Available online at www.sciencedirect.com
ScienceDirect
Progress in Natural Science 19 (2009) 1327–1331
Progress in Natural Science
www.elsevier.com/locate/pnsc

Remarks on the bounds for cryptanalysis of low private key RSA

Haijian Zhou [a,*], Ping Luo [b], Daoshun Wang [a], Yiqi Dai [a]

[a] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[b] School of Software, Tsinghua University, Beijing 100084, China

Received 20 July 2008; received in revised form 21 September 2008; accepted 22 September 2008

Abstract

Boneh and Durfee have developed a cryptanalytic algorithm on low private key RSA. The algorithm is based on lattice basis reduction and breaks RSA with private key $d < N^{0.292}$. Later on, an improved version by Blömer and May enhanced the efficiency, while reaching approximately this same upper bound. Unfortunately, in both the algorithms, there is a critical error in theoretical analysis, leading to the overestimated upper bound $N^{0.292}$. In this paper we present a more precise analytical model, with which the theoretical upper bound on d is modified to approximately $d < N^{0.277}$ for ordinary RSA systems with a 1024-bit public key (N, e).

Keywords: RSA; Cryptanalysis; Low private key; Lattice basis reduction

1. Introduction

The RSA public key cryptosystem [1] has been applied in many fields since its first introduction, and research on cryptanalysis of RSA is still in progress. Recently, a special case of RSA with low private key d has been studied. Wiener and Dujella have developed a cryptanalytic algorithm [2,3] using continued fraction and broke RSA with a private key $d < N^{0.25}$. Boneh and Durfee proposed in Ref. [4] that this bound could be increased to $d < N^{0.292}$ by applying the LLL [5] lattice basis reduction algorithm on a specifically constructed lattice. Later Blömer and May [6] improved the efficiency of the algorithm by reducing the dimension of the lattice used in the reduction procedure. Meanwhile, a similar bound $d < N^{0.290}$ is achieved in this modified version. In addition, Ernst et al. proposed partial key exposure attacks on low private RSA [7], in which techniques similar to Boneh–Durfee's algorithm were used.

In Boneh and Durfee's algorithm (as well as in Blömer and May's improved version) given public exponent e of RSA and carefully selected parameters m and t, the following inequality

$$\det(L) < e^{m(w-1)}/\gamma \tag{1}$$

holds. Here, $\det(L)$ is the determinant of the lattice L involved in the algorithm, and w is the dimension of this lattice, satisfying

$$w \approx (1-\delta)m^2 + o(m^2) \tag{2}$$

γ is a constant only depending on the dimension w:

$$\gamma = (w2^w)^{(w-1)/2} \tag{3}$$

Boneh and Durfee took γ as a small factor compared with e^{mw} in the inequality, and ignored it when analyzing the upper bound on the private exponent d; hence, leading to the result $d < N^{0.292}$. We argue that this result should be reconsidered, if we take into account the effect due to this "small" constant γ. As a matter of fact, their algorithm achieves the upper bound $N^{0.292}$ by taking sufficiently large m and w, and in this case, γ may be approximating or even

* Corresponding author. Tel.: +86 13581816780.
E-mail address: zhouhj04@mails.tsinghua.edu.cn (H. Zhou).

1002-0071/$ - see front matter © 2009 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.
doi:10.1016/j.pnsc.2008.09.013

larger than e^{mw}. On the other hand, for smaller m and w, the lower order components $o(m^2)$ should be considered. All these factors lead to a modification on the bounds.

Hinek et al. have noticed this same problem [8] and gave a modified upper bound on d by taking $\gamma = (w2^w)^{(w-1)/2}$. For example, $d < N^{0.271}$ with 1000-bit N, using Blömer and May's lattice. Hinek's results could still be improved. By applying the newest research results [9,10], for almost all lattices with sufficiently high dimension w, γ is modified to

$$\gamma \approx (w1.02^w)^{(w-1)/2} \tag{4}$$

so we can build a more precise analytical model for the upper bound on d. Evaluation results show that $d < N^{0.292}$ can be achieved only with a sufficiently large (e.g., 100,000-bit) public key (N, e). For common RSA systems with a 1024-bit public key, the theoretical upper bound on d is modified to approximately $d < N^{0.277}$; and for a 10,240-bits public key, the bound is modified to approximately $d < N^{0.288}$.

