A New Efficient Asymmetric Cryptosystem for large data sets

M.R.K.Ariffin[1,2,[a]], M.A.Asbullah[1,2,[b]], and N.A.Abu[1,3,[c]]

[1] Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Universiti Putra Malaysia (UPM), Selangor, Malaysia
[2] Department of Mathematics, Faculty of Science, Universiti Putra Malaysia (UPM), Selangor, Malaysia
[3] Faculty of Information Technology and Communication, Unitversiti Teknikal Malaysia (UTeM), Melaka, Malaysia
[a]rezal@putra.upm.edu.my,[b]ma_asyraf@putra.upm.edu.my,[c]nura@utem.edu.my

Abstract. The Diophantine Equation Hard Problem (DEHP) is a potential cryptographic problem on a Diophantine equation. The DEHP has been in existence for "worst case scenario" of the RSA, Diffie-Hellman and El-Gammal schemes. However, the DEHP emerges after the exponentiation and modular reduction process. The proposed scheme (known as the AA_{β}-cryptosystem) is an asymmetric cryptographic scheme that utilizes this concept (without any prior mathematical operation) together with the factorization problem of two large primes. Its encryption speed has a complexity order faster than the Diffie-Hellman Key Exchange, El-Gammal, RSA and ECC. It can encrypt large data sets than its key size. It has a simple mathematical structure. Thus, it would have low computational requirements and would enable communication devices with low computing power to deploy secure communication procedures efficiently.

Keywords: Diophantine equation hard problem (DEHP), integer factorization problem, asymmetric cryptography

1 Introduction

The discrete log problem (DLP) and the elliptic curve discrete log problem (ECDLP) has been the source of security for cryptographic schemes such as the Diffie Hellman key exchange (DHKE) procedure, El-Gamal cryptosystem and elliptic curve cryptosystem (ECC) respectively [2], [7]. As for the world renowned RSA cryptosystem, the inability to find the e-th root of the ciphertext C modulo N from the congruence relation $C \equiv M^e(modN)$ coupled with the inability to factor $N = pq$ for large primes p and q is its fundamental source of security [8]. It has been suggested that the ECC is able to produce the same level of security as the RSA with shorter key length. Thus, ECC should be the preferred

Supported by Fundamental Research Grant Scheme #5523934 and Prototype Research Grant Scheme #5528100 Ministry of Higher Education, MALAYSIA.

asymmetric cryptosystem when compared to RSA [12]. Hence, the notion "cryptographic efficiency" is conjured. That is, to produce an asymmetric cryptographic scheme that could produce security equivalent to a certain key length of the traditional RSA but utilizing shorter keys. However, in certain situations where a large block needs to be encrypted, RSA is the better option than ECC because ECC would need more computational effort to undergo such a task [10]. Thus, adding another characteristic toward the notion of "cryptographic efficiency" which is it must be less "computational intensive" and be able to transmit large blocks of data (when needed). In 1998 the cryptographic scheme known as NTRU was proposed with better "cryptographic efficiency" relative to RSA and ECC [3] [4] [5]. NTRU has a complexity order of $O(n^2)$ for both encryption and decryption as compared to DHKE, EL-Gammal, RSA and ECC (all have a complexity order of $O(n^3)$). As such, in order to design a state-of-the-art public key mechanism, the following are characteristics that must be "ideally" achieved (apart from other well known security issues):

1. Shorter key length. If possible shorter than ECC 160-bits.
2. Speed. To have speed of complexity order $O(n^2)$ for both encryption and decryption.
3. Able to increase data set to be transmitted asymmetrically. That is, not to be restricted in size because of the mathematical structure.
4. Simple mathematical structure for easy implementation.

The Diophantine Equation Hard Problem (DEHP) as mentioned in this work has been described as the "worst case scenario" for the RSA, El-Gammal and DHKE schemes. Each of these schemes have their "worst case scenario" in some form of Diophantine equation. Observe the following:

1. For the case of the RSA problem, from $C = M^e - pqj$ for $j \in \mathbb{Z}$, the DEHP emerges after exponentiation and modular process. The power modulo complexity is $O(n^3)$.
2. The same goes for the DHKE (and El-Gammal). From $A = g^a - pj$ for $j \in \mathbb{Z}$, the DEHP also emerges after exponentiation and modular process. The power modulo complexity is $O(n^3)$.

However, in this work the DEHP is utilized in the first instance of the ciphertext representation without any prior "expensive" mathematical operation. Only basic multiplication is required without division or modulo operation.

