Cryptanalysis of SFLASH

Henri Gilbert and Marine Minier

France Télécom R&D, 38-40, rue du Général Leclerc,
92794 Issy les Moulineaux Cedex 9 – France,
henri.gilbert@francetelecom.com

Abstract. SFLASH [Spec] is a fast asymmetric signature scheme intended for low cost smart cards without cryptoprocessor. It belongs to the family of multivariate asymmetric schemes. It was submitted to the call for cryptographic primitives organised by the European project NESSIE, and successfully passed the first phase of the NESSIE selection process in September 2001. In this paper, we present a cryptanalysis of SFLASH which allows an adversary provided with an SFLASH public key to derive a valid signature of any message. The complexity of the attack is equivalent to less than 2^{38} computations of the public function used for signature verification. The attack does not appear to be applicable to the FLASH companion algorithm of SFLASH and to the modified (more conservative) version of SFLASH proposed in October 2001 to the NESSIE project by the authors of SFLASH in replacement of [Spec].

Keywords: asymmetric signature, cryptanalysis, multivariate polynomials, SFLASH.

1 Introduction

SFLASH [Spec] is a an asymmetric signature scheme which was submitted to the call for cryptographic primitives organized by the European project NESSIE, together with a more conservative companion algorithm named FLASH. SFLASH was selected in September 2001 for phase II of the NESSIE project (whereas FLASH was not, probably because its longer key size makes it less attractive than SFLASH, assuming equivalent security levels [Nes01a]). No weakness of the SFLASH and FLASH algorithms was reported in the NESSIE security evaluation [Nes01b].

SFLASH and FLASH both belong to the family of multivariate asymmetric schemes [Pa00, Cou01], and do both represent particular instances of C^{*--}, a variant of the C^* scheme [MI88] in which a sufficient number r of public equations of the C^* trapdoor permutation are withdrawn in order to withstand Patarin's cryptanalysis of C^* [Pa95]. Both schemes are based on the difficulty of solving large systems of quadratic multivariate polynomials over a finite field K. Their trapdoor essentially consists in hiding a monomial transformation over an extension L of K, using two affine transformations s and t of the K-vector space K^n.

L.R. Knudsen (Ed.): EUROCRYPT 2002, LNCS 2332, pp. 288–298, 2002.

One of the distinctive properties of SFLASH and FLASH is that unlike most standard public key signature schemes (e.g. RSA, DSA, ECDSA, etc.), they are sufficiently fast to be well suited for implementation on low cost smart cards without cryptographic coprocessor. SFLASH and FLASH produce rather short signatures (259 bits in the case of SFLASH). The moderate public key size of SFLASH (2.2 Kbytes, versus 18 Kbytes for FLASH) represents an additional advantage for such applications.

In this paper, we present an attack of SFLASH which takes advantage of some special features introduced in SFLASH in order to save a substantial factor in the public key size as compared with more general instances of C^{*--} such as FLASH. This attack allows an adversary provided with an SFLASH public key to derive a valid signature, for that public key, of any message M. The complexity of the attack is well under the security target of 2^{80}: it is equivalent to less than 2^{38} computations of the SFLASH public function used for signature verification. Although the attack was not fully implemented, the essential parts were confirmed by computer experiments.

Our attack does not appear to be applicable to FLASH and to the modified (more conservative, at the expense of a larger public key size) version of SFLASH proposed in October 2001 to the NESSIE project by the authors of SFLASH, in replacement of [Spec].

This paper is organised as follows. Section 2 describes SFLASH and its connection to C^*. Section 3 gives an overview of the attack. Section 4 details the two most essential steps of the attack.

2 Outline of C^*, C^{*--}, and SFLASH

In this Section, we briefly outline those features of C^* and its cryptanalysis which are relevant for the attack presented here, and then provide a short description of SFLASH.

