Improved Algorithms for Isomorphisms of Polynomials

Jacques Patarin, Louis Goubin
Bull Smart Cards and Terminals
68 route de Versailles - BP 45
78431 Louveciennes Cedex - France
e-mail : {J.Patarin,L.Goubin}@frlv.bull.fr

Nicolas Courtois
Modélisation et Signal
Université de Toulon et du Var - BP 132
83957 La Garde Cedex - France
e-mail : courtois@univ-tln.fr

Abstract. This paper is about the design of improved algorithms to solve Isomorphisms of Polynomials (IP) problems. These problems were first explicitly related to the problem of finding the secret key of some asymmetric cryptographic algorithms (such as Matsumoto and Imai's C^* scheme of [12], or some variations of Patarin's HFE scheme of [14]). Moreover, in [14], it was shown that IP can be used in order to design an asymmetric authentication or signature scheme in a straightforward way. We also introduce the more general Morphisms of Polynomials problem (MP). As we see in this paper, these problems IP and MP have deep links with famous problems such as the Isomorphism of Graphs problem or the problem of fast multiplication of $n \times n$ matrices.
The complexities of our algorithms for IP are still not polynomial, but they are much more efficient than the previously known algorithms. For example, for the IP problem of finding the two secret matrices of a Matsumoto-Imai C^* scheme over $K = \mathbf{F}_q$, the complexity of our algorithms is $\mathcal{O}(q^{n/2})$ instead of $\mathcal{O}(q^{(n^2)})$ for previous algorithms. (In [13], the C^* scheme was broken, but the secret key was not found). Moreover, we have algorithms to achieve a complexity $\mathcal{O}(q^{\frac{3}{2}n})$ on any system of n quadratic equations with n variables over $K = \mathbf{F}_q$ (not only equations from C^*). We also show that the problem of deciding whether a polynomial isomorphism exists between two sets of equations is <u>not</u> NP-complete (assuming the classical hypothesis about Arthur-Merlin games), but solving IP is at least as difficult as the Graph Isomorphism problem (GI) (and perhaps much more difficult), so that IP is unlikely to be solvable in polynomial time. Moreover, the more general Morphisms of Polynomials problem (MP) is NP-hard. Finally, we suggest some variations of the IP problem that may be particularly convenient for cryptographic use.

Note: An extended version of this paper can be obtained from the authors.

1 Introduction

This paper presents new algorithms for the Isomorphism of Polynomials (IP) problem. IP was explicitly described in [14], where it was shown that if some efficient algorithm exists for IP, then the secret key of some asymmetric cryptosystems (such as the C^* scheme of [12] or some variations of the HFE scheme of [14]) would be found. (The cryptanalysis of C^* given in [13] breaks the scheme without finding the secret key). No polynomial algorithm is known for IP. On the other side, in [14], it was also shown that if no efficient algorithm exists for IP, then this IP problem can be used to design some zero-knowledge authentication schemes. These schemes can also be transformed in order to have an asymmetric signature scheme.

This paper is divided into two parts. In part I, we present different variations of IP and we study a closely related problem, called MP for "Morphisms of Polynomials". We then show that these problems are closely related to other and more famous problems such as the Graph Isomorphism problem, and Fast Matrix Multiplication. In part II, we present our improved algorithms for IP. These algorithms are not polynomial (so the schemes based on IP are not broken), but they are much more efficient than the previously known schemes.

Part I: Presentation and general properties of the IP and MP problems

2 The IP and MP problems

IP was presented in [14]. Let us recall what this problem is, in the particular case of quadratic forms (the problem can be generalized without difficulties to cubic forms, as well as forms of higher degree).

Let u and n be two integers. Let $\mathbf{F}_q$ be a finite field. Let $(\mathcal{A})$ be a public set of u quadratic equations with n variables $a_1, \ldots, a_n$ over the field $\mathbf{F}_q$. We can write these equations as follows:

$$b_k = \sum_i \sum_j \gamma_{ijk} a_i a_j + \sum_i \mu_{ik} a_i + \delta_k \qquad (1 \le k \le u). \qquad (\mathcal{A})$$

Now let s be a bijective and affine transformation of the variables a_i, $1 \le i \le n$, and let t be a bijective and affine transformation of the variables b_k, $1 \le k \le u$.

We denote $s(a_1, \ldots, a_n) = (x_1, \ldots, x_n)$ and $t(b_1, \ldots, b_n) = (y_1, \ldots, y_n)$.

From $(\mathcal{A})$ we obtain another set $(\mathcal{B})$ of k equations that give the y_k values from the x_i values:

$$y_k = \sum_i \sum_j \gamma'_{ijk} x_i x_j + \sum_i \mu'_{ik} x_i + \delta'_k \qquad (1 \le k \le u). \qquad (\mathcal{B})$$

We say that (s, t) is an *isomorphism* from $(\mathcal{A})$ to $(\mathcal{B})$, and we say that $(\mathcal{A})$ and $(\mathcal{B})$ are *isomorphic*.

The IP problem is the following: if $(\mathcal{A})$ and $(\mathcal{B})$ are two public sets of u quadratic equations, and if $(\mathcal{A})$ and $(\mathcal{B})$ are isomorphic, find an isomorphism (s,t) from $(\mathcal{A})$ to $(\mathcal{B})$.

When s and t are not necessary bijective, we call the corresponding problem the "Morphism of Polynomials" problem (MP).

3 Examples

We now present three "toy examples" in order to become more familiar with the IP and MP problems.

Example of IP with two secrets
Let $K = \mathbf{F}_2$ be the field in which all the variables x_0, ..., x_4, y_0, ..., y_4, a_0, ..., a_4, b_0, ..., b_4 lie.

