The security of cryptographic primitives based on multivariate problems over finite fields MQ, IP, MinRank, HFE

Nicolas T. Courtois
courtois@minrank.org
www.nicolascourtois.net

SIS, Toulon University

Road map

- ♢ Cryptography and Multivariate Cryptography
- ♢ Quadratic Equations and MQ problem
- ♢ Hidden Field Equations (HFE) and HFE problem
- ♢ Attacks on HFE
- ♢ MinRank problem
- ♢ Attacks on MinRank
- ♢ Applications of HFE : Short signatures
- ♢ Applications of MinRank : Zero-knowledge authentication

Why cryptography is becoming important in our societies

$$\text{Freedom} \to \infty \quad \Rightarrow \quad \text{Security} \to 0$$

The goal of cryptography : Add security to the information technologies that are **by nature** insecure.

♢ Privacy, Anonymity.

♢ Authenticity, Integrity, Non-repudiation

♢ Fair play and robustness in multiparty protocols.

Main tool

Main tool to protect information : the secret (secrecy).

Evolution of protections.

♢ Protections that are **secret** : E.g. Enigma, DVD.

♢ Based on a **secret** key : E.g. DES, AES.

♢ **Private-public** key solutions : E.g. PKI certificates.

Public key cryptography

a.k.a **Asymmetric** or **Private-public** key cryptography.

No private or secure channel, prior transmission of an **authentic** public key.

♢ Public key encryption : Privacy

♢ (Public key) Authentication.

♢ Digital signatures : Authenticity + Non-repudiation.

Multivariate and univariate P.K. cryptography

Main stream PK-cryptography :

♢ 1970s.. **Univariate** equations, mostly exponentiation over finite fields or rings. Subexponential algorithms ⇒ huge blocks (1024 bits).

♢ 1986.. **Bivariate** equations of algebraic curves. Exponential, $\sqrt{\text{exh. search}}$. Blocks of about (160 bits).

Multivariate cryptography :

♢ 1970s : McEliece and knapsacks, 1980s-90s Many schemes proposed in several countries, many broken

1996 : invention of HFE.

Multivariate equations over finite fields.

Goal to achieve : $\mathcal{O}(\text{exh. search})$. Small blocks $\approx 80 bits$.

Multivariate cryptography cont^d.

A. Algebraic schemes : algebraic trapdoor embedded in a basic one-way primitive in an indistinguishable way.

1. Linear branch.
   ♢ EC codes : McEliece, Niederreiter, SD.
   ♢ Rank-distance codes : GPT, Chen, MinRank.
2. Quadratic branch.
   ♢ **M**ultivariate **Q**uadratic equations (MQ) :
   Matsumoto-Imai (C^*), D^*, [C] (HM), HFE, HFEv-, TPM , TTM, Flash, Sflash, Quartz, Shamir's birational signatures, etc..

B. Combinatorial schemes : combinatorial trapdoor.

3. Combinatorial schemes : PKP, CLE, IP, GI, UOV etc..
4. Combinatorial transformations based on schemes from 1 or 2 : HFEv, HFE-, HFEv-, GPT with right scrambler, C^*- etc.

Modern cryptography

Given **a cryptosystem**, a security level is defined by a triple :

1. Adversarial **objective** (the weakest possible)
2. Adversarial **ressources** (the strongest possible).
   Example 1. Probabilistic Turing machine running in time $T(n)$ and working with success probability $\varepsilon(n)$; n = security parameter.
   Example 2. Quantum computer with 1000 binary quantum states and running in 1 s (Crypto science-fiction).
3. **Access** to the system (i.e. initial penetration).
   The strongest are usually adaptive attacks from inside.

Given a cryptosystem and a security level, we may have proofs, reductions, arguments or beliefs that determine the confidence level in this security level.

Security notions in multivariate cryptography

Multivariate cryptography attempts to achieve **the strongest security definitions known** for each respective application :

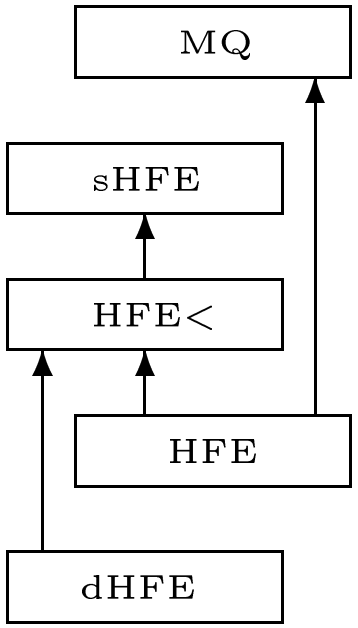
1. Adversarial goal.
   ♢ PK encryption : Distinguishability (Semantic security) :
   REACT-HFE [Pointcheval] is a secure encryption function based on a single one-wayness of HFE (called HFE problem).
   ♢ PK authentication : Learn something about the Skey.
   MinRank scheme is proven Zero-knowledge.
   ♢ PK signature : Resistance against existential forgery :
   Quartz based on HFE, proven for PK-only attacks [2nd Nessie workshop], conjectured in general (slight modifications).
2. Adversarial ressources : Asymptotic and concrete security close to exhaustive search.
3. Access (initial penetration) : Most general adaptive insider attacks.
   Proven for MinRank, conjectured for Quartz.