The layout of this paper is as follows. In Section 2, the DEHP will be described. The mechanism of the AA_β-cryptosystem will be detailed in Section 3. In Section 4, the authors detail the decryption process and provide a proof of correctness. An example will also be presented. Continuing in Section 5, we will discuss a congruence attack, a Coppersmith type attack and a Euclidean division attack. An analysis of lattice based attack will be given in Section 6. Section 7 will be about the underlying security principles of the AA_β scheme together with security reduction results. Indistinguishability results of the AA_β scheme will be provided in Section 8. The ability to transmit large data sets will be described

in Section 9. A table of comparison between the AA_β scheme against RSA,ECC and NTRU is given in Section 10. Finally, we shall conclude in Section 11.

2 The Diophantine Equation Hard Problem

Definition 1. *Let $Y = \sum_{i=1}^{u} A_i x_i$ be a summation of unknown integers x_i which are of the same bit length and significantly greater than the bit length of A_i by n-bits where A_i is a public sequence of constants and $gcd(A_i, A_j) = 1$ where $i \neq j$. We define the DEHP is solved when Y is prf-solved. That is, the preferred integer set x_i^* is found from the set of all possible integers x_i such that $Y = \sum_{i=1}^{j} A_i x_i$.*

Proposition 1. *Consider the linear equation $Y = A_1x_1 + A_2x_2$ and $gcd(A_1, A_2) = 1$. Suppose the bit length of unknown integers x_1 and x_2 are significantly greater than the bit length of A_1 and A_2 by at least n-bits, then there exist exponentially many pairs of (x_1, x_2) that satisfy the equation.*

Proof. Let x_1 and x_2 be of length $(\nu n + n)$-bits long each and A_1 and A_2 be of length νn bits. Let x_1 and x_2 be the *prf*-solution for $Y = A_1x_1 + A_2x_2$ and $\gcd(A_1, A_2) = 1$. Let

$$x_1 = x_1^0 + A_2 t \tag{1}$$

and

$$x_2 = x_2^0 - A_1 t \tag{2}$$

for some $t \in \mathbb{Z}$. Since $\gcd(A_1, A_2) = 1$ and Y is of size $(2\nu + 1)n$-bits we have both x_1^0 and x_2^0 of size $(2\nu + 1)n$-bits. Since both A_1 and A_2 are νn-bits and x_1 and x_2 are $(\nu + 1)n$-bits, then the best case scenario is to choose from a set of possible $t's$ which consists of $2^{(\nu+1)n}$ elements. □

3 The AA_β Public Key Cryptosystem

Let us begin by stating that the communication process is between A (Along) and B (Busu), where Busu is sending information to Along after encrypting the plaintext with Along's public key.

• Key Generation by Along

INPUT: The size n of the prime numbers.
OUTPUT: A public key tuple (n, e_{A1}, e_{A2}) and a private key pair (pq, d).

1. Generate two random and distinct n-bit strong primes p and q satisfying
$$\begin{cases} p \equiv 3 (\text{mod } 4),\ 2^n < p < 2^{n+1}, \\ q \equiv 3 (\text{mod } 4),\ 2^n < q < 2^{n+1}. \end{cases}$$

2. Choose random d such that $d > (p^2q)^{\frac{4}{9}}$.
3. Choose random integer e such that $ed \equiv 1 (\text{mod } pq)$ and add multiples of pq until $2^{3n+4} < e < 2^{3n+6}$ (if necessary).
4. Set $e_{A1} = p^2q$. We have $2^{3n} < e_{A1} < 2^{3n+3}$.
5. Set $e_{A2} = e$.
6. Return the public key tuple (n, e_{A1}, e_{A2}) and a private key pair (pq, d).

We also have the fact that $2^{2n} < pq < 2^{2n+2}$.

• Encryption by Busu

INPUT: The public key tuple (n, e_{A1}, e_{A2}) and the message $\mathbf{M}$.
OUTPUT: The ciphertext C.

1. Represent the message $\mathbf{M}$ as a $4n$-bit integer m with $m = m_1 \cdot 2^n + m_2$ where m_1 is a $3n+1$-bit integer and m_2 is a $n-1$-bit integer.
2. Choose a random n-bit integer k_1 and compute $U = m_1 \cdot 2^n + k_1$. We have $2^{4n} < U < 2^{4n+1}$.
3. Choose a random n-bit integer k_2 and compute $V = m_2 \cdot 2^n + k_2$. We have $2^{2n-2} < V < 2^{2n-1}$.
4. Compute $C = Ue_{A1} + V^2 e_{A2}$.
5. Send ciphertext C to Along.

4 Decryption

Proposition 2. *Decryption by Along is conducted in the following steps:*

INPUT: The private key (pq, d) and the ciphertext C.
OUTPUT: The plaintext $\boldsymbol{M}$.

