Differential Cryptanalysis for Multivariate Schemes II

Jacques Stern

Joint work with
Vivien Dubois, Pierre-Alain Fouque and Adi Shamir

Ecole normale supérieure, Paris

Multivariate Schemes

- A family of asymmetric schemes
- Hard problems involve MQ polynomials over a finite field $\mathbb{F}_q$
- e.g. solving an MQ system is NP-hard and currently requires exponential time and memory on average

The Generic Multivariate Construction

- Hiding an easily invertible function using linear transforms

$$\boldsymbol{P} = T \circ P \circ S$$

- Schemes differ from the type of easy function embedded

Famous Examples of Multivariate Schemes

- C^* [MI88] (broken by Patarin in 95)
- HFE [Pat96]
- SFLASH [PGC01] selected by NESSIE for fast signatures

FGS05 : Differential Cryptanalysis for Multivariate Schemes

The differential of a quadratic function P at a is :

$$DP(a,x) = P(a+x) - P(x) - P(a) + P(0)$$

- DP is bilinear in (a,x)
- If $\mathbf{P} = T \circ P \circ S$ then $D\mathbf{P} = T \circ DP(S,S)$

Consider linear properties of the *pointwise* differential $DP(a,\cdot)$

e.g. the dimension of the kernel, intersections etc...

- New cryptanalysis of C^*, cryptanalysis of PMI [D04,FGS05]
- A quasipolynomial distinguisher for HFE [DGS06]
- Cryptanalysis of IPHFE [DGS07]

A New Approach

- *Functional* properties of the differential seen as a bilinear map. e.g. we consider skew-symmetric maps M w.r.t DP :

$$DP(M(a), x) + DP(a, M(x)) = 0$$

- Cryptanalysis of SFLASH and other C^{*-} schemes

Description of SFLASH

- SFLASH belongs to the family of C^{*-} schemes [PGC98]
- C^{*-} schemes are C^* schemes with a truncated public key

Construction of a C^{*-} scheme

(n, θ, r) are the parameters of the scheme

1. Generate a C^* with parameters (n, θ) : $P(x) = x^{1+q^\theta}$
2. Remove the last r polynomials from the public key

$$T \circ P \circ S = \begin{cases} \boldsymbol{p}_1(x_1, \ldots, x_n) \\ \vdots \\ \vdots \\ \boldsymbol{p}_n(x_1, \ldots, x_n) \end{cases} \overset{\Pi}{\longmapsto} \begin{cases} \boldsymbol{p}_1(x_1, \ldots, x_n) \\ \vdots \\ \boldsymbol{p}_{n-r}(x_1, .., x_n) \end{cases} = \Pi \circ \boldsymbol{P}$$

Signing with a C^{*-} scheme

1. Append r random bits k to the message m to be signed
2. Find a preimage σ of (m, k) by $\boldsymbol{P} = T \circ P \circ S$
3. σ is a valid signature since $\Pi \circ \boldsymbol{P}(\sigma) = m$

Choosing Parameters

- $\gcd(q^\theta + 1, q^n - 1) = 1$ for C^* bijectivity. This condition is equivalent to n/d odd where $d = \gcd(n, \theta)$
- $q^r \geq 2^{80}$ to avoid a possible recomposing attack from [PGC98]

Proposed parameters

	q	n	θ	d	r	Length	PubKey Size
FLASH	2^8	29	11	1	11	296 bits	18 Ko
SFLASHv2 [NESSIE]	2^7	37	11	1	11	259 bits	15 Ko
SFLASHv3	2^7	67	33	1	11	469 bits	112 Ko

Basic Strategy

- A recomposing attack using a family $\mathcal{F}$ of linear commuting maps. For any M in $\mathcal{F}$, there exists N in $\mathcal{F}$ such that

$$P \circ M = N \circ P$$

[Not obvious since P is quadratic]. Let $\boldsymbol{M} = S^{-1} \circ M \circ S$

$$\begin{aligned} (\Pi \circ T \circ P \circ S) \circ \boldsymbol{M} &= \Pi \circ T \circ (P \circ M) \circ S \\ &= \Pi \circ T \circ (N \circ P) \circ S \\ &= (\Pi \circ T \circ N) \circ P \circ S \end{aligned}$$

Use of $\boldsymbol{M}$ recovers enough coordinates of the public key :

$$\left.\begin{array}{r} (\Pi \circ T) \circ P \circ S \\ (\Pi \circ T \circ N) \circ P \circ S \end{array}\right\} \longmapsto C^*$$

- In C^*, multiplications $x \mapsto \xi.x$ are a commuting family.
- **Goal** : Discover maps $\boldsymbol{M}$ where M is a multiplication.

