Cryptanalysis of

Multivariate Public Key Schemes

Adi Shamir

Department of Computer Science and Applied Mathematics

Weizmann Institute of Science

Joint work with

Vivien Dubois, Pierre-Alain Fouque, and Jacques Stern

École normale supérieure

The Bottom Line of This Talk:

The Bottom Line of This Talk:

- RSA - the original is still the best! Don't use substitutes!

Univariate vs Multivariate Cryptography

- The security of univariate schemes depends on the difficulty of solving a single algebraic equation in a single variable over a large domain, such as: $y = x^2 \ (mod n)$

Univariate vs Multivariate Cryptography

- The security of univariate schemes depends on the difficulty of solving a single algebraic equation in a single variable over a large domain, such as: $y = x^2 \ (mod n)$
- The security of this problem is related to the difficulty of factoring n, which is at most subexponential.

Univariate vs Multivariate Cryptography

- The security of multivariate schemes depends on the difficulty of solving many algebraic equations in many variables over a small domain, such as:
- $y_1 = x_1x_2 + x_2x_2 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_1 + x_2x_3 + x_3x_3$
  $y_3 = x_1x_2 + x_1x_3 + x_2x_3$

Univariate vs Multivariate Cryptography

- The security of multivariate schemes depends on the difficulty of solving many algebraic equations in many variables over a small domain, such as:
- $y_1 = x_1x_2 + x_2x_2 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_1 + x_2x_3 + x_3x_3$
  $y_3 = x_1x_2 + x_1x_3 + x_2x_3$
- This problem is NP-complete even for a system of quadratic equations modulo 2.

Univariate vs Multivariate Cryptography

- The security of multivariate schemes depends on the difficulty of solving many algebraic equations in many variables over a small domain, such as:
- $y_1 = x_1x_2 + x_2x_2 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_1 + x_2x_3 + x_3x_3$
  $y_3 = x_1x_2 + x_1x_3 + x_2x_3$
- This problem is NP-complete even for a system of quadratic equations modulo 2.
- Such schemes are faster than RSA on weak processors, and in addition no polynomial time quantum algorithms are known for solving large systems of algebraic equations.

The Matsumoto-Imai scheme (MI)

- One of the earliest proposals was the Matsumoto-Imai (MI) scheme from 1988.

The Matsumoto-Imai scheme (MI)

- One of the earliest proposals was the Matsumoto-Imai (MI) scheme from 1988.
- It was based on the dual representation of elements of a large extension field $\mathbb{F}_{q^n}$ as vectors of values over the base field $\mathbb{F}_q$:

$$x \leftrightarrow (x_1, x_2, \ldots, x_n), y \leftrightarrow (y_1, y_2, \ldots, y_n)$$

The Matsumoto-Imai scheme (MI)

- One of the earliest proposals was the Matsumoto-Imai (MI) scheme from 1988.
- It was based on the dual representation of elements of a large extension field $\mathbb{F}_{q^n}$ as vectors of values over the base field $\mathbb{F}_q$:

$$x \leftrightarrow (x_1, x_2, \ldots, x_n), y \leftrightarrow (y_1, y_2, \ldots, y_n)$$

- Any single univariate f over $\mathbb{F}_{q^n}$ can thus be represented by n multivariate algebraic mappings $y_i = f_i(x_1, x_2, \ldots, x_n)$ over $\mathbb{F}_q$.

The Matsumoto-Imai scheme (MI)

- The starting point of the MI scheme is the RSA-like univariate exponentiation $x \mapsto y = x^e$ over the extension field $\mathbb{F}_{q^n}$, which has an easily computable inverse mapping of the same form $y \mapsto x = y^d$.

The Matsumoto-Imai scheme (MI)

- The starting point of the MI scheme is the RSA-like univariate exponentiation $x \mapsto y = x^e$ over the extension field $\mathbb{F}_{q^n}$, which has an easily computable inverse mapping of the same form $y \mapsto x = y^d$.
- Problem: the representation of such a univariate f over $\mathbb{F}_{q^n}$ by their components f_i's over $\mathbb{F}_q$ may contain huge multivariate polynomials of very high degree, which are very hard to represent and to evaluate.

The Matsumoto-Imai scheme (MI)

- The main observation of Matsumoto and Imai was that for some values of e the representation of f over the base field is particularly simple.

The Matsumoto-Imai scheme (MI)

- The main observation of Matsumoto and Imai was that for some values of e the representation of f over the base field is particularly simple.
- Consider exponents of the special form $e = q^{\theta} + q^{\eta}$.

The Matsumoto-Imai scheme (MI)

- The main observation of Matsumoto and Imai was that for some values of e the representation of f over the base field is particularly simple.
- Consider exponents of the special form $e = q^\theta + q^\eta$.
- Raising x to any power of the form q^i in $\mathbb{F}_{q^n}$ is a linear operation since $(a + b)^q = a^q + b^q$.