1. *Compute $W \equiv Cd (mod\ pq)$.*
2. *Compute $M_1 \equiv q^{-1} (mod\ p)$ and $M_2 \equiv p^{-1} (mod\ q)$.*
3. *Compute*
$$x_p \equiv W^{\frac{p+1}{4}} (mod\ p), x_q \equiv W^{\frac{q+1}{4}} (mod\ q).$$
4. *Compute*
$$V_1 \equiv x_p M_1 q + x_q M_2 p\ (mod\ pq),$$
$$V_2 \equiv x_p M_1 q - x_q M_2 p\ (mod\ pq),$$
$$V_3 \equiv -x_p M_1 q + x_q M_2 p\ (mod\ pq),$$
$$V_4 \equiv -x_p M_1 q - x_q M_2 p\ (mod\ pq).$$
5. *For $i = 1, 2, 3, 4$ compute $U_i = \frac{C - V_i^2 e_{A2}}{e_{A1}}$.*
6. *Sort the pair (U_j, V_j) for integer U_j.*
7. *Compute integral part $m_1 = \lfloor \frac{U_j}{2^n} \rfloor$.*
8. *Compute integral part $m_2 = \lfloor \frac{V_j}{2^n} \rfloor$.*
9. *Form the integer $m = m_1 \cdot 2^n + m_2$.*
10. *Transform the number m to the message $\boldsymbol{M}$.*
11. *Return the message $\boldsymbol{M}$.*

We now proceed to give a proof of correctness.

Along will begin by computing $W \equiv Cd \equiv V^2 (\text{mod } pq)$. Along will then have to solve $W \equiv V^2 (\text{mod } pq)$ using the Chinese Remainder Theorem.

Lemma 1. *Let p and q be two different primes such that $p \equiv 3(mod\ 4)$ and $q \equiv 3(mod\ 4)$. Define x_p and x_q by*

$$x_p \equiv W^{\frac{p+1}{4}} (mod\ p), x_q \equiv W^{\frac{q+1}{4}} (mod\ q).$$

Then the solutions of the equation $x^2 \equiv W(mod\ p)$ are $\pm x_p (mod\ p)$ and the solutions of the equation $x^2 \equiv W(mod\ q)$ are $\pm x_q (mod\ q)$.

Lemma 2. *Let p and q be two different primes such that $p \equiv 3(mod\ 4)$ and $q \equiv 3(mod\ 4)$. Define x_p and x_q by*

$$x_p \equiv W^{\frac{p+1}{4}} (mod\ p), x_q \equiv W^{\frac{q+1}{4}} (mod\ q).$$

Define $M_1 \equiv q^{-1} (mod\ p)$ and $M_2 \equiv p^{-1} (mod\ q)$. Then the solutions of the equation $V^2 \equiv W(mod\ pq)$ are

$$V_1 \equiv x_p M_1 q + x_q M_2 p\ (mod\ pq),$$

$$V_2 \equiv x_p M_1 q - x_q M_2 p\ (mod\ pq),$$
$$V_3 \equiv -x_p M_1 q + x_q M_2 p\ (mod\ pq),$$
$$V_4 \equiv -x_p M_1 q - x_q M_2 p\ (mod\ pq).$$

Proof. To solve the equation $V^2 \equiv W(\text{mod } pq)$, we use the Chinese Remainder Theorem. Consider the equations $x_p^2 \equiv W(\text{mod } p)$ and $x_q^2 \equiv W(\text{mod } q)$. Then the solution of the equation $V^2 \equiv W(\text{mod } pq)$ are the four solutions of the four systems

$$\begin{cases} V \equiv \pm x_p (\text{mod } p) \\ V \equiv \pm x_q (\text{mod } q) \end{cases}$$

Define $M_1 \equiv q^{-1} (\text{mod } p)$ and $M_2 \equiv p^{-1} (\text{mod } q)$. We will get explicitly

$$V_1 \equiv x_p M_1 q + x_q M_2 p\ (\text{mod } pq),$$

$$V_2 \equiv x_p M_1 q - x_q M_2 p\ (\text{mod } pq),$$
$$V_3 \equiv -x_p M_1 q + x_q M_2 p\ (\text{mod } pq),$$
$$V_4 \equiv -x_p M_1 q - x_q M_2 p\ (\text{mod } pq).$$

It can be seen that solving $V^2 \equiv W(\text{mod } pq)$, we will get four solutions V_i for $i = 1, 2, 3, 4$.

We prove below that only one of them leads to the correct decryption and consequently, there is no decryption failure.