2.1 C^*

C^* is a trapdoor permutation based on hidden monomial field equations proposed by Matsumoto and Imai in 1988 [MI88]. An efficient attack of C^* was found by Patarin [Pa95]. It is sufficient for the sequel to only consider the basic version of C^*, which can be summarised as follows:

- K denotes a finite field of characteristic 2: $K = F_{2^m} = F_q$, where $q = 2^m$.
- L denotes an extension of K of degree n: $L = F_{q^n}$. The representation of L associated with a $P(X)$ irreducible polynomial of degree n of $K[X]$ is used in the various computations. Thus any element a of L can be represented as the $\sum_{i=0}^{n-1} a_i X^i$ element of $K[X]/P(X)$. We will denote in the sequel by φ the one to one mapping from K^n to L associated with this representation: $\forall c = (c_0, c_1, .., c_{n-1}) \in K^n$, $\varphi(c) = \sum_{i=0}^{n-1} c_i X^i \bmod P(X)$.

- The public key of C^* consists of a set of n quadratic functions from K^n to K which together define a G function from K^n to K^n

  $$\begin{array}{rcl} G: \quad K^n & \to & K^n \\ x = (x_0, .., x_{n-1}) & \mapsto & y = (y_0, .., y_{n-1}) \end{array}$$

  where

  $$y_i = \sum_{0 \leq j < k \leq n-1} \rho_{ijk} x_j x_k + \sum_{0 \leq j \leq n-1} \sigma_{ij} x_j + \tau_i$$

  (in other words, the public key is made up of the ρ_{ijk}, σ_{ij} and τ_i coefficients in K of the n public equations).
- The private key consists of two secret affine one to one functions of K^n: s and t (each determined by $n(n+1)$ K coefficients). The knowledge of s and t provides a secret representation of G as:

  $$G = t \circ \varphi^{-1} \circ F \circ \varphi \circ s$$

  where

  $$\begin{array}{rcl} F : L & \to & L \\ a & \mapsto & b = a^{q^\theta + 1} \end{array}$$

  (θ being a public or private integer such that $q^\theta + 1$ be co-prime with $q^n - 1$). Note that since F is the pointwise product of the two L automorphisms $a \mapsto a$ and $a \mapsto a^{q^\theta}$, $\varphi^{-1} \circ F \circ \varphi$ (and thus G) is quadratic. Moreover, F is one to one, and its inverse F^{-1} is the monomial function $a \mapsto a^h$, where h is the inverse of $q^\theta + 1$ modulo $q^n - 1$.

The knowledge of the private key allows to compute the inverse of the G function. Thus G is a trapdoor permutation which was initially conjectured to be one way, and proposed as a public key encryption or signature function.

2.2 Attack of C^*

The main attack of C^* described in [Pa95] is based upon the following observation: the $b = a^{q^\theta + 1}$ equation of the F function implies $a^{q^{2\theta}} \cdot b = a \cdot b^{q^\theta}$ as can be seen is multiplying the former equation by $a^{q^{2\theta}}$. But the latter equation has the property that both the left and the right terms are “bilinear” in a and b. As a consequence, there exists “bilinear” equations of the form

$$\sum_{0 \leq j \leq n-1, 0 \leq k \leq n-1} \gamma_{jk} x_j y_k + \sum_{0 \leq j \leq n-1} \delta_j x_j + \sum_{0 \leq j \leq n-1} \epsilon_j y_j + \eta = 0$$

relating the $(x_0, .., x_{n-1})$ and $(y_0, .., y_{n-1})$ K^n input and output vectors of the G public function (i.e. equations of total degree 2 without any $x_j x_k$ or $y_j y_k$ term). It is shown in [Pa95] that the linear equations in γ_{jk}, δ_j, ϵ_j and η provided by a sufficient number of G input-output pairs allow to recover these unknown coefficients, and that once this has been done, the obtained vector space of solutions can be used to compute the inverse by G of any K^n element y at the expense of solving a small K-linear system. The complexity of the attack is about $m^2n^4\log n$.

