Isomorphism of Polynomials

Nicolas T. Courtois
Paris 6 and Toulon University

Jacques Patarin, Louis Goubin
Bull Smart Cards and Terminals, France

Summary

1. What is Isomorphism of Polynomials **(IP)** ?
2. Cryptographic relevance.
3. Related problems:
   - ◇ It generalises the Graph Isomorphism **(GI)**.
   - ◇ It generalizes to Morphism of Polynomials **(MP)**.
4. How difficult it is ?
5. Advances in attacks: $q^{n^2} \rightsquigarrow q^{n\sqrt{n}} \rightsquigarrow q^{\mathcal{O}(n)} \rightsquigarrow q^{n/2}$

Isomorphism of Polynomials (**IP**)

Given two sets of **u** multivariate polynomials

with **n** variables over a finite field $\mathbf{F}_q$.

$$b_k = \delta_k + \sum_i \mu_{ik} a_i + \sum_{i,j} \gamma_{ijk} a_i a_j + [\ldots] \qquad (1 \le k \le u). \qquad (\mathcal{A})$$

$$y_k = \delta'_k + \sum_i \mu'_{ik} x_i + \sum_{i,j} \gamma'_{ijk} x_i x_j + [\ldots] \qquad (1 \le k \le u). \qquad (\mathcal{B})$$

IP: Find two affine bijections S and T such that:

$$\mathcal{B} = T \circ \mathcal{A} \circ S.$$

An example with $u = n = 5$ quadratic equations over $\mathbf{F}_2$:

Our new methods allow to solve it by hand:

IP and asymmetric cryptography

public key: $\mathcal{B}$.

secret key: T, S and $\mathcal{A}$.

Some special properties allow to compute $\mathcal{A}^{-1}$.

Secret key allows to compute $\mathcal{B}^{-1} = S^{-1} \circ \mathcal{A}^{-1} \circ T^{-1}$.

Main unbroken candidate: HFE [Patarin Eurocrypt'96].

Solving IP is not enough to break HFE.

Many other schemes are in some way related to IP.

IP in cryptanalysis

1. **Many** schemes have been broken without recovering the secret key. (no IP solving).
   - 2 Shamir schemes. [Stern, Coppersmith, Vaudenay]
   - Matsumoto and Imai's C^* and $[C]$ schemes [Patarin]
   - Patarin's D^*, Little Dragon, S-boxes, Scotch [authors]
2. **Few** schemes have been broken **with** the underlying IP problem.
   - D^* [Courtois 97].
   - 'Oil and Vinegar' [Kipnis, Shamir Crypto'98]

IP in authentication

(associated decision problem)

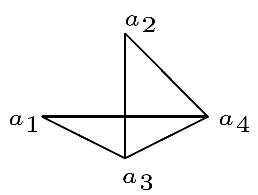
- IP has a zero-knowledge interactive proof.
  $\rightsquigarrow$ authentication algorithm.
- It can be transformed into a signature scheme.

IP = Harder Graph Isomorphism generalization:

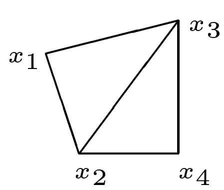
Graph Isomorphism

Graph Isomorphism = very particular case of IP:

- $x_i x_j$ says that vertices i, j are connected:



$a_1a_4 + a_1a_3 + a_2a_3 + a_2a_4$



$x_1x_2 + x_1x_3 + x_2x_4 + x_3x_4$

- Isomorphisms of Polynomials that permute variables are Graphs Isomorphisms.
- Other IP solutions do not proceed from a Graph Isomorphism.
- **Construction:** A particular instance of IP equivalent to finding a graph isomorphism. (extended version of the paper)

Conclusion

IP is at least as difficult as Graph Isomorphism.
(not likely to be polynomial ?!)

Deciding IP is not $\mathcal{NP}$-hard ?!

Non-IP problem has a constant-round interactive proof:

- P: produces equations isomorphic to either $\mathcal{A}$ or $\mathcal{B}$.
- V: guesses which one.

Theorem: If Deciding(IP) is $\mathcal{NP}$-complete, the polynomial hierarchy collapses to ($\mathcal{P}$,$\mathcal{NP}$,$\mathcal{IP}$).

Proof: As for **GI** [Boppana, Håstad, Zachos 87].

Morphism of Polynomials (MP)

IP with S and T that are no longer bijective.

Non-commutative MP version: K is a ring.

Example:

$$\mathcal{A}: \begin{cases} b_1 = a_1 a'_1 \\ \vdots \\ b_7 = a_7 a'_7 \end{cases}$$

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

Get $\mathcal{B} = T \circ \mathcal{A} \circ S$ (or how to multiply 2x2 matrices with only 7 multiplications).

