Inoculating Multivariate Schemes Against Differential Attacks

Jintai Ding and Jason E. Gower

Department of Mathematical Sciences
University of Cincinnati
Cincinnati, OH 45221-0025 USA

Email: ding@math.uc.edu, gowerj@math.uc.edu

Abstract. We demonstrate how to prevent differential attacks on multivariate public key cryptosystems using the Plus (+) method of external perturbation. In particular, we prescribe adding as few as 10 Plus polynomials to the Perturbed Matsumoto-Imai (PMI) cryptosystem when $g = 1$ and $r = 6$, where θ is the Matsumoto-Imai exponent, n is the message length, $g = \gcd(\theta, n)$, and r is the internal perturbation dimension; or as few as $g + 10$ when $g \neq 1$. The external perturbation does not significantly decrease the efficiency of the system, and in fact has the additional benefit of resolving the problem of finding the true plaintext among several preimages of a given ciphertext. We call this new scheme the Perturbed Matsumoto-Imai-Plus (PMI+) cryptosystem.

Keywords: multivariate, public key, cryptography, Matsumoto-Imai, perturbation, plus, differential.

1 Introduction

Though number theory based cryptosystems such as RSA are currently nearly ubiquitous, they are not appropriate for all implementations. Most notably, such schemes are not well-suited for use in small devices with limited computing resources. Multivariate public key cryptography provides one alternative since computations in small finite fields can be faster than working with large numbers. Furthermore, solving systems of multivariate quadratic polynomial equations over a finite field appears to be a difficult problem (analogous to integer factorization, though it is unknown precisely how difficult either problem actually is), so it seems reasonable to expect that we will be able to build secure multivariate public key cryptosystems and signature schemes from systems of quadratic polynomials that appear to be randomly chosen. Indeed, such systems may even resist future quantum computer attacks. In the last ten years, there has been significant effort put into realizing practical implementations, such as Matsumoto-Imai, HFE, HFEv, Sflash, Oil & Vinegar, Quartz, TTM, and TTS, to name but a few. So far the most secure encryption scheme seems to be HFE [13], though such an implementation with 2^{80} security would be very slow. On the other hand, Sflash [1] has been recommended by the New European Schemes for Signatures, Integrity, and Encryption (NESSIE, [11]) as a signature scheme for constrained environments.

Internal perturbation was recently introduced by Ding [3] as a general method to improve the security of multivariate public key cryptosystems. Roughly speaking, the idea is to "internally perturb" the system using a randomly chosen subspace of small dimension to create "noise" to be added to the system so that the resulting system still works efficiently and is much more difficult to break. The first application of this method was to the Matsumoto-Imai (MI) cryptosystem [10], a system that is otherwise vulnerable to the linearization attack [12]. The resulting system, called the perturbed Matsumoto-Imai cryptosystem (PMI), is slower as one needs to go through a search process on the perturbation space, though it is much faster than a 1024-bit implementation of RSA [15]. However, the recent attack of Fouque, Granboulan, and Stern [8] has shown that PMI is insecure. The basic idea of this attack is to use differentials to create a test for membership in the subset $\mathcal{K}$ of plaintexts that produce no noise. Once $\mathcal{K}$ is known, one can effectively "denoise" the system and thereby eliminate the internal perturbation. The linearization attack can then be applied to break the system as in the case of MI.

1.1 Our Results

In this paper we will show that PMI is easily protected from this attack by adding a small amount of external perturbation in the form of Plus (+) polynomials [14]. To put things in more concrete terms, let $g = \gcd(\theta, n)$, where θ is the Matsumoto-Imai exponent and n is the message length. Then by adding as few as 10 Plus polynomials to PMI when $g = 1$ and $r = 6$, or as few as $g + 10$ when $g \neq 1$, we will have a new scheme that resists the differential attack. The resulting scheme, called the Perturbed Matsumoto-Imai-Plus (PMI+) cryptosystem, uses the externally added random quadratic polynomials to create a situation in which almost all plaintexts satisfy the test for membership used in the differential attack on PMI. Not only is PMI+ then protected from the differential attack, we can use the theory of Markov chains to pick an optimal amount of perturbation so that the resulting efficiency degradation is slight. Moreover, the extra Plus polynomials can be used to solve the problem of finding the true plaintext from among several preimages of a given ciphertext.

1.2 Outline of the Paper

The remainder of this paper is organized as follows. After briefly recalling MI and PMI in Section 2.2, we describe the differential attack on PMI in Section 3. We show how to protect PMI from the differential attack in Section 4, and discuss how to use the theory of Markov chains to choose the optimal amount of external perturbation in the form of Plus polynomials. We conclude the paper in Section 5.

