Reduction in Lossiness of RSA Trapdoor Permutation

Santanu Sarkar

Indian Statistical Institute, 203 B T Road, Kolkata 700 108, India sarkar.santanu.bir@gmail.com

Abstract. We consider the lossiness of RSA trapdoor permutation studied by Kiltz, O'Neill and Smith in Crypto 2010. In Africacrypt 2011, Herrmann improved the cryptanalytic results of Kiltz et al. In this paper, we improve the bound provided by Herrmann, considering the fact that the unknown variables in the central modular equation of the problem are not balanced. We provide detailed experimental results to justify our claim. It is interesting that in many situations, our experimental results are better than our theoretical predictions. Our idea also extends the weak encryption exponents proposed by Nitaj in Africacrypt 2012.

Keywords: Multi-Prime Φ -Hiding Problem, Lattice, Modular Equation.

1 Introduction

1.1 Multi-Prime Φ -Hiding Assumption

Multi-Prime RSA is a generalization of the RSA public key cryptosystem [13] where the modulus is a product of more than two primes, i.e., $N = p_1 \cdots p_m$, with p_i (for $1 \le i \le m$) primes of same bitsize. Note that for a fixed bit length of Multi-Prime RSA modulus N, the number of primes m can not be very large since, in that case one may efficiently extract the smallest factor of N using the Elliptic Curve Method for factorization [10].

 Φ -Hiding Assumption is one of the most well known assumptions in modern cryptography. It is used in various applications to produce secure primitives. For an RSA modulus N=pq and a prime e, the Φ -Hiding Assumption states that

"it is hard to decide whether
$$e$$
 divides $\Phi(N) = (p-1)(q-1)$,"

where $\Phi(\cdot)$ denotes the Euler's totient function. So the Φ -Hiding Problem is to deterministically predict whether a given prime e is a factor of $\Phi(N)$ or not, where only the knowledge of e and N is available.

It is well known that Φ -Hiding problem can be solved efficiently using the idea of Coppersmith [1] if $e \geq N^{0.25}$. In Asiacrypt 2008, Schridde and Freisleben [14] proved that the Φ -Hiding Assumption does not hold for the composite integers of the form $N = pq^{2k}$ for k > 0. These kind of moduli are known to be used in a variant of RSA called Takagi's RSA [15], which provides faster decryption.

A. Bogdanov and S. Sanadhya (Eds.): SPACE 2012, LNCS 7644, pp. 144–152, 2012.

[©] Springer-Verlag Berlin Heidelberg 2012

Multi-Prime Φ -Hiding Assumption has been proposed by Kiltz et al [8] in Crypto 2010, where they obtained standard model instantiations of RSA-OAEP by constructing a lossy trapdoor permutation from RSA, based on the multi-prime generalization of the Φ -Hiding Assumption.

In their protocol, they considered Multi-Prime RSA with modulus $N=p_1\cdots p_m$. The prime e is chosen such that e divides $p_1-1,\ldots,p_{m-1}-1$. The lossy trapdoor permutation then relies on the Multi-Prime Φ -Hiding Assumption, which states that

"it is hard to decide whether e divides $p_i - 1$ for all but one prime factor of N".

1.2 Cryptanalysis of Multi-Prime Φ -Hiding Assumption

Kiltz et al. [8] present a cryptanalysis of the Multi-Prime Φ -Hiding Assumption using the idea of Herrmann et al. [3]. Note that if e divides all p_i-1 for $1 \le i \le m$, $N \equiv 1 \mod e$. It gives a polynomial time distinguisher. To decide if e is Multi-Prime Φ -Hidden in N, consider the system of equations

$$ex_1 + 1 \equiv 0 \mod p_1$$
, $ex_2 + 1 \equiv 0 \mod p_2$, ..., $ex_{m-1} + 1 \equiv 0 \mod p_{m-1}$.

Kiltz et al. [8] construct a polynomial equation

$$e^{m-1}\left(\prod_{i=1}^{m-1} x_i\right) + \dots + e\left(\sum_{i=1}^{m-1} x_i\right) + 1 \equiv 0 \bmod \prod_{i=1}^{m-1} p_i \tag{1}$$

by multiplying all given equations. Then they linearize the polynomial and solve it using a result due to Herrmann and May [3]. However, the work of [3] provides an algorithm with runtime exponential in the number of unknown variables. So for large m, the idea of [3] will not be efficient.