MP is $\mathcal{NP}$-hard

♣ Proven for finite fields and $\mathbb{Q}$.

Idea of proof: It allows to compute the rank of a tensor. Tensor rank problem is $\mathcal{NP}$-complete [Håstad 90].

♣ Non-commutative **MP** solving would lead to better algorithms, e.g. fast matrix multiplication.

It also seems extremely hard in practice.

Real example

Matrices 2x2 [Strassen 69]:

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

3x3 matrices $\leadsto$ at most 23 multiplications [Laderman 76].

Massive parallelism - no success [Gustafson, Alaru 96].

How secure is **IP** ?

Between hard and easy $\mathcal{NP}$ problems.

easy $\uparrow$

GI	$n >> 1000$ [Fortin 96]
IP	$n > 80$ [present work]
MP	open problem for $n = 9$

$\downarrow$ hard

Solving **IP**

1. Exhaustive search $\rightsquigarrow q^{n^2}$ ($q =$ base field size).
2. Improved method $\rightsquigarrow q^{n\sqrt{n}}$.
3. Advanced methods.
   - Inversion attack for non bijective forms $\rightsquigarrow q^{\mathcal{O}(n)}$.
   - The TO AND FRO attack $\rightsquigarrow q^{\mathcal{O}(n)}$
   - Combined power attack: as low as $\rightsquigarrow q^{n/2}$ (S,T linear and with quadratic equations.)

The main idea

♣ We start from some initial equation(s) on S or T.

♣ We use equations of $\mathcal{A}$ and $\mathcal{B}$ to deduce some other equations on S or T.

To And Fro

Starting equations on inputs:

$$\left\{ \begin{array}{ccc} \mathcal{B} & & \mathcal{A} \\ s(1) & = & 1 \\ s(2) & = & 7 \end{array} \right.$$

We get 3 dependent equations on inputs:

$$\left\{\begin{array}{rcr} \mathcal{B} & & \mathcal{A} \\ s(1) & = & 1 \\ s(2) & = & 7 \\ s(3) & = & 6 \end{array}\right.$$

Equations on inputs give equations on outputs:

$$\begin{array}{cccccccc|c} & T & \circ & \mathcal{A} & \circ & S & = & \mathcal{B} & \\ & & & & 1 & 1 & & & S(1) = 1 \\ & & & \swarrow & & & & \searrow & \Downarrow \\ 5 & 1 & & & & & & 5 & T(1) = 5 \end{array}$$

It gives 3 **independent** equations on outputs (!).

$$\left\{\begin{array}{rcl} \mathcal{B} & & \quad\mathcal{A} \\ 5 & = & t(1) \\ 16 & = & t(4) \\ 24 & = & t(23) \end{array}\right.$$

$\boxed{\text{Miracle:}}$ 2 equations $\leadsto$ 3 equations (!).

We use non-linearity to 'boost' the initial knowledge.

$\vdots$

n such equations $\leadsto$ give S or T.

Even better algorithms

$q^{n/2}$ Algorithm ?

Two problems in doing better that q^n:

Problem 1

Find only 1 equation on $S \rightsquigarrow \mathcal{O}(q^n)$.

We have designed a birthday-paradox approach.

Problem 2

The 'FRO' part requires to compute $\mathcal{A}^{-1}$ and $\mathcal{B}^{-1}$ - $\mathcal{O}(q^n)$.

Idea of 'Boosting Function' that amplifies an initial information on **inputs** and gives still information on **inputs** of $\mathcal{A}$ of $\mathcal{B}$.

Differential Solving:

Given a **quadratic** form $\mathcal{A}$ and $\triangle x = c$ and $\triangle y = d$, it is **easy** to find x and x' such that:

$$\left\{ \begin{array}{rcl} x - x' & = & c \\ \mathcal{A}(x) - \mathcal{A}(x') & = & d \end{array} \right.$$

Conclusion

Isomorphism of Polynomials is an important problem in both cryptography and cryptanalysis.

It's difficulty lies in between two famous problems: **GI** (easy but not polynomial) and **MP** (hard).

Questions:

⊗ Even better attacks for **IP** ?

⊗ How difficult are different variations of **IP** and **MP** ? (in both theoretical and practical aspects).

influence of $\frac{u}{n}$ value, only S is secret, commutative/not

⊗ Can **IP** algorithms be generalized to solve **MP** ?

⊗ Is **MP** really **that** hard ?

⊗ Asymmetric cryptosystems based on **MP** problem ?