Lemma 3. *Let C be an integer representing a ciphertext encrypted by the AA_β algorithm. The equation $C = Ue_{A1} + V^2 e_{A2}$ has only one solution satisfying $V < 2^{2n-1}$.*

Proof. Suppose for contradiction that there are two couples of solutions (U_1, V_1) and (U_2, V_2) of the equation $C = Ue_{A1} + V^2 e_{A2}$ with $V_1 \neq V_2$ and $V_i < 2^{2n-1}$. Then $U_1 e_{A1} + V_1^2 e_{A2} = U_2 e_{A1} + V_2^2 e_{A2}$. Using $e_{A1} = p^2 q$, this leads to

$$(U_2 - U_1)p^2 q = (V_1 + V_2)(V_1 - V_2)e_{A2}.$$

Since $\gcd(p^2 q, e_{A2}) = 1$, then $p^2 q | (V_1 + V_2)(V_1 - V_2)$ and the prime numbers p and q satisfy one of the conditions

$$p^2 | (V_1 \pm V_2) \text{ or } \begin{cases} pq | (V_1 \pm V_2) \\ p \nmid (V_1 \mp V_2) \end{cases}$$

Observe that $p^2 > 2^{2n}$ and $pq > 2^{2n}$ while $|V_1 \pm V_2| < 2 \cdot 2^{2n-1} = 2^{2n}$. This implies that none of these conditions is possible. Hence the equation $C = Ue_{A1} + V^2 e_{A2}$ has only one solution with the parameters of the scheme.□

4.1 Example

Let $n = 16$. Along will choose the primes $p = 62683$ and $q = 62483$. The public keys will be

1. $e_{A1} = 245505609868187$
2. $e_{A2} = 4106878163802480$

The private keys will be

1. $pq = 3916621889$
2. $d = 2486483$

Busu's message will contain the following parameters

1. $m_1 = 544644664056570$
2. $m_2 = 21777$

Busu will also generate the following ephemeral random session keys

1. $k_1 = 54433$
2. $k_2 = 33079$

Busu will then generate

1. $U = 35693832703611425953$
2. $V = 1427210551$ and consequently $V^2 = 2036929956885723601$

The ciphertext will be $C = 171284593275622664566002243879187691$.
To decrypt Along will first compute $W = 3215349249$. Along will then obtain the following root values

$$V_1 = 318887097,$$

$$V_2 = 2489411338,$$

$$V_3 = 1427210551,$$

and

$$V_4 = 3597734792.$$

Only $U_3 = \frac{C - V_3^2 e_{A2}}{e_{A1}}$ will produce an integer value. That is $U_3 = 35693832703611425953$. Finally, m_1 and m_2 can be obtained. □

5 Basic Attacks

5.1 Congruence attack

From $C = Ue_{A1} + V^2 e_{A2}$ and since $\gcd(e_{A1}, e_{A2}) = 1$ we have

$$U \equiv Ce_{A1}^{-1} \equiv a \pmod{e_{A2}}.$$

Hence $U = a + e_{A2}j$ for some $j \in \mathbb{Z}$. Replacing into C we have

$$C = Ue_{A1} + V^2 e_{A2} = (a + e_{A2}j)e_{A1} + V^2 e_{A2}.$$

Then,

$$V^2 = \frac{C - (a + e_{A2}j)e_{A1}}{e_{A2}} = \frac{C - e_{A1}a}{e_{A2}} - e_{A1}j,$$

where $\frac{C - e_{A1}a}{e_{A2}} = b \in \mathbb{Z}$. It follows that the equation $C = Ue_{A1} + V^2 e_{A2}$ has the parametric solutions

$$U = a + e_{A2}j \text{ and } V^2 = b - e_{A1}j.$$

• Computing with U

To find $U = a + e_{A2}j$, we should find an integer j such that $2^{4n} < U < 2^{4n+1}$. This gives

$$\frac{2^{4n} - a}{e_{A2}} < j < \frac{2^{4n+1} - a}{e_{A2}}.$$

We know that $2^{3n+4} < e_{A2} < 2^{3n+6}$. Then the difference between the upper and the lower bound is

$$\frac{2^{4n+1} - a}{e_{A2}} - \frac{2^{4n} - a}{e_{A2}} = \frac{2^{4n}}{e_{A2}} > \frac{2^{4n}}{2^{3n+6}} = 2^{n-6}.$$

Hence the difference is very large and finding the correct j is infeasible.