2 Matsumoto-Imai and Perturbed Matsumoto-Imai

In this section we provide a brief description of the Matsumoto-Imai cryptosystem, its variant, the Perturbed Matsumoto-Imai cryptosystem, and the most serious non-differential attacks on each.

2.1 Matsumoto-Imai

Let k be a finite field of size q and characteristic two, and fix an irreducible polynomial of $g(x) \in k[x]$ of degree n. Then $K = k[x]/(g(x))$ is an extension of degree n over k. We also have a k-vector space isomorphism $\phi : K \longrightarrow k^n$ defined by $\phi(a_0 + \cdots + a_{n-1}x^{n-1}) = (a_0, \ldots, a_{n-1})$. Fix θ so that $\gcd(1 + q^\theta, q^n - 1) = 1$ and define $F : K \longrightarrow K$ by

$$F(X) = X^{1+q^\theta}.$$

Then F is invertible and $F^{-1}(X) = X^t$, where $t(1 + q^\theta) \equiv 1 \bmod q^n - 1$. Define the map $\tilde{F} : k^n \longrightarrow k^n$ by $\tilde{F}(x_1, \ldots, x_n) = \phi \circ F \circ \phi^{-1}(x_1, \ldots, x_n) = (\tilde{F}_1, \ldots, \tilde{F}_n)$. In this case, the $\tilde{F}_i(x_1, \ldots, x_n)$ are quadratic polynomials in the variables $x_1, \ldots, x_n$. Finally, let L_1 and L_2 be two randomly chosen invertible affine transformation over k^n and define $\bar{F} : k^n \longrightarrow k^n$ by

$$\bar{F}(x_1, \ldots, x_n) = L_1 \circ \tilde{F} \circ L_2\,(x_1, \ldots, x_n) = (\bar{F}_1, \ldots, \bar{F}_n).$$

The public key of the Matsumoto-Imai cryptosystem (referred to as C^* or MI) consists of the polynomials $\bar{F}_i(x_1, \ldots, x_n)$. See [10] for more details.

2.2 Perturbed Matsumoto-Imai

Fix a small integer r and randomly choose r invertible affine linear functions $z_1, \ldots, z_r$, written

$$z_j(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_{ij} x_i + \beta_j,$$

for $j = 1, \ldots, r$. This defines a map $Z : k^n \longrightarrow k^r$ by $Z(x_1, \ldots, x_n) = (z_1, \ldots, z_r)$. Randomly choose n quadratic polynomials $f_1, \ldots, f_n$ in the variables $z_1, \ldots, z_r$. The f_i define a map $f : k^r \longrightarrow k^n$ by $f(z_1, \ldots, z_r) = (f_1, \ldots, f_n)$. Define $\tilde{f} : k^n \longrightarrow k^n$ by $\tilde{f} = f \circ Z$, and $\bar{\bar{F}} : k^n \longrightarrow k^n$ by

$$\bar{\bar{F}} = \tilde{F} + \tilde{f}.$$

The map $\bar{\bar{F}}$ is called the perturbation of $\tilde{F}$ by $\tilde{f}$, and as with MI, its components are quadratic polynomials in the variables $x_1, \ldots, x_n$. Finally, define the map $\hat{F} : k^n \longrightarrow k^n$ by

$$\hat{F}(x_1, \ldots, x_n) = L_1 \circ \bar{\bar{F}} \circ L_2(x_1, \ldots, x_n) = (y_1, \ldots, y_n),$$

where the L_i are randomly chosen invertible affine maps on k^n. The public key of the Perturbed Matsumoto-Imai (PMI) cryptosystem consists of the components y_i of $\hat{F}$. See [3] for more details.

Although for MI there is a bijective correspondence between plaintext and ciphertext, PMI does not enjoy this property. Indeed, for a given ciphertext $c \in k^n$, $\hat{F}^{-1}(c)$ may have as many as q^r elements, though we may add some redundancy to the plaintext in order to distinguish it from the other preimages.

2.3 Non-Differential Attacks on MI and PMI

Patarin's linearization attack [12] is the most successful attack against MI, and it is clear that it cannot be used to attack a general PMI with a reasonable r. However, Gröbner bases algorithms, such as Faugère's F_4 [6], can be used to attack any multivariate scheme. Though the exact running time complexity is unknown, there is evidence [5] which strongly suggests that PMI is resistant to attacks using F_4. More specifically, experiments from [5] indicate that within a reasonably range of n, a polynomial model is appropriate for predicting the security of PMI with $r < 6$, while an exponential model is appropriate for $r \geq 6$. For example, the exponential model is used to predict a security level of 2^{160} against F_4 for instances of PMI with parameters $(q, n, r, \theta) = (2, 136, 6, 40)$.

