ON THE HEURISTIC GUESS OF 2-DIMENSION LATTICE ATTACK ON LOW PRIVATE EXPONENT RSA

TRAN DINH LONG*, NGUYEN DINH THUC**, TRAN DAN THU**

ABSTRACT

In two dimension lattice attack on low private exponent RSA cryptosystem, the reasonable and non-provable guess shows that the private exponent d could be recovered by finding a shortest vector of a 2-dimension lattice by Gaussian reduction algorithm. The paper considers the determination of the attack by giving a precise interval of private d where the heuristic guess in 2-dimension lattice attack on RSA holds and gives a proof for that heuristic guess.

Keywords: lattice, lattice reduction algorithm, RSA cryptosystem.

TÓM TẮT

Về dự đoán trong cách tấn công dùng dàn hai chiều vào hệ mã RSA có khóa riêng nhỏ

Trong việc tấn công bằng dàn hai chiều vào hệ mã RSA có khóa riêng nhỏ,một dự đoán hợp lí nhưng không được chứng minh chỉ ra rằng khóa riêng d của hệ mã RSA có thể tìm được bằng cách tìm một vector ngắn nhất của một dàn hai chiều bởi thuật toán Gauss. Bài viết này khảo sát tính tất định của việc tấn công trên bằng cách chỉ ra một khoảng chính xác sao cho nếu khóa riêng d nằm trong khoảng đó thì việc tấn công RSA bằng dàn hai chiều luôn thành công, đồng thời đưa ra cách chứng minh chặt chẽ cho điều này.

Từ khóa: dàn, thuật toán tìm cơ sở thu gọn của dàn, hệ mã hóa RSA.

1. Introduction

Besides constructing new variants of RSA, cryptanalysing on RSA cryptosystem has been concerned by many authors. Some early attacks on RSA had been considered by G.J.Simmons [7], J.M.DeLaurentis [4]… A remarkable result was made by M. Wiener in 1990; by considering the continued fraction expansion of $\frac{e}{n}$, Wiener showed in [8] that one can recover d in the case $d < \frac{1}{3}n^{\frac{1}{4}}$, where e, d and n are public key, private key and the modulus of the cryptosystem, respectively. Lattice reduction based attacks on RSA was first presented by Coppersmith at Eurocrypt '96 [3]. Lattice reduced algorithms such as Gauss or LLL algorithms can be applied to recover the private exponent d in low exponent private key RSA cryptosystem. D. Boneh and G.

[*] MSc, Faculty of Mathematics, College of Science, Hue University;
Email: trandinhlong1963@yahoo.com.vn

[**] Assoc, PhD, Faculty of Information Technology, Ho Chi Minh University of Science

Durfee [2] considered the case where $d < n^{0.292}$, then by solving small inverse problem using LLL algorithm, one can recover d. Modifying the attack of D. Boneh and G. Durfee, Blomer and May [1] had improved 0.292 to $\frac{\sqrt{6}-1}{5} - \varepsilon$, where ε is the term can be made arbitrary small by considering sufficiently large modulus n. High dimension lattice attacks are based on LLL algorithm while two dimension lattice attacks are based on Gaussian algorithm. Lattice now is an effective tool in cryptanalysing on RSA.

We wish to investigate the heuristic attack on low private exponent RSA using two dimension lattice. This attack is indeed mounted from Wiener attack (see [5]) and based on Gaussian algorithm. Section 2 is devoted to some basic properties of lattices. The heuristic attack will be recalled in Section 3 together with our work, which considers the determination of the attack. The last section gives comment about our approach to the problem.

2. Lattices

2.1. Background

A lattice of $\mathbb{R}^n$ is a discrete subgroup of $(\mathbb{R}^n, +)$, that is a subgroup of $(\mathbb{R}^n, +)$ which has the discreteness property. Like vector spaces, a lattice has a basis and each element in lattice can be represented as a integral linear combination of vectors in basis. If $\{b_1, b_2, \dots, b_m\}$ is a basis of the lattice $L \subset \mathbb{R}^n$, then

$$L = \{\sum_{i=1}^{m} n_i b_i : n_1, n_2, \dots, n_m \in \mathbb{Z}\}.$$

The *fundamental domain* for L corresponding to the basis $\{b_1, b_2, \dots, b_m\}$ is the set

$$\mathcal{F}(b_1, b_2, \dots, b_m) = \{t_1 b_1 + t_2 b_2 + \cdots + t_m b_m : t_i \in \mathbb{R}, 0 \le t_i < 1\}.$$

The n-dimension volume of $\mathcal{F}(b_1, b_2, \dots, b_m)$ is called the *determinant of* L and denoted by $\det(L)$. We have the *Hadamard's Inequality* as follows

$$\det(L) \le \|b_1\|\|b_2\| \dots \|b_m\|,$$

where $\|v\|$ is the Euclidean norm of a vector $v \in \mathbb{R}^n$.