Note that the coefficients of the polynomial in Equation (1) are all powers of e. In Africacrypt 2011, Herrmann [4] used this fact to improved the attack of [8], by considering a different linearization to reduce the number of variables. Suppose we have $(ex_1 + 1)(ex_2 + 1)(ex_3 + 1) \equiv 0 \mod p_1p_2p_3$. Then instead of considering the polynomial equation

$$e^3x_1x_2x_3 + e^2(x_1x_2 + x_1x_3 + x_2x_3) + e(x_1 + x_2 + x_3) + 1 \equiv 0 \mod p_1p_2p_3,$$
 (2)

Herrmann [4] considered the polynomial equation

$$e^2x + ey + 1 \equiv 0 \bmod p_1 p_2 p_3,$$
 (3)

where $x = ex_1x_2x_3 + x_1x_2 + x_1x_3 + x_2x_3$ and $y = x_1 + x_2 + x_3$ are the unknowns. One positive aspect of Equation (3) is that it has only two variables x, y instead of the original three x_1, x_2, x_3 . On the negative side, the size of the variable x is increased by a factor of e compared to the original unknown variables x_1, x_2, x_3 . However, the problem remains similar, as finding x, y allows one to factor N. In [4], it has been proved that considering Equation (3) provides an advantage in terms of better upper bounds on x_i than in the case with Equation (2).

In the general case, instead of considering the polynomial $e^{m-1}y_{m-1} + e^{m-2}y_{m-2} + \cdots + ey_1 + 1$ over the variables y_1, \ldots, y_{m-1} with root

$$(y_1, \dots, y_{m-1}) = \left(\prod_{i=1}^{m-1} x_i, \dots, \sum_{i=1}^{m-1} x_i\right),$$

Herrmann [4] considered the polynomial $e^2x + ey + 1$ over the variables x, y with root

$$(x_0, y_0) = \left(e^{m-3} \prod_{i=1}^{m-1} x_i + \dots + \sum_{j>i} x_i x_j, \sum_{i=1}^{m-1} x_i\right)$$
(4)

to obtain the improvement over the work of Kiltz et al. [8].

1.3 Our Contribution

In summary, we obtain the following attempts in analyzing the Multi-Prime Φ -Hiding problem for an RSA modulus with m factors:

- Kiltz et al. [8] used linear modular equation over m-1 variables. As in this case dimension of lattice will be exponential in m-1, method will not be efficient at all for large value of m.
- Herrmann [4] considered a bivariate modular polynomial. This makes the
 method efficient for larger values of m too. Also this gives better theoretical
 bound than the work of [8].

However, note that in Equation (4) used by Herrmann [4], the variable y_0 is much smaller than x_0 . It was already indicated in [3] that one may get better bound for these unbalanced variables. However this option has not been analyzed systematically in the literature till date. In this work we analyzed this issue carefully, and use the unbalanced property of the variables x, y to get further improvement over the result of Herrmann [4].

Our improvement originates from providing extra shifts over the variable y in the same bivariate scenario as Herrmann has considered. This reduces the lossiness of the work of Kiltz et al. [8] even further. In Table 1, we present the impact of our result on the work of Kiltz et al.

Table 1. Impact of our results on the lossiness of Kiltz et al. [8] for different values of m, with 2048 bit N and for 80 bit security.

Value	Lossiness in the work of Kiltz et al. [8]					
of m	Before the work of [4]	After the work of [4]	After our work			
4	806	778	768			
5	872	822	778			

We present the main technical result, our attack on the Multi-Prime Φ -Hiding Assumption, in Section 2, and the respective experimental results in Section 3. But before proceeding with the main content of this paper, let us state the following two existing results on lattices that will be required for our work. We first state the following due to Howgrave-Graham [5].