J. Patarin, L. Goubin and N. Courtois investigated in [PGC98] the simple variant of C^* obtained by removing r of the public equations, say the r last ones. Thus the public key now consists of a G function from K^n to K^{n-r} given by $n-r$ quadratic equations over K. They came to the conclusion that the obtained variant of C^* (denoted by C^{*-}) can still be attacked if r is sufficiently small. However, attacks investigated in [PGC98] are not applicable when q^r is larger than say 2^{64}. The C^{*--} name was introduced to refer to C^{*-} variants for which q and r satisfy this condition. Unlike C^*, C^{*--} can only be used for signature purposes, not for encryption purposes.

2.3 Description of SFLASH

SFLASH is a special instance of C^{*--}, in which a particular choice of the s and t functions (and of the polynomials associated with the representation of K and L) enables to considerably shorten the public key size.

More precisely:

- K is chosen equal to $GF(2^7)$, i.e. $m = 7$ and $q = 2^7$. We denote by K' the $GF(2) = \{0, 1\}$ subfield of K. K elements are represented as 7-tuples of K' elements, using the representation of $GF(2^7)$ associated with the X^7+X+1 irreducible polynomial of $K'[X]$.
- L is chosen equal to $K[X]/P(X)$, where $P(X)$ is publicly known and equal to the $X^{37}+X^{12}+X^{10}+X^2+1$ irreducible polynomial of $K[X]$. (Note that all coefficients of $P(X)$ belong to K'). Thus n is equal to 37 and L elements can be represented as 37-tuples of K elements.
- The F monomial function of L involved in the secret representation of G is taken equal to $a \mapsto a^{128^{11}+1}$; in other words, θ is public and equal to 11.
- The number r of withdrawn equations is equal to 11. Thus $q^r = 2^{77} > 2^{64}$ and the C^{*--} condition is satisfied.
- The two secret affine functions s and t of $K^n = K^{37}$ are taken from a small subset of the bijective affine functions from K^n to K^n, namely those which can be represented by an $n \times n$ matrix and a $n \times 1$ column vector which $n \times (n + 1)$ coefficients do all belong to the K' subfield.

It is easy to see that as a consequence of the special choice of the s and t functions (and of the K and L representations), all the coefficients of the $n - r = 26$ public quadratic equations of the public function G belong to the $K' = GF(2)$ subfield. This results in a gain by a factor of approximately $m = 7$ in the length of the SFLASH public key.

In addition to the above mentioned s and t affine mappings, an SFLASH private key also contains a 80-bit secret key Δ, which acts as a pseudo-random generation seed in the signature generation process.

In order to sign a message M, the owner of a (s, t, Δ) private key performs the following operations.

- The $M1 = SHA - 1(M)$ and $M2 = SHA - 1(M1)$ 160-bit strings, the 182-bit string $V = M1_{0\to 159}||M2_{0\to 21}$ and the 77-bit string $W = SHA - 1(V||\Delta)_{0\to 76}$ are computed.
- V is divided into $n - r = 26$ strings $y_0, ..., y_{25}$ of length 7 bits each, representing 26 elements of K, and W is divided into $r = 11$ strings $y_{26}, .., y_{36}$ of length 7 bits each representing 11 elements of K. Let us denote the $(y_0, .., y_{25})$ 26-tuple by y, and the $(y_0, .., y_{25}, y_{26}, \cdots y_{36})$ 37-tuple by y^*.
- The secret function $s^{-1} \circ \varphi^{-1} \circ F^{-1} \circ \varphi \circ t^{-1}$ is applied to y^*. The obtained 37-tuple x of K elements represents the signature of M. In order to check that the x signature of an M message is valid, a verifier just needs to compute $G(x)$, using the 26 public quadratic equations of G, and to make sure that the obtained value is equal to y.