Some problems on lattices sush as finding shortest vector problem, finding closest vector problem… can be easily solved when an orthogonal basis of lattice is determined. Unfortunately, a lattice may not have an orthogonal basis. Therefore, finding a "near orthogonal" basis, or an "optimal basis" is a problem has been concerned by many authors. Two famous algorithms for finding such basis are Gaussian and LLL algorithms, we call those algorithms as *lattice reduction* algorithm.

2.2. Gaussian algorithm

We recall Gaussian algorithm in this section. For a vector $\mathrm{v} \in \mathbb{R}^2$, we denote $\|\mathrm{v}\|$ for the Euclidean norm of v and $\langle \mathrm{v}_1, \mathrm{v}_2 \rangle$ for the inner product of two vectors $\mathrm{v}_1, \mathrm{v}_2 \in \mathbb{R}^2$.

Let v_1, v_2 be two independent vectors in $\mathbb{R}^2$ and $L \subset \mathbb{R}^2$ be the lattice spanned by v_1, v_2. Gaussian algorithm is applied to basis v_1, v_2 and yields a good basis $\overline{v_1}, \overline{v_2}$ for L.

```
Input: a basis $\{v_1, v_2\}$ of a lattice $L \subset \mathbb{R}^2$.
loop
    if$\|v_2\| < \|v_1\|$then
        swap $v_1$ and $v_2$
    end if
Compute $m = \frac{\langle v_1, v_2 \rangle}{\|v_1\|^2}$
                    $v_2 = v_2 - \lfloor m + 0.5 \rfloor v_1$
until $\|v_1\| < \|v_2\|$
                    $\overline{v_1} = v_1, \overline{v_2} = v_2$
Output: a reduced basis $\{\overline{v_1}, \overline{v_2}\}$ of $L$.
```

Gaussian algorithm

$\overline{v_1}$ is a shortest vector in L and the angle θ between $\overline{v_1}$ and $\overline{v_2}$ satisfies $|cos\theta| \leq \frac{\|\overline{v_1}\|}{2\|\overline{v_2}\|}$, so in particular we have $\frac{\pi}{3} \leq \theta \leq \frac{2\pi}{3}$ or $\left|\frac{\langle \overline{v_1}, \overline{v_2} \rangle}{\|v_1\|^2}\right| \leq \frac{1}{2}$. The Gaussian algorithm will terminate in at most $\left[\log_{1+\sqrt{2}}\left(\frac{\|v_1\|}{\lambda_2}\right)\right] + 3$ iterations [9], where λ_2 is the second minima of L.For more details on Gaussian algorithm, we refer the reader to [8].

2.3. Properties of reduced basis in two dimension lattice case

Suppose that $\{\overline{v_1}, \overline{v_2}\}$ is the reduced basis of lattice L when applying Gaussian algorithm to a basis $\{v_1, v_2\}$ of L. We first show that $\overline{v_2}$ is the shortest vector which is independent to $\overline{v_1}$, it means that there is no $v \in L$ such that $\|v\| < \|v_2\|$ and $\overline{v_1}, v$ are independent.

Proposition 1. *Suppose that $L \subset \mathbb{R}^2$ is the lattice spanned by two independent vectors $v_1, v_2 \in \mathbb{R}^2$. Apply Gaussian algorithm to basis $\{v_1, v_2\}$ of L and yield basis $\{\overline{v_1}, \overline{v_2}\}$. If $v \in L, v \neq 0$ satisfying $\|v\| < \|\overline{v_2}\|$ then $v = s\overline{v_1}$ with $s \in \mathbb{Z}$.*

Proof. Since $v \in L$, then $v = s\overline{v_1} + r\overline{v_2}$ with $s, r \in \mathbb{Z}$. Assume the contrary that $r \neq 0$, consider three following cases.