• Computing with V^2

To find $V^2 = b - e_{A1}j$, we should find an integer j such that $2^{4n-4} < V < 2^{4n-2}$. This gives

$$\frac{2^{4n-4} - b}{-e_{A1}} > j > \frac{2^{4n-2} - b}{-e_{A1}}.$$

We know that $2^{3n} < e_{A1} < 2^{3n+3}$. Then the difference between the upper and the lower bound is

$$\frac{2^{4n-4} - b}{-e_{A1}} - \frac{2^{4n-2} - b}{-e_{A1}} = \frac{3 \cdot 2^{4n-4}}{e_{A1}} = 3 \cdot 2^{n-7}.$$

Hence the difference is very large and finding the correct j is infeasible.

5.2 Coppersmith type attack

Theorem 1. *Let N be an integer of unknown factorization. Furthermore, let $f_N(x)$ be an univariate, monic polynomial of degree δ. Then we can find all solutions x_0 for the equation $f_N(x) \equiv 0(mod\ N)$ with*

$$|x_0| < N^{\frac{1}{\delta}}.$$

in time polynomial in $(logN, \delta)$.

Theorem 2. *Let N be an integer of unknown factorization, which has a divisor $b > N^{\beta}$. Furthermore let $f_b(x)$ be an univariate, monic polynimial of degree δ. Then we can find all solutions x_0 for the equation $f_b(x) \equiv 0(mod\ b)$ with*

$$|x_0| \leq \frac{1}{2} N^{\frac{\beta^2}{\delta} - \epsilon}$$

in polynomial time in $(logN, \delta, \frac{1}{\epsilon})$.

• Attacking V

With reference to Theorem 1. Let $N = e_{A1} = p^2 q$ and $d' \equiv e^{-1} (\text{mod } N)$. Compute $W \equiv Cd' \equiv V^2 (\text{mod } N)$. Let $f_N(x) \equiv x^2 - W \equiv 0 (\text{mod } N)$. Hence, $\delta = 2$. Thus the root $x_0 = V$ can be recovered if $V < N^{\frac{1}{2}} \approx 2^{1.5n}$. But since $V \approx 2^{2n}$, this attack is infeasible.

• Attacking d

With reference to Theorem 2. We begin by observing $f_b(x) = ex - 1 \equiv 0 (\text{mod } pq)$ where pq in an unknown factor of $N = e_{A1} = p^2 q$. Since $pq > N^{\frac{2}{3}}$ we have $\beta = \frac{2}{3}$. From $f_b(x)$ we also have $\delta = 1$. By the Coppersmith theorem, the root $x_0 = d$ can be found if $|x_0| < N^{\frac{4}{9}}$.But since $d > N^{\frac{4}{9}}$, this attack is infeasible.

5.3 Euclidean division attack

From $C = Ue_{A1} + V^2 e_{A2}$, the size of each public parameter within C ensures that Euclidean division attacks does not occur. This can be easily deduced as follows:

1. $\lfloor \frac{C}{e_{A1}} \rfloor \neq U$
2. $\lfloor \frac{C}{e_{A2}} \rfloor \neq V^2$

6 Analysis on lattice based attack

The square lattice attack has been an efficient and effective means of attack upon schemes that are designed based on Diophantine equations. The AA_{β} scheme has gone through analysis regarding lattice attacks while it went through the design process. Let $C = Ue_{A1} + V^2 e_{A2}$ be an

AA_β ciphertext. Consider the diophantine equation $e_{A1}x_1 + e_{A2}x_2 = C$. Introduce the unknown x_3 and consider the diophantine equation

$$e_{A1}x_1 + e_{A2}x_2 - Cx_3 = 0.$$

Then $(U, V^2, 1)$ is a solution of the equation. Next let T be a number to fixed later. Consider the lattice $\mathcal{L}$ spanned by the matrix:

$$M = \begin{pmatrix} 1 & 0 & e_{A1}T \\ 0 & 1 & e_{A2}T \\ 0 & 0 & -CT \end{pmatrix}$$

Observe that

$$(x_1, x_2, x_3)M = (x_1, x_2, T(e_{A1}x_1 + e_{A2}x_2 - Cx_3)).$$

This shows that the lattice $\mathcal{L}$ contains the vectors $(x_1, x_2, T(e_{A1}x_1 + e_{A2}x_2 - Cx_3))$ and more precisely the vector-solution $V_0 = (U, V^2, 0)$. Observe that the length of V_0 satisfies

$$\|V_0\| = \sqrt{U^2 + V^4} \approx 2^{4n}.$$

On the other hand, the determinant of the lattice is $\det(\mathcal{L}) = CT$ and the Gaussian heuristics for the lattice $\mathcal{L}$ asserts that the length of its shortest non-zero vector is usually approximately $\sigma(\mathcal{L})$ where

$$\sigma(\mathcal{L}) = \sqrt{\frac{dim(\mathcal{L})}{2\pi e}} \det(\mathcal{L})^{\frac{1}{\dim(\mathcal{L})}} = \sqrt{\frac{3}{2\pi e}} (CT)^{\frac{1}{3}}.$$