Building secure multivariate cryptosystems :

- ♢ Security reductions / proofs / arguments / or relations.
- ♢ Study the hardness of the basic problems.

Hierarchy of problems for a trapdoor function

A trapdoor is embedded in a basic one-way primitive in an indistinguishable way. Example :



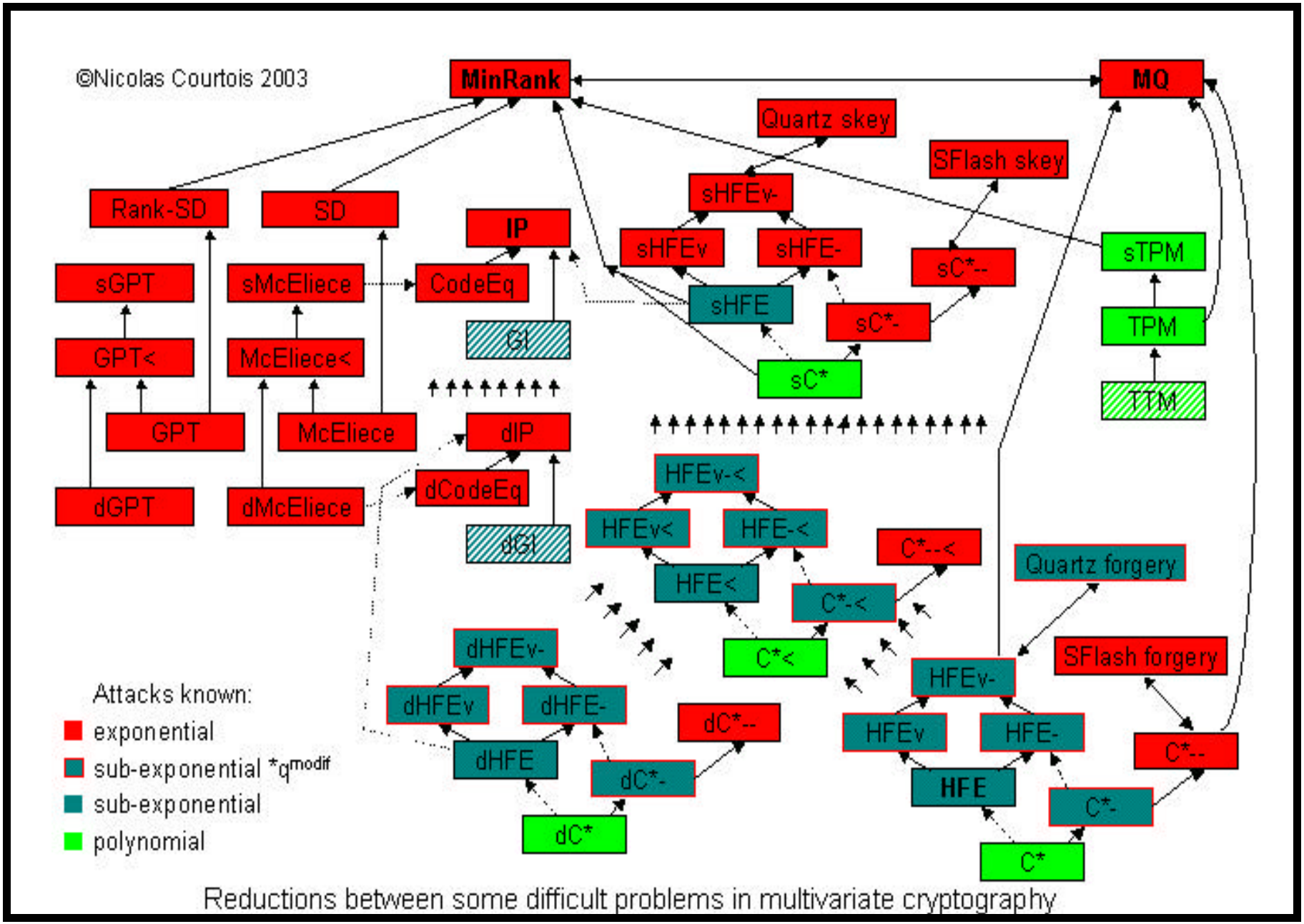
base one-way problem

recover the **s**ecret key, computational **s**tructural attack

inversion attack faster than for base problem

inversion attack - given y, find x, $y = F(x)$

distinguish from a random MQ, **d**ecisional problem



Reductions between some difficult problems in multivariate cryptography

Multivariate **Q**uadratic one-way functions

The **MQ** problem over a ring K : Find (one) solution to a system of **m** quadratic equations with **n** variables in K.