Let:

$$(\mathcal{A}) \begin{cases} b_1 = a_1 + a_1a_5 + a_2a_3 + a_2a_4 + a_3a_4 \\ b_2 = a_3 + a_4 + a_5 + a_1a_2 + a_1a_4 + a_4a_5 \\ b_3 = a_5 + a_1a_2 + a_1a_3 + a_1a_5 + a_2a_3 + a_3a_5 + a_4a_5 \\ b_4 = a_2 + a_3 + a_4 + a_5 + a_1a_5 + a_3a_4 + a_3a_5 \\ b_5 = a_4 + a_1a_3 + a_1a_5 + a_2a_3 + a_2a_4 + a_2a_5 + a_3a_4 + a_3a_5 + a_4a_5 \end{cases}$$

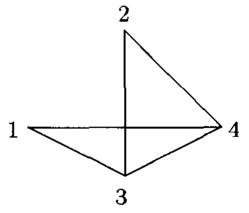
and let:

$$(\mathcal{B}) \begin{cases} y_1 = x_1 + x_3 + x_4 + x_5 + x_1x_2 + x_1x_3 + x_1x_5 + x_2x_4 + x_3x_5 + x_4x_5 \\ y_2 = x_1x_4 + x_1x_5 + x_2x_3 + x_2x_4 + x_1x_5 + x_4x_5 \\ y_3 = x_1 + x_3 + x_4 + x_1x_2 + x_1x_5 + x_2x_3 + x_3x_4 \\ y_4 = x_3 + x_4 + x_1x_2 + x_1x_5 + x_3x_4 + x_3x_5 + x_4x_5 \\ y_5 = x_2 + x_4 + x_5 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_2x_5 \end{cases}$$

The problem is to find two bijective and linear transformations s and t such that $(x_1, ..., x_5) = s(a_1, ..., a_5)$, $(y_1, ..., y_5) = t(b_1, ..., b_5)$ and such that $(\mathcal{A})$ becomes $(\mathcal{B})$ with these transformations. (This is sometimes called an "IP problem with two secrets" since here there are two secret affine transformations to find: s and t).

Note 1: This "toy example" comes from the "toy example" given in [12] for C^*, when the 8 variables are separated in $3+5$ variables, as explained in [13]. The equation $(\mathcal{A})$ come from the equation $b = a^3$ in $\mathbf{F}_{2^5}$.

Note 2: It is possible to show that, when s is found, then t is easy to find. However, an exhaustive search on s would require $2^{n^2} = 2^{25}$ computations in this toy example. Our aim is to improve this complexity.



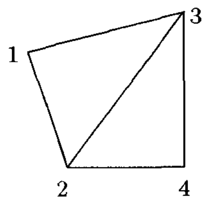


Fig. 1. Graph (I) Graph (II)

Example of IP with one secret

Let us consider the problem of finding an isomorphism between graphs (I) and (II) of figure 1.

Let K be the finite field $\mathbf{F}_2$.

Let $(\mathcal{A})$ and $(\mathcal{B})$ be the two following sets of equations:

$$(\mathcal{A})\begin{cases} y_1 = a_1a_3 + a_1a_4 + a_2a_3 + a_2a_4 + a_3a_4 \\ y_2 = a_1^2 + a_2^2 + a_3^2 + a_4^2 \end{cases}$$

and

$$(\mathcal{B})\begin{cases} y_1 = x_1x_2 + x_1x_3 + x_2x_3 + x_2x_4 + x_3x_4 \\ y_2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 \end{cases}$$

The problem is to find a bijective and linear transformation s such that $(x_1, ..., x_4) = s(a_1, ..., a_4)$, and such that – with this change of variables – the system $(\mathcal{A})$ becomes the system $(\mathcal{B})$ (*i.e.* such that y_1 becomes as written in $(\mathcal{B})$, and y_2 also becomes as written in $(\mathcal{B})$).

If we are able to find all the solutions of this IP problem (this is sometimes called an "IP problem with one secret" since here there is one affine transformation s to find), then some of these solutions will be Graph Isomorphisms from graph (I) to graph (II). This comes from the fact that – in $(\mathcal{A})$ – y_1 was written such that the monomial a_ia_j is in y_1 if and only if the point i is linked with the point j in graph (I). Similarly – in $(\mathcal{B})$ – y_1 was written such that the monomial x_ix_j is in y_1 if and only if the point i is linked with the point j in graph (II). Moreover, y_2 was chosen to be a symmetric polynomial of the variables, so that any graph isomorphism between graph (I) and graph (II) will be a solution to this particular IP problem with one secret. In section 4, we will generalize this construction to study more precisely the links between Graph Isomorphism and IP with one secret.

Example of MP

The variables belong to a ring or to a field K. Let $(\mathcal{A})$ and $(\mathcal{B})$ be the two following sets of equations:

$$(\mathcal{A})\begin{cases} b_1 = a_1 a'_1 \\ b_2 = a_2 a'_2 \\ b_3 = a_3 a'_3 \\ b_4 = a_4 a'_4 \\ b_5 = a_5 a'_5 \\ b_6 = a_6 a'_6 \\ b_7 = a_7 a'_7 \end{cases} \quad \text{and} \quad (\mathcal{B})\begin{cases} y_1 = x_1 x'_1 + x_3 x'_2 \\ y_2 = x_2 x'_1 + x_4 x'_2 \\ y_3 = x_1 x'_3 + x_3 x'_4 \\ y_4 = x_2 x'_3 + x_4 x'_4 \end{cases}$$

The problem is to find two (non bijective) linear transformations s and t such that $(a_1, ..., a_7, a'_1, ..., a'_7) = s(x_1, ..., x_4, x'_1, ..., x'_4)$, $(y_1, ..., y_4) = t(b_1, ..., b_7)$, s of rank 8, t of rank 4, and such that when $(\mathcal{A})$ is satisfied and when these two transformations s and t are done, then $(\mathcal{B})$ is satisfied.