The Matsumoto-Imai scheme (MI)

- The main observation of Matsumoto and Imai was that for some values of e the representation of f over the base field is particularly simple.
- Consider exponents of the special form $e = q^\theta + q^\eta$.
- Raising x to any power of the form q^i in $\mathbb{F}_{q^n}$ is a linear operation since $(a+b)^q = a^q + b^q$.
- $x^e = x^{q^\theta + q^\eta} = x^{q^\theta} \cdot x^{q^\eta}$ is thus the product of two linear mappings, which can be represented by a system of n quadratic expressions in the n variables x_i.

The Matsumoto-Imai scheme (MI)

- The main observation of Matsumoto and Imai was that for some values of e the representation of f over the base field is particularly simple.
- Consider exponents of the special form $e = q^\theta + q^\eta$.
- Raising x to any power of the form q^i in $\mathbb{F}_{q^n}$ is a linear operation since $(a+b)^q = a^q + b^q$.
- $x^e = x^{q^\theta + q^\eta} = x^{q^\theta} \cdot x^{q^\eta}$ is thus the product of two linear mappings, which can be represented by a system of n quadratic expressions in the n variables x_i.
- When the base field is $\mathbb{F}_2$: raising x to the powers 2, 4, 8, etc are represented by linear polynomials f_i; raising x to the powers 3, 5, 6 etc are represented by quadratic polynomials; raising x to the powers 7, 11, 14 etc are represented by cubic polynomials; and so on.

The Matsumoto-Imai scheme (MI)

- Matsumoto and Imai proposed using the binary base field $\mathbb{F}_q$ with $q = 2$ or $q = 2^t$ and encryption exponents of the form $e = q^\theta + 1$.

The Matsumoto-Imai scheme (MI)

- Matsumoto and Imai proposed using the binary base field $\mathbb{F}_q$ with $q = 2$ or $q = 2^t$ and encryption exponents of the form $e = q^\theta + 1$.
- Unlike the RSA scheme, the size $q^n - 1$ of the multiplicative group of $\mathbb{F}_{q^n}$ is known, and thus anyone can compute d from e. MI thus based the security of the scheme on the different principle of mapping obfuscation.

The MI scheme: Mapping Obfuscation

- For example, assume that the univariate exponentiation $F = x^5$ has the multivariate quadratic form:
- $y_1 = x_1x_1 + x_1x_3 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_2 + x_2x_2 + x_3x_3$
  $y_3 = x_1x_1 + x_1x_2 + x_2x_3 + x_3x_3$

The MI scheme: Mapping Obfuscation

- For example, assume that the univariate exponentiation $F = x^5$ has the multivariate quadratic form:
- $y_1 = x_1x_1 + x_1x_3 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_2 + x_2x_2 + x_3x_3$
  $y_3 = x_1x_1 + x_1x_2 + x_2x_3 + x_3x_3$
- Perform a linear invertible input transformation:
- $x_1 \leftarrow x_1 + x_3$, $x_2 \leftarrow x_1$, $x_3 \leftarrow x_2 + x_3$

The MI scheme: Mapping Obfuscation

- For example, assume that the univariate exponentiation $F = x^5$ has the multivariate quadratic form:
- $y_1 = x_1x_1 + x_1x_3 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_2 + x_2x_2 + x_3x_3$
  $y_3 = x_1x_1 + x_1x_2 + x_2x_3 + x_3x_3$
- Perform a linear invertible input transformation:
- $x_1 \leftarrow x_1 + x_3, x_2 \leftarrow x_1, x_3 \leftarrow x_2 + x_3$
- Perform a linear invertible output transformation:
- $y_1 \leftarrow y_2 + y_3, y_2 \leftarrow y_1 + y_2 + y_3, y_3 \leftarrow y_1 + y_2$

The MI scheme: Mapping Obfuscation

- For example, assume that the univariate exponentiation $F = x^5$ has the multivariate quadratic form:
- $y_1 = x_1x_1 + x_1x_3 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_2 + x_2x_2 + x_3x_3$
  $y_3 = x_1x_1 + x_1x_2 + x_2x_3 + x_3x_3$
- Perform a linear invertible input transformation:
- $x_1 \leftarrow x_1 + x_3, x_2 \leftarrow x_1, x_3 \leftarrow x_2 + x_3$
- Perform a linear invertible output transformation:
- $y_1 \leftarrow y_2 + y_3, y_2 \leftarrow y_1 + y_2 + y_3, y_3 \leftarrow y_1 + y_2$
- to get a different looking system of quadratic relations:
- $y_1 = x_1x_2 + x_2x_2 + x_2x_3 + x_3x_3$
  $y_2 = x_1x_1 + x_1x_3 + x_2x_3 + x_3x_3$
  $y_3 = x_1x_2 + x_1x_3 + x_2x_3$

The MI scheme: Mapping Obfuscation

- More formally, MI make the very specific quadratic mapping $\boldsymbol{F}$ corresponding to exponentiation look like a random quadratic mapping by mixing its input and output variables by two secret invertible linear mappings $\boldsymbol{U}$ and $\boldsymbol{T}$.