$$f : \begin{cases} b_k = \sum_{i=0}^{n} \sum_{j=i}^{n} \lambda_{ijk} \; a_i a_j \\ \text{with } k = 1..m, \quad a_0 = 1 \end{cases}$$

MQ - Univariate case : $n = 1$

For $K = \mathbb{Z}_N$, as hard as factoring N [Rabin].
For $K = GF(q)$, solved in polynomial time [Berlekamp 1967].

MQ - Multivariate case

NP-hard for **any field** K [Garey-Johnson, Patarin-Goubin].

Solving MQ

Case $m > \frac{n^2}{2}$ **:** MQ is solved by linearization (folklore) :

- New variables $y_{ij} = x_i x_j$.
- At least m linear equations with m variables.

Case $m = \varepsilon \frac{n^2}{2}$ **:** MQ is expected to be polynomial in $n^{\mathcal{O}(1/\sqrt{\varepsilon})}$.

First claimed by Shamir and Kipnis at Crypto'99.

Demonstrated by Courtois, Patarin, Shamir and Klimov at Eurocrypt 2000, a better and simpler algorithm :

XL algorithm

For a given output $y \in K^m$ we put

$$l_k = f_k(x_1, \ldots, x_n) - y_k$$

The instance to solve is :

$$\begin{cases} l_1(x_1, \ldots, x_n) = 0 \\ \vdots \\ l_m(x_1, \ldots, x_n) = 0 \end{cases}$$

E**X**tended **L**inerization or Multiply(**X**) and **L**inearize.

$$\begin{cases} x_1\ l_1 & = & 0 \\ x_2\ l_1 & = & 0 \\ & \vdots & \\ x_n\ l_1 & = & 0 \\ & \vdots & \\ x_n\ l_m & = & 0 \\ & \vdots & \\ x_1 x_2\ l_1 & = & 0 \\ & \vdots & \end{cases}$$

Eliminate all terms in **all but a small number** of variables.

Special case of Gröbner bases - important part of applied mathematics : Buchberger algorithm [1965]....F_5 by Jean-Charles Faugère.
It is unclear if they give better results in practice that the simple XL.

The complexity of XL

Let ω be the exponent of Gaussian reduction. $2 \leq \omega < 3$.

Question : Given n, m, up to what maximum degree D, XL should be applied ? Our estimations confirmed simulations :

$$D \approx \frac{n}{\sqrt{m}} \approx \lceil \frac{1}{\sqrt{\varepsilon}} \rceil$$

XL is expected to solve $m = \varepsilon n^2$ equations with n variables in **polynomial** time :

$$WF \quad = \quad T^{\omega} \quad \approx \quad \mathcal{O}\left(n^{\frac{\omega}{\sqrt{\varepsilon}}}\right)$$

Our discovery : When $m = n + 1, n + 2, ..$

When m becomes slightly greater than n, XL works **much** better.

Application to systems with $m \approx n$

If q is small, guess some variables and $m - n$ increases !

FXL algorithm

Hypothesis : It is enough to fix $\sqrt{n}$ variables.
Then FXL is expected to solve a system of n quadratic equations with n unknowns over $GF(q)$ in **subexponential** time :

$$WF \approx q^{\sqrt{n}} n^{\omega\sqrt{n}}$$

Lack of evidence

In practice, the best known algorithms for solving multivariate equations over a very small finite field are still close to the exhaustive search.

HFE Challenge 1

We ignore the internal structure of HFE cryptosystem.
The public key is $m = 80$ quadratic equations with $n = 80$ variables over $GF(2)$, $q = 2$.

$$WF \approx q^{\sqrt{n}} n^{\omega\sqrt{n}} \approx 2^{179}$$

FXL exceeds by far the exhaustive search in 2^{80}.

Trapdoors in MQ

General principles od design :

[Base] A function such that
◇ It is invertible due to some algebraic properties.
◇ Can be re-written as MQ.

[A] Hidden function - a basic (algebraic) version of a trapdoor.
Conceals algebraic structure with invertible affine variable changes (e.g. basic HFE).