● Case 1 of $|s| = 1$ and $|r| = 1$: In this case we have

$\|v\|^2 = \|s\overline{v_1} + r\overline{v_2}\|^2 = \|\overline{v_1} \pm \overline{v_2}\|^2$

$= \|\overline{v_1}\|^2 + \|\overline{v_2}\|^2 \pm 2\langle\overline{v_1}, \overline{v_2}\rangle$

$= \|\overline{v_2}\|^2 + 2\|\overline{v_1}\|^2\left(\frac{1}{2} \pm \frac{\langle\overline{v_1},\overline{v_2}\rangle}{\|v_1\|^2}\right).$

Since $\left|\frac{\langle\overline{v_1},\overline{v_2}\rangle}{\|v_1\|^2}\right| \le \frac{1}{2}$, then $\frac{1}{2} \pm \frac{\langle\overline{v_1},\overline{v_2}\rangle}{\|v_1\|^2} \ge 0$. Hence,

$$\|v\|^2 = \|\overline{v_2}\|^2 + 2\|\overline{v_1}\|^2\left(\frac{1}{2} \pm \frac{\langle\overline{v_1},\overline{v_2}\rangle}{\|v_1\|^2}\right) \ge \|\overline{v_2}\|^2.$$

● Case 2 of $|s| > 1$ or $|r| > 1$:

If $|s| = |r|$ then

$\|v\|^2 = s^2\|\overline{v_1} \pm \overline{v_2}\|^2 \ge s^2\|\overline{v_2}\|^2 \ge \|\overline{v_2}\|^2.$

If $|s| \ne |r|$ then

$\|v\|^2 = s^2\|\overline{v_1}\|^2 + r^2\|\overline{v_2}\|^2 + 2st\langle\overline{v_1}, \overline{v_2}\rangle$

$\ge s^2\|\overline{v_1}\|^2 + r^2\|\overline{v_2}\|^2 - 2|st|.|\langle\overline{v_1}, \overline{v_2}\rangle|$

$\|\overline{v_2}\|^2 + s^2\|\overline{v_1}\|^2 + (r^2 - 1)\|\overline{v_2}\|^2 - 2|st|.|\langle\overline{v_1}, \overline{v_2}\rangle|$

$\ge \|\overline{v_2}\|^2 + s^2\|\overline{v_1}\|^2 + (r^2 - 1)\|\overline{v_2}\|^2 - |st|.\|\overline{v_1}\|^2$

$= \|\overline{v_2}\|^2 + (s^2 + r^2 - |sr| - 1)\|\overline{v_1}\|^2$

$= \|\overline{v_2}\|^2 + ((|s| - |r|)^2 + |sr| - 1)\|\overline{v_1}\|^2$

$\ge \|\overline{v_2}\|^2$

since $(|s| - |r|)^2 - 1 \ge 0$.

● Case 3 of $s = 0$: In this case, $\|v\| = \|r\overline{v_2}\| \ge \|\overline{v_2}\|$ since $r \ne 0$.

Thus, all three cases above lead to $\|v\| \ge \|\overline{v_2}\|$, a contradiction. Therefore, we must have $r = 0$ or $v = s\overline{v_1}$. ■

3. Two dimension lattice attack on RSA cryptosystem

3.1. The heuristic attack

Consider the RSA cryptosystem, where the modulus n is the product of two distinct primes p and q, e and d are public and private keys, respectively. We recall the argument of reasonable guess in 2-dimension lattice attack on RSA in the case $d < n^{\frac{1}{4}}$ in [6] as follows. Suppose that p and q are balanced, then $p = O(\sqrt{n})$ and $q = O(\sqrt{n})$, therefore $\varphi(n) = (p - 1)(q - 1) = n + O(\sqrt{n})$. Since $ed \equiv 1(\bmod\, \varphi(n))$ then there exists $k = O(d)$ such that $ed = 1 + k\varphi(n) = 1 + k(n + O(\sqrt{n}))$. It deduces that $ed - kn = kO(\sqrt{n})$. Denote $l = ed - kn$ then $l = O(d\sqrt{n})$. Consider the lattice $L \subset \mathbb{R}^2$ spanned by two vector $v_1 = (e, \sqrt{n})$ and $v_2 = (n, 0)$, then L contains $t = dv_1 - kv_2 = (l, d\sqrt{n})$. Since $\|t\| = \sqrt{l^2 + nd^2} \approx d\sqrt{n}$ and $(vol(L))^{\frac{1}{2}} = n^{\frac{3}{4}}$, then t could be a shortest vector in L if

$d\sqrt{n} < n^{\frac{3}{4}}$, or $d < n^{\frac{1}{4}}$. So in the case $d < n^{\frac{1}{4}}$, one can find out t by Gaussian reduced basis algorithm and hence, the private key d could be recovered.

3.2. *Experimental study*

In our experiments, two balanced primes pand q are generated then both shortest vector in L as well t are computed. We discovered many cases where the heuristic guess above does not holds. In the argument in section 3.1, the relation O could miss some constants, then some factor in the conditiond $< n^{\frac{1}{4}}$ could be ignored. We are thus led to the following problem: find a constant α such that if $d < \alpha n^{\frac{1}{4}}$ then t is a shortest vector in L.