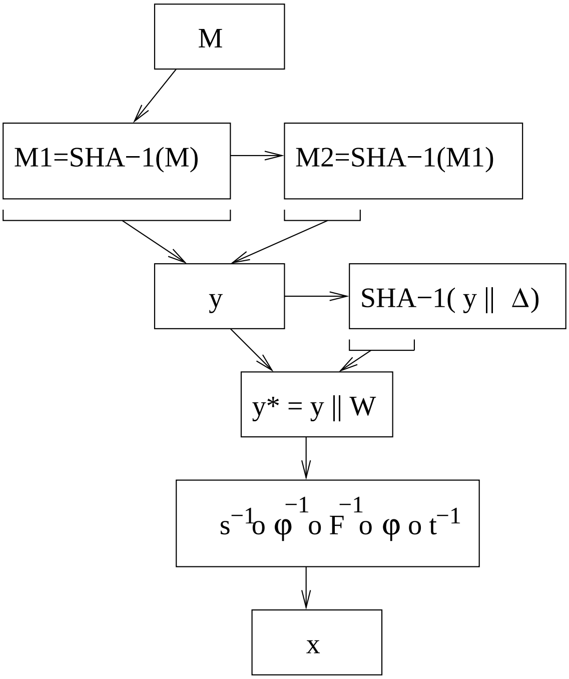

Fig. 1. SFLASH signature scheme

3 Overview of Our Attack

The following simple observation represents the starting point for our attack. Let us consider the $G^* = t \circ \varphi^{-1} \circ F \circ \varphi \circ s$ untruncated SFLASH transformation of K^n from which G is derived (G^* is given by the $n - r$ quadratic equations of G and r additional quadratic equations). Since the s, t and $\varphi^{-1} \circ F \circ \varphi$ mappings are constructed as to leave the $K'^n = GF(2)^{37}$ subset of K^n invariant (this is the price to pay for having very compact public key equations), G^* and its secret inverse $s^{-1} \circ \varphi^{-1} \circ F^{-1} \circ \varphi \circ t^{-1}$ used in the signature computations also leave K'^n invariant. In other words, the restriction of G^* (resp G) to K'^n induces a g^* (resp g) mapping of K'^n to K'^n (resp K'^n to K'^{n-r}) and since G^* is one to one, g^* is also one to one. [1] It is also worth noticing that G^* and g^*, though they are defined over distinct vector spaces (K^n and K'^n), are described by exactly the same set of n quadratic equations which coefficients belong by construction to K'.

Moreover, due to the fact that $K'^n = GF(2)^{37}$ is a small set, it is computationally easy to "invert" the public function g, i.e. given any 26-tuple y of K' elements, to determine the $class(y)$ set of all the $2^r = 2^{11}$ x values in K'^{37} such that $g(x) = y$. Our attack makes an extensive use of this property.

The purpose of our attack is to find r additional quadratic equations of the form

$$z_i(x) = \sum_{0 \leq j < k \leq n-1} \alpha_{ijk} x_j x_k + \sum_{0 \leq j \leq n-1} \beta_{ij} x_j$$

(where the α_{ijk} and β_{ij} coefficients are K' elements) which, together with the $n - r$ G quadratic equations

$$y_i(x) = \sum_{0 \leq j < k \leq n-1} \rho_{ijk} x_j x_k + \sum_{0 \leq j \leq n-1} \sigma_{ij} x_j + \tau_i$$

represent a full C^* instance consistent with G. More formally, we want to find r additional quadratic equations such that there exists a t' one to one affine mapping of K^n with coefficients in K' such that

$$\begin{gathered} \forall x = (x_0, x_1, .., x_{n-1}) \in K^n, \\ (y_0(x), .., y_{n-r-1}(x), z_0(x), .., z_{r-1}(x)) = t' \circ \varphi^{-1} \circ F \circ \varphi \circ s(x) \end{gathered} \qquad (1)$$

Once any such set of n equations over K^n satisfying (1) have been determined, then the C^* attack of [Pa95] can be applied to compute the preimage of any K^n element in few operations, so that a valid signature of any message M can then be computed by the adversary, using the following procedure:

[1] The following even stronger property of g deserves being mentioned: the public function g represents a "restricted SFLASH" induced over K'^n by the initial SFLASH , with distinct parameters ($q' = 2$ whereas $q = 2^7$, $\theta' = 3$ whereas $\theta = 11$, $n' = n = 37$, $r' = r = 11$). The $q'^{r'} > 2^{64}$ condition of C^{*--} is not satisfied by this restricted SFLASH, since q'^r is only equal to 2^{11}. This mere property is sufficient to make the security of SFLASH suspicious, as first pointed out by Nicolas Courtois, Louis Goubin and Jacques Patarin in a discussion we had with them at an early stage of this work. However, we did not manage to apply the attacks of C^{*-} described in [PGC98] to the g function, so we are unsure that this property is sufficient to draw firm conclusions concerning the security of SFLASH. Therefore we mounted a different attack dedicated to SFLASH, which takes advantage of the small value of $q'^n = 2^{37}$, as explained in the rest of this paper.

- $V = SHA - 1(M)_{0\rightarrow 159} || SHA - 1(M1)_{0\rightarrow 21}$, is computed and is divided into $n - r = 26$ 7-bit strings $y_0, ..., y_{25}$, and $r = 11$ arbitrary additional 7-bit values $z_0, ..., z_{10}$ are selected ;
- The preimage of $(y_0, ..., y_{25}, z_0, ..., z_{10})$, which is computed using the C^* attack of [Pa95], is a valid signature of M.

It is easy to see (one simply to consider t' and t) that the

$$z_i(x) = \sum_{0 \leq j < k \leq n-1} \alpha_{ijk} x_j x_k + \sum_{0 \leq j \leq n-1} \beta_{ij} x_j$$

quadratic equations satisfying requirement (1) are those linear combinations of the $n - r$ public quadratic equations $y_i(x)$ (without their τ_i constants) and the r additional hidden quadratic equations (again without their τ_i constants) such that in addition the n quadratic functions $y_0(x), \cdots, y_{n-r-1}(x), z_0(x), \cdots, z_{r-1}(x)$ be linearly independent.

So, each of the $r = 11$ additional quadratic functions z_i we are trying to determine, belongs to the same 37-dimensional K'-vector space E of quadratic functions, generated by the 37 public and hidden (constant less) quadratic equations. Our attack from now on consists in determining this partly unknown vector space E. (Once E has been found, any $z_0(x), .., z_{r-1}(x)$ functions of E such the n quadratic equations $y_0(x), .., y_{n-r-1}(x), z_0(x), .., z_{r-1}(x)$ be linearly independent can be used to mount the rest of the attack, using the C^* cryptanalysis of [Pa95].) There are two main steps in the determination of E:

The first step consists of an initial (partial) characterization of the coefficients of the $z_i(x)$ equations by expressing the fact that g^* is one to one. This first phase allows to reduce the set of $z_i(x)$ candidates from the K'-vector space of all quadratic functions constant less with K' coefficients, which dimension is $n(n-1)/2 + n = 703$, to a smaller K'-vector space E' of dimension $4 * 37 = 148$.

The second step consists of an enhanced characterization of the $z_i(x)$ coefficients. We are using the knowledge of E' to express additional conditions reflecting the a priori knowledge by the adversary of the degree in the y_0 to y_{n-1} variables of the quadratic functions of E'. Our computer experiments indicated that these additional conditions allow to fully determine the E set.

4 Detail of the Two Main Steps of the Attack

As said before, we attempt to characterize the 703 $GF(2)$-coefficients of any quadratic functions of E

$$z(x) = \sum_{0 \leq j \leq k \leq n-1} \alpha_{jk} x_j x_k + \sum_{0 \leq j \leq n-1} \beta_j x_j$$

(representing any of the $z_0(x)$ to $z_{10}(x)$ functions we are try to determine in order to extend the G set of 26 public equations to a complete set G^* of 37 equations representing a C^* instance.)