[B] Added perturbations - an extended (combinatorial) version of a trapdoor function - destroys the algebraic structure by non-invertible operations (e.g. HFEv-).

Analogous to diffusion / confusion principles [Shannon, Feistel] :
A is linear, global and invertible, B are local perturbations.

K - finite field $K = GF(q)$, q prime or $q = p^\alpha$

$\exists$ a (unique) finite field $GF(q^n) = K[X]/P(X)$

with P being a degree n irreducible polynomial over K.

$GF(q^n) \equiv K^n$, vector space, dimension n over K :

$x \in GF(q^n)$ is encoded as $(x_1, \ldots, x_n)$, n-tuple of coeffs. of a polynomial from $K[X]$ modulo P.

Multivariate and univariate representations.

Every function $f : GF(q)^n \to GF(q)^n$ can be written as :

$\diamond$ a univariate polynomial.

$\diamond$ n multivariate polynomials with n variables over K.

Multivariate and univariate degree.

If $b = f(a) = a^{q^s}$ then all the $b_i = f_i(a_1, \ldots, a_n)$ are K-linear.
If $f(a) = \sum a^{q^s + q^t}$ then the f_i are quadratic.

Example over $GF(2)$.

$b = f(a) = a + a^3 + a^5 =$
$(a_2X^2 + a_1X + a_0) + (a_2X^2 + a_1X + a_0)^3 + (a_2X^2 + a_1X +$
$a_0)^5 \ mod \ X^3 + X^2 + 1 =$
$(a_2+a_2a_1+a_2a_0+a_1)X^2+(a_2a_1+a_1a_0+a_2)X+(a_0+a_2+a_1a_0+a_2a_0)$

$$\begin{cases} b_2 &= a_2 + a_2a_1 + a_2a_0 + a_1 \\ b_1 &= a_2a_1 + a_1a_0 + a_2 \\ b_0 &= a_0 + a_2 + a_1a_0 + a_2a_0 \end{cases}$$

Hidden Field Equation (HFE).

$$f(a) = \sum_{q^s+q^t \leq d} \gamma_{st} \quad a^{q^s+q^t}$$

– Re-write as n **multivariate** quadratic equations :

$$f : \left\{ \quad b_i = f_i(a_1, \ldots, a_n) \quad \right\}_{i=1..n}$$

– Conceal the algebraic structure of f :
Apply two affine invertible variable changes S and T.

$$g = T \circ f \circ S$$

$$g : x \overset{S}{\mapsto} a \overset{f}{\mapsto} b \overset{T}{\mapsto} y$$

Using HFE

public key : n quadratic polynomials

$$g : \left\{ \quad y_i = g_i(x_1, \ldots, x_n) \quad \right\}_{i=1..n}$$

private key : Knowledge of T, S and f.

Since f is bounded degree and univariate, we can invert it :

Several methods for factoring univariate polynomials over a finite field are known since [Berlekamp 1967]. Shoup's NTL library.

Quite slow, example n=128, d=25, 0.17s on PIII-500.

Computing g^{-1} using the private information

$$x \xleftarrow{S^{-1}} a \xleftarrow{f^{-1}} b \xleftarrow{T^{-1}} y$$

The HFE problem

A restriction of MQ to the trapdoor function g defined above.

Given the multivariate representation of $\mathbf{g}$ and a random $\mathbf{y}$.

Find a solution $\mathbf{x}$ such that g(x)=y.

It is **not** about recovering the secret key (problem sHFE $\neq$ HFE).

Claim

The HFE problem is necessary and sufficient to achieve secure encryption and secure signature schemes with basic HFE.

HFE problem $\neq$ HFE cryptosystem

basic HFE - algebraical, $\exists$ subexponential attacks on the trapdoor.

HFE-, HFEv, .. combinatorial versions - no attacks known.

How to recover S and T.

If f were known, $\exists$ algo in $q^{n/2} = \sqrt{\text{exhaustive search}}$.
the **IP** problem [Courtois, Goubin, Patarin, Eurocrypt'98].

Remark [Shamir] : f is 'known in 99%' because $d << q^n - 1$

The weakness of HFE identified [Shamir-Kipnis, Crypto'99].

The homogenous quadratic parts of g (and f) can be written in the univariate representation and represented by a using a symmetric matrix G (resp. F) :

$$g(x) = \sum_{i=0}^{n-1} \sum_{j=i}^{n-1} G_{ij} \; x^{q^i+q^j}$$

rank(G)= supposedly n, and rank(F)=r avec $r = \log d$.

$$T^{-1} \circ g \stackrel{?}{=} f \circ S$$

Lemma 1 [Shamir-Kipnis] : The matrix representation of $f \circ S$ is of the form $G' = WGW^t$. Same rank r.

