], it can be proved that the linear method constructs intrinsically a Lex-Gröbner basis. As explained previously, we avoid to compute directly Lex Gröbner bases. We prefer to compute a DRL-Gröbner basis, and then use FGLM to obtain the Lex-Gröbner basis. Thus, the practical gain of F_5+FGLM versus the linear method will be even more important. This was already pointed out in [2].

Finally, the main idea of the linear method is to remove linear dependencies induced by trivial relation of the form $f \cdot g - g \cdot f$. This is actually the basic idea of F_5. Note that in F_5 the matrices are constructed "incrementally" to be sure of removing all the trivial linear dependencies. This is not the case for the linear method. To summarize one can say that the linear method is a degraded/devalued version of F_5 using the worst strategy for computing varieties.

We now come back to the complexity of F_5. An important parameter for evaluating this complexity is the *degree of regularity* which is defined as follows :

Definition 4. *We shall call* **degree of regularity** *of homogeneous polynomials* $p_1, \ldots, p_u \in \mathbb{K}[x_1, \ldots, x_n]$*, denoted* $d_{\mathrm{reg}}(p_1, \ldots, p_m)$*, the smallest integer* $d \geq 0$

such that the polynomials of degree d in $\mathcal{I} = \langle p_1, \ldots, p_u \rangle$ generate – as a $\mathbb{K}$ vectorial space – the set of all monomials of degree d in n variables (the number of such monomials $\#\mathrm{M}_d(n)$ is [3] C_{n+d-1}^{d}*). In other words :*

$$\min\left\{d \geq 0 : \dim_{\mathbb{K}}\left(\left\{f \in \mathcal{I} : \deg(f) = d\right\}\right) = \#\mathrm{M}_d(n)\right\}.$$

For non-homogeneous polynomials $p_1, \ldots, p_u \in \mathbb{K}[x_1, \ldots, x_n]$, the degree of regularity is defined by the degree of regularity of the homogeneous components of highest degree of the polynomials $p_1, \ldots, p_u$.

This degree of regularity corresponds to the maximum degree reached during a Gröbner basis computation. The overall complexity of F_5 is dominated by the cost of computing the row echelon form of the last matrix $A_{d_{\mathrm{reg}}}$, leading to a complexity :

$$\mathcal{O}\left((m \cdot \mathrm{C}_{n+d_{\mathrm{reg}}-1}^{d_{\mathrm{reg}}})^{\omega}\right),$$

with $\omega, 2 \leq \omega \leq 3$ being the linear algebra constant.

In general, it is a difficult problem to know *a priori* the degree of regularity. However, for *semi-regular sequences* [3, 5, 4] – that we are going to introduce – the behavior of the degree of regularity is well mastered.

Definition 5. *[3, 5, 4] Let $p_1, \ldots, p_u \in \mathbb{K}[x_1, \ldots, x_n]$ be homogeneous polynomials of degree $d_1, \ldots, d_u$ respectively. This sequence is* **semi-regular** *if :*

– $\langle p_1, \ldots, p_u \rangle \neq \mathbb{K}[x_1, \ldots, x_n]$,
– *for all $i, 1 \leq i \leq u$ and $g \in \mathbb{K}[x_1, \ldots, x_n]$:*

$$\deg(g \cdot p_i) \leq d_{\mathrm{reg}} \text{ et } g \cdot p_i \in \langle p_1, \ldots, p_{i-1} \rangle \Rightarrow g \in \langle p_1, \ldots, p_{i-1} \rangle.$$

We can extend the notion semi-regular sequence to non-homogeneous polynomials by considering the homogeneous components of highest degree of theses polynomials. We mention that the semi-regularity has been introduced by Bardet, Faugère, Salvy and Yang to generalize the notion of regularity [3, 5, 4].