The MI scheme: Mapping Obfuscation

- More formally, MI make the very specific quadratic mapping F corresponding to exponentiation look like a random quadratic mapping by mixing its input and output variables by two secret invertible linear mappings U and T.
- The public key of the MI scheme is $P = T \circ F \circ U$ where $F(x) = x^{q^\theta+1}$ over $\mathbb{F}_{q^n}$. WLG we can assume that $e = q^\theta + 1$ is known, since θ can have at most n possible values.

The MI scheme: Mapping Obfuscation

- More formally, MI make the very specific quadratic mapping F corresponding to exponentiation look like a random quadratic mapping by mixing its input and output variables by two secret invertible linear mappings U and T.
- The public key of the MI scheme is $P = T \circ F \circ U$ where $F(x) = x^{q^\theta+1}$ over $\mathbb{F}_{q^n}$. WLG we can assume that $e = q^\theta + 1$ is known, since θ can have at most n possible values.
- The secret key is the pair of linear mappings U and T; Since it is easy to invert both the exponentiation and the linear mappings, anyone who knows the secret key can easily invert P, but for someone who does not know the secret key the problem seems to be hard.

Changing MI to SFLASH

- The MI scheme was broken by a very clever attack developed by Patarin in 1995

Changing MI to SFLASH

- The MI scheme was broken by a very clever attack developed by Patarin in 1995
- Based on an idea of Shamir from 1993, Patarin et al proposed to avoid their attack by deleting ***r*** out of the ***n*** equations from the MI public key, and called the resulting scheme SFLASH.

Changing MI to SFLASH

- The MI scheme was broken by a very clever attack developed by Patarin in 1995
- Based on an idea of Shamir from 1993, Patarin et al proposed to avoid their attack by deleting r out of the n equations from the MI public key, and called the resulting scheme SFLASH.
- If we denote the final truncation Π, the SFLASH public key is: $P_\Pi = \Pi \circ T \circ F \circ U$

Changing MI to SFLASH

- The MI scheme was broken by a very clever attack developed by Patarin in 1995
- Based on an idea of Shamir from 1993, Patarin et al proposed to avoid their attack by deleting r out of the n equations from the MI public key, and called the resulting scheme SFLASH.
- If we denote the final truncation Π, the SFLASH public key is: $P_\Pi = \Pi \circ T \circ F \circ U$
- Such truncated keys can be used in signature schemes but not in encryption schemes, since they cannot be inverted uniquely.

Recommended Parameters for SFLASH

- The SFLASH scheme was selected in 2003 by the NESSIE European Consortium as one of only three recommended public key signature schemes, and as the best known solution for low cost smart cards

Recommended Parameters for SFLASH

- The SFLASH scheme was selected in 2003 by the NESSIE European Consortium as one of only three recommended public key signature schemes, and as the best known solution for low cost smart cards
- The first version of SFLASH, called SFLASH^{v1}, had a subtle bug which was discovered by Gilbert and Minier. It was replaced by two versions (SFLASH^{v2} and SFLASH^{v3}). They differ only in their recommended security parameters:
- for SFLASH^{v2} : $q = 2^7$, $n = 37$, $\theta = 11$ and $r = 11$
- for SFLASH^{v3} : $q = 2^7$, $n = 67$, $\theta = 33$ and $r = 11$

The Smile of the Cheshire Cat

- The original MI key can be viewed as a Cheshire cat. Each deleted equation eliminates some organ of the cat. It fades away, but its smile persists...

The Smile of the Cheshire Cat

- The original MI key can be viewed as a Cheshire cat. Each deleted equation eliminates some organ of the cat. It fades away, but its smile persists...
- Our goal is to reconstruct the original cat from its smile!

The Smile of the Cheshire Cat

- The original MI key can be viewed as a Cheshire cat. Each deleted equation eliminates some organ of the cat. It fades away, but its smile persists...
- Our goal is to reconstruct the original cat from its smile!
- This is information theoretically possible, since $\boldsymbol{T}$ and $\boldsymbol{U}$ contain a quadratic number of unknown entries, whereas the MI public key contains a cubic number of coefficients. In fact, the first few published quadratic equations should suffice to uniquely determine $\boldsymbol{T}$ and $\boldsymbol{U}$.

The Smile of the Cheshire Cat

- The original MI key can be viewed as a Cheshire cat. Each deleted equation eliminates some organ of the cat. It fades away, but its smile persists...
- Our goal is to reconstruct the original cat from its smile!
- This is information theoretically possible, since $\boldsymbol{T}$ and $\boldsymbol{U}$ contain a quadratic number of unknown entries, whereas the MI public key contains a cubic number of coefficients. In fact, the first few published quadratic equations should suffice to uniquely determine $\boldsymbol{T}$ and $\boldsymbol{U}$.
- The problem is how to recover the missing equations in a computationally efficient way.