In the next section we will recall the new differential attack of Fouque, Granboulan, and Stern [8]. Both MI and PMI as previously described are susceptible to this attack. In particular, it is claimed that this attack applied to PMI will have a computation complexity of at most 2^{49} binary operations.

3 Differential Attack on PMI

We begin by establishing the notation used in the sequel; see [8] for proofs of quoted results. For each plaintext message $v \in k^n$, define the differential

$$L_v(x) = \hat{F}(x + v) + \hat{F}(x) + \hat{F}(v) + \hat{F}(0),$$

for a given instance of PMI. It is straightforward to show that L_v is linear in x.

Let $\mathcal{K}$ be the "noise kernel," the kernel of the linear part of the affine transformation $Z \circ L_2$. Then it can also be shown that

$$v \in \mathcal{K} \implies \dim(\ker(L_v)) = \gcd(\theta, n).$$

The differential attack amounts to computing a basis for $\mathcal{K}$, followed by q^r MI-type attacks, each attack being against PMI restricted to one of the q^r affine planes parallel to $\mathcal{K}$. For the MI-type attack to begin, $\mathcal{K}$ must be computed. In order to more clearly see how to thwart this attack, we now recall the particulars of this computation.

3.1 Testing for Membership in $\mathcal{K}$

For each $v \in k^n$, define the function T by

$$T(v) = \begin{cases} 1, & \text{if } \dim(\ker(L_v)) \neq \gcd(\theta, n); \\ 0, & \text{otherwise.} \end{cases}$$

Let $\alpha = P[T(v) = 0]$ and $\beta = P[v \in \mathcal{K}] = q^{-r}$; in other words, α is the probability that $T(v) = 0$, and β is the probability that $v \in \mathcal{K}$. We can use T to devise a test for detecting whether or not a given v is very likely to be in $\mathcal{K}$, assuming the following proposition: If for many different v'_i such that $T(v'_i) = 0$ we have $T(v + v'_i) = 0$, then $v \in \mathcal{K}$ with high probability. Suppose we pick N vectors $v'_1, \ldots, v'_N$ such that $T(v'_i) = 0$. Define $p(v) = P[T(v+v'_i) = 0 \,|\, T(v'_i) = 0]$. If v is chosen at random, then $p(v) = \alpha$; otherwise, $p(v) = \frac{\beta}{\alpha} + \frac{(\alpha-\beta)^2}{\alpha(1-\beta)}$. In this latter case it is not hard to show that $\frac{p(v)}{\alpha} - 1 = \frac{\beta}{1-\beta}(\frac{1}{\alpha} - 1)^2 \doteq \beta(\frac{1}{\alpha} - 1)^2$, where $\frac{\beta}{1-\beta} = \beta + \beta^2 + \beta^3 + \cdots \doteq \beta$ if β is very small. Thus we have the approximation $p(v) \doteq \alpha + \alpha\beta(\frac{1}{\alpha} - 1)^2$ whenever $v \in \mathcal{K}$. It follows that one way to decide whether or not $v \in \mathcal{K}$ is to approximate $p(v)$ and decide whether it is closer to α or $\alpha + \alpha\beta(\frac{1}{\alpha} - 1)^2$.

At this point we note that it seems more natural to consider the function $T'(v + v'_i) = \frac{1-T(v+v'_i)}{\alpha} - 1$, which has expected value $E[T'(v + v'_i)] = \frac{p(v)}{\alpha} - 1$, and then consider the average $\frac{1}{N}\sum_{i=1}^{N} T'(v + v'_i)$, which we expect to be close to $\frac{p(v)}{\alpha} - 1$, for large enough N by the Central Limit Theorem (see [7]). Then our task would be to determine whether this average is closer to 0 or $\beta(\frac{1}{\alpha} - 1)^2$.

The new function T' is defined as above in terms of T, and is such that

$$T'(v + v'_i) = \begin{cases} \frac{1}{\alpha} - 1, & \text{with probability } p(v); \\ -1, & \text{with probability } 1 - p(v). \end{cases}$$

Also $\mu = E[T'(v + v'_i)] = \frac{p(v)}{\alpha} - 1$ and $\sigma^2 = Var[T'(v + v'_i)] = \frac{p(v)(1-p(v))}{\alpha^2}$. Let X_i be independent and identically distributed random variables with the same distribution as T', and define $S_N = \sum_{i=1}^{N} X_i$. Then the Central Limit Theorem states that

$$P\left[\frac{S_N - N\mu}{\sigma\sqrt{N}} < x\right] \longrightarrow \mathfrak{N}(x) \quad \text{as} \quad N \longrightarrow \infty,$$

where

$$\mathfrak{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dx$$

is the standard normal distribution function. In other words, the Central Limit Theorem implies that the following approximation is valid for large N:

$$A_N \approx \mu + \frac{\sigma}{\sqrt{N}}\,\chi,$$

where $A_N = \frac{1}{N} S_N$ and χ is a random variable with standard normal distribution.