2. Boneh–Durfee's algorithm

2.1. The main procedure

We only give the main procedure of Boneh and Durfee's algorithm here. The readers may find a detailed description in Ref. [4].

Recall that an RSA public key is a pair of integers (N, e) where $N = pq$ is the product of two n-bit primes. The corresponding private key is an integer d satisfying $ed \equiv 1 \bmod (\phi(N)/2)$, where $\phi(N) = N - p - q + 1$ is the Euler function. It follows that there exists an integer x such that

$$ed + x\left(\frac{N+1}{2} - \frac{p+q}{2}\right) = 1 \tag{5}$$

Writing $y = -\frac{p+q}{2}$ and $A = \frac{N+1}{2}$, we have

$$x(A + y) \equiv 1 \bmod e \tag{6}$$

Denote $e = N^\alpha$ and $d = N^\delta$. Typically, e is the same order of magnitude as N and so $\alpha \approx 1$. Thereby, we can compute $|x| < 3e^{1+(\delta-1)/\alpha} \approx e^\delta$ and $|y| < 2e^{1/(2\alpha)} \approx e^{0.5}$. Denote

$$f(x, y) = x(A + y) - 1 \tag{7}$$

The algorithm is trying to find out (x_0, y_0) as a root of $f(x, y) = 0$, such that $|x_0| < e^\delta$ and $|y_0| < e^{0.5}$. Given the parameter pair (m, t), and define the following polynomials:

$$\begin{aligned} g_{i,k} &:= x^i f^k(x, y) e^{m-k} \\ h_{j,k} &:= y^j f^k(x, y) e^{m-k} \end{aligned} \tag{8}$$

where $k = 0, \ldots, m$; for each k we use $g_{i,k}(x, y)$ for $i = 0, \ldots, m - k$ and use $h_{j,k}(x, y)$ for $j = 1, \ldots, t$. Here, the $g_{i,k}(x, y)$ polynomials are referred to as x-shifts and $h_{j,k}(x, y)$ polynomials as y-shifts. Observe that (x_0, y_0) is the root of all these polynomials modulo e^m for $k = 0, \cdots, m$. The authors construct a lattice L on the matrix spanned by the corresponding coefficient vectors of the polynomials, and apply the LLL lattice basis reduction algorithm to find two linearly independent bivariate polynomials $g_1, g_2 \in \mathbb{Z}[x, y]$, satisfying

$$\begin{aligned} g_1(x_0, y_0) &= 0 \\ g_2(x_0, y_0) &= 0 \end{aligned} \tag{9}$$

By computing the resultant $h(y) = \mathrm{Res}(g_1, g_2)$ and solving $h(y) = 0$, one root of $h(y)$ will expose $y_0 = -(p + q)/2$ and facilitate the factorization of $N = pq$.

Remark. The bivariate polynomials g_1, g_2 are not guaranteed to be algebraically independent though they are proven to be linearly independent. In this case, the resultant is identically zero, so that Boneh–Durfess's algorithm fails.