Lemma 2 [Shamir-Kipnis] : $T^{-1} \circ g = \sum_{k=0}^{n-1} t_k G^{*k}$ with G^{*k} being the **public** matrix representations of g^{q^k}.

The attack focuses on finding a transformation T such that the matrix representation of $T^{-1} \circ g$ is of small rank. Find such $t_k \in K^n$ that

$$Rank(\sum_{k=0}^{n-1} t_k G^{*k}) = r$$

Recovering the secret key of HFE is reduced to MinRank : sHFE $\rightarrow$ MinRank over $GF(q^n)$.

The problem MinRank

MinRank$(n \times n, m, r, K)$

Given : m matrices $n \times n$ over a ring K : $M_1, \ldots M_m$.

Find a linear combination α of M_i of rank $\leq r$.

$$Rank(\sum_i \alpha_i M_i) \leq r.$$

Fact : MinRank is NP-complete [Shallit, Frandsen, Buss 1996].

MinRank can encode any set of multivariate equations.

MinRank contains syndrome decoding, probably exponential.
Also reduction from rank-distance syndrome decoding.

MinRank attacks

- ♢ First put $m := Min(m,\ \eta n + r^2 - (\eta + n)r + 1)$
  Exceptions : constructed small number of solutions.
- ♢ If $\eta < n$, transpose all matrices. Now $\eta \geq n$.
- ♢ **Brute force attack** $WF \approx q^m$.
- ♢ **Random square MinRank r≈n** [Schnorr] Let $n = r - s$, then $m := Min(m,\ s^2)$, broken in q^{s^2}. Fails for rectangular matrices.
- ♢ **Sub-matrices Attack r≪n** [Coppersmith-Stern-Vaudenay, C'93]. All the sub-matrices $(r+1)$x$(r+1)$ are singular. $WF \approx m^{\omega r}$.
- ♢ **MinRank → overdefined MQ r≪n** [Shamir-Kipnis, Crypto'99] Express the fact that columns $r+1..n$ are dependent. $WF \approx n^{\omega r}$
- ♢ **The Kernel Attack** [Goubin, Asiacrypt 2000] Guess the kernel of the matrix. The best attack for small q, $WF \approx q^{\lceil \frac{m}{\eta} \rceil r}$.
- ♢ **The Big m Attack m≫n** [Courtois], $WF \approx q^{Max(0, \eta(n-r)-m)}$.
- ♢ **The Syndrome Attack m≫n** [Gabidulin, Courtois]. Uses syndromes and linearity. $WF \approx q^{Max(\frac{\eta n - m - 1}{2}, (\eta+n)r/2 - m - r^2/4)}$

$\Rightarrow$ All attacks are exponential

MinRank attacks on HFE in practice

Reference point : 80-bit trapdoor HFE Challenge 1.

Solving this MinRank using :

♢ All the attacks with q^{sth} fail as here $q = 2^{80}$.

♢ **MinRank $\rightarrow$ overdefined MQ** [Shamir-Kipnis, Crypto'99]
$n(n-r)$ quadratic equations with $r(n-r)$ variables over $GF(2^{80})$, solve by XL.

$$\mathbf{2^{152}}$$

♢ Use **Sub-matrices Attack** [Courtois RSA 2001]

$$\mathbf{2^{97}}$$

♢ Worse than the exhaustive search on the underlying HFE

$$\mathbf{2^{80}}$$

Do we need to recover the secret key ?

Some cryptanalyses of multivariate schemes :

1. For **some** the secret key is computed :
   - D^* [Courtois 97].
   - 'Balanced Oil and Vinegar' [Kipnis, Shamir Crypto'98]
   - HFE [Kipnis, Shamir Crypto'99].
2. In **many** cases the attack does not compute the secret key :
   - Matsumoto and Imai C^* and $[C]$ schemes [Patarin]
   - Shamir birational signat. [Coppersmith, Stern, Vaudenay]
   - D^*, L. Dragon, S-boxes, C^{*-} [Patarin, Goubin, Courtois]
   - Equational attacks on HFE [Courtois]

What characterizes functions g that can(not) be inverted ?

◇ Symmetric cryptography - there should be **no** simple way to relate x and $g(x)$ with some equations [Shannon's thoughts]
Idea of unpredictability, pseudorandomness.

◇ Asymmetric cryptography - usually explicit equations g(x).
The pseudorandomness paradigm can hardly be applied.

Every deterministic attack can be seen as a series of transformations that start with some **complex** and **implicit** equations $G(x_i) = 0$.
It gives at the end some equations that are **explicit** and **simple**, e.g. $x_i = 0$ ou 1.