We can notice that the system $(\mathcal{B})$ is the system of equations of a product of two 2×2 matrices:

$$\begin{pmatrix} y_1 & y_3 \\ y_2 & y_4 \end{pmatrix} = \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix} \cdot \begin{pmatrix} x'_1 & x'_3 \\ x'_2 & x'_4 \end{pmatrix}$$

In the system $(\mathcal{A})$, we have exactly seven multiplications, so that solving the MP problem is solving the problem: "How to multiply 2×2 matrices with 7 (instead of 8) multiplications". (The number of additions, subtractions and multiplications with constants of K can be high, but the number of multiplications of terms of the matrices is at most 7).

V. Strassen found in 1969 how to multiply 2×2 matrices with 7 multiplications (cf [16]). His solution is:

$$\begin{cases} a_1 = x_3 - x_4 \\ a_2 = x_1 + x_4 \\ a_3 = x_1 - x_2 \\ a_4 = x_1 + x_3 \\ a_5 = x_1 \\ a_6 = x_4 \\ a_7 = x_2 + x_4 \end{cases} \quad \begin{cases} a'_1 = x'_2 + x'_4 \\ a'_2 = x'_1 + x'_4 \\ a'_3 = x'_1 + x'_3 \\ a'_4 = x'_4 \\ a'_5 = x'_3 - x'_4 \\ a'_6 = x'_2 - x'_1 \\ a'_7 = x'_1 \end{cases} \quad \text{and} \quad \begin{cases} y_1 = b_1 + b_2 - b_4 + b_6 \\ y_2 = b_6 + b_7 \\ y_3 = b_4 + b_5 \\ y_4 = b_2 - b_3 + b_5 - b_7 \end{cases}$$

4 IP with one secret is at least as difficult as Graph Isomorphism

This section is given in appendix 1.

5 MP is NP-hard

In this section, we prove that the Morphism of Polynomials (MP) problem, defined in section 2, is NP-hard for any finite field, and for the rational numbers.

The proof uses some properties of three-dimensional tensors. Let us first recall some basic definitions:

Definitions:

1. A *three-dimensional tensor* is a three-dimensional array $T = (t_{ijk})$ of numbers.
2. It has *rank 1* iff it can be written as the outer product of three vectors (*i.e.* iff there exist three vectors x, y and z such that $m_{ijk} = x_i y_j z_k$ for all indices i, j, k).
3. The rank of a general tensor T is the minimal number of rank 1 tensors T_ν such that $T = \sum\limits_{\nu} T_\nu$.

The following result about the complexity of finding the rank of a three-dimensional tensor was proved by Johan Håstad in [10]:

Theorem 5.1 (Håstad) *Tensor rank is NP-complete for any finite field and NP-hard for the rational numbers.*

Let us now see how this fundamental property can be applied to our MP problem.

Let us suppose we have an algorithm Φ that solves the MP-problem in polynomial time (in particular, this algorithm can be used to know whether there exist a solution or not for some given instance of the MP-problem).

Let $T = (t_{ijk})$ be a $m \times n \times \ell$ (three-dimensional) tensor. It is well known (see for instance [17]) that the rank of T is exactly equal to the minimal number of multiplications needed to compute the following corresponding set of bilinear forms by a bilinear non-commutative algorithm:

$$(\mathcal{B}) \begin{cases} y_1 = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} t_{ij1} x_i x'_j \\ \quad \vdots \\ y_\ell = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} t_{ij\ell} x_i x'_j . \end{cases}$$

Let r ($= rank(T)$) be this minimal value. By definition, r can be thought as the smallest integer u such that two linear transformations $s : (x_1, ..., x_m, x'_1, ..., x'_n) \mapsto (a_1, ..., a_u, a'_1, ..., a'_u)$ and $t : (b_1, ..., b_u) \mapsto (y_1, ..., y_\ell)$ exist, that transform

$$(\mathcal{A}) \begin{cases} b_1 = a_1 \cdot a'_1 \\ \quad \vdots \\ b_u = a_u \cdot a'_u . \end{cases}$$

into $(\mathcal{B})$.

This is a particular instance of the MP problem. For $u = mn$, finding a solution (s, t) is quite easy. Namely, we can define s and t by the following formulas:

$$\forall i,\ 1 \le i \le m,\ \forall j,\ 1 \le j \le n,\ \begin{cases} a_{(i-1)n+j} = x_i \\ a'_{(i-1)n+j} = x'_j . \end{cases}$$

$$\forall k,\ 1 \leq k \leq \ell,\ y_k = \sum_{i=1}^{m}\sum_{j=1}^{n} t_{ijk} b_{(i-1)n+j}.$$

It is therefore easy to build an algorithm that outputs the correct value of r after at most mn calls to the (polynomial) Φ algorithm, so that we have built an algorithm that computes the rank of a three-dimensional tensor in polynomial time. According to the result of Johan Håstad mentioned above, we can thus conclude that the MP-problem is NP-hard for any finite field, and for the rational numbers.