4. The determination of the heuristic attack

Consider the RSA cryptosystem as mentioned in 3.1. Assume that p and q are balanced, as in [2] we use the condition $\frac{1}{2}\sqrt{n} < p, q < 2\sqrt{n}$ for this. Typically, we can suppose that $1 < e, d < \varphi(n) = (p-1)(q-1)$. Since $ed \equiv 1(\mathrm{mod}\ \varphi(n))$, then $ed = 1 + k\varphi(n)$ with $k \in \mathbb{Z}$. We firstly estimate k and $ed - kn$ as follows.

Proposition 2. *Suppose that $n = pq$ is the product of two distinct primes p and q, e and d are positive integers satisfying $1 < e, d < \varphi(n)$ and $ed = 1 + k\varphi(n)$. Then*

a) $k < d$.

b) $|ed - kn| < \frac{5}{2}d\sqrt{n}.$

Proof The proof is straightforward as follows.

a) Since $e < \varphi(n)$ then $1 + k\varphi(n) = ed < d\varphi(n)$. Hence, $k < d - \frac{1}{\varphi(n)} < d$.

b) We have $ed = 1 + k\varphi(n) = 1 + k(p-1)(q-1) = 1 + k(n + 1 - p - q)$.

Denote $a = \frac{p}{\sqrt{n}}$ then $p = a\sqrt{n}$, $q = \frac{1}{a}\sqrt{n}$ and $\frac{1}{2} < a < 2$. Then

$|ed - kn| = |k(p + q - 1) - 1| < |k(p + q)| = k\sqrt{n}(a + \frac{1}{a})$.

It is easy to check that $a + \frac{1}{a} < \frac{5}{2}$ for all $a \in (\frac{1}{2}, 2)$. Therefore,

$|ed - kn| < k\sqrt{n}\left(a + \frac{1}{a}\right) < \frac{5}{2}k\sqrt{n} < \frac{5}{2}d\sqrt{n}$. ■

As in 3.1, from now on we denote $v_1 = \left(e, \sqrt{n}\right), v_2 = (n, 0)$ and consider the lattice $L \subset \mathbb{R}^2$ spanned by v_1, v_2. Then $t = dv_1 - kv_2 = (ed - kn, d\sqrt{n})$ is a vector in L. Apply Gaussian algorithm for basis $\{v_1, v_2\}$ of L then yield a basis $\{\overline{v_1}, \overline{v_2}\}$. The following proposition estimates the norms of t and $\overline{v_2}$.

Proposition 3. *Let n, e, d be the integers as in Proposition 2, L denote the lattice in $\mathbb{R}^2$ spanned by two vectors $v_1 = \left(e, \sqrt{n}\right)$, $v_2 = (n, 0)$ and $t = dv_1 - kv_2 = (ed - kn, d\sqrt{n}) \in L$. Suppose that $\{\overline{v_1}, \overline{v_2}\}$ is the reduced basis when applying Gaussian* *algorithm to basis* $\{v_1, v_2\}$ *of* L. *Then*

a) $\|t\| < \frac{\sqrt{29}}{2} d\sqrt{n}$.

b) $\|\overline{v_2}\| \geq n^{\frac{3}{4}}$.

Proof.

a) It follows from the Proposition 2 that

$$\|t\| = \sqrt{(ed - kn)^2 + (d\sqrt{n})^2} < \sqrt{\left(\frac{5}{2} d\sqrt{n}\right)^2 + \left(d\sqrt{n}\right)^2} = \frac{\sqrt{29}}{2} d\sqrt{n}.$$

b) We have $\det(L) = \left| det \begin{pmatrix} e & \sqrt{n} \\ n & 0 \end{pmatrix} \right| = n\sqrt{n}$.

According Hadamard inequality, $\det(L) \leq \|\overline{v_1}\|.\|\overline{v_2}\|$. It yields that

$n\sqrt{n} \leq \|\overline{v_1}\|.\|\overline{v_2}\| \leq \|\overline{v_2}\|^2$.

Therefore, $n^{\frac{3}{4}} \leq \|\overline{v_2}\|$. ■

Proposition 4. *Under the assumptions in Proposition 3, if* v *is a vector in* L *satisfying* $t = sv$ *with* $s \in \mathbb{Z}$ *then* $s = \pm 1$.

Proof.

Note that $\gcd(d, k) = 1$ since $ed = 1 + k(p - 1)(q - 1)$.

Since $v \in L$ then $v = av_1 + bv_2 = (ae + bn, a\sqrt{n})$ with $a, b \in \mathbb{Z}$.