The Basic Idea of the Attack

- Consider the public key

$$P_\Pi = \Pi \circ T \circ F \circ U$$

Each row in T randomly samples the linear space V spanned by the n quadratic expressions generated by $F \circ U$, but only $n - r$ of the n samples are published.

The Basic Idea of the Attack

- Consider the public key

$$P_\Pi = \Pi \circ T \circ F \circ U$$

Each row in T randomly samples the linear space V spanned by the n quadratic expressions generated by $F \circ U$, but only $n - r$ of the n samples are published.

- If we could replace T by another random T' before the truncation, we would get additional samples from V. By repeating this with several random T', we could transform even a single given equation into n linearly independent quadratic equations in V, which correspond to some other possible obfuscating matrix T'', and then apply Patarin's attack to this reconstructed Cheshire cat!

Finding Additional Elements in V

- We want to recreate the missing parts in T, but we can not do it from the output side. Instead, we will use laparoscopy, operating on the output transformation T in $P_\Pi = \Pi \circ T \circ F \circ U$ from the input side!

Finding Additional Elements in V

- We want to recreate the missing parts in T, but we can not do it from the output side. Instead, we will use laparoscopy, operating on the output transformation T in $P_\Pi = \Pi \circ T \circ F \circ U$ from the input side!
- Assume that we could multiply the input of F by some element $\xi \in \mathbb{F}_{q^n}$. Since $F = x^{q^\theta+1}$, this would multiply the output of F by another constant $\eta = \xi^{q^\theta+1} \in \mathbb{F}_{q^n}$.

Finding Additional Elements in V

- We want to recreate the missing parts in T, but we can not do it from the output side. Instead, we will use laparoscopy, operating on the output transformation T in $P_\Pi = \Pi \circ T \circ F \circ U$ from the input side!
- Assume that we could multiply the input of F by some element $\xi \in \mathbb{F}_{q^n}$. Since $F = x^{q^\theta+1}$, this would multiply the output of F by another constant $\eta = \xi^{q^\theta+1} \in \mathbb{F}_{q^n}$.
- Since multiplications by a field constant is a linear operation, its effect over $\mathbb{F}_q$ can be described by matrix multiplication. Denote these matrices for the input and output multiplications by M and K, respectively:

$$\Pi \circ T \circ F \circ U \circ M = \Pi \circ T \circ K \circ F \circ U$$

Finding Additional Elements in V

- We want to recreate the missing parts in T, but we can not do it from the output side. Instead, we will use laparoscopy, operating on the output transformation T in $P_\Pi = \Pi \circ T \circ F \circ U$ from the input side!
- Assume that we could multiply the input of F by some element $\xi \in \mathbb{F}_{q^n}$. Since $F = x^{q^\theta+1}$, this would multiply the output of F by another constant $\eta = \xi^{q^\theta+1} \in \mathbb{F}_{q^n}$.
- Since multiplications by a field constant is a linear operation, its effect over $\mathbb{F}_q$ can be described by matrix multiplication. Denote these matrices for the input and output multiplications by M and K, respectively:

$$\Pi \circ T \circ F \circ U \circ M = \Pi \circ T \circ K \circ F \circ U$$

- This randomizes T into $T' = T \circ K$!

The Basic Idea of the Attack

- Let Q be the space containing all the possible quadratic expressions, let V be the linear subspace spanned by the quadratic expressions in $T \circ F \circ U$, and let V_Π be the linear subspace spanned by the quadratic expressions in $\Pi \circ T \circ F \circ U$.

The Basic Idea of the Attack

- Let Q be the space containing all the possible quadratic expressions, let V be the linear subspace spanned by the quadratic expressions in $T \circ F \circ U$, and let V_Π be the linear subspace spanned by the quadratic expressions in $\Pi \circ T \circ F \circ U$.
- The dimension of Q is about $n^2/2$, the dimension of V is n, and the dimension of V_Π is $n - r$. For SFLASH^{v3}, the parameters are $n = 67$ and $r = 11$, and thus the dimensions of the three subspaces are:

$$V_\Pi(dim = 56) \subset V(dim = 67) \subset Q(dim = 2278)$$

The Basic Idea of the Attack

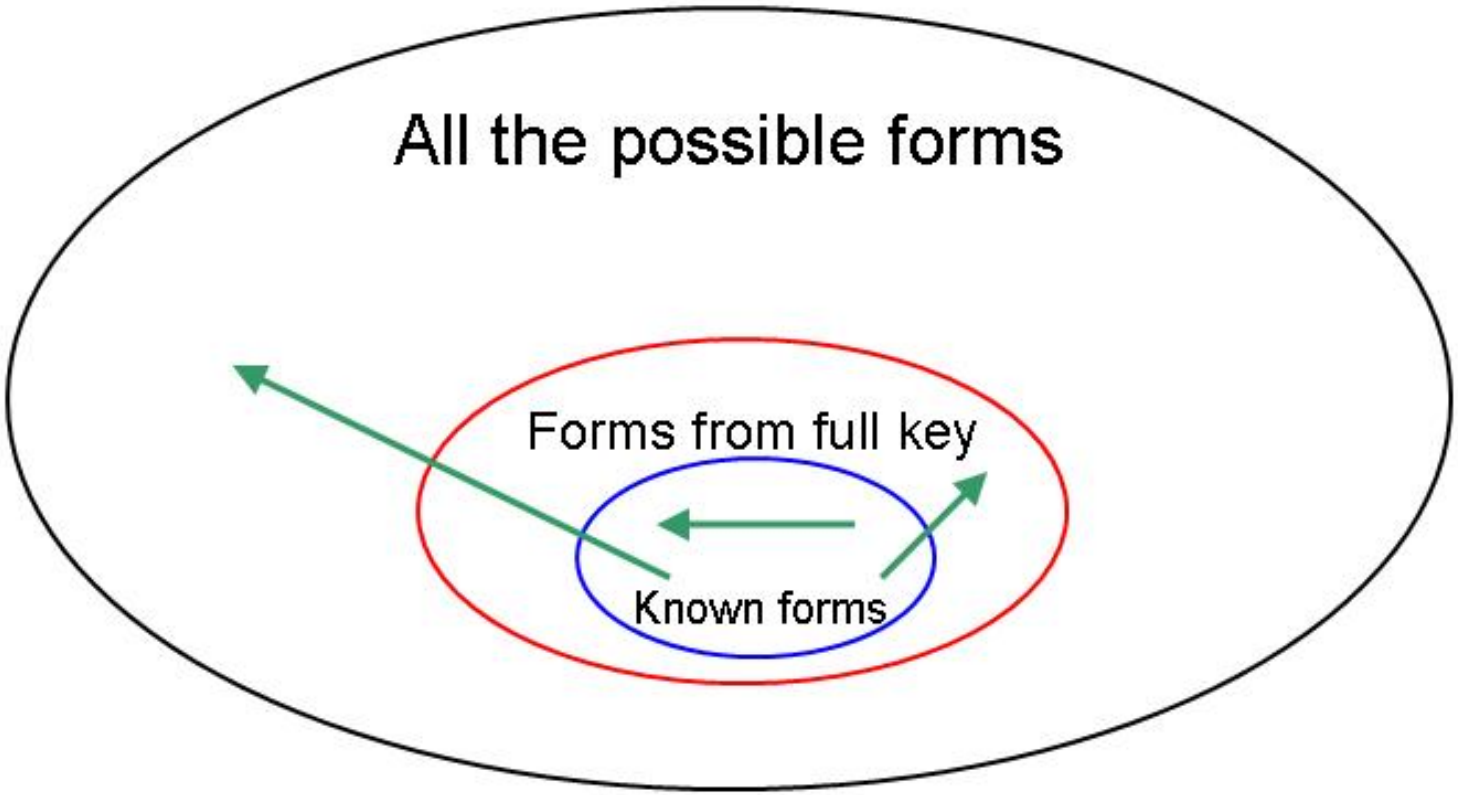
- Let Q be the space containing all the possible quadratic expressions, let V be the linear subspace spanned by the quadratic expressions in $T \circ F \circ U$, and let V_Π be the linear subspace spanned by the quadratic expressions in $\Pi \circ T \circ F \circ U$.
- The dimension of Q is about $n^2/2$, the dimension of V is n, and the dimension of V_Π is $n - r$. For SFLASH^{v3}, the parameters are $n = 67$ and $r = 11$, and thus the dimensions of the three subspaces are:

$$V_\Pi(dim = 56) \subset V(dim = 67) \subset Q(dim = 2278)$$

- Note that V_Π is a large subspace of V, but both of them are tiny subspaces of Q.

The Basic Idea of the Attack

The effect of various linear input transformations:



Finding Additional Elements in V

- Our goal now is to find a "good matrices" M whose effect on the inputs is equivalent to multiplication by some field constant $\xi \in \mathbb{F}_{q^n}$ after its obfuscation by the unknown U.

Finding Additional Elements in V

- Our goal now is to find a "good matrices" M whose effect on the inputs is equivalent to multiplication by some field constant $\xi \in \mathbb{F}_{q^n}$ after its obfuscation by the unknown U.
- Note that there are q^{n^2} possible matrices M over $\mathbb{F}_q$, but only q^n elements in $\mathbb{F}_{q^n}$, and thus the "good matrices" corresponding to extension field multiplications form a tiny linear subspace in the space of all matrices.