Remarks:

1. MP is clearly a very important problem in mathematics: an efficient algorithm would give the computation of the minimum number of standard multiplications to compute the product of two 3×3, 4×4 or $k \times k$ matrices, for small k, and – from this – improved algorithms for Gaussian reductions and related problems may be found. (The best known asymptotic algorithms are at the present in $\mathcal{O}(n^c)$, where $c \simeq 2.3755$, see [4]. It is also interesting to see what kind of brute-force computations have been done so far to solve such problems, see [9].)
2. The fact that MP is NP-hard, and moreover the fact that MP is a very important problem that seems to be very difficult even with very small parameters, are strong motivations to design cryptographic algorithms based on this problem. From [2] and [6], it is known that any problem of NP can be used to design an asymmetric authentication scheme (with the help of "good" hash functions). (The proof is extendable to design asymmetric signature schemes. again with the help of "good" hash functions). Since MP is in NP, we can apply these general results. However, these very general constructions are not very practical, and it may be more difficult to design efficient schemes from MP than from IP.

6 Deciding IP is not NP-complete

We call "Deciding IP" the problem of finding whether there exist an isomorphism between two sets of multivariate polynomials equations. In this section, we prove that the Deciding IP problem is not NP-complete, under the classical hypothesis that the so-called "polynomial-time hierarchy" does not collapse.

The proof is based on the following general results (see [8] and [3] for proofs):

Theorem 6.1 (Goldwasser, Sipser) *If a problem has a constant-round interactive proof, then it also has a constant-round Arthur-Merlin protocol.*

Theorem 6.2 (Boppana, Håstad, Zachos) *If the complement of a problem Π has an Arthur-Merlin protocol with a constant number of rounds, and if Π is NP-complete, then the polynomial-time hierarchy collapses.*

Note: Arthur-Merlin protocols have been studied by Babai and Moran in [1], but we do not need to go into further details for our proof.

What remains to do is building a constant-round interactive proof for the "Non Isomorphism of Polynomials" (Non-IP) problem. For that purpose, we use techniques discovered by Goldwasser, Micali and Rackoff ([7]) for the analogous "Quadratic Non-Residuosity" problem, and also used by Goldreich, Micali and Wigderson ([6]) for "Graph Non-Isomorphism", and by Petrank and Roth ([15]) for "Code Non-Equivalence".

(Due to the lack of space, details are omitted, but can be found in the extended version of this paper.)

Part II: New algorithms for IP

We use the same notations as in section 2 (for u, n, q, $(\mathcal{A})$ and $(\mathcal{B})$).

In the three next sections, we limit ourselves to the central case of two secrets s and t, and $u = n$. We call this case central because the case $u = 1$ is easily solved as most cases $u > (n+1)(n+2)/2$. As a matter of fact it becomes easy as soon as some $(n+1)(n+2)/2$ equations are independent (as a formal sum of $x_i x_j$ coefficients), so that the u quadratic equations form a generating set. The probability of being so rapidly tends to 1 as u grows. In that case **any** invertible s allow to find at least one t by Gaussian reduction and **every** two generating sets of equations are isomorphic. We also found interesting to separate the affine part of applications s and t, as the linear case seems a bit easier. However, since it is always possible to find the constant terms by exhaustive search in $\mathcal{O}(q^n)$, any algorithm for the IP problem with linear s and t with complexity $\mathcal{O}(q^{\alpha n})$ can obviously be transformed into a general algorithm for affine s and t with complexity $\mathcal{O}(q^{(\alpha+1)n})$.

The present part is divided into sections 7, 8 and 9 (in these three sections, s and t are assumed to be linear). First, we will see a simple algorithm solving most cases of the IP in $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ and using $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ memory. Then we will explain a very different approach which uses only polynomial memory and runs in between $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ and $n^{\mathcal{O}(1)}\mathcal{O}(q^{2n})$ for all IP cases. Finally we will try to combine the ideas of these two attacks in a very powerful, 'birthday paradox'-based attack which to the best of our knowledge runs for every IP in $n^{\mathcal{O}(1)}\mathcal{O}(q^{n/2})$ with $n^{\mathcal{O}(1)}\mathcal{O}(q^{n/2})$ memory. Only the last attack uses the fact that equations given in IP are quadratic forms. The first two can operate on any function of indifferent degree.

7 A simple attack in $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ based on inversion

This attack operates in $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ and uses $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ of memory space used essentially to form a complete table of $\mathcal{A}$ and $\mathcal{B}$ functions.

Let $(\mathcal{A})$ be a randomly chosen quadratic equations set. We studied a probability p_i of a random value b of the $(\mathcal{A})$ form to have i possible corresponding entries a. We made a lot of computer simulations and we have found that the probability distribution of p_i for a randomly chosen set of quadratic equations is the same as for any randomly chosen function, which can be easily shown to be $p_i = \frac{1}{ei!}$.

It allows to divide entries into classes, following their image's inversion degree. The elements of each class can be summed up and the corresponding values for $\mathcal{A}$ and $\mathcal{B}$ must correspond through the linear transformation s. They do not necessarily sum up to 0, since as $p_i \to 0$ we have very small classes and for a random $(\mathcal{A})$ they should not sum to 0.

It is easy to show that if $n \leq q$ we would get more than n equations and we could recover s by Gaussian reduction. With less than n equations we iterate the process as follows:

If we have found one equation $s(x^{(0)}) = a^{(0)}$ we can increase the number of classes in our partition of initial values space. We use the fact that if $s(x) = a$, we also have $s(x + x^{(0)}) = a + a^{(0)}$. Therefore we classify a not only by the inversion degree of its image $\mathcal{A}(a)$ as before, but we will also consider the class of $a + a^{(0)}$. The number of classes will systematically increase.

This attack will work for any sensible non-bijective (> 2 classes) quadratic equations set. Yet we cannot solve a toy example given at the beginning of the article, since it is bijective. In order to solve these cases we hit upon another approach.