4.1 First Step of the Attack: Derivation of E'

For that purpose, we are expressing the fact that since g^* is one to one, each $class(y)$ of 2^{11} x preimages by g of any arbitrary element $y = (y_0, .., y_{26})$ of K'^{26} necessarily contains exactly $2^{r-1} = 2^{10}$ x values such that $z(x) = 0$ and $2^{r-1} = 2^{10}$ x values such that $z(x) = 1$, so that

$$\sum_{x \in class(y)} z(x) \equiv 0 \bmod 2$$

So any arbitrary y value provides one $GF(2)$-linear equation in the 703 coefficients of the quadratic function $z(x)$.

In order to compute the coefficients of the equation associated with y, one first needs to determine $class(y)$. This can be done with a total of less than 2^{37} computations and a limited amount of memory if we first select once for all the N arbitrary $y = (y_0, \cdots, y_{25})$ values for which we want to determine $class(y)$ and if we then perform an exhaustive computation of the g public function for all 2^{37} possible x input values, and store the $N2^r$ x preimages of the N selected y values. Once $class(y)$ has been determined, the $GF(2)$-coefficients of the corresponding equation are easy to compute, and equal to $\sum_{x \in class(y)} x_j x_k$ for each α_{jk} coefficient, and to $\sum_{x \in class(y)} x_j$ for each β_j coefficient.

We collect a little bit more than 703 such equations (say $N = 1000$ for instance) thus obtaining a $N \times 703$ matrix representing a system of N $GF(2)$-linear equations which right terms are equal to zero, and compute the kernel of this matrix using gaussian elimination.

Instead of the initially anticipated 37-dimensional $GF(2)$ vector space E spanned by the 26 public equations and the 11 hidden public equations without their constant terms, we found a much larger $GF(2)$-vector space E' of solutions, of dimension $37 * 4 = 148$. Unsurprisingly, E' is a superset of E.

4.2 Explanation of the Above Phenomenon

The reason why E' contains parasitic solutions distinct from the quadratic functions of the E set appears to be the following: $z(x) = \sum_{x \in class(y)} z(x)$ can be regarded as a $z(y^*)$ function of the actual (partly hidden) $y^* = (y_0, ..., y_{36}) = g^*(x)$ value. For any fixed $y = (y_0, \cdots, y_{25})$ value, let us denote by $V_{11}(y)$ the $(y_0, \cdots, y_{25}) \times GF(2)^{11}$ affine subset of $GF(2)^{37}$. We can write

$$\sum_{x \in class(y)} z(x) = \sum_{y^* \in V_{11}} z(s^{-1} \circ \varphi^{-1} \circ F^{-1} \circ \varphi \circ t^{-1}(y^*)) =_{def} \sum_{y^* \in V_{11}} z(y^*)$$

In other words, $\sum_{x \in class(y)} z(x)$ can be expressed as an 11th order derivative of the $z(y^*)$ function of y^* induced by $z(x)$. Therefore, the equations of the previous Section are satisfied if $z(y^*)$ can be expressed as a boolean function of total degree at most 10 of the components of y^*.

Now, let us consider any $g_i(x) = \varphi^{-1} \circ f_i \circ \varphi \circ s(x)$ quadratic function of K'^{37} associated with any f_i monomial function $a \mapsto a^{2^i+1}$ of $L' = GF(2^{37})$. Let us use the q' and θ' notation of the footnote of Section 3 to refer to the parameters of the restricted SFLASH g. Since $b = t^{-1}(y^*)$ is equal to $g_3(x)$, x is equal to $b^{h'}$, where h' is the inverse of $q'^{\theta'} + 1 = 2^3 + 1$ modulo $2^{37} - 1$. Therefore, if $z(x)$ is equal to any linear combination of the outputs of $g_i(x)$, $z(x)$ can be expressed as a linear combination of the 37 $GF(2)$-components of $b^{h'.(2^i+1)}$. Thus the degree of $z(x)$ as seen as a $z(y^*)$ function of y^* is then bounded above by the Hamming weight of $h'.(2^i + 1)$ mod $2^{37} - 1$.