3.2 Efficiency of the Test

Suppose $v \in K$. In this case $\mu = \frac{p(v)}{\alpha} - 1 = \beta(\frac{1}{\alpha} - 1)^2$, and $\sigma^2 = \frac{p(v)(1-p(v))}{\alpha^2}$, which can be computed in terms of α and β. We also take $N = \frac{1}{(\alpha\beta)^2}$, as in

[8]. We first consider the probability that the question "$A_N > \beta(\frac{1}{\alpha} - 1)^2$?" will return true. Equivalently, we consider the probability that

$$\mu + \frac{\sigma}{\sqrt{N}}\chi > \beta\left(\frac{1}{\alpha} - 1\right)^2 = \mu,$$

which is the probability that $\chi > 0$. But this probability is $1 - \mathfrak{N}(0) = 1 - 0.5 = 0.5$. In other words, the "efficiency" of this test is such that it detects a vector $v \in \mathcal{K}$ (which is actually in $\mathcal{K}$) roughly half of the time. If we are to collect $n - r$ linearly independent vectors in $\mathcal{K}$, then we must perform on average $2(n - r)q^r$ tests.

3.3 Reliability of the Test

Let us now compute the probability that this question returns a false-positive; i.e., the question "$A_N > \beta(\frac{1}{\alpha} - 1)^2$?" returns true for $v \notin \mathcal{K}$. Here we must consider the probability that

$$\mu + \frac{\sigma}{\sqrt{N}}\chi > \beta\left(\frac{1}{\alpha} - 1\right)^2, \tag{1}$$

where now $\mu = 0$ and $\sigma^2 = \frac{1-\alpha}{\alpha}$. For example, if we take $\alpha = 0.59$ and $\beta = 2^{-6}$ as in the examples given in [8], then this is the probability that $\chi > 0.9819$, which is $1 - \mathfrak{N}(0.9819) \doteq 1 - 0.8369 = 0.1631$. This quantity gives us a measure of the "reliability" of this test in the sense that it tells us that roughly 16% of the $n-r$ vectors that our test leads us to believe are in $\mathcal{K}$ actually are *not* in $\mathcal{K}$. Though this might seem like a serious problem, it can be remedied by repeating the test a few times, each time with a different set of vectors $v'_1, \ldots, v'_N$. In the example above, by taking $8N$ vectors v'_i, performing the test 8 times with a new set of N vectors each time, and rejecting the vector v if any of the 8 tests fails, the probability that we correctly conclude that $v \in \mathcal{K}$ is $1 - (.1631)^8 \doteq 0.9999995$. This in turn means that the probability that there are no false-positives among our final set of $n - r$ vectors is $(1 - (.1631)^8)^{130} \doteq 0.9999349$. Therefore, if we perform 8 tests on $\frac{2(n-r)q^r}{0.1631}$ vectors, then the probability that we have $n - r$ vectors in $\mathcal{K}$ is 0.9999349.

We note that the above is a description of a modified version of Technique 1 for which a much higher degree of reliability is obtained. The authors in [8] do not necessarily require such a high level of reliability from Technique 1 since they also use Technique 2, which we have not yet addressed, as a filter to find those elements from Technique 1 which are actually in $\mathcal{K}$. Later in this paper we will show that Technique 2 will not be practical once we add external perturbation in the form of the Plus method. Therefore, we have presented Technique 1 is it must be implemented to be used without filters.

4 Preventing Differential Attacks on PMI

One way to prevent the differential attack is to perturb the system so that the dimension of the kernel of the differential L_v is the same for nearly every vector in k^n. This can be achieved by adding a sufficient number of randomly chosen quadratic polynomials according to the Plus method [14].

4.1 Perturbed Matsumoto-Imai-Plus

We now present the Perturbed Matsumoto-Imai-Plus cryptosystem. We will use the same notation as before. In particular, let L_2 and $\bar{\bar{F}}$ be as defined in Section 2.2. Randomly pick a quadratic polynomials $q_i(x_1, \ldots, x_n)$ and define the map $\bar{\bar{F}}^+ : k^n \longrightarrow k^{n+a}$

$$\bar{\bar{F}}^+ = \left(\bar{\bar{F}}_1, \bar{\bar{F}}_2, \ldots, \bar{\bar{F}}_n, q_1, ..., q_a\right).$$

Let $\hat{L}_1$ be a randomly chosen invertible affine map on k^{n+a} and define the map $\hat{F}^+ : k^n \longrightarrow k^{n+a}$ by

$$\hat{F}^+(x_1, \ldots, x_n) = \hat{L}_1 \circ \bar{\bar{F}}^+ \circ L_2(x_1, \ldots, x_n) = (\hat{y}_1, \ldots, \hat{y}_{n+a}),$$

The public key of the Perturbed Matsumoto-Imai-Plus (PMI+) cryptosystem consists of the $n + a$ quadratic polynomial components $\hat{y}_i$ of $\hat{F}$. Clearly PMI+ is simply PMI with a additional random quadratic polynomials (externally) mixed into the system by $\hat{L}_1$.