Finding Additional Elements in V

- Our goal now is to find a "good matrices" M whose effect on the inputs is equivalent to multiplication by some field constant $\xi \in \mathbb{F}_{q^n}$ after its obfuscation by the unknown U.
- Note that there are q^{n^2} possible matrices M over $\mathbb{F}_q$, but only q^n elements in $\mathbb{F}_{q^n}$, and thus the "good matrices" corresponding to extension field multiplications form a tiny linear subspace in the space of all matrices.
- The last part of the attack will find some good matrices by using the fact that they preserve membership of the output quadratic expressions in V. Since V is so sparse in Q, "bad matrices" (which do not make algebraic sense over $\mathbb{F}_{q^n}$) are extremely unlikely to have this property.

Characterizing Good Matrices with Algebraic Equations

- We would like find some good matrices M by solving a system of sufficiently many algebraic equations in its n^2 entries m_{ij}. Such equations are actually easy to derive, but they are quadratic equations, which cannot be solved efficiently.

Characterizing Good Matrices with Algebraic Equations

- We would like find some good matrices M by solving a system of sufficiently many algebraic equations in its n^2 entries m_{ij}. Such equations are actually easy to derive, but they are quadratic equations, which cannot be solved efficiently.
- To overcome this problem, we have to use calculus in addition to algebra, by using a differential operator to reduce the degree of the resultant algebraic equations.

The Differential Operator

- The differential DF of any univariate mapping $F(x)$ is defined as the bivariate mapping $DF(a, x)$ which is derived from $F(x)$ by the linear operator :

$$DF(a, x) = F(a + x) - F(a) - F(x) + F(0)$$

The Differential Operator

- The differential DF of any univariate mapping $F(x)$ is defined as the bivariate mapping $DF(a, x)$ which is derived from $F(x)$ by the linear operator :

$$DF(a, x) = F(a + x) - F(a) - F(x) + F(0)$$

- This operation can be carried out either over $\mathbb{F}_{q^n}$ or over $\mathbb{F}_q$. For example, if one of the expressions in the public key is:

$$x_1x_2 + x_2x_2 + x_2x_3 + x_3x_3$$

then its differential can be easily computed as:

$$(a_1+x_1)(a_2+x_2)+(a_2+x_2)(a_2+x_2)+(a_2+x_2)(a_3+x_3)+(a_3+x_3)$$
$$-(a_1a_2+a_2a_2+a_2a_3+a_3a_3)-(x_1x_2+x_2x_2+x_2x_3+x_3x_3)$$
$$= a_1x_2 + a_2x_1 + a_2x_3 + a_3x_2$$

The Differential Operator

- When $F(x) = x^{q^\theta+1}$, we get for all $a, x \in \mathbb{F}_{q^n}$

$$DF(a,x) = ax^{q^\theta} + a^{q^\theta}x.$$

The Differential Operator

- When $F(x) = x^{q^\theta+1}$, we get for all $a, x \in \mathbb{F}_{q^n}$

$$DF(a,x) = ax^{q^\theta} + a^{q^\theta}x.$$

- Note that this differential transformed the univariate quadratic expression $F(x)$ into a bivariate bilinear expression $DF(a,x)$, which is symmetric in a and x.

The Multiplicative Property of the Differential

- The differential $DF(a,x) = ax^{q^\theta} + a^{q^\theta}x$ of the internal polynomial $F(x) = x^{q^\theta+1}$ has a very interesting multiplicative property: For all $\xi \in \mathbb{F}_{q^n}$

$$DF(\xi \cdot a, x) + DF(a, \xi \cdot x) = (\xi + \xi^{q^\theta}) \cdot DF(a,x) \quad (1)$$

- Proof:

$$DF(\xi \cdot a, x) = \xi \cdot ax^{q^\theta} + \xi^{q^\theta} a^{q^\theta} x$$

$$DF(a, \xi \cdot x) = \xi^{q^\theta} \cdot ax^{q^\theta} + \xi \cdot a^{q^\theta} x$$

The sum of these expressions is:

$$(\xi + \xi^{q^\theta}) \cdot (ax^{q^\theta} + a^{q^\theta} x)$$

Using the Multiplicative Property of the Differential

- Note that this is a functional identity and not an equation, so it is true for any choice of $\boldsymbol{a}$, $\boldsymbol{x}$, and $\boldsymbol{\xi}$.

Using the Multiplicative Property of the Differential

- Note that this is a functional identity and not an equation, so it is true for any choice of $\boldsymbol{a}$, $\boldsymbol{x}$, and ξ.
- Due to the linearity of $\boldsymbol{D}$, $\boldsymbol{T}$ and $\boldsymbol{\Pi}$, the differential of the obfuscated public key (which can be easily computed by the attacker) has the same multiplicative property as the secret internal exponentiation.

Using the Multiplicative Property of the Differential

- Each symmetric bilinear form can be seen as a $n(n+1)/2$-dimensional vector, where the $a_i x_j$ products act as the dimensions. The linear space of all such forms is denoted by B.