8 The "to and fro" attack

We will describe the attack on a toy example, showing directly how it works.

Let $q = 2$ and $n = 5$. We consider

$$(\mathcal{A}) : \begin{cases} b_0 = a_0 + a_0a_4 + a_1a_2 + a_1a_3 + a_2a_3 \\ b_1 = a_2 + a_3 + a_4 + a_0a_1 + a_0a_3 + a_3a_4 \\ b_2 = a_4 + a_0a_1 + a_0a_2 + a_0a_4 + a_1a_2 + a_2a_4 + a_3a_4 \\ b_3 = a_1 + a_2 + a_3 + a_4 + a_0a_4 + a_2a_3 + a_2a_4 \\ b_4 = a_3 + a_0a_2 + a_0a_4 + a_1a_2 + a_1a_3 + a_1a_4 + a_2a_3 + a_2a_4 + a_3a_4 \end{cases}$$

and

$$(\mathcal{B}) : \begin{cases} y_0 = x_0 + x_2 + x_3 + x_4 + x_0x_1 + x_0x_2 + x_0x_4 + x_1x_3 + x_2x_4 + x_3x_4 \\ y_1 = x_0x_3 + x_0x_4 + x_1x_2 + x_1x_3 + x_1x_4 + x_3x_4 \\ y_2 = x_0 + x_2 + x_3 + x_0x_1 + x_0x_4 + x_1x_2 + x_2x_3 \\ y_3 = x_2 + x_3 + x_0x_1 + x_0x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\ y_4 = x_1 + x_3 + x_4 + x_0x_2 + x_0x_3 + x_1x_2 + x_1x_3 + x_1x_4 \end{cases}$$

as functions $\mathbf{F}_{2^5} \to \mathbf{F}_{2^5}$.

The main idea is to say that if we have k equations on s:

$$\begin{cases} s(x^{(1)}) = a^{(1)} \\ \vdots \\ s(x^{(k)}) = a^{(k)} \end{cases}$$

that are linearly independent, we have $q^k - 1$ dependent equations

$$\begin{cases} s(\sum \sim x^{(i)}) = \sum \sim a^{(i)} \\ \qquad\qquad \vdots \end{cases}$$

– where $\sim$ denotes some coefficients. But since $\mathcal{A}$ is generally non-linear we can get from that as much as (or almost) $q^k - 1$ independent equations 'on the other side':

$$\begin{cases} \mathcal{A}\left(s(\sum \sim x^{(i)})\right) = \mathcal{A}\left(\sum \sim a^{(i)}\right) \\ \qquad\qquad \vdots \end{cases}$$

Then we apply t (formally), which changes nothing as to linear independence of these equations:

$$\begin{cases} t\left(\mathcal{A}(s(\sum \sim x^{(i)}))\right) = t\left(\mathcal{A}(\sum \sim a^{(i)})\right) \\ \qquad\qquad \vdots \end{cases}$$

And finally, since $\mathcal{B} = t \circ \mathcal{A} \circ s$ we have about $q^k - 1$ implicit equations on t:

$$\begin{cases} \mathcal{B}(\sum \sim x^{(i)}) = t\left(\mathcal{A}(\sum \sim a^{(i)})\right) \\ \qquad\qquad \vdots \end{cases}$$

We can do that the other direction, start from equations on t and get s in a very similar way, but it is less convenient since we need to compute $\mathcal{A}^{-1}$ or $\mathcal{B}^{-1}$ and sum up all possible solutions. In our example it was easy, since it is bijective.

The whole method is called "to and fro" as it proceeds toing and froing from one side to another and ends when we get n independent equations on s and another n on t.

If $q \neq 2$ we have $q^k - 1 >> k$ and the attack complexity is about $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$ and $n^{\mathcal{O}(1)}\mathcal{O}(q^{2n})$ at most to find initial 1 or 2 equations.

If $q = 2$ it is $n^{\mathcal{O}(1)}\mathcal{O}(q^{2n})$ in most cases (no exceptions are known) since we can start with 2 equations.

Let us see how it works on our example:

We start with the two following equations:

$$\begin{cases} s(1) = 1 \\ s(2) = 7 \end{cases}$$

Our equation set $(\mathcal{A})$ comes from $b = a^3$ in the $\mathbf{F}_{2^n}$, $n = 5$ field, and it has $n2^n$ automorphisms of the form:

$$\begin{cases} s : x \mapsto \alpha x^{1+2^{\beta}} \\ t : x \mapsto \alpha^{-1-2^{\theta}} x^{2^{-\beta}} \end{cases} \text{with } \alpha \neq 0,\ \beta \in \{0, 1, \ldots, n-1\}.$$

Therefore there are at least $n2^n$ solutions s and t (in fact we have found there are no more). Thus the probability of our 2 equations being right is $\frac{n2^n}{2^{2n}} = \frac{n}{2^n}$

and the whole attack complexity is only $n^{\mathcal{O}(1)}\mathcal{O}(q^n)$. Moreover, for $n = 5$ one gets $\frac{n}{2^n} \sim 0.15$ and we admit to have been lucky choosing two starting equations right. Then we get 3 dependent equations:

$$\begin{cases} s(1) = 1 \\ s(2) = 7 \\ s(3) = 6 \end{cases}$$

as – in $\mathbf{F}_{32}$ – $1 + 2 = 3$ and $1 + 7 = 6$.

Then we use the table to 'transfer' equations on the other side:

$$\begin{cases} t(1) = 5 \\ t(4) = 16 \\ t(23) = 24 \end{cases}$$

For example with $s(2) = 7$, we get $\mathcal{A}(7) = 4$ and $\mathcal{B}(2) = 16$ and all that implies that $t(4) = 16$.