Lemma 1. Let $h(x_1, x_2) \in \mathbb{Z}[x_1, x_2]$ be the sum of at most ω monomials. Suppose that $h(x_1^{(0)}, x_2^{(0)}) \equiv 0 \pmod{N^m}$ where $|x_1^{(0)}| \leq X_1, |x_2^{(0)}| \leq X_2$ and

$$||h(x_1X_1,x_2X_2)||<\frac{N^m}{\sqrt{\omega}}.$$

Then $h(x_1^{(0)}, x_2^{(0)}) = 0$ over the integers.

We also note that the basis vectors of an LLL-reduced basis fulfill the following property (as explained in [9]).

Lemma 2. Let L be an integer lattice of dimension ω . The LLL algorithm applied to L outputs a reduced basis of L spanned by $\{v_1, \ldots, v_{\omega}\}$ with

$$||v_1|| \le ||v_2|| \le 2^{\omega/4} \det(L)^{1/(\omega-1)}$$

in polynomial time of dimension ω and the bit size of the entries of L.

Now we move on to the main technical content of this paper.

2 Our Attack on the Multi-Prime Φ -Hiding Assumption

Note that from Equation (4), value of y_0 is much smaller than x_0 . We use this fact to get the improvement over [4]. Our approach is exactly the same as [3] except that we use extra shifts over the variable y.

Theorem 1. Let $N = p_1 \cdots p_m$ be a Multi-Prime RSA modulus where p_i are of same bit size for $1 \le i \le m$. Let e be a prime such that $e > N^{\frac{1}{m} - \delta}$. Then one can solve Multi-Prime hidden Φ problem in polynomial time if there exist two non-negative real numbers τ_1, τ_2 such that

$$\Psi(\tau_1, \tau_2, \delta, m) = 3\tau_1 \tau_2^2 m - \tau_2^3 m + 3\tau_1^2 \delta m - 6\tau_1 \tau_2 m + 3\tau_2^2 m + 9\tau_1 \delta m + 6\tau_1 \tau_2 + 3\tau_1 m - 3\tau_2 m + 3\delta m - 9\tau_1 + 3\tau_2 + m - 3 < 0.$$

Proof. To decide if e is Multi-Prime Φ -hidden in N, consider the system of equations

$$ex_1 + 1 \equiv 0 \mod p_1, \quad \dots, \quad ex_{m-1} + 1 \equiv 0 \mod p_{m-1}$$

As p_i are of same bit size and $e > N^{\frac{1}{m} - \delta}$, we have $|x_i| \leq N^{\delta}$ for $1 \leq i \leq m - 1$. Denote $P = \prod_{i=1}^{m-1} p_i$. Now consider the polynomial $g(x,y) = e^2x + ey + 1$. It is clear that $g(x_0, y_0) \equiv 0 \mod P$ where

$$(x_0, y_0) = \left(e^{m-3} \prod_{i=1}^{m-1} x_i + \dots + \sum_{j>i} x_i x_j, \sum_{i=1}^{m-1} x_i\right).$$

From g(x, y), one can obtain a polynomial f(x, y) of the form $x + a_1y + a_2$ such that $f(x_0, y_0) \equiv 0 \mod P$. It is clear that the size of x_0 is dominated by the term $e^{m-3}x_1 \ldots x_{m-1}$. Hence we have

$$|x_0| \le N^{(m-3)(\frac{1}{m}-\delta)+(m-1)\delta} = N^{\frac{m-3}{m}+2\delta}$$
 and $|y_0| \le (m-1)N^{\delta}$.

As $m < \log_2 N$, we can assume $|y_0| \le N^{\delta}$, neglecting m-1 term.

Take two integers $X = N^{\frac{m-3}{m}+2\delta}$ and $Y = N^{\delta}$. Clearly X, Y is an upper bound on x_0, y_0 respectively.

Now consider the set of polynomials

$$g_{k,i}(x,y) = y^i f^k(x,y) N^{\max\{s-k,0\}},$$

for k = 0, ..., u, i = 0, ..., u - k + t where u is a positive integer and s, t are non-negative integers. Note that $g_{k,i}(x_0, y_0) \equiv 0 \mod P^s$.