Using the Multiplicative Property of the Differential

- Each symmetric bilinear form can be seen as a $n(n+1)/2$-dimensional vector, where the $a_i x_j$ products act as the dimensions. The linear space of all such forms is denoted by B.
- DP is a vector of n symmetric bilinear forms, and the space of all their linear combinations is denoted by W. We know only the $n - r$ of these bilinear forms which we can derive by differentiating the public key, and the space of their linear combinations is denoted by W_Π. B, W, and W_Π are the bilinear analogs of the quadratic spaces Q, V, and V_Π, respectively, and their dimensions for SFLASH^{v3} are:

$$W_\Pi(dim = 56) \subset W(dim = 67) \subset B(dim = 2278)$$

.

Using the Multiplicative Property of the Differential

- Matrices $\boldsymbol{M}$ which correspond to obfuscated multiplications satisfy the following multiplicative property:

$$\boldsymbol{T} \circ \boldsymbol{DF}(\boldsymbol{UMa}, \boldsymbol{Ux}) + \boldsymbol{T} \circ \boldsymbol{DF}(\boldsymbol{Ua}, \boldsymbol{UMx}) = \boldsymbol{C} \circ \boldsymbol{DF}(\boldsymbol{Ua}, \boldsymbol{Ux})$$

Using the Multiplicative Property of the Differential

- Matrices M which correspond to obfuscated multiplications satisfy the following multiplicative property:

$$T \circ DF(UMa, Ux) + T \circ DF(Ua, UMx) = C \circ DF(Ua, Ux)$$

- To demonstrate how such equations look like, consider the known differential of the first published expression which was $a_1x_2 + a_2x_1 + a_2x_3 + a_3x_2$, and assume that M and C have unspecified entries m_{ij} and c_{ij}. The multiplicative property can then be written as:

$$\begin{aligned}
&(m_{11}a_1 + m_{12}a_2 + m_{13}a_3)x_2 + (m_{21}a_1 + m_{22}a_2 + m_{23}a_3)x_1 \\
&+(m_{21}a_1 + m_{22}a_2 + m_{23}a_3)x_3 + (m_{31}a_1 + m_{32}a_2 + m_{33}a_3)x_2 \\
&+a_1(m_{21}x_1 + m_{22}x_2 + m_{23}x_3) + a_2(m_{11}x_1 + m_{12}x_2 + m_{13}x_3) \\
&+a_2(m_{31}x_1 + m_{32}x_2 + m_{33}x_3) + a_3(m_{21}x_1 + m_{22}x_2 + m_{23}x_3) \\
&= c_{11}DP_1 + c_{12}DP_2 + c_{13}DP_3
\end{aligned}$$

Using the Multiplicative Property of the Differential

- Rearranging these expressions, we get:

$$(m_{11}+m_{13})a_1x_1+(m_{12}+m_{23}+m_{33})a_1x_2+(m_{13}+m_{23})a_1x_3+\ldots$$

$$=(c_{11}+c_{13})a_1x_1+(c_{12}+c_{13})a_1x_2+(c_{11}+c_{12}+c_{13})a_1x_3+\ldots$$

Using the Multiplicative Property of the Differential

- Rearranging these expressions, we get:

$$(m_{11}+m_{13})a_1x_1+(m_{12}+m_{23}+m_{33})a_1x_2+(m_{13}+m_{23})a_1x_3+\ldots$$

$$=(c_{11}+c_{13})a_1x_1+(c_{12}+c_{13})a_1x_2+(c_{11}+c_{12}+c_{13})a_1x_3+\ldots$$

- When the a_ix_j products are considered as formal place holders, we can equate the coefficients of each product in order to get from this single functional relationship a quadratic number of linear equations in the quadratic number of unknown values of the m_{ij} and c_{ij} values!

How to Deal with a Truncated Public Key

- When we drop r expressions from the public key of SFLASH, the expression

$$T \circ DF(UMa, Ux) + T \circ DF(Ua, UMx)$$

should definitely be in unknown W, but has a probability of only q^{-r} to be in its known subspace W_Π defined by the $n - r$ expressions derived from the public key.

How to Deal with a Truncated Public Key

- When we drop r expressions from the public key of SFLASH, the expression

$$T \circ DF(UMa, Ux) + T \circ DF(Ua, UMx)$$

should definitely be in unknown W, but has a probability of only q^{-r} to be in its known subspace W_Π defined by the $n - r$ expressions derived from the public key.

- This probability (q^{-11} for the recommended parameters) is small, but since there are q^n (q^{67} for the recommended parameters) possible values of ξ, there will still be many solutions left. When r is small compared to n, we can use several (**6** for the recommended parameters) such conditions for different published expressions without losing all the "good matrices" M.

Using the Multiplicative Property of the Differential

- Consider once again the expression: $T \circ DF(UMa, Ux) + T \circ DF(Ua, UMx) = C \circ DF(Ua, Ux)$. When M is bad, the probability that it satisfies this relationship is about $q^{-n^2/2}$. Since there are about q^{n^2} bad matrices, we expect that all of them will be eliminated after considering 2-3 conditions.