Now we combine them again to get $2^3 - 1 = 7$ dependent equations:

$$\begin{cases} t(1) = 5 \\ t(4) = 16 \\ t(5) = 21 \\ t(18) = 13 \\ t(19) = 8 \\ t(22) = 29 \\ t(23) = 24 \end{cases}$$

And using $\mathcal{A}^{-1}$ and $\mathcal{B}^{-1}$ we get 7 equations on s:

$$\begin{cases} s(1) = 1 \\ s(2) = 7 \\ s(3) = 6 \\ s(4) = 17 \\ s(8) = 18 \\ s(20) = 14 \\ s(30) = 27 \end{cases}$$

6 from those 7 equation are actually independent, and yet it is enough to recover s by Gaussian reduction. Similarly we get t and verify the correctness of our solution.

$$(a_0, a_1, a_2, a_3, a_4) = \begin{bmatrix} 1\,1\,1\,0\,1 \\ 0\,1\,0\,1\,1 \\ 0\,1\,0\,0\,1 \\ 0\,0\,0\,0\,1 \\ 0\,0\,1\,1\,1 \end{bmatrix} (x_0, x_1, x_2, x_3, x_4)$$

$$(y_0, y_1, y_2, y_3, y_4) = \begin{bmatrix} 1\,1\,0\,0\,0 \\ 0\,1\,0\,1\,1 \\ 1\,0\,0\,0\,1 \\ 0\,0\,0\,1\,1 \\ 0\,1\,1\,1\,1 \end{bmatrix} (b_0, b_1, b_2, b_3, b_4)$$

9 Combined power attack

In order to further improve our attacks we found out that it would be nice to have a function which we call "boosting function" with the following properties: on an entry x we compute in polynomial time $x' = F(x)$ such that F is preserved by the isomorphism of polynomials:

If $s(x) = a,\ F_{\mathcal{B}}(x) = x'$ *and* $F_{\mathcal{A}}(a) = a'$, *then* $s(x') = a'$.

However:

Theorem 9.1 *There is no non-trivial boosting function which works for all the quadratic equations sets.*

Proof: The "to and fro" attacks will almost always, apart from improbable cases, allow to recover in a unique, deterministic way, the entire s and t starting from 2 equations on s. Yet, a non-trivial boosting function would allow, with some non-zero probability, to do the same starting with only one equation.

Now, in the very precise equations set derived from $b = a^3$ it is not possible. From the structure of $b = a^3$ automorphisms shown in the last chapter we can easily see that there are at least n possible solutions s and t satisfying any given equation on s.

Therefore we must look for less powerful functions and in some way relax its axioms. We have found few ways of doing so:

1. We may consider more general functions, dealing with different sets of information on entries and results of a function.
2. Moreover we may use different boosting functions depending on the particularities of $\mathcal{A}$.
3. We ask the condition of $x' = F(x)$ being preserved by the isomorphism to be true only with a certain noticeable probability.
4. We can have a function F probabilistic.
5. It is enough to suppose that $x' = F(x)$ is preserved no more by **all**, but by at least by **one** possible isomorphism that satisfies the initial data.

The last two additions are important, since they are those which do demolish our proof of non-existence of boosting function.

We call a boosting function or "weak boosting function" a function which satisfies 1.-5., and the previous definition is referred to as a "strong boosting function".

The combined power attack This attack is designed to solve **all** the IP cases, $u = n$ and with linear s and t transformations, with the complexity $n^{\mathcal{O}(1)}q^{n/2}$ and using $n^{\mathcal{O}(1)}q^{n/2}$ of space.

We do not know any example of quadratic equations for which all the three variations of the attack would fail altogether. The starting point consists in picking at random $(\log n)^{\mathcal{O}(1)}q^{n/2}$ entries $z^{(i)}$ and $(\log n)^{\mathcal{O}(1)}q^{n/2}$ entries $c^{(i)}$ (the same entries can be used). Thus by the well known 'birthday principle' there are at least $(\log n)^{\mathcal{O}(1)}$ collisions such that $s(z^{(i)}) = c^{(j)}$. Then the "boosting functions" can be used to generalize the "to and fro" attack (details are available in the extended version of this paper).

10 Suggestions of IP variations

This section is given in appendix 2.

11 Conclusion

In this paper, we have shown that the MP and IP problems are not only related to the problem of finding the secret key of esoteric cryptographic algorithms, but also to very general problems such as Graph Isomorphism and Fast Matrix Multiplication. From their definitions, MP and IP look very similar but, as we have seen in this paper, Deciding MP is NP-complete whereas Deciding IP is not NP-complete (assuming the usual hypothesis about Arthur-Merlin games).

If $\mathcal{O}(n^2)$ is the length of the secrets to be found, then our new algorithms for IP ("with two secrets") have a complexity in $\mathcal{O}(q^n)$ or $\mathcal{O}(q^{n/2})$ when s and t are linear (so at most $\mathcal{O}(q^{\frac{3}{2}n})$ when s and t are affine). This is not polynomial, but this is clearly much better than the complexity $\mathcal{O}(q^{(n^2)})$ of the previous algorithms. For example, for the toy example given in [12], we have found the secret key with only about 7 computations (!), instead of 2^{25} for previously known algorithms (or 2^{50} for exhaustive search on the two secret matrices s and t). This means that the schemes based on IP problems with two secrets (such as some the schemes described in [14]) should have larger values of the parameters than expected for security. However, there are a lot of possible variations for IP and the choice of the variations and the choice of the parameters can still be done in order to have both efficient asymmetric schemes and no practical attacks.