Skew-symmetric Maps w.r.t the Differential

Definition

M is skew-symmetric with respect to the bilinear map DP iff

$$DP(M(a), x) + DP(a, M(x)) = 0$$

Theorem

When P is the C^* monomial x^{1+q^θ}, the skew-symmetric maps w.r.t to DP are multiplications by ξ with $\xi + \xi^{q^\theta} = 0$.

Proof.

Since $M(x) = \sum_{k=0}^{n-1} \lambda_k x^{q^k}$, $DP(M(a), x) + DP(a, M(x))$ is written on the basis of monomials $a^{q^i} x^{q^j}$. Equaling to zero all coefficients gives the wanted condition. The converse is easily checked. □

- Dimension of the space of skew-symmetric maps = $\dim(\ker L)$ where $L(\xi) = \xi + \xi^{q^\theta}$.

$$\xi \neq 0, L(\xi) = 0 \iff \xi^{q^\theta - 1} = 1$$

So : $\dim(\ker L) = d := \gcd(n, \theta)$.

- Non-trivial maps only exist when $d > 1$.
- Skew-symmetric maps w.r.t the C^* public key $\boldsymbol{P}$ are :

$$\boldsymbol{M}_\xi = S^{-1} \circ M_\xi \circ S \qquad \text{where} \quad M_\xi(x) = \xi.x$$

- They can be recovered through linear algebra from :

$$D\boldsymbol{P}(\boldsymbol{M}(a), x) + D\boldsymbol{P}(a, \boldsymbol{M}(x)) = 0$$

which is a system of $\simeq n^3$ linear equations in n^2 unknowns :
We might not need all coordinates of $\boldsymbol{P}$ to recover the $\boldsymbol{M}_\xi$!

- If we are only given the first $n - r$ coordinates of $\boldsymbol{P}$:

$$\Pi \circ D\boldsymbol{P}(\boldsymbol{M}(a), x) + \Pi \circ D\boldsymbol{P}(a, \boldsymbol{M}(x)) = 0$$

  gives $(n-r)n(n-1)/2$ linear equations in n^2 unknowns
- The skew-symmetric maps $\boldsymbol{M}_\xi$ are solutions.
- We expect no other solutions when :

$$(n-r)\frac{n(n-1)}{2} \geq n^2 - d$$

- Hence, heuristically, the $\boldsymbol{M}_\xi$ are the only solutions up to :

$$r^*_{max} = n - \left\lceil 2\frac{n^2-d}{n(n-1)} \right\rceil = n - 3$$

- The actual value r_{max} is very close to the heuristical r^*_{max} :

n	36	36	38	39	39	40	42	42	44
θ	8	12	10	13	9	8	12	14	12
d	4	12	2	13	3	8	6	14	4
r_{max}	33	32	35	35	36	37	39	38	41

In Brief

- The skew-symmetric maps can be recovered from as few as 3 or 4 coordinates of the public key.
- These maps form a subspace of dimension d and some are non-trivial when $d > 1$.

Recovering a Full C^* Public Key

Using a single non-trivial $\boldsymbol{M}_\xi$, up to $r = n/2$

1. We complete $\Pi \circ \boldsymbol{P}$ using r coordinates of $\Pi \circ \boldsymbol{P} \circ \boldsymbol{M}_\xi$.
2. We can check that this is a full C^* public key since Patarin's attack works again.

n	36	36	38	39	39	40	42	42	44
θ	8	12	10	13	9	8	12	14	12
d	4	12	2	13	3	8	6	14	4
r	11	11	11	12	12	12	13	13	13
$C^{*-} \mapsto C^*$	57s	57s	94s	105s	90s	105s	141s	155s	155s

Note : parameters are close to those of SFLASHv2, with the same $q = 2^7$.

Recovering a Full C^* Public Key

Using a whole basis of $\boldsymbol{M}_\xi$

Since we have $d(n-r)$ coordinates available, the overall bound is :

$$r \leq \min\left\{ r_{max} \; ; \; n\left(1 - \frac{1}{d}\right)\right\}$$

n	36	36	38	39	39	40	42	42	44
θ	8	12	10	13	9	8	12	14	12
d	4	12	2	13	3	8	6	14	4
r	27	32^*	19	35^*	26	35	35	38^*	33
$C^{*-} \mapsto C^*$	$65s$	$51s$	$112s$	$79s$	$107s$	$95s$	$134s$	$117s$	$202s$

Note : the star symbol means $r = r_{max}$, and $r = n(1 - 1/d)$ otherwise.