To decrypt, we must first invert $\hat{L}_1$. After we set aside the last a components, we can then apply the decryption process for the associated PMI. We note that the extra a components can be used to determine the true plaintext from among the (possibly q^r) preimages of the given ciphertext. We now study the effect that the Plus polynomials have on the computation of $\mathcal{K}$ using the differential attack.

4.2 PMI+ and the Effect on $\mathcal{K}$

We begin with the case where $\gcd(\theta, n) = 1$. Here $\dim(\ker(L_v)) = 1$ for every $v \in \mathcal{K}$. The fact that $\dim(\ker(L_v)) \neq 1$ for many $v \notin \mathcal{K}$ is the very fact that Technique 1 exploits in computing $\mathcal{K}$. So our task is to perturb PMI so that $\dim(\ker(L_v)) = 1$ for nearly every $v \notin \mathcal{K}$.

Consider the effect on the linear differential $L_v(x)$ upon adding Plus polynomials. We write $M_{v,a}$ for the matrix associated with the linear differential obtained after adding a Plus polynomials, and in particular, $M_{v,0}$ for the matrix associated with the linear differential L_v with no Plus polynomials. Let $R(a)$ be the rank of the matrix $M_{v,a}$. Note that $R(a) < n$, since $M_{v,a}\, v^T = 0$ for any a.

Suppose we add one more Plus polynomial (increase a by one). What is the probability that $R(a+1) = R(a) + 1$? Note that if $R(a) = n - 1$, then this probability is zero since $R(a) < n$. So let's assume $R(a) = n - i$, where $i = 2, 3, \ldots, n-1$. This probability is equivalent to the probability that we choose a new row-vector to be added to form $M_{v,a+1}$ from $M_{v,a}$, which is orthogonal to v and is not in the span of the row-vectors of $M_{v,a}$. The space of vectors orthogonal to v is of dimension $n-1$, and the span of the row-vectors of $M_{v,a}$ is of dimension $n-i$, hence the probability that $R(a+1) = R(a)+1$ will be $1-2^{1-i}$, where $i = 2, 3, ..., n-1$. Thus, if $n_{\delta,a}$ is the number of vectors v with $\dim(\ker(M_{v,a})) = \delta$, for a given a and $\delta = 1, 2, \ldots, n-1$, then we expect:

$$n_{\delta,a+1} = n_{\delta,a} \cdot 2^{1-\delta} + n_{\delta+1,a} \cdot \left(1 - 2^{-\delta}\right)$$

In order to obtain the distribution for $n_{\delta,a}$ when $a = 0$, and to predict how large we must choose a in order to protect PMI+ from the differential attack, we will use the language of Markov chains [9]. Let $P = (p_{ij})$ be the $n \times n$ matrix with entries given by:

$$p_{ij} = \begin{cases} 2^{-i+1}, & \text{if } i = j; \\ 1 - 2^{-i+1}, & \text{if } i = j+1; \\ 0, & \text{otherwise.} \end{cases}$$

Then for a fixed vector $v \in k^n$, p_{ij} gives the 1-step transition probability from state s_i to s_j upon appending a randomly chosen row vector to $M_{v,a}$, where state s_i corresponds to $\text{nullity}(M_{v,a}) = i$. Here s_1 is an absorbing state and for all other $i \neq 1$, s_i is a transient state.

Let $\mathcal{M}_v$ be the matrix associated with MI for a given v. Without loss of generality, assume that L_2 is chosen so the the perturbation Z is a function only of r variables, say $x_1, \ldots, x_r$. Adding the perturbation then is analogous to removing the first r columns of $\mathcal{M}_v$ and replacing them with r randomly chosen column vectors. Deleting r columns will increase the nullity to either $r+1$ with probability $\binom{n-1}{r}/\binom{n}{r} = 1 - \frac{r}{n}$, or r with probability $\binom{n-1}{r-1}/\binom{n}{r} = \frac{r}{n}$. If we then add r random column vectors to this matrix one at a time, the nullity will increase according to r-step transition probability matrix P_r^r, where P_r is the top-left $(r+1) \times (r+1)$ submatrix of P. In particular, if we let $\pi_0 = (0, 0, \ldots, 0, \frac{r}{n}, 1 - \frac{r}{n})$ be the initial state distribution vector, then $\pi_0 P_r^r$ can be used to calculate the probability that $\text{nullity}(M_{v,0}) = i$. For example, if $n = 31$ and $r = 6$, then these probabilities are given by:

$$\pi_0 P_6^6 = \begin{pmatrix} 0.350125 \\ 0.539086 \\ 0.106813 \\ 3.94582 \times 10^{-3} \\ 3.01929 \times 10^{-5} \\ 4.67581 \times 10^{-8} \\ 1.17354 \times 10^{-11} \end{pmatrix}$$