Definition [very informal] : One-way function in PKC
All 'basic' combinations of given equations do not give equations that are explicit or 'simpler'.

We denote by G_j the expressions in the x_i of public equations of g s.t. the equations to solve are $G_j = 0$.

We can generate other (multivariate) equations (true for x) by low degree combination of the G_j and the x_i.

We require that such 'trivial' combinations of public equations remain 'trivial'

Definition [informal] : A trivial equation is small degree combination of the G_j and the x_i, with terms containing at least one G_j and such that it's complexity (e.g. multivariate degree) does not collapse.

Soundness of the definition : One such equation, substituted with the values of $G_j = 0$ gives a new low degree equation in the x_i.

Remark : The same notion applies to block and stream ciphers.

Implicit equations attacks [Patarin, Courtois].

Several attacks that use several types of equations.
Common properties :

◇ We can only predict the results in very simple cases.

◇ Experimental equations can be found with no apparent theoretical background.

◇ The equations are detected **only** beyond some threshold (e.g. 840 Mo).

HFE Challenge 1

We found equations of type $1 + x + y + x^2y + xy^2 + x^3y + x^2y^2$.

Gives an attack in 2^{62} published at RSA 2001, 390 Gb.

In 2002 Faugère solved it in 2^{48} using Gröbner bases F5/2 algo.

Asymptotic security of HFE

Attack	Cxty	$d = n^{\mathcal{O}(1)}$
Shamir-Kipnis Crypto'99 HFE $\rightsquigarrow$ sHFE $\rightsquigarrow$ MinRank $\rightsquigarrow$ MQ	$n^{\log^2 \mathbf{d}}$	$e^{\log^3 \mathbf{n}}$
Shamir-Kipnis-Courtois HFE $\rightsquigarrow$ sHFE $\rightsquigarrow$ MinRank	$n^{\mathbf{3}\log \mathbf{d}}$	$e^{\log^2 \mathbf{n}}$
Courtois HFE $\rightsquigarrow$ Implicit Equations	$n^{\frac{\mathbf{3}}{\mathbf{2}}\log \mathbf{d}}$	$e^{\log^2 \mathbf{n}}$

HFE problem is **polynomial if** d fixed (not HFEv etc..).

The degree d can be quite big in practice.

It is **subexponential**, in general : $d = n^{\mathcal{O}(1)}$.

sHFE is probably **not** polynomial in general,
as MinRank is believed exponential.

State of Art on HFE security

◇ The asymptotic complexity of breaking the algebraical HFE (HFE problem) is currently $e^{log^2 n}$.

◇ In practice basic HFE with $d > 128$ is still quite secure ($\mathcal{O}(n^{10})$).

◇ Modified, combinatorial versions of HFE improve the security :
-HFE$^-$ [Eurocrypt'96, Asiacrypt'98],
-HFEv [Eurocrypt'99], Quartz, Flash and Sflash [RSA 2001, Nessie].
However Courtois, Daum and Felke showed that they can still be attacked [PKC 2003].

◇ Combinatorial versions of HFE can be **either** :
-hundreds of times faster than RSA and be implemented on smart cards (Flash, Sflash), **or**
-give very short signatures for memory cards (Quartz).

Digital signatures.

F - a trapdoor function, $GF(q^n) \to GF(q^m)$ bits.

Usual method : $\boxed{\sigma = F^{-1}(\ H(M)\ )}$ H - cryptographic hash.

A generic attack : Existential Forgery in $q^{m/2}$.

1. Precompute a list $q^{m/2}$ outputs $F(\sigma_j)$, for random σ_j.
2. It allows to compute an inverse of F, $x = F^{-1}(y)$, with probability $q^{m/2}/q^m = q^{-m/2}$, over y.
3. Try $q^{m/2}$ **versions** of the message to be signed $H(M_1), \ldots, H(M_{q^{m/2}})$, adding spaces, commas, addenda etc..
4. For roughly about one of them we are able to compute an inverse.
5. We have then a valid pair (message, signature) : $F(\sigma_j) = H(M_i)$

Classical scheme : signatures of m bits are broken in $q^{m/2}$.

Public key cryptosystems : very few candidates known.
PKC with very short block sizes : much fewer, only recently known.