It follows from $t = sv$ that

$$\begin{cases} ed - kn = s(ae + bn) \\ \quad d\sqrt{n} = sa\sqrt{n}. \end{cases}$$

Thus,

$$ed - kn = sae + sbn, \tag{1}$$

and

$$d = sa. \tag{2}$$

Replace d from (2) into (1) implies that

$esa - kn = sae + sbn$,

or

$$k = -sb. \tag{3}$$

It deduces from (2) and (3) that s is a common divisor of d and k. Combining this with $\gcd(d, k) = 1$ leads to $s = \pm 1$.■

Proposition 5. *Under the assumptions in Proposition 2, if* $d < \frac{2}{\sqrt{29}} n^{\frac{1}{4}}$ *then* t *is a shortest vector in* L.

Proof. According to Proposition 2 and Proposition 3 we have

$$\|t\| = \sqrt{(ed - kn)^2 + \left(d\sqrt{n}\right)^2}$$

$$< \sqrt{\left(\frac{5}{2} d\sqrt{n}\right)^2 + \left(d\sqrt{n}\right)^2}$$

$$= \frac{\sqrt{29}}{2} d\sqrt{n}$$

$$\leq \frac{\sqrt{29}}{2} . \frac{2}{\sqrt{29}} n^{\frac{1}{4}} \sqrt{n}$$

$$= n^{\frac{3}{4}}$$

$$\leq \|\overline{v_2}\|.$$

It follows from Proposition 1 that $t = s\overline{v_1}$ and then deduces from Proposition 2 that $s = \pm 1$. Therefore, $t = \pm\overline{v_1}$ is a shortest vector in L.■

5. Conclusions

The paper shows that in the case $d < \frac{2}{\sqrt{29}} n^{\frac{1}{4}}$ then the private key d in RSA crytpsystem can be recovered from the vector $t = (ed - kn, d\sqrt{n})$ which is found by Gaussian algorithm. The constant $\frac{2}{\sqrt{29}}$ can be larged depending on some conditions. If we use the condition $p < q < 2p$ for the balance of p and q then we obtain $p < \sqrt{n}, q < \sqrt{2n}$ and $p + q < (1 + \sqrt{2})\sqrt{n}$. By similar argument, if $d < \frac{1}{\sqrt{4+2\sqrt{2}}} n^{\frac{1}{4}}$ then t is a shortest vector in L.

As mentioned above, if $d < n^{\frac{1}{4}}$ then the heuristic guess in 2-dimension lattice attack on RSA does not always holds. However, experiments have showned that if $d \approx n^{\frac{1}{4}}$ then that heuristic guess still holds in many cases. We constructed RSA cryptosystems where p, q are two consecutive 32-bit primes and the private exponent d satisfying $\frac{3}{4} n^{\frac{1}{4}} < d < n^{\frac{1}{4}}$ then the percentage of the cases where the heuristic guess holds is 65%. This arises an following open problem: find out some extra condition which ensures the heuristic guess in 2-dimension lattice attack on RSA.

REFERENCES

1. J. Blomer and A. May (2003), “New partial key explosure attacks on RSA”, CRYPTO, Vol. 2729 of Lecture Notes in Computer Science, pp. 27-43, Springer.
2. D. Boneh and G. Durfee (1999), “Cryptanalysis of RSA with private key d less than $n^{0.292}$”, *Proceedings of Eurocrypt'99*.
3. D. Coppersmith, M. Franklin, J. Patarin, and M. Reiter (1996), “Low exponent RSA with related messages”, *Proceedings of Eurocrypt 96*.
4. J. M. DeLaurentis (1984), “A further weakness in the common modulus protocol for the RSA crypto algorithm”, *Cryptologia*, 8(3):253-259.
5. M. Jason Hinek (2009), *Cryptanalysis of RSA and its variants*, Chapman and Hall_CRC, pp.71-72.
6. Phong Q. Nguyen (2008), “Public key cryptanalysis”, Recent trends in cryptography, *Contemporary Mathematics series*, AMS-RSME.
7. G. J. Simmons (1983), “A weak privacy protocol using the RSA crypto algorithm”, *Cryptologia*, 7(2):180-182.
8. M. Wiener (1990), “Cryptanalysis of short RSA secret exponents”, *IEEE Transactions on Information Theory*, 36:553-558.
9. C. P. Schorr, *Gittertheori und Kryptographie (1994)*, Ausarbreitung, Johann-Wolfgang-Goethe-Univesitat Franfurt, Main.

(*Received: 14/01/2015; Revised: 28/01/2015; Accepted: 12/02/2015)*