It can be proved that no useless reduction to zero is performed by F_5 on semi-regular (resp. regular) sequences [3, 5, 4, 19] , i.e. all the matrices A_d ($d < d_{\mathrm{reg}}$) generated in F_5 are of full rank. Moreover, the degree of regularity of a semi-regular sequence $(p_1, \ldots, p_u)$ of degree $d_1, \ldots, d_u$ respectively is given [3, 5, 4] by the index of the first non-positive coefficient of :

$$\sum_{k \geq 0} c_k \cdot z^k = \frac{\prod_{i=1}^{m}(1 - z^{d_i})}{(1 - z)^n}.$$

For instance, it has been proved that the degree [3, 5] of regularity of a semi-regular system of $n - 1$ variables and n equations is asymptotically equivalent to $\left\lceil \frac{(n+1)}{2} \right\rceil$. The authors have recovered here a result obtained, with a different technique, by Szanto [35]. For a semi-regular system of n variables and n equations, we obtain a degree of regularity equal to $n + 1$, which is the well-known Macaulay bound. More details on these complexity analyses, and further complexity results can be found in [3, 5, 4].

[3] C_n^d if you consider the field equations.

4 Description of the Attack

In this part, we present our attack against TRMS [9]. Our goal is to forge a valid signature $\mathbf{s}' \in \mathbb{K}^n$ for a given digest $\mathbf{m} = (m_1, \ldots, m_u) \in \mathbb{K}^u$. In other words, we want to find an element of the variety :

$$V_{\mathbb{K}}(p_1 - m_1, \ldots, p_u - m_u) \subseteq \mathbb{K}^n,$$

with $p_1, \ldots, p_u \in \mathbb{K}[x_1, \ldots, x_n]$ the polynomials of a TRMS public-key. We recall that the parameters are $\mathbb{K} = \mathbb{F}_{2^8}, n = 28$ and $u = 20$.

Following the zero dim-solving strategy presented in Section 3, one can directly try to compute this variety. Unfortunately, there is at least two reasons for which such a direct approach cannot be efficient in this context. First, we have explicitly supposed that the field equations are included in the signature-forgery system. In our context, $\mathbb{K}$ is relatively large; leading to field equations of high degree. In particular, the degree of regularity of the system will be at least equal to $\#\mathbb{K}$. Thus, the computation of a Gröbner basis is impossible in practice.

Another limitation is due to the fact that the number of equations (u) is smaller that the number of variables (n). As a consequence, there is at least $(\#\mathbb{K})^{n-u}$ valid solutions to the signature-forgery system. Hence, even if you suppose that you have been able to compute a DRL-Gröbner basis, you will probably not be able to recover efficiently the Lex-Gröbner basis using FGLM.

A natural way to overcome these practical limitations is to randomly specialize (i.e. fix) $n - u$ variables, and remove the field equations. We will have to solve a system having the same number of variables and equations (u). For each specification of the $n - u$ variables, we can always find a solution of the new system yielding to a valid signature. We also mention that the specialized system will have very few solutions in practice. Thus, the cost of computing the variety will be now essentially the cost of computing a Gröbner basis.

The important observation here is that – after having specified $n - u$ variables – the new system will behave like a semi-regular system. We will present latter in this section experimental results supporting this claim. Note that such a behavior has been also observed, in a different context, in [38]. The degree of regularity of a semi-regular system of u variables and equations is equal to $u + 1$. In our context ($u = 20$), this remains out of the scope of the F_5 algorithm.

To decrease this degree of regularity, we can specialize $r \geq 0$ more variables (in addition of the $n - u$ variables already fixed). Thus, we will have to solve a systems of u equations with $u - r$ variables, which behave like semi-regular systems. This allows to decrease the degree of regularity, and thus the complexity of F_5. For instance, the degree of regularity of a semi-regular system of $u - 1$ variables and u equations is approximately equal to $\left\lceil \frac{(u+1)}{2} \right\rceil$. More generally, the degree of regularity is given by the index of the first non-positive coefficient of the series :

$$\frac{\prod_{i=1}^{u}(1 - z^2)}{(1 - z)^{u-r}}.$$

In the following table, we have quoted the degree of regularity observed in our experiments. Namely, the maximum degree reached during F_5 on systems obtained by fixing $n-u+r$ variables ($r \geq 0$) on signature-forgery systems. We have also quoted the theoretical degree of regularity of a semi-regular system of u equations in $u-r$ variables. These experiments strongly suggest that the systems obtained when mounting a specify+solve signature forgery attack against TRMS behave like semi-regular systems.