Now we construct the lattice L spanned by the coefficient vectors of the polynomials $g_{k,i}(xX, yY)$. One can check that the dimension of the lattice L is

$$\omega = \sum_{k=0}^{u} \sum_{i=0}^{u-k+t} 1 \approx \frac{u^2}{2} + tu.$$

The determinant of L is

$$det(L) = \prod_{k=0}^{u} \prod_{i=0}^{u-k+t} X^k \cdot Y^i \cdot N^{\max\{s-k,0\}} = X^{s_X} Y^{s_Y} N^{s_N},$$
 (5)

where
$$s_X = \sum_{k=0}^{u} \sum_{i=0}^{u-k+t} k \approx t \frac{u^2}{2} + \frac{u^3}{6}$$
,
 $s_Y = \sum_{k=0}^{u} \sum_{i=0}^{u-k+t} i \approx \frac{t^2 u}{2} + \frac{t u^2}{2} + \frac{u^3}{6}$,
 $s_N = \sum_{k=0}^{u} \sum_{i=0}^{u-k+t} \max\{s-k,0\} \approx \frac{u s^2}{2} + \frac{t s^2}{2} - \frac{s^3}{6}$ assume $t \leq u$.

Using Lattice reduction on L by LLL algorithm [9], one can find two non-zero vectors b_1, b_2 such that $||b_1|| \leq ||b_2|| \leq 2^{\frac{\omega}{4}} (\det(L))^{\frac{1}{\omega-1}}$. The vectors b_1, b_2 are the coefficient vector of the polynomials $h_1(xX, yY), h_2(xX, yY)$ with

$$||h_1(xX, yY)|| = ||b_1||$$
 and $||h_2(xX, yY)|| = ||b_2||,$

where $h_1(x,y), h_2(x,y)$ are the integer linear combinations of the polynomials $g_{k,i}(x,y)$. Hence

$$h_1(x_0, y_0) \equiv h_2(x_0, y_0) \equiv 0 \mod P^s$$
.

To find two polynomials $h_1(x, y), h_2(x, y)$ which share the root (x_0, y_0) over integers, using Lemma 1 we get the condition

$$2^{\frac{\omega}{4}} (det(L))^{\frac{1}{\omega - 1}} < \frac{P^s}{\sqrt{\omega}}.$$
 (6)

Note that ω is the dimension of the lattice which we may consider as small constant with respect to the size of P and the elements of L. Thus, neglecting $2^{\frac{\omega}{4}}$ and $\sqrt{\omega}$, we can rewrite (6) as $det(L) < (P^s)^{\omega-1}$. In general [7], it is considered that the condition $det(L) < (P^s)^{\omega}$ is sufficient to find two polynomials $h_1(x,y), h_2(x,y)$ such that $h_1(x_0,y_0) = h_2(x_0,y_0) = 0$.

Under the assumption that $gcd(h_1, h_2) = 1$, we can collect the root (x_0, y_0) using resultant method. Let $t = \tau_1 u$ and $s = \tau_2 u$ where τ_1, τ_2 are non-negative reals. Now putting the value of t, s in the condition $det(L) < P^{s\omega}$, we get the required condition.

Remark 1. For fixed δ and m, we will take the partial derivative of Ψ with respect to τ_1, τ_2 and equate each of them to 0 to get non-negative solutions of τ_1, τ_2 . Given any pair of such non-negative solutions, if Ψ is less than zero, then for that δ , x_0, y_0 can be obtained efficiently.

Comparison with [4] and [8]: In the work of [4], the variable τ_1 was not involved. The bound on δ in [4] is presented as

$$\delta < \frac{2}{3\sqrt{m^3}}.$$

The bound on δ in the work of [8] is

$$\delta < \frac{2\left(m^{-1/(m-1)} - m^{-m/(m-1)}\right)}{m(m-1)}.$$

In Table 2, we present a comparison of the upper bounds of δ as in our case (Theorem 1) with those in [4] and [8], for different values of m.

From Table 2, it is clear that upper bound of δ in our case is higher than that of [4]. Hence our new attack solves the Multi-Prime Φ -Hidding Problem for more values of e. Also note that when m becomes larger, difference between the upper bound of δ in Theorem 1 and the upper bound of [4] increases.

Recently Tosu and Kunihiro [16] have studied Multi-Prime Φ -Hiding Problem. In [16, Section 4.4], authors have mentioned that their bound is same as Herrmann Method for m=3,4,5. Hence for m=4,5, our method is better than that of [16]. Also when m=10 with 4096 bit modulus, attack of [16] works when size of e is more than 314. However, in our case lower bound on size of e is $(0.1-0.0248)\times 4096=308$. Hence in this situation too, our method is better.