If we choose T such that $\sigma(\mathcal{L}) > \|V_0\|$, then V_0 can be among the short non-zero vectors of the lattice $\mathcal{L}$. To this end, T should satisfy

$$T > (\frac{\pi e}{2})^{\frac{3}{2}} \cdot \frac{2^{12n}}{C} \tag{3}$$

Next, if we apply the LLL algorithm to the lattice $\mathcal{L}$, we will find a basis (b_1, b_2, b_3) such that $\|b_1\| \leq \|b_2\| \leq \|b_3\|$ and

$$b_i \leq 2^{\frac{n(n-1)}{4(n+1-i)}} \det(\mathcal{L})^{\frac{1}{n+1-i}}, \text{for } i = 1, ..., 4 \text{ and } n = 3.$$

For $i = 1$, we choose T such that $\|V_0\| \leq \|b_1\| \leq 2^{\frac{1}{2}}(CT)^{\frac{1}{3}}$. Using the approximation $\|V_0\| \approx 2^{4n}$, this is satisfied if

$$V > 2^{-\frac{1}{2}} \cdot \frac{2^{12n}}{C},$$

which follows from the lower bound of equation (3). We experimented this result to try to find $(U, V^2, 0)$. The LLL algorithm outputs a basis with a matrix in the form

$$ML = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & T \end{pmatrix}$$

If $(U, V^2, 0)$ is a short vector, then $(U, V^2, 0) = (x_1, x_2, x_3)ML$ for some short vector (x_1, x_2, x_3). We then deduce the system

$$\begin{cases} a_{11}x_1 + a_{21}x_2 = U \\ a_{12}x_1 + a_{22}x_2 = V^2 \end{cases}$$

from which we can deduce that $x_3 = 0$. If we compute $(Ue_{A1} - V^2 e_{A2})/C$, we get $x_2 = 1$ for some x_1. It follows that

$$\begin{cases} a_{11}x_1 + a_{21} = U \\ a_{12}x_1 + a_{22} = V^2 \end{cases}$$

This situation is similar to the congruence attack. We can also observe that this is a system of two equations with three unknowns (i.e. x_1, U, V).

6.1 Example with lattice based attack

We will use the parameters in the earlier example in Section 4. Observe the lattice $\mathcal{L}$ spanned by the matrix:

$$M = \begin{pmatrix} 1 & 0 & e_{A1}T \\ 0 & 1 & e_{A2}T \\ 0 & 0 & -CT \end{pmatrix}$$

the length of the vector $V = (U, V^2, 0)$ is $\| V \| \approx 35751905917344588937$. We will use $T = 2^{20n}$ which would result in the length of the vector V is shorter than the gaussian heuristic of the lattice $\mathcal{L}$.

The LLL algorithm outputs:

$$ML = \begin{pmatrix} -4106878163802480 & 245505609868187 & 0 \\ 247367271832221073 & 4155888875658045598 & 0 \\ -1118395942494397 & 66856738131713 & T \end{pmatrix}$$

Observe that element $(1, 1)$ within the above matrix is $-e_{A2}$ while element $(1, 2)$ is e_{A1}. The rest of the process as observed in Section 6 can be obtained trivially.

7 Underlying security principles

In this section we will view four underlying security principles applied either directly or indirectly. We opine that each principle can be viewed independently from one another.

• The AA_β-DEHP

To find the unknown parameters U and V^2 from the public "summation composite" C. That is, to prf-solve C.

• The integer factorization problem

To find the unknown composite p and q such that $e_{A1} = p^2q$.

• The square root modulo problem

Since $\gcd(e_{A1}, e_{A2}) = 1$, one can obtain $V^2 \equiv \alpha (\text{mod } e_{A1})$. Since $e_{A1} = p^2q$, then this is equivalent to calculating square roots modulo composite integers with unknown factorization which is infeasible.

• The modular reduction problem

Since $\gcd(e_{A1}, e_{A2}) = 1$, one can obtain $U \equiv \beta (\text{mod } e_{A2})$. Since $U \gg e_{A2}$, to compute U prior to modular reduction by e_{A2} is infeasible.

7.1 Security Reduction

Proposition 3. *Let C be the AA_β ciphertext. Then, decryption$\equiv_T$ prf-solve C.*

Proof. Let θ_1 be an oracle that is able to *prf*-solve C. Call $\theta_1(C)$ to obtain U and V^2. Thus, $\mathbf{M}$ can be obtained. Hence, decryption has occurred. If the decryption has occurred then C has been *prf*-solved.