Finally, to obtain the probability that $\text{nullity}(M_{v,a}) = i$, we let $\pi' = \pi_0 P_r^r$ and compute $\pi' P^a$

We performed experiments to test the validity of our model. Each experiment was characterized by an instance of PMI defined by the parameters (q, n, r, θ), the number of Plus polynomials a, and κ randomly chosen test vectors. For each test vector v, we computed $\dim(\ker(M_{v,a}))$. Tables 1 and 2 report the observed (predicted) values of $n_{\delta,a}$ for two experiments performed with parameters $(q, n, r, \theta, \kappa) = (2, 31, 6, 2, 2^{15})$ and $(2, 36, 6, 4, 2^{15})$, respectively, each with $a = 0, 1, 2, \ldots, 11$. The predictions for $a = 0$ are obtained from the matrix $\pi' = \pi_0 P_r^r$, while the predictions for $a > 0$ are obtained by using the observed distribution from $a-1$ and the 1-step transition matrix P_r.

We note that although the predictions for $a = 0$ are not as accurate as those for $a > 0$, this is likely due to the fact that we chose the perturbation variables $z_1, \ldots, z_r$ in a simplified way for the experiments.

	$v \notin \mathcal{K}$				$v \in \mathcal{K}$
a	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$	$\delta = 1$
0	19003 (11304)	12182 (17404)	1081 (3448)	19 (127)	483
1	25081 (25094)	6906 (6902)	298 (287)	0 (2)	483
2	28548 (28534)	3660 (3676)	77 (74)	0 (0)	483
3	30366 (30378)	1896 (1888)	23 (19)	0 (0)	483
4	31334 (31314)	944 (965)	7 (6)	0 (0)	483
5	31810 (31806)	473 (477)	2 (2)	0 (0)	483
6	32040 (32046)	244 (238)	1 (0)	0 (0)	483
7	32154 (32162)	130 (123)	1 (0)	0 (0)	483
8	32208 (32219)	77 (66)	0 (0)	0 (0)	483
9	32246 (32246)	39 (38)	0 (0)	0 (0)	483
10	32263 (32266)	22 (20)	0 (0)	0 (0)	483
11	32278 (32274)	7 (11)	0 (0)	0 (0)	483

Table 1. Observed (predicted) values of $n_{\delta,a}$ for $(q, n, r, \theta, \kappa) = (2, 31, 6, 2, 2^{15})$ and $a = 0, 1, \ldots, 11$.

It remains to predict how large a must be in order to protect PMI+ against a differential attack. As was previously stated, the effect of adding Plus polynomials is to increase the value of α. In the example given in [8] $\alpha \doteq 0.59$ and so the question "$A_N > \beta(\frac{1}{\alpha} - 1)^2$?" is answered with a false-positive with the probability that $\chi > 0.9819$, which is 0.1631. Now suppose the attacker is willing to do as much as 2^{2w} work to correctly decide the answer to this test with this same probability. Then Inequality (1) becomes

$$\chi > \frac{\sqrt{N}}{\sigma}\left[\beta\left(\frac{1}{\alpha} - 1\right)^2 - \mu\right] = 2^{w-r}\left(\frac{1-\alpha}{\alpha}\right)^{3/2}.$$

If we assume that we are using Technique 1 as described in Section 4, then our total work (for the entire attack) will be

$$8N \cdot \frac{n^3}{6} \cdot \frac{2(n-r)q^r}{0.1631} \quad \doteq \quad 2^{2w+38.32},$$