u	$u-r$	r	d_{reg} (theoretical)	d_{reg} (observed)
20	19	1	11	
20	18	2	9	9
20	17	3	8	8
20	16	4	7	7
20	15	5	6	6

By fixing variables, we obtain a significant gain on the complexity the F_5. On the other hand, as soon as $r > 0$, each specification of the r variables will not necessarily lead to an algebraic system whose set of solutions is not empty . But, we know that there exists a least one guess of the r variables (in practice exactly one) leading to a system whose zeroes allow to construct a valid signature. Thus, we have to perform an exhaustive search on the r new variables. In other words, instead of computing one Gröbner basis of a system of u equations and variables, we compute $(\#\mathbb{K})^r$ Gröbner bases of "easier" systems (u equations with $u-r$ variables). The complexity of this hybrid approach is bounded from above by :

$$\mathcal{O}\left((\#\mathbb{K})^r\left(m\cdot \mathrm{C}_{u+d_{\mathrm{reg}}-1}^{d_{\mathrm{reg}}}\right)^{\omega}\right),$$

with $\omega, 2 \leq \omega \leq 3$ being the linear algebra constant. We have then to find an optimal tradeoff between the cost of F_5 and the number of Gröbner basis that we have to compute.

In the following table, practical results that we have obtained with F_5 when solving systems obtained by fixing $n-u+r$ variables ($r \geq 0$) on signature-forgery systems. We have quoted the experimental complexity of this approach (T) for different values of r (for that, we assumed that the r guesses are correct). We included the timings we obtained with F_5 ($\mathrm{T}_{\mathrm{F}_5}$) for computing one Gröbner basis, and the maximum number $\left((\#\mathbb{K})^r\right)$ of Gröbner bases that we have to compute. We also included the corresponding number of operations (field multiplications) $\mathrm{Nop}_{\mathrm{F}_5}$ performed by F_5 for computing, and the total number N of operations of our attack (i.e. the cost of computing $2^{8\cdot r}$ Gröbner bases). Finally, we have quoted the maximum memory, denoted Mem, used during the Gröbner basis computation. The experimental results have been obtained using a bi-pro Xeon 2.4 Ghz with 64 Gb. of Ram.

u	$u-r$	r	$(\#\mathbb{K})^r$	$\mathrm{T}_{\mathrm{F}_5}$	Mem	$\mathrm{Nop}_{\mathrm{F}_5}$	T
20	18	2	2^{16}	51h	41940 Mo	2^{41}	2^{57}
20	17	3	2^{24}	2h45min.	4402 Mo	2^{37}	2^{61}
20	16	4	2^{32}	626 sec.	912 Mo	2^{34}	2^{66}
20	15	5	2^{40}	46 sec.	368 Mo.	2^{30}	2^{70}

We observe that the optimal choice is for $r = 2$, for which you obtain a complexity bounded from above by 2^{57} to actually forge a signature on the parameters proposed by the designers of TRMS [9]. We also would like to emphasize that this approach is fully parallelizable (each computation of the $(\#\mathbb{K})^r$ Gröbner basis are totally independent). For instance, assuming an access to 2^{16} processors (which is very reasonable), the computation can be done in two days.

By extrapolating – from these experiments – the practical behavior of our approach for $r = 1$, we have estimated that one can forge a signature in approximately in 2^{53} (in terms of fields operations). As the consequence, the parameters of TRMS [9] should be increased to achieve a reasonable level of security. Further works need to be done for finding the optimal set of parameters.