3 Experimental Results

We have implemented the programs in SAGE 3.1.1 over Linux Ubuntu 8.04 on a laptop with Dual CORE Intel(R) Pentium(R) D CPU 1.83 GHz, 2 GB RAM and 2 MB Cache. The results are as follows.

Value	Upper bound on δ					
of m	Our result (Theorem 1)	Herrmann [4]	Kiltz et al. [8]			
3	0.1283	0.1283	0.1283			
4	0.0835	0.0833	0.0787			
5	0.0608	0.0596	0.0535			
6	0.0475	0.0454	0.0388			
7	0.0387	0.0360	0.0295			
8	0.0327	0.0295	0.0232			
9	0.0283	0.0247	0.0188			
10	0.0248	0.0211	0.0154			

Table 2. Comparison of upper bound on δ between our result and those of [4] and [8]

In Table 3, we present few experimental results for different values of m.

m	δ	u	t	s	$\dim(L)$	time (sec.)
3	0.120	8	2	4	63	209.35
4	0.085	7	3	4	60	206.90
5	0.065	7	3	4	60	140.42
6	0.054	6	4	4	56	87.34
7	0.044	6	4	4	56	73.97
8	0.039	5	4	5	51	45.41
9	0.034	5	4	5	51	39.05
10	0.029	5	4	5	51	30.43

Table 3. Experimental results for different values of m with 2048 bit N

From Table 3, it may be noted that for $m \geq 4$ we get much better results in the experiments than the theoretical bounds. This is because, for the parameters we consider here, the shortest vectors may belong to some sublattice. However, the theoretical calculation in Theorem 1 cannot capture that and further, identifying such optimal sublattice seems to be difficult.

In [8, Proposition 5.3], Kiltz et al. proved that their construction provides $(m-1)(1/m-\delta-\epsilon)\log_2 N$ bits of lossiness for $\epsilon\log_2 N$ bit security. One can achieve 80 bit security by taking $\epsilon=0.04$ for 2048 bit modulus. In this case for m=3,4,5, Kiltz et al. showed that one can obtain 676, 778 and 822 bits lossiness respectively, considering the upper bound of δ as $\frac{2}{3\sqrt{m^3}}$. For m=4, we achieve the bound of δ as 0.085. This implies that actual lossiness in this case is less than $3\times(0.25-0.085-0.04)\times2048=768$ instead of 778. Similarly for m=5, actual lossiness is less than 778 instead of 822.

In many situations (like [6]), experimental results provide better bound than the theoritical prediction. Thus, any concrete parameters given in [8] for instantiating RSA-OAEP that depends on the Multi-Prime Φ -Hidding problem should need experimental verification.

3.1 Weak Encryption Exponents

Recently in Africacrypt 2012, Nitaj [12] proposed a new class of weak Encryption Exponents for RSA. The flow of the algorithm in [12] that exploits these weak keys is as follows.

- Consider that the public exponent e satisfies $ex+y \equiv 0 \mod p$ with $|x| < N^{\gamma}$, $|y| < N^{\delta}$ and $ex + y \neq 0 \mod N$.
- Use the idea of [3] to find x, y.
- Calculate $p = \gcd(N, ex + y)$

Nitaj [12] proved that when $\gamma + \delta \leq \frac{\sqrt{2}-1}{2} \approx 0.207$, one can find x,y in the above algorithm. He also estimated that the number of such encryption exponent is at least $N^{0.707-\epsilon}$ where $\epsilon \to 0$.

Note that when γ and δ are not same, i.e., x,y are of different bitsizes, we can improve the upper bound of $\gamma + \delta$ using our idea as in Theorem 1. In fact, when either $\gamma \to 0$ or $\delta \to 0$, it is already mentioned in [3] that the upper bound of $\gamma + \delta$ would be 0.25. Hence we have the following result, using an approach similar to that of [12, Theorem 5].