Using the Multiplicative Property of the Differential

- Consider once again the expression: $T \circ DF(UMa, Ux) + T \circ DF(Ua, UMx) = C \circ DF(Ua, Ux)$. When M is bad, the probability that it satisfies this relationship is about $q^{-n^2/2}$. Since there are about q^{n^2} bad matrices, we expect that all of them will be eliminated after considering 2-3 conditions.
- The attack ends as a Shakespearean tragedy, in a major bloodbath: Almost everyone is killed, but there is a race between the few "good matrices" which are eliminated slowly as we add more conditions, and the many "bad matrices" which are eliminated quickly. Our hope is that all the "bad matrices" will be eliminated before we inadvertently kill off all the "good matrices", since we need at least one good matrix to finish our attack.

Dimension of the Solution Spaces

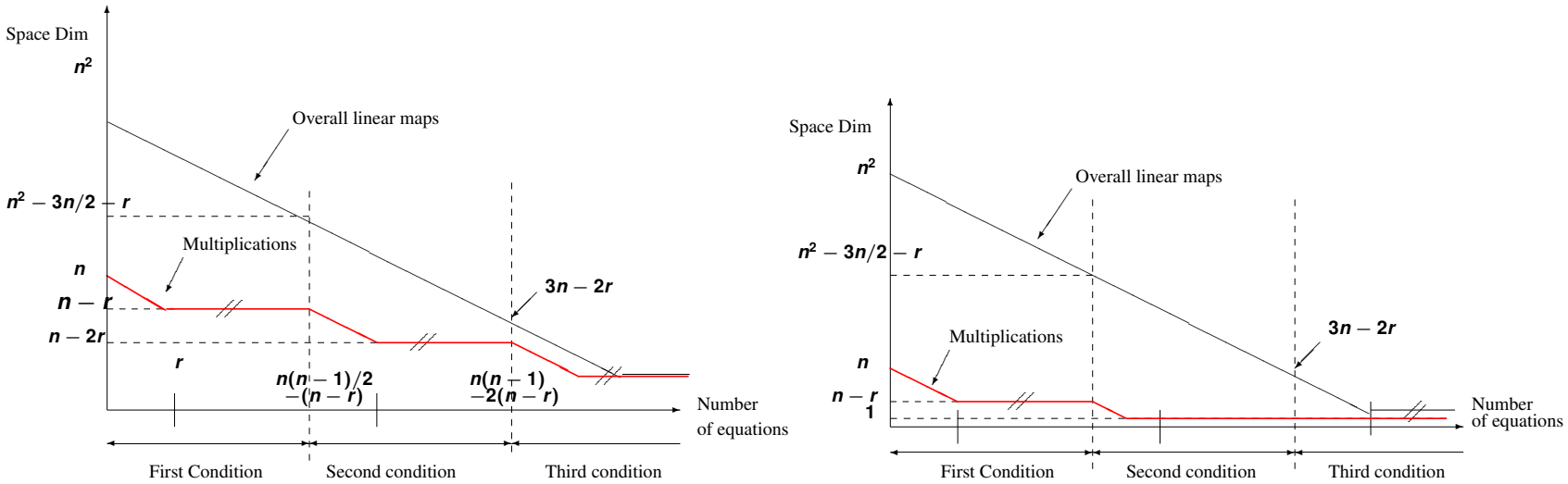


Figure: Evolution of the dimensions of the solution spaces of good and bad matrices when $r < n/3$ and when $r \geq n/3$, as we add more conditions.

Dimension of the Solution Spaces

- The attack described so far can deal with any $r < n/3$. In the paper we show how to extend it to deal with any $r < n/2$.

Dimension of the Solution Spaces

- The attack described so far can deal with any $r < n/3$. In the paper we show how to extend it to deal with any $r < n/2$.
- Note that even without this improvement, our technique is already sufficient to recover non-trivial multiplications for the recommended parameters of SFLASH^{v2} and SFLASH^{v3}, since $r = 11$ is smaller than both $37/3$ and $67/3$

Practical Results

- We carried our experiments on a 2GHz AMD Opteron PC using different parameters. The following table provides the time to recover a non-trivial multiplication and the time to recover the missing equations which replace those that were were deleted from the public key.

Practical Results

- We carried our experiments on a 2GHz AMD Opteron PC using different parameters. The following table provides the time to recover a non-trivial multiplication and the time to recover the missing equations which replace those that were were deleted from the public key.
- Note that both of these computations have to be carried out only once per public key, and then Patarin's attack requires about one second to forge an actual signature for any given message.

Practical results

n	37	37	67	67	131
θ	11	11	33	33	33
q	2	128	2	128	2
r	11	11	11	11	11
M Recovery	4s	70s	1m	50m	35m
MI Recovery	7.5s	22s	2m	10m	7m
Sig Forgery	0.01s	0.5s	0.02s	2s	0.1s

Conclusions

- Natural mathematical structures are very persistent, and once again we have shown that it is very hard to obfuscate them!