Corollary 1. *prf-solve $C \leq_T$ Factoring p^2q.*

Remark 1. The converse of Corollary 1 is unknown.

7.2 Equivalence with integer factorization

From $C = Ue_{A1} + V^2e_{A2}$ we have

$$C \equiv V^2 (\text{mod } e_{A1})$$

where $e_{A1} = p^2q$ is of unknown factorization. We show here that solving this congruence relation is equivalent to factoring e_{A1}. If we know the factorization of e_{A1}, then it is easy to solve the congruence relation. Conversely, suppose that we know all the solutions. By Lemma 2, the four solutions are

$$V_1 \equiv x_pM_1q + x_qM_2p \text{ (mod } pq),$$
$$V_2 \equiv x_pM_1q - x_qM_2p \text{ (mod } pq),$$
$$V_3 \equiv -x_pM_1q + x_qM_2p \text{ (mod } pq),$$
$$V_4 \equiv -x_pM_1q - x_qM_2p \text{ (mod } pq).$$

and are such that $V_i < pq$ for $i = 1, 2, 3, 4$. We have $V_1 + V_3 = 2x_qM_2p + \alpha pq$ for some integer α. Then $V_1 + V_3 \equiv 0 (\text{mod } p)$. On the other hand, $V_1 + V_3 < 2pq < p^2q$. Hence $V_1 + V_3 \not\equiv 0 (\text{mod } p^2q)$. Therefore

$$p = gcd(e_{A1}, V_1 + V_3) = gcd(p^2q, V_1 + V_3).$$

Hence $q = \frac{p^2q}{p^2}$.

8 Indistinguishability [1]

8.1 IND-CPA

Since the random parameter k_1 and k_2 are ephemeral keys, we achieve IND-CPA.

8.2 IND-CCA2

• IND-CCA2 Design - Encryption

Let E and D denote the encryption and decryption oracles respectively. Also let H be a random oracle where it is a mapping from j_0+j_1 to j_2 bit strings and $j_2 < j_0 + j_1$. Prior to sending the ciphertext, the encryption oracle will compute:

1. $C_1 = E(M) = Ue_{A1} + V^2 e_{A2}$
2. $C_2 = H(U + V)$
3. The ciphertext to be sent is (C_1, C_2).

• IND-CCA2 Design - Decryption

The decryption oracle will begin by computing $D(C) = M$. From M the decryption oracle can obtain U and V. The random oracle will then compute $h = H(U + V)$. If $h \oplus C_2 \neq 0$ then output $\perp$ which means the ciphertext is illegal. Otherwise, the decryption oracle outputs the correct plaintext.

Remark 2. In order to successfully to substitute C_2 the adversary has to be able to *prf*-solve C_1 successfully in order to obtain (U, V).

9 Secure Transmission of large data sets

9.1 Motivation

Initially, key distribution issues surrounding the implementation of symmetric cryptographic solutions triggered research for an asymmetric model. However, since the first asymmetric cryptographic model was first discovered, it has been the traditional role of the asymmetric scheme to encrypt the session key to be utilize by its corresponding symmetric scheme. While this seems practical, the authors opine that the initial objective to design an asymmetric cryptosystem to encrypt data has only been partially achieved.
Embedded within the construction of the RSA is the condition that the message, $\mathbf{M} \in \mathbb{Z}_N$. As for the DHKE, El-Gammal and ECC, the message is an element of $\mathbb{Z}_p$ (where p is referring to their respective prime structures). This means, the amount of data to be transmitted securely is limited. Thus, the objective to design an asymmetric scheme that could overcome this barrier.

9.2 Transmitting large data asymmetrically

From the LLL algorithm, the vector $\lambda_1 = (-e_{A2}, e_{A1}, 0)$ is produced (see Section 6). Let $V = (v_1, v_2, 0) \in \mathcal{L}$ be an arbitrary vector. We also can see that the vector $\lambda_2 = (v_1 - te_{A2}, v_2 - te_{A1}, 0) \in \mathcal{L}$ and $\| \lambda_2 \| < \| (v_1, v_2, 0) \|$. This implies:

$$e_{A1}^2 + e_{A2}^2 + \frac{2e_{A2}}{t} v_2 < \frac{2e_{A1}}{t} v_1$$

For $|t| \approx 2^n$ and the fact that $e_{A1} \approx 2^{3n}$, v_1 should be $\geq 2^{4n}$ for the above inequality to hold. Hence, U could be as large as possible, which in turn implies that m_1 can be as large as possible. The above inequality also implies that there exists vector shorter than $(U, V^2, 0)$. Hence, the inability of the LLL algorithm to detect the vector $(U, V^2, 0)$.