2.2. Bounds analysis

For the lattice L used in the reduction procedure, given selected parameters (m, t), we can compute its dimension

$$w = \frac{(m+1)(m+2)}{2} + t(m+1) \tag{10}$$

The determinant of the lattice is denoted by $\det(L)$, which comprises two parts, corresponding to the x-shifts and y-shifts, respectively, i.e.,

$$\det(L) = \det_x \cdot \det_y \tag{11}$$

where

$$\det_x = e^{m(m+1)(m+2)(5+4\delta)/12} \tag{12}$$

and

$$\det_y = e^{tm(m+1)(1+\delta)/2 + t(m+1)(m+t+1)/4} \tag{13}$$

To produce the bivariate polynomials g_1 and g_2 by the LLL algorithm, we have to satisfy the following inequalities

$$\det(L) < e^{m(w-1)}/\gamma \tag{14}$$

where

$$\gamma = (w2^w)^{(w-1)/2} \tag{15}$$

Boneh and Durfee considered that γ is a constant only depending on the dimension w and negligible when compared with $e^{m(w-1)}$. When m is large enough (hence $o(m^2)$ is negligible), the inequality above turns out to be

$$m^2(-1 + 4\delta) - 3tm(1 - 2\delta) + 3t^2 < 0 \tag{16}$$

For every m the left-hand side is minimized at $t = m(1 - 2\delta)/2$. Plug in this value and simplify the inequality, we have

$$m^2(-7 + 28\delta - 12\delta^2) < 0 \tag{17}$$

which implies

$$\delta < \frac{7}{6} - \frac{1}{3}\sqrt{7} \approx 0.284 \tag{18}$$

For m being large enough, whenever $d < N^{0.284}$, we can break RSA by factoring the modulus N, using the algorithm in Section 2.1. Note that the LLL lattice reduction algorithm ends in polynomial time, Boneh–Durfee's algorithm is also of polynomial-time complexity.

2.3. Improving the bounds

Boneh and Durfee improved the upper bound on δ by eliminating some "damaging" rows in the matrix spanned by the coefficients vectors, and constructed a new lattice L_1. The readers may find detailed analysis in Ref. [4].

In the improved algorithm, we take parameter pair (m, t) satisfying $t = [(1 - 2\delta)m]$, where $[x]$ denotes the integer closest to x. The dimension of the new lattice L_1 is

$$w = (1 - \delta)m^2 + o(m^2) \tag{19}$$

and its determinant is bounded by

$$\det(L_1) \leqslant e^{\left(\frac{5}{6} - \frac{\delta}{3} - \frac{\delta^2}{3}\right)m^3 + o(m^3)} \tag{20}$$

Again we consider the inequality

$$\det(L_1) < e^{m(w-1)}/\gamma \tag{21}$$

where $\gamma = (w2^w)^{(w-1)/2}$. Ignore the "small" constant γ and take large enough m (hence $o(m^3)$ is negligible), we must have

$$e^{\left(\frac{5}{6} - \frac{\delta}{3} - \frac{\delta^2}{3}\right)m^3 + o(m^3)} < e^{m(w-1)}/\gamma \tag{22}$$

which can be simplified to

$$m^3\left(-\frac{1}{6} + \frac{2\delta}{3} - \frac{\delta^2}{3}\right) < 0 \tag{23}$$

implying $2\delta^2 - 4\delta + 1 > 0$. Hence, for all

$$\delta < 1 - \frac{\sqrt{2}}{2} \approx 0.292 \tag{24}$$

The RSA cryptosystem is vulnerable to attacks by lattice reduction.

Remark. Boneh–Durfee's result $d < N^{0.292}$ is a best theoretical bound on private key d now. Besides, Blömer and May proposed an improved algorithm [6], which enhances the efficiency by reducing the dimension of the lattice used; meanwhile, reaching a close upper bound $d < N^{0.290}$.

3. Problem with Boneh–Durfee's algorithm

Both of the algorithms above achieve the upper bounds on δ by taking large enough m, and ignoring the "small" constant γ. We find that these two conditions are actually conflicting. To show more clearly, denote $B(n) = \log_2 n$, where $B(n)$ is a real number close to the number of bits for integer n. We have the following Theorem.

Theorem 1. *For an RSA cryptosystem with public key (N, e) and small private key $d = N^\delta$, if we choose large enough $m > 2B(e)/(1 - \delta)$ in Boneh–Durfee's algorithm, then $\gamma = (w2^w)^{w