References

1. W.W. Adams and P. Loustaunau. *An Introduction to Gröbner Bases.* Graduate Studies in Mathematics, Vol. 3, AMS, 1994.
2. G. Ars, J.-C. Faugère, H. Imai, M. Kawazoe, and M. Sugita. *Comparison Between XL and Gröbner Basis Algorithms.* Advances in Cryptology – ASIACRYPT 2004, Lecture Notes in Computer Science, vol. 3329, pp. 338-353, 2004.
3. M. Bardet. *Etude des systèmes algébriques surdéterminés. Applications aux codes correcteurs et à la cryptographie.* Thèse de doctorat, Université de Paris VI, 2004.
4. M. Bardet, J-C. Faugère, B. Salvy *On the complexity of Grbner basis computation of semi-regular overdetermined algebraic equations.* In Proc. International Conference on Polynomial System Solving (ICPSS), pp. 71–75, 2004. Available at *http://www-calfor.lip6.fr/ICPSS/papers/43BF/43BF.htm.*
5. M. Bardet, J-C. Faugère, B. Salvy and B-Y. Yang. *Asymptotic Behaviour of the Degree of Regularity of Semi-Regular Polynomial Systems.* In Proc. of MEGA 2005, Eighth International Symposium on Effective Methods in Algebraic Geometry, 2005.
6. B. Buchberger, G.-E. Collins, and R. Loos. *Computer Algebra Symbolic and Algebraic Computation.* Springer-Verlag, second edition, 1982.
7. B. Buchberger. *Gröbner Bases : an Algorithmic Method in Polynomial Ideal Theory.* Recent trends in multidimensional systems theory. Reider ed. Bose, 1985.
8. C. Berbain, H. Gilbert, J. Patarin. *QUAD: A Practical Stream Cipher with Provable Security.* Advances in Cryptology – EUROCRYPT 2006, Lecture Notes in Computer Science, vol. 4004, Springer–Verlag, pp. 109–128, 2006.
9. C.-Y. Chou, Y.-H. Hu, F.-P. Lai, L.-C. Wang, and B.-Y. Yang. *Tractable Rational Map Signature.* International Workshop on Theory and Practice in Public Key Cryptography (PKC'05), Lecture Notes in Computer Science, vol. 3386, Springer–Verlag, pp. 244–257, 2005.
10. N. Courtois, A. Klimov, J. Patarin, and A. Shamir. *Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations.* Advances in Cryptology – EUROCRYPT 2000, Lecture Notes in Computer Science, vol. 1807, Springer–Verlag, pp. 392-407, 2000.