	$v \notin \mathcal{K}$					$v \in \mathcal{K}$			
a	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$	$\delta = 5$	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$
0	14602 (101)	14942 (2274)	2610 (16272)	120 (1865)	2 (37)	0	0	0	492
1	21975 (22073)	9550 (9428)	722 (758)	28 (17)	1 (0)	0 (0)	0 (0) (0)	433 (430)	59 (62)
2	26693 (26750)	5367 (5316)	210 (205)	6 (4)	0 (0)	0 (0)	322 (325)	165 (160)	5 (7)
3	29380 (29376)	2838 (2841)	58 (58)	0 (1)	0 (0)	167 (161)	273 (285)	52 (46)	0 (1)
4	30810 (30799)	1457 (1462)	9 (14)	0 (0)	0 (0)	295 (304)	180 (176)	17 (13)	0 (0)
5	31519 (31538)	756 (735)	1 (2)	0 (0)	0 (0)	383 (385)	106 (103)	3 (4)	0 (0)
6	31916 (31897)	359 (379)	1 (0)	0 (0)	0 (0)	433 (436)	57 (55)	2 (1)	0 (0)
7	32095 (32096)	181 (180)	0 (0)	0 (0)	0 (0)	460 (462)	30 (30)	2 (0)	0 (0)
8	32205 (32186)	71 (90)	0 (0)	0 (0)	0 (0)	470 (475)	21 (16)	1 (0)	0 (0)
9	32246 (32240)	30 (36)	0 (0)	0 (0)	0 (0)	481 (480)	11 (11)	0 (0)	0 (0)
10	32258 (32261)	18 (15)	0 (0)	0 (0)	0 (0)	487 (486)	5 (6)	0 (0)	0 (0)
11	32270 (32267)	6 (9)	0 (0)	0 (0)	0 (0)	490 (490)	2 (2)	0 (0)	0 (0)

Table 2. Observed (predicted) values of $n_{\delta,a}$ for $(q, n, r, \theta, \kappa) = (2, 36, 6, 4, 2^{15})$ and $a = 0, 1, \ldots, 11$.

which if we want less than 2^{80} then we must have $w < 20.84$. This implies that we must take $2^{14.84} \left(\frac{1-\alpha}{\alpha}\right)^{3/2} < 0.9819$, or $\alpha > 0.998962$ if we wish to thwart this attack. To compute the value of a necessary to insure $\alpha > 0.998962$, we use the matrix P. In particular, we must compute a so that the first entry of $\pi' P^a$ is greater than 0.998962. If we take $n = 136$, $r = 6$, and $\gcd(\theta, n) = 1$, then we must take $a \geq 10$.

Finally, we consider $\gcd(\theta, n) \neq 1$. Let $g = \gcd(\theta, n)$. If $v \in \mathcal{K}$, then $\text{nullity}(M_{v,0}) = g$; otherwise $\text{nullity}(M_{v,0}) \in \{g - r, \ldots, g + r\}$. We must now add roughly g Plus polynomials just to get to a situation similar to the $g = 1$ case. Thus, by taking $a \doteq g + 10$, we can protect the special case of $g \neq 1$ from the differential attack.

4.3 Using Filters with the Differential Attack and Other Security Concerns

We now address Technique 2 of [8]. The idea of this technique is to look for a maximal clique in the graph with vertices $v \in k^n$ such that $T(v) = 0$, where two vertices v, v' are connected if $T(v + v') = 0$. Since $\mathcal{K}$ is a subspace of k^n, the elements of $\mathcal{K}$ form a clique. The hypothesis underlying Technique 2 is that if we look at a big enough subgraph then the maximal clique in this subgraph will consist almost exclusively of vectors from $\mathcal{K}$. However, by increasing the value of α near one, this clique is now very likely to have many elements *not* in $\mathcal{K}$ (in fact almost *every* element of k^n is in the clique) and therefore membership in this clique cannot be used as a filter to Technique 1.

We must be careful not to add too many extra polynomials since otherwise we may create a weakness to Gröbner bases attacks [2, 16]. From [5], we know that if we choose $r = 6$ and $n > 83$, then we can expect the PMI cryptosystem to have the security of 2^{80} against such an attack using F_4. In order to create a secure PMI+ scheme from these parameters, we suggest $(q, n, r, \theta) = (2, 84, 6, 4)$ and $a = 14$. Since we have added relatively very few extra polynomials, the attack complexity of F_4 will be essentially the same as it is for the corresponding PMI. Other secure implementations include the now-salvaged scheme $(q, n, r, \theta) = (2, 136, 6, 8)$ with $a = 18$, or any (q, n, r, θ) with $a = 11$, $g = 1$, $r = 6$ and $n > 84$. In summary, when designing PMI+, one must be careful with the choice of $g = \gcd(\theta, n)$, as $g + 10$ extra polynomials will be needed in order to defend against the differential attack, but if g is too large the extra polynomials may increase the vulnerability to a Gröbner basis attack.

Of course, it may also be possible to attack PMI+ by looking for ways to somehow separate the PMI polynomials from the Plus polynomials. If this was possible, the differential attack could then proceed as with PMI alone. However this approach has yet to be successfully applied to the MI-Minus-Plus cryptosystem [14], as we have no such method to differentiate between MI polynomials and random polynomials. Therefore, it seems unlikely that such an approach will be successfully applied to PMI+.