Fact

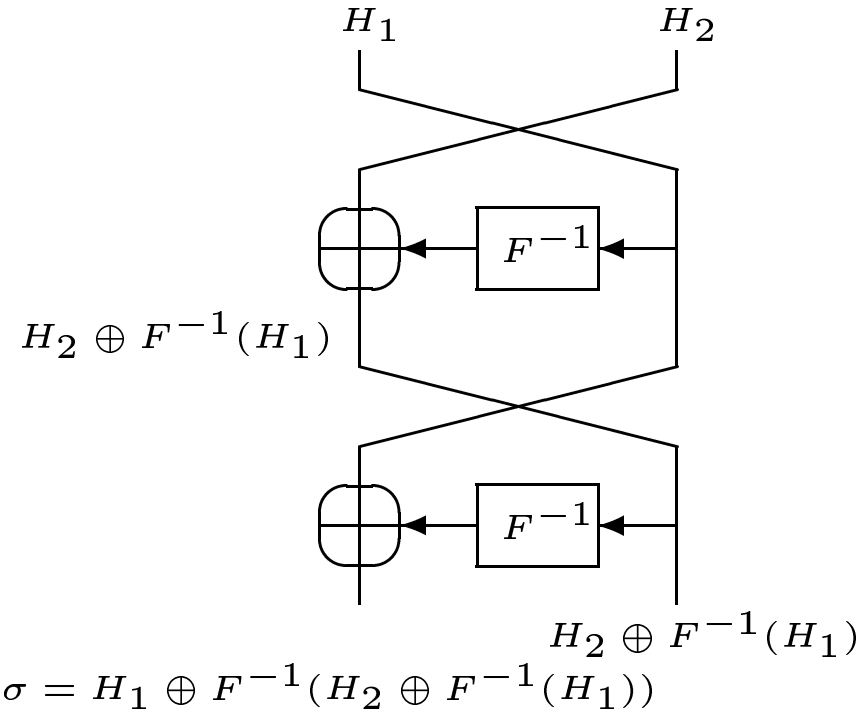
Before HFE - published in 1996,
no non-broken trapdoor function known,
without an attack in $q^{m/2}$
due to some linearity/group structure.

$\Downarrow$

Nobody dreamed about achieving a security better than $q^{m/2}$.

Short signatures

Solution for $m = n$, and 2 inverses.



$|(H_1, H_2)| = 160$ bits

$|\sigma| = 80$ bits.

Generalization for K inverses

$$\sigma = F^{-1}(H_K \oplus \ldots \oplus F^{-1}(H_3 \oplus F^{-1}(H_2 \oplus F^{-1}(H_1))\ldots)$$

If $m > n$: for each inverse add $m - n$ bits to signature.

$\sigma \leftarrow 0$
for $i = 1$ to K do
{
$\sigma \leftarrow \sigma \oplus H_i(M)$
$U \in F^{-1}(\sigma)$
$\sigma \leftarrow U_{1 \rightarrow m}$
$Add_{i1}||\ldots||Add_{i(n-m)} \leftarrow U_{(m+1) \rightarrow n}$
}
return $\sigma||Add_{11}||\ldots||Add_{K(n-m)}$

The generalized Feistel-Patarin construction

A generic attack : Existential Forgery in $q^{\frac{K}{K+1}m}$.

1. Precompute a list $q^{\frac{K}{K+1}m}$ outputs $F(\sigma_j)$, for random σ_j.
2. It allows to compute an inverse of F, $x = F^{-1}(y)$, with probability $q^{\frac{K}{K+1}m}/q^m = q^{-\frac{1}{K+1}m}$ over y.
3. It allows to compute K inverses of F, with probability $\left(q^{-\frac{1}{K+1}m}\right)^K = q^{-\frac{K}{K+1}m}$.
4. Try $q^{\frac{K}{K+1}m}$ **versions** of the message.
5. We have then a valid pair (message, signature) : $F(\sigma_j) = H(M_i)$

The converse, or is to possible to prove the security of Quartz ?

Conjecture : If H behaves as a random oracle and if there are no algorithms better than the exhaustive search for F, then, **given the public key**, the signature **cannot be forged** in less than

$$T \geq q^{\frac{K}{K+1}m}.$$

For chosen-message attacks, false in general, ambiguity on F^{-1}.
Proven in a paper to appear at PKC 2003, but **not** for Quartz construction. For Quartz the proof is not tight, $T \geq 2^{50}$ instead of 2^{80}.

Signature length for given security

$$|\sigma| = \lceil (m + K \cdot (n - m)) \log_2 q \rceil \quad \text{bits}$$

Application to Quartz :
$K = 4$, $|\sigma| = 128$ bits, $Security = 2^{80}$.