We computed $h'_i = h'.(2^i + 1)$ mod $2^{37} - 1$ for the i values between 0 and 36, and found exactly 4 h'_i values of weight at most 10 [2], of weights 1, 4, 7 and 10 respectively, namely $i = 3$, 9, 15 and 21. Thus the output bits of $g_3(x)$, $g_9(x)$ $g_{15}(x)$ and $g_{21}(x)$ are quadratic in x and can all be expressed as functions of degree at most 10 of y^* , so that any $z(x)$ linear combination of these 148 output bits satisfies the equations of the previous Section.

So in summary E' is the 148-dimensional vector space spanned by the 37 components of each of the g_3, g_9 g_{15} and g_{21} functions of $GF(2)^{37}$.

4.3 Second Step of the Attack: Derivation of *E*

We select an arbitrary $B = (\zeta_0(x), .., \zeta_{147}(x))$ basis of E' provided by the gaussian elimination of step 1, and now attempt to characterize the 148 $GF(2)$-coordinates γ_i in this basis of any $z(x) = \sum_{0 \leq i \leq 147} \gamma_i \zeta_i(x)$ element of E, in order to eliminate the $E' \backslash E$ set of “parasitic solutions”.

As said in the previous Section, each $\zeta_i(x)$ quadratic function and their $z(x)$ linear combination can be seen as a $\zeta_i(y^*)$ and a $z(y^*)$ boolean function of the y^* 37-tuple. We can notice that due to the structure of E', the total degree in $y_0, .., y_{36}$ of each of the $\zeta_i(y^*)$ functions is very likely to be equal to 10. If $z(x) \in E$ the total degree in $y_0, .., y_{36}$ of $z(y^*)$ is by definition equal to 1. Therefore, if $z(x)$ belongs to E, then any 12th degree derivative of each of the $z(y^*) \cdot \zeta_i(y^*)$ functions (which degree is at most $10 + 1 = 11$) is equal to zero. On the other hand, if $z(x)$ belongs to $E' \backslash E$, the degree of at least one of the $z(y^*) \cdot \zeta_i(y^*)$ functions (in practice of one of the $z(y^*) \cdot \zeta_0(y^*)$ and $z(y^*) \cdot \zeta_1(y^*)$ functions) can be expected to be at least $4 + 10 = 14$, due to the structure of E', so that the 12th degree derivative of $z(y^*) \cdot \zeta_0(y^*)$ or $z(y^*) \cdot \zeta_1(y^*)$ (or both) can then be expected to differ from the null function. This provides one non trivial linear equation in the γ_i unknown coefficients of $z(x)$.

For any fixed $y = (y_0, \cdots, y_{24})$ value, let us denote by $V_{12}(y)$ the affine subset of $GF(2)^{37}$ defined by $(y_0, \cdots, y_{24}) \times GF(2)^{12}$. For any arbitrary $(y_0, \cdots, y_{24})$ value we have

$$\sum_{y^* \in V_{12}(y)} z(y^*) \cdot \zeta_0(y^*) = 0 \text{ and } \sum_{y^* \in V_{12}(y)} z(y^*) \cdot \zeta_1(y^*) = 0.$$

For each y value, each of these two equations provides the cryptanalyst with a $GF(2)$-linear equation in the 148 unknown $GF(2)$ coefficients γ_i, as can be

[2] up to circular rotations of h'_i. Indeed, if $i1$ and $i2$ are such that $h'_{i1} = 2^\delta h'_{i2}$ mod $2^{37} -$ 1, then g_{i1} and g_{i2} are equal up to a linear monomial transformation, and span the same set of quadratic functions.

seen in rewriting the first equation (associated with ζ_0) as

$$\sum_{0\leq i\leq 147} \gamma_i \left(\sum_{x\in class(y_0,..,y_{24},0)\cup class(y_0,..,y_{24},1)} \zeta_i(x)\cdot\zeta_0(x) \right) \equiv 0 \bmod 2$$

We collect a little bit more than 148 such equations (say $N' = 200$, some of which being associated with $z(x)\cdot\zeta_0(x)$ and the other being associated with $z(x)\cdot\zeta_1(x)$), thus obtaining a $N' \times 148$ matrix representing a system of N' 148-bit vectors corresponding to $GF(2)$-linear equations which right terms are equal to zero. We compute the kernel of this matrix using gaussian elimination. It was confirmed by computer experiments that we obtain a kernel of dimension only 37, equal to the E subspace of E'. This completes step 2 of our attack.