Acknowledgements: We want to thank Jean-Pierre Seifert for pointing out references [10] and [15] to us, and also Henri Gilbert and Frédéric Cherbonnier for useful comments.

References

1. László Babai, Shlomo Moran, *Arthur-Merlin games: A randomized proof system, and a hierarchy of complexity classes*, JCSS, vol. 36, 1988, pp. 254-276.

2. Manuel Blum, *How to prove a theorem so no one else can claim it*, Proceeedings of the International Congress of Mathematics, Berkeley CA, 1986, pp. 1444-1451.

3. Ravi B. Boppana, Johan Håstad, Stathis Zachos, *Does co-NP have short interactive proofs*, Information Proc. Letters, vol. 25, 1987, pp. 127-132.

4. Don Coppersmith, Shmuel Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Computation (1990), **9**, pp. 251-280.

5. Scott Fortin, *The Graph Isomorphism Problem*, Technical Report 93-20, University of Alberta, Edmonton, Alberta, Canada, July 1996. This paper is available at ftp://ftp.cs.ualberta.ca/pub/TechReports/1996/TR96-20/TR96-20.ps.gz

6. Oded Goldreich, Silvio Micali, Avi Wigderson, *Proofs that yield nothing but their validity and a methodology of cryptographic protocol design*, Journal of the ACM, v. 38, n. 1, Jul. 1991, pp. 691-729.

7. Shafi Goldwasser, Silvio Micali, Charles Rackoff, *The knowledge complexity of interactive proofs*, SIAM J. Comput., vol. 18, 1989, pp. 186-208.

8. Shafi Goldwasser, Michael Sipser, *Private coins vs. public coins in interactive proof systems*, Advances in Computing Research, S. Micali (Ed.), vol. 5, 1989, pp. 73-90.

9. John Gustafson, Srinivas Aluru, *Massively Parallel Searching for Better Algorithms or, How to Do a Cross Product with Five Multiplications*, Ames Laboratory, Department of Energy, ISU, Ames, Iowa. This paper is available at http://www.scl.ameslab.gov/Publications/FiveMultiplications/Five.html

10. Johan Håstad, *Tensor Rank is NP-Complete*, Journal of Algorithms, vol. 11, pp. 644-654, 1990.

11. Rudolf Lidl, Harald Niederreiter, *"Finite Fields", Encyclopedia of Mathematics and its applications*, Volume 20, Cambridge University Press.

12. Tsutomu Matsumoto, Hideki Imai, *Public quadratic polynomial-Tuples for efficient Signature-Verification and Message-Encryption*, EUROCRYPT'88, Springer-Verlag, pp. 419-453.

13. Jacques Patarin, *Cryptanalysis of the Matsumoto and Imai public Key Scheme of Eurocrypt'88*, CRYPTO'95, Springer-Verlag, pp. 248-261.

14. Jacques Patarin, *Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): two new Families of asymmetric Algorithms*, EUROCRYPT'96, Springer-Verlag, pp. 33-48.

15. Erev Petrank, Ron M. Roth, *Is Code Equivalence Easy to Decide ?*, IEEE Transactions on Information Theory, 1997.

16. Volker Strassen, *Gaussian elimination is not optimal*, Numerische Mathematik 13, 1969, pp. 354-356.

17. Volker Strassen, *The asymptotic spectrum of tensors*, J. Reine Angew. Math., vol. 384, pp. 102-152, 1988.

Appendix 1

IP with one secret is at least as difficult as Graph Isomorphism

First links between the two problems

We first generalize the construction of section 3 about IP with one secret. Let (I) and (II) be two n-vertices graphs, and let $(\mathcal{A})$ and $(\mathcal{B})$ the following systems:

$$(\mathcal{A}) \begin{cases} y_1 = \sum\limits_{i,j} \gamma_{ij} a_i a_j \\ y_2 = \sum\limits_{i=1}^{n} a_i^2 \end{cases} \quad \text{and} \quad (\mathcal{B}) \begin{cases} y_1 = \sum\limits_{i,j} \mu_{ij} x_i x_j \\ y_2 = \sum\limits_{i=1}^{n} x_i^2, \end{cases}$$

where all the γ_{ij} and μ_{ij} coefficients all lie in $\{0,1\}$ and satisfy:

$$\gamma_{ij} = 1 \Leftrightarrow \textit{The } i \textit{ and } j \textit{ vertices of graph } (I) \textit{ are linked together}$$

$$\mu_{ij} = 1 \Leftrightarrow \textit{The } i \textit{ and } j \textit{ vertices of graph } (II) \textit{ are linked together.}$$

It is easy to see that each graph isomorphism between (I) and (II) corresponds to a morphism of polynomials between $(\mathcal{A})$ and $(\mathcal{B})$. This morphism of polynomials is a very special one, since it can be defined by $x_i = a_{\varphi(i)}$, where φ is some permutation of $\{1,...,n\}$.

Reciprocally, can *any* isomorphism of polynomials between $(\mathcal{A})$ and $(\mathcal{B})$ be characterized by $x_i = a_{\varphi(i)}$ for some permutation φ of $\{1,...,n\}$? Of course not, and explicit examples can be built to be convinced. Nevertheless, we can consider the following argument: there exist $q^{n(n+1)}$ possible affine transformations between the a_i and the x_i, among which $q^n(q^n-1)(q^n-q)(q^n-q^2)...(q^n-q^{n-1})$ are bijective. The probability that such a transformation leaves y_1 invariant is *a priori* $\leq \frac{1}{q^{\frac{n(n+1)}{2}+1}}$ (because there are $\frac{n(n+1)}{2}$ monomials $x_i x_j$ $(i \leq j)$, and a constant term). The same property is *a priori* true for y_2. Moreover, since

$$q^n(q^n-1)(q^n-q)(q^n-q^2)...(q^n-q^{n-1}) < \left(q^{\frac{n(n+1)}{2}+1}\right)^2,$$

we can think that – "most of the time" – there are very few solutions for the IP problem between $(\mathcal{A})$ and $(\mathcal{B})$ that do not correspond to a solution of Graph Isomorphism. Of course, it is a rough evaluation. Nevertheless, it suggests that if a method is found to solve the IP problem with one secret, then we should be able to use this method to also solve the graph isomorphism problem. We will now study how to obtain a real proof of this property.