As we mentioned before, the extra Plus polynomials can be used to identify the true plaintext from among all preimages of a given ciphertext. Though the Plus polynomials slightly decrease the efficiency and increase the key sizes of the scheme, they do serve to both protect against the differential attack and aid in finding the true plaintext during the decryption process.

Recently, the perturbation method was also applied to the HFE cryptosystem to improve its security and efficiency [4]. Our preliminary experiments suggest that the differential analysis attack cannot be used to attack HFE, though further experiments and theoretical arguments are needed to confirm this hypothesis.

5 Conclusion

We have presented a method for preventing differential attacks against multivariate schemes. In particular, we have shown that by externally adding as few as 10 Plus polynomials in the case where $\gcd(\theta, n) = 1$, we create a new scheme (PMI+) which is resistant to the differential attack. Since very few extra polynomials are needed, the threat posed by Gröbner bases attacks is not significantly increased. If $g = \gcd(\theta, n) \neq 1$, then as few as $g + 10$ Plus polynomials will be needed to protect PMI+, though we do not claim PMI+ will be secure against Gröbner bases attacks if g is large. In any case, as long as the external perturbation is not too large, the efficiency of PMI+ will not be significantly degraded. In fact, the extra Plus polynomials can be used to identify the true plaintext from among all pre-images of a given ciphertext. For use in practical implementations, which will enjoy a security level of 2^{80}, we suggest that $n \geq 83$, $r = 6$ and $a = 14$ whenever $g \leq 4$. In particular the scheme $(q, n, r, \theta) = (2, 136, 6, 8)$ with $a = 18$ will be both very efficient and have a security level of 2^{80}. Sizes for the public keys of these implementations are roughly 41 kilobytes and 175 kilobytes, respectively.

References

1. M.-L. Akkar, N. T. Courtois, R. Duteuil, and L. Goubin. *A Fast and Secure Implementation of Sflash.* In *PKC 2003*, LNCS 2567:267–278.
2. N. Courtois, A. Klimov, J. Patarin, and A. Shamir. *Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations.* In *Eurocrypt 2000*, LNCS 1807:392–407.
3. Jintai Ding. *A New Variant of the Matsumoto-Imai Cryptosystem Through Perturbation.* In *PKC 2004*, LNCS 2947:305–318.
4. J. Ding and D. Schmidt. *Cryptanalysis of HFEv and Internal Perturbation of HFE.* In *PKC 2005*, LNCS 3386:288–301.
5. J. Ding, J. E. Gower, D. Schmidt, C. Wolf, and Z. Yin. *Complexity Estimates for the F_4 Attack on the Perturbed Matsumoto-Imai Cryptosystem.* In the proceedings of the *Tenth IMA International Conference on Cryptography and Coding*, LNCS, 3796:262–277.
6. Jean-Charles Faugère. *A New Efficient Algorithm for Computing Gröbner Bases (F_4).* In *Journal of Applied and Pure Algebra*, 139:61–88, June 1999.
7. William Feller. An Introduction to Probability Theory and Its Applications. Third edition, vol. I, Wiley & Sons, 1968.
8. P.-A. Fouque, L. Granboulan, and J. Stern. *Differential Cryptanalysis for Multivariate Schemes.* In *Eurocrypt 2005*, LNCS 3494:341–353.
9. J. G. Kemeny and J. L. Snell. Finite Markov Chains. D. Van Nostrand Company, Inc., 1960.
10. T. Matsumoto and H. Imai. *Public Quadratic Polynomial-Tuples for Efficient Signature-Verification and Message-Encryption.* In *Eurocrypt 1988*, LNCS 330:419–453.
11. NESSIE. European project IST-1999-12324 on New European Schemes for Signature, Integrity and Encryption. `http://www.cryptonessie.org`.
12. Jacques Patarin. *Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88.* In *Crypto 1995*, LNCS 963:248–261.
13. Jacques Patarin. *Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms.* In *Eurocrypt 1996*, LNCS 1070:33–48. Extended version: `http://www.minrank.org/hfe.pdf`.
14. J. Patarin, L. Goubin, and N. Courtois. *C^*_{-+} and HM: Variations Around Two Schemes of T. Matsumoto and H. Imai.* In *Asiacrypt 1998*, LNCS 1514:35–50.
15. B.-Y. Yang, J.-M. Chen, and Y.-H. Chen. Private communication.
16. B.-Y. Yang, J.-M. Chen, and N. Courtois. *On Asymptotic Security Estimates in XL and Gröbner Bases-Related Algebraic Cryptanalysis.* In *ICICS 2004*, LNCS 3269:410–413.