Once E has been recovered with the above method, a complete G^* set of 37 K^n-quadratic functions with K' coefficient can be obtained (one just needs to complete the 26 public equations of G as to obtain a basis of E), and the C^* attack of [Pa95] can be applied to compute the inverse by G^* of any K^{37} element, so that a valid signature of any message M can be produced by the adversary.

4.4 Complexity of the Attack

The most complex calculation required by the attack is the exhaustive computation of the 2^{37} values of the public function g, which is needed to obtain the (at most) $N + 2N'$ sets of 2^{11} preimages required for the computations of step 1 and step 2.

The computations of step 1 are essentially the derivation of the $N = 1000$ linear equations in 703 variables and the gaussian elimination of the resulting $N \times 703$ system in step 1. So, the complexity of step 1 is bounded above by $N.703.2^{11} + \frac{N^3}{3} \leq 2^{31}$. In the same way, the complexity of the derivation of the $N' = 148$ linear equations in 703 variables and the gaussian elimination of the resulting $N' \times 148$ system in step 2 are bounded above by 2^{27}. Both complexities are far lower than 2^{37} computations of the SFLASH public function. Moreover the complexity of the attack of C^* presented in [Pa95] is here about 2^{27} computations. In summary, the overall complexity of the attack is bounded above by 2^{38}.

5 Conclusion

The attack presented in this paper uses extensively the fact that the SFLASH public function over K^{37} induces a restricted scale function over the much smaller vector space $GF(2)^{37}$.

Our attack does not seem applicable to more conservative instances of C^{*--} such as FLASH, because a more sophisticated method than the one used in our attack would then have to be found to determine complete sets of 2^r preimages of some C^{*--} outputs.

Acknowledgements

We thank Mehdi-Laurent Akkar, Nicolas Courtois, Louis Goubin, and Jacques Patarin for interesting discussions and helpful remarks, and for having provided us with the reference implementation of SFLASH we used for computer experiments.

References

[MI88] T. Matsumoto and H. Imai, “Public Quadratic Polynomial-tuples for efficient signature-verification and message encryption”. In *Advances in Cryptology – Eurocrypt'88*, pp. 419-453, LNCS 330, Springer Verlag, May 1988.

[Nes01a] NESSIE Phase I: Selection of Primitives, september 2001, available at `http://www.cryptonessie.org/`.

[Nes01b] Security Evaluation of NESSIE First Phase, september 2001, available at `http://www.cryptonessie.org/`.

[Pa95] J. Patarin, “Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88”. In *Advances in Cryptology – Crypto'95*, pp. 248-261, LNCS 963, Springer Verlag, August 1995.

[Pa00] J. Patarin, “La Cryptographie Multivariable”, mémoire d'habilitation à diriger des recherches, Université Paris VII, France, 2000.

[PGC98] N. Courtois, L. Goubin and J. Patarin, “C^{*-+} and HM: Variations around two Schemes of T. Matsumoto and H. Imai”. In *Advances in Cryptology – Asiacrypt 98*, pp. 35-49, LNCS 1514, Springer-Verlag, October 1998.

[Cou01] N. Courtois, “La Sécurité des Primitives Cryptographiques Basées sur des Problèmes Algébriques Multivariables: MQ, IP, MinRank, HFE”, PhD. dissertation, Université Paris VI, France, September 2001, available at `http://www.minrank.org/phd.pdf`.

[Spec] Specifications of SFLASH, NESSIE documentation, available at `https://www.cosic.esat.kuleuven.ac.be/nessie/workshop/`.