Theorem 2. Let N=pq be an RSA modulus with q . Let public exponent <math>e satisfy $ex + y \equiv 0 \mod p$ with $|x| < N^{\epsilon_1}$ and $|y| < N^{\delta}$. If $ex + y \neq 0 \mod N$ and $\delta < 0.25$, one can factor N in polynomial time where $\epsilon_1 \to 0$. The number of such encryption exponents is at least $N^{0.75-\epsilon}$, where $\epsilon \to 0$.

4 Conclusion

In this paper we consider Multi-Prime Φ -Hidding problem and provide better theoretical results than what were obtained by Herrmann [4]. For $m \geq 4$, the experimental results are better than our theoretical prediction. In this direction, an interesting open problem would be to provide a theoretical model for constructing the sublattice.

References

- Coppersmith, D.: Small Solutions to Polynomial Equations and Low Exponent Vulnerabilities. Journal of Cryptology 10(4), 223–260 (1997)
- Fujioka, A., Okamoto, T., Miyaguchi, S.: ESIGN: An Efficient Digital Signature Implementation for Smart Cards. In: Davies, D.W. (ed.) EUROCRYPT 1991. LNCS, vol. 547, pp. 446–457. Springer, Heidelberg (1991)

- 3. Herrmann, M., May, A.: Solving Linear Equations Modulo Divisors: On Factoring Given Any Bits. In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 406–424. Springer, Heidelberg (2008)
- 4. Herrmann, M.: Improved Cryptanalysis of the Multi-Prime ϕ Hiding Assumption. In: Nitaj, A., Pointcheval, D. (eds.) AFRICACRYPT 2011. LNCS, vol. 6737, pp. 92–99. Springer, Heidelberg (2011)
- Howgrave-Graham, N.: Finding Small Roots of Univariate Modular Equations Revisited. In: Darnell, M.J. (ed.) Cryptography and Coding 1997. LNCS, vol. 1355, pp. 131–142. Springer, Heidelberg (1997)
- Jochemsz, E., May, A.: A Polynomial Time Attack on RSA with Private CRT-Exponents Smaller Than N^{0.073}. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 395–411. Springer, Heidelberg (2007)
- Jochemsz, E., May, A.: A Strategy for Finding Roots of Multivariate Polynomials with New Applications in Attacking RSA Variants. In: Lai, X., Chen, K. (eds.) ASIACRYPT 2006. LNCS, vol. 4284, pp. 267–282. Springer, Heidelberg (2006)
- 8. Kiltz, E., O'Neill, A., Smith, A.: Instantiability of RSA-OAEP under Chosen-Plaintext Attack. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 295–313. Springer, Heidelberg (2010), http://eprint.iacr.org/2011/559
- Lenstra, A.K., Lenstra Jr., H.W., Lovász, L.: Factoring polynomials with rational coefficients. Mathematische Annalen 261, 515–534 (1982)
- Lenstra Jr., H.W.: Factoring integers with elliptic curves. Annals of Mathematics 126, 649–673 (1987)
- 11. May, A.: Secret Exponent Attacks on RSA-type Schemes with Moduli $N=p^rq$. In: Bao, F., Deng, R., Zhou, J. (eds.) PKC 2004. LNCS, vol. 2947, pp. 218–230. Springer, Heidelberg (2004)
- Nitaj, A.: A New Attack on RSA and CRT-RSA. In: Mitrokotsa, A., Vaudenay, S. (eds.) AFRICACRYPT 2012. LNCS, vol. 7374, pp. 221–233. Springer, Heidelberg (2012)
- Rivest, R.L., Shamir, A., Adleman, L.: A Method for Obtaining Digital Signatures and Public Key Cryptosystems. Communications of ACM 21(2), 158–164 (1978)
- Schridde, C., Freisleben, B.: On the Validity of the Ф-Hiding Assumption in Cryptographic Protocols. In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 344–354. Springer, Heidelberg (2008)
- 15. Takagi, T.: Fast RSA-type Cryptosystem Modulo p^kq . In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 318–326. Springer, Heidelberg (1998)
- Tosu, K., Kunihiro, N.: Optimal Bounds for Multi-Prime Φ-Hiding Assumption. In: Susilo, W., Mu, Y., Seberry, J. (eds.) ACISP 2012. LNCS, vol. 7372, pp. 1–14. Springer, Heidelberg (2012)