Multiplicative Property of the Differential

- A more general property of multiplications :

$$DP(M_\xi(a), x) + DP(a, M_\xi(x)) = M_{L(\xi)} \circ DP(a, x)$$

where $M_\xi(x) = \xi.x$ and $L(\xi) = \xi + \xi^{q^\theta}$.

- Let us denote :

$$S_M(a, x) = DP(M(a), x) + DP(a, M(x))$$

- Coordinates of $S_M(a, x)$ and $DP(a, x)$ are bilin. symm. forms.
- Let us call V the span of the coordinates of $DP(a, x)$.
- Characterization of the M_ξ : Any coordinate of S_{M_ξ} is in V.

Implications in the Public World

We are only given the first $(n-r)$ coordinates of $D\boldsymbol{P}$.

$$\tilde{\boldsymbol{V}} = Span(d\boldsymbol{p}_1, \ldots, d\boldsymbol{p}_{n-r}) \quad \subseteq \quad \boldsymbol{V} := Span(D\boldsymbol{P})$$

We express partial conditions :

For a fixed coordinate i among the first $(n-r)$, what is the dimension of solutions of the equation :

$$\boldsymbol{S}_{\boldsymbol{M}}[i] \in \tilde{\boldsymbol{V}}$$

- which are multiplications ?
- in all ?

Solutions which are multiplications

- For all $\boldsymbol{M}_\xi$ (an n-dimensional space) : $\boldsymbol{S}_{\boldsymbol{M}_\xi}[i] \in \boldsymbol{V}$.
- Enforcing

$$\boldsymbol{S}_{\boldsymbol{M}_\xi}[i] \in \tilde{\boldsymbol{V}}$$

results in r linear constraints.

The dimension of Multiplications is $n - r$

Overall solution space

- For a general $\boldsymbol{M}$, $\boldsymbol{S}_{\boldsymbol{M}}[i]$ is some vector of length $n(n-1)/2$.
- Enforcing

$$\boldsymbol{S}_{\boldsymbol{M}}[i] \in \tilde{\boldsymbol{V}}$$

results in $n(n-1)/2 - (n - r)$ linear constraints.

The overall dimension of solutions is $n^2 - (n(n-1)/2 - (n - r))$

- The overall dimension is lower-bounded by the dimension of multiplications, which itself contain those in $\ker(L)$ $(d = 1)$.
- The dimension of the solutions is :

$$\max\left\{n^2 - (n(n-1)/2 - (n-r))\ ;\ n-r\ ;\ 1\right\}$$

- More generally, for k coordinates, this dimension is :

$$\max\left\{n^2 - k(n(n-1)/2 - (n-r))\ ;\ n-kr\ ;\ 1\right\}$$

Recovering Non-Trivial Multiplications

$$\dim(\text{Solutions}[k]) = \max\left\{n^2 - k(n(n-1)/2 - (n-r))\ ;\ n - kr\ ;\ 1\right\}$$

When $r \leq (n-2)/3$

- At $k = 3$, the first term is negative.
- Only multiplications are expected, with dimension :

$$\max\left\{n - 3r\ ;\ 1\right\}$$

- It contains non-trivial multiplications as soon as :

$$n - 3r > 1 \iff r \leq \frac{n-2}{3}$$

When $r \leq (n-2)/2$

- At $k = 2$, the solution space has dimension :

$$n^2 - 2(n(n-1)/2 - (n-r)) = 3n - 2r \ll n^2/2$$

- The dimension of multiplications in it is : $n - 2r < \epsilon.n$.

We use sum and intersection to refine a multiplication subspace :

- Consider $k = \frac{1}{\epsilon}$ solutions spaces $E_1, \ldots, E_k$ for different pairs of coordinates.
- $(\sum_k E_k) \cap E_{k+1}$ contains only multiplications, and some are non-trivial when $r \leq (n-2)/2$.

Experimental Results

1. Multiplications Recovery : for the 3 proposed schemes :
   - SFLASHv2, FLASH : $r \simeq n/3$
   - SFLASHv3 : $r \simeq n/6$
2. Full C^* recovery : works as for the first attack.
3. Signature Forgery : uses Patarin's attack over C^*.

n	37	**37**	67	**67**	131
θ	11	**11**	33	**33**	33
q	2	**128**	2	**128**	2
r	11	**11**	11	**11**	11
Mult. Recovery	4s	**70s**	1m	**50m**	35m
C^* Recovery	7.5s	**22s**	2m	**10m**	7m
Forgery	0.01s	**0.5s**	0.02s	**2s**	0.1s

Note : parameters in bold are those of SFLASHv2 and SFLASHv3.