11. N. Courtois, L. Goubin, and J. Patarin. *SFLASH, a Fast Symmetric Signature Scheme for low-cost Smartcards – Primitive Specification and Supporting documentation.* Available at `www.minrank.org/sflash-b-v2.pdf`.
12. D. A. Cox, J.B. Little and D. O'Shea. *Ideals, Varieties, and algorithms: an Introduction to Computational Algebraic Geometry and Commutative algebra.* Undergraduate Texts in Mathematics. Springer-Verlag. New York, 1992.
13. V. Dubois, P.-A. Fouque, and J. Stern. *Cryptanalysis of SFLASH with Slightly Modified Parameters.* Advances in Cryptology – EUROCRYPT 2007, to appear.
14. V. Dubois, P.-A. Fouque, A. Shamir, and J. Stern. *Practical Cryptanalysis of SFLASH.* Advances in Cryptology – CRYPTO 2007.
15. W. Diffie, and H. J. Fell. *Analysis of a Public Key Approach Based on Polynomial Substitution.* Advances in Cryptology – CRYPTO 1985, Lecture Notes in Computer Science, vol. 218, pp. 340–349, 1986.
16. W. Diffie, and M.E. Hellman. *New Directions in Cryptography.* IEEE Transactions on Information Theory, IT–22(6), pp. 644–654, 1976.
17. J. C. Faugère, P. Gianni, D. Lazard, and T. Mora. *Efficient Computation of Zero-Dimensional Gröbner Bases by Change of Ordering.* Journal of Symbolic Computation, 16(4), pp. 329–344, 1993.
18. J.-C. Faugère. *A New Efficient Algorithm for Computing Gröbner Basis:* F_4. Journal of Pure and Applied Algebra, vol. 139, pp. 61–68, 1999.
19. J.-C. Faugère. *A New Efficient Algorithm for Computing Gröbner Basis without Reduction to Zero:* F_5. Proceedings of ISSAC, pp. 75–83. ACM press, July 2002.
20. J.-C. Faugère, and A. Joux. *Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems using Gröbner bases.* Advances in Cryptology - CRYPTO 2003, Lecture Notes in Computer Science, vol. 2729, Springer-Verlag, pp. 44–60, 2003.
21. J.-C. Faugère, and L. Perret. *Polynomial Equivalence Problems: Algorithmic and Theoretical Aspects.* Advances in Cryptology – EUROCRYPT 2006, Lecture Notes in Computer Science, vol. 4004, pp. 30–47, 2006.
22. J.-C. Faugère, and L. Perret. *Cryptanalysis of* $2R^-$ *schemes.* Advances in Cryptology – CRYPTO 2006, Lecture Notes in Computer Science, vol. 4117, pp. 357–372, 2006.
23. M.R. Fellows, N. Koblitz. *Combinatorial cryptosystems galore! Contemporary Math.* **168** 51–61 (1994).
24. M. R. Garey, and D. B. Johnson. *Computers and Intractability. A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.
25. A. Joux, S. Kunz-Jacques, F. Muller, and P.-M. Ricordel. *Cryptanalysis of the Tractable Rational Map Cryptosystem.* International Workshop on Theory and Practice in Public Key Cryptography (PKC'05), Lecture Notes in Computer Science, vol. 3386, Springer–Verlag, pp. 258–274, 2005.
26. A. Kipnis, J. Patarin, and L. Goubin. *Unbalanced Oil and Vinegar Signature Schemes.* Advances in Cryptology – EUROCRYPT 1999, Lecture Notes in Computer Science, vol. 1592 , Springer-Verlag, pp. 206–222,1999.
27. N. Koblitz. *Algebraic Aspects of Cryptography.* Algorithms and Computation in Mathematics, Volume 3, Springer (1998).
28. F. Levy–dit–Vehel, T. Mora, L. Perret, and C. Traverso . *A Survey of Polly Cracker Systems.* To appear.

29. T. Matsumoto, and H. Imai. *Public Quadratic Polynomial-tuples for Efficient Signature-Verification and Message-Encryption.* Advances in Cryptology – EUROCRYPT 1988, Lecture Notes in Computer Science, vol. 330, Springer–Verlag, pp. 419–453, 1988.
30. F. S. Macaulay. *The Algebraic Theory of Modular Systems.* Cambrige University Press, Cambrige, 1916.
31. J. Patarin. *Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): two new families of asymmetric algorithms.* Advances in Cryptology – EUROCRYPT 1996, Lecture Notes in Computer Science, vol. 1070, Springer–Verlag, pp. 33–48, 1996.
32. J. Patarin, N. Courtois, L. Goubin. *QUARTZ, 128-Bit Long Digital Signatures.* Topics in Cryptology - CT-RSA 2001, Lecture Notes in Computer Science, vol. 2020, Springer–Verlag, pp. 282–297, 2001.
33. R. Rivest, A. Shamir and L. Adleman. *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems.* Communications of the ACM, 21(2), pp. 120–126, 1978.
34. P. W. Shor. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer.* SIAM J. Computing 26, pp. 1484-1509 (1997).
35. A. Szanto. *Multivariate subresultants using jouanolous resultant matrices.* Journal of Pure and Applied Algebra, to appear.
36. L. Wang, and F. Chang. *Tractable Rational Map Cryptosystem.* Cryptology ePrint archive, Report 2004/046, available at http://eprint.iacr.org.
37. C. Wolf. *Multivariate Quadratic Polynomials in Public Key Cryptography.* Ph.D. thesis, Katholieke Universiteit Leuven, B. Preneel (supervisor), 156+xxiv pages, November 2005.
38. B.-Y. Yang, J.-M. Chen, and N. T. Courtois. *On Asymptotic Security Estimates in XL and Gröbner Bases-Related Algebraic Cryptanalysis.* In proc. of ICICS 2004, Lecture Notes in Computer Science, vol. 3269, Springer–Verlag, pp. 401413, 2004.