10 Table of Comparison

The following is a table of comparison between RSA, ECC, NTRU and AA_β. Let $|E|$ denote public key size. The AA_β cryptosystem has the ability to encrypt large data sets (i.e. arbitrary size). The ratio of $M : |E|$ suggests better economical value per public key bit being used. At the same time the ratio between the message and the ciphertext is $\approx 1 : 1$ which implies that message expansion due to encryption is negligible.

Algorithm	Encryption Speed	Decryption Speed	Ratio $M : C$	Ratio $M : \|E\|$
RSA	$O(n^3)$	$O(n^3)$	$1 : 1$	$1 :\approx 1$
ECC	$O(n^3)$	$O(n^3)$	$1 : 2$	$1 : 1$
NTRU	$O(n^2)$	$O(n^2)$	Varies [4]	N/A
AA_β	$O(n^2)$	$O(n^3)$	$\approx 1 : 1$	$t : 6$ where $t \geq 4$

Table 1. Comparison table for input block of length n

11 Conclusion

The DEHP, is proposed to be another source of cryptographic primitive, that if utilized correctly could give rise to other possible asymmetric algorithms differing from classical algorithms based on the difficulty of solving the integer factorization problem, discrete log problem (DLP), elliptic curve discrete log problem (ECDLP), e-th root modulo N, square root problem and others. The AA_β cryptosystem is capable of encrypting large data sets using a fixed key size. This will give a significant contribution in a niche area for implementation of asymmetric type security in transmitting large data sets.
The scheme is also comparable to the Rabin cryptosystem with the advantage of having a unique decryption result. It has achieved an encryption speed with complexity order of $O(n^2)$ and it also has a simple mathematical structure for easy implementation.

Acknowledgments

The authors would like to thank Prof. Dr. Abderrahmane Nitaj of Départe ment de Mathématiques, Université de Caen, France, Dr. Yanbin Pan of Key Laboratory of Mathematics Mechanization Academy of Mathematics and Systems Science, Chinese Academy of Sciences Beijing, China and Dr. Gu Chunsheng of School of Computer Engineering, Jiangsu Teachers University of Technology, Jiangsu Province, China for valuable comments and discussion on all prior AA_{β} designs.

References

[1] M. R. K. Ariffin, A Proposed IND-CCA2 Scheme for Implementation on an Asymmetric Cryptosystem Based on Diophantine Equation Hard Problem, *Proc. Third International Conference on Cryptology and Computer Security 2012*. (2012), pp. 193–197.

[2] W. Diffie and M. E. Hellman, New Directions in Cryptography, *IEEE Transactions on Information Theory*. vol. 22, no. 26 (1976), pp. 644–654.

[3] J. Hoffstein, J. Pipher and J. H. Silverman, An Introduction to Mathematical Cryptography. *New York: Springer*. (2008), pp. 352–358.

[4] J. Hoffstein, D.Lieman, J. Pipher and J. H. Silverman, NTRU : A Public Key Cryptosystem, NTRU Cryptosystems Inc. (2008, April 27) [Online]. Available: http://grouper.ieee.org/groups/1363/lattPK/submissions/ntru.pdf

[5] J. Hermans *et. al.*, Speed Records for NTRU, *CT-RSA 2010, LNCS 5985*. (2010), pp. 73–88.

[6] J. Hoffstein, J. Pipher, J. H. Silverman. NTRU: A Ring Based Public Key Cryptosystem in Algorithmic Number Theory (ANTS III) Lecture Notes in Computer Science 1423, *Springer-Verlag, Berlin*. (1998), pp. 267–288.

[7] N. Koblitz, Elliptic Curve Cryptosystems, *Math. Comp*. vol. 48, no. 177 (1987), pp. 203–209.

[8] R. L. Rivest, A. Shamir and L. Adleman, A method for obtainning digital signatures and public key cryptosystems, *Commun. ACM*. vol. 21, issue 2 (1978), pp. 120–126.

[9] B. Schneier, Key length in Applied Cryptography. *New York: John Wiley & Sons*. (1996), pp. 151–168.

[10] M. Scott, When RSA is better than ECC. (2008, November 15) [Online]. Available: http://www.derkeiler.com/Newsgroups/sci.crypt/2008-11/msg00276.html

[11] S. S. Wagstaff, Cryptanalysis of Number Theoretic Ciphers, *Divisibility and Arithmetic*. (2003), pp. 27–42.

[12] S. Vanstone, ECC holds key to next generation cryptography. (2006, March 18) [Online].Available: http://www.design-reuse.com/articles/7409/ecc-hold-key-to-next-gen-cryptography.html