The real proof

We first prove a general property about permutations:

Theorem 11.1 *Any permutation σ of $\{1,...,n\}$ can be written in a unique way as follows:*

$$\sigma = \tau_{i_n,n} \circ \tau_{i_{n-1},n-1} \circ \ldots \circ \tau_{i_1,1},$$

where $i_k \in \{1,...,k\}$ for all k, $1 \leq k \leq n$, and where $\tau_{i,j}$ is the permutation that exchanges i and j (by convention, $\tau_{i,i}$ =Id).

The proof is easy by induction on n (details are available from the authors).

We now see how to use this theorem to "translate" the fact that the transformation s involved in the NP-problem corresponding to a Graph Isomorphism instance is characterized by $x_i = a_{\varphi(i)}$ for some permutation $\varphi \in S_n$. According to theorem 4.1, such a permutation can be written as the product of (at most) n permutations $\tau_{i,j}$.

To simplify, let us first give a "translation" of the fact that an affine transformation $s : a \mapsto x$ is characterized either by s =Id, or by $x_i = a_{\varphi(i)}$ (for all i), with $\varphi = \tau_{ij}$ for two given indices i and j. It is equivalent to saying that s is an isomorphism of polynomials between the two following systems:

$$(\mathcal{A})\begin{cases} y_0 = (X-a_i)(X-a_j) & \\ y_k = a_k & (1 \leq k \leq n,\ k \neq i,j) \\ y_{n+1} = X & \end{cases}$$

and

$$(\mathcal{B})\begin{cases} y_0 = (A-x_i)(A-x_j) & \\ y_k = x_k & (1 \leq k \leq n,\ k \neq i,j) \\ y_{n+1} = A & \end{cases}$$

By using this argument several times, it is possible to obtain a "translation" of a Graph Isomorphism problem into an IP problem: s corresponds to an isomorphism of the graphs (*i.e.* in particular is of the form $x_i = a_{\varphi(i)}$ for some $\varphi \in S_n$) iff it is an isomorphism of polynomials between two systems of equations (that can be explicitely written). Each of these two systems contains $\frac{2n^3-3n^2-n+4}{2}$ equations over $\frac{(n+1)(n^2-2n+2)}{2}$ variables.

Conclusion: By solving IP with one secret on a set of $\mathcal{O}(n^3)$ quadratic equations we can solve any Graph Isomorphism problem with n vertices. Therefore, IP is at least as hard as GI.

Remark: We do not pretend that this construction gives a new and more efficient way to solve the Graph Isomorphism problem. In fact, although no polynomial algorithm is known for the Graph Isomorphism problem, in practice very efficient algorithms are known: for example it is feasible to find an isomorphism for graphs even for "hard" instances of 1000 vertices graphs in less than 10 minutes on a personal computer (see [5] p. 22). So the main interest in this construction is to show that the IP problem with one secret is probably not solvable with a probabilistic algorithm of polynomial complexity. (Because GI was carefully studied and many people think that GI is not solvable in polynomial complexity).

Appendix 2

Suggestions of IP variations

IP with one secret

As we have seen above, some improved algorithms exist for IP with two secrets. Therefore, when IP is used for authentication or signature as explained in [14], it might be suggested to use IP with one secret instead of IP with two secrets in order to have a more efficient scheme. It may look surprising that IP with one secret might be a more difficult problem (with practical values of the parameters) than IP with two secrets. However, this is not so surprising: in IP with two secrets, we have about $2n^2$ unknown coefficients of K (the secret values of the s and t matrices) and about $\frac{n^3}{2}$ quadratic equations on these unknown values (when we formally identify the two sets of equations $(\mathcal{A})$ and $(\mathcal{B})$), *i.e.* much more equations than unknowns. However, in IP with one secret, the parameters can be chosen in order that the number of equations (given by a formal identification) will be about approximately equal to the number of unknowns. This occurs for instance when $(\mathcal{A})$ and $(\mathcal{B})$ are two sets of two quadratic equations. As a result, despite the fact that there is only one affine change of variables, such a problem might be more difficult than the IP with two secrets.

Subgroups of $GL_n(K)$

Another possible idea would be to choose the secret transformations s and t of IP with two secrets in a subgroup G of $GL_n(K)$. ($GL_n(K)$ is the set of all linear bijective transformations from K^n to K^n). This way, n may be chosen rather large (in order to make the problem difficult) and the length of the asymmetric signature (when the problem is used as explained in [14] for asymmetric signatures) might still be of reasonable size. For instance, G might be the orthogonal group of some quadratic form q (*i.e.* the set of all $g \in GL_n(K)$ such that $\forall x \in K^n$, $q(g(x)) = q(x)$), or G might be the group of all matrices of the form $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$, where A and B are two $\frac{n}{2} \times \frac{n}{2}$ matrices. It is not clear whether choosing s and t in such a subgroup G makes it easier to solve the IP problem or not.

Remark: A similar idea is used in the DSS or in Schnorr's algorithm, where – in order to have shorter length for the signatures – a subgroup is chosen for the exponents used in these schemes. However, here, the schemes are very different.