The shortest signatures known for security $\approx 2^{80}$:

RSA	$\rightsquigarrow$	1024 bits	
improved DSA	$\rightsquigarrow$	240 bits	[Rump session Crypto 2001]
EC + Weil pairing	$\rightsquigarrow$	160 bits	[Boneh et al. Asiacrypt 2001]
HFEv-, Quartz	$\rightsquigarrow$	128 bits	[Courtois, Goubin, Patarin 2000]
HFEf+	$\rightsquigarrow$	92 bits	My PhD thesis, sec. 19.4.2.
McEliece	$\rightsquigarrow$	87 bits	[Courtois, Finiasz, Sendrier 2001]

What signatures are the best ?

Bad question

Use several algorithms and issue several certificates.

Programs, terminals and devices will have at least one common algorithm for many years.

Pro-active scenario : Invalidate some algorithms and introduce new ones.

Example, when 768-bit RSA is broken, the 1024-bit RSA expires.

Un example of combined certificate :

RSA + EC + HFE = $1024 + 320 + 128$ bits.

RSA is slow and signatures are so long that all the rest is for free !

Zero-knowledge Identification

The breakthrough invention of Zero-knowledge [Goldwasser, Micali, Rackoff 1985].

A Zero-knowledge identification protocol :

A protocol between two units the **P**rover and the **V**erifier.

- **Goal** At the end of interaction **V** says *Accept* or *Refuse*.
- **Correctness** An honest Prover is always accepted.
- **Soundness** No one can impersonate the Prover with an overwhelming probability.
- **Zero-knowledge** An (even malicious) verifier cannot extract from the Prover any information about Prover's secret knowledge (or ability) that he can't find out himself.

Zero-knowledge Identification schemes :

[Goldwasser, Micali, Rackoff 1985]
Provably secure entity authentication based on a difficult problem.

Known solutions

The best practical Zero-knowledge protocols are arithmetical :
Fiat-Shamir, Guillou-Quisquater, Schnorr.

Still, there are **practical** schemes based on a NP complete problem :
- PKP [Shamir]
- CLE [Stern]
- PPP [Pointcheval]
- Schemes related to coding [Harari, Girault, Veron, Stern, Chen].

New algorithm in this branch - MinRank.

Protocol settings

The public key are m matrices $n \times n$ over a finite field $GF(q)$, $M_1, \ldots, M_m$.

The secret key is $\alpha \in GF(q)^n$, such that $M = \sum \alpha_i \cdot M_i$ has the rank $r < n$.

Proposed instances of MinRank - example :

$K = GF(65521)$, $n = 7$, $m = 10$, $r = 4 \Rightarrow$ best attack in 2^{122}.
9 random matrices 7×7, 10th matrix is a sum of a random matrix of rank $r = 4$ and some linear combination of the other 9 matrices.

The Prover computations

P chooses :

– two random non-singular matrices S and T.

– a random $n \times n$ matrix X.

– a random combination β_1 of M_i :

$$N_1 = \sum \beta_{1i} \cdot M_i$$

P uses the expression of M to get an expression of $N_2 = M + N_1$ as :

$$N_2 = \sum \beta_{2i} \cdot M_i$$

Now $N_2 - N_1 = M$, each of them is just a random combination.

$$\begin{aligned} (TN_2S) - (TN_1S) &= T(N_2 - N_1)S = TMS \\ (TN_2S + X) - (TN_1S + X) &= T(N_2 - N_1)S = TMS \end{aligned}$$

One round of identification

$$\longrightarrow$$

$$H(X),\ H(TN_1S+X),\ H(TN_2S+X),\ H(S,T)$$

$$\longleftarrow$$

$$q \in \{0,1,2\}$$

If $q = 0$ the prover reveals :

$$\longrightarrow$$

$$(TN_1S+X),\ (TN_2S+X)$$

If $q = 1, 2$ the prover reveals :

$$\longrightarrow$$

$$S,\ T,\ \beta_q, X$$

Conclusion : how to design new multivariate schemes

The security should be [provably] reduced to some difficult problem. Such reductions already exist for Quartz, Flash, Sflash, MinRank authentication, McEliece, and partly for sHFE.

Further work on the problems

♣ **MQ** and **MinRank** are the **foundation** of many cryptographic schemes. All attacks are currently exponential.
→ If broken, most multivariate schemes are broken.

♣ There are many other open problems to study : for example the security of Sflash/C^{*--}. Currently at least exponential.
→ If broken, most other schemes are **still secure**.

Perspectives

Multivariate cryptography is very rich :

♢ Each cryptographic problem probably has a solution in multivariate cryptography. (e.g. ring signatures with MinRank).

♢ In most cases, a small modification of a scheme that is broken gives schemes that are probably very secure.

Applications in practice

The schemes such as Sflash designed for performance are fragile and could be broken.

Other schemes such as Quartz or MinRank, will probably always be secure if the parameters are well chosen, and can be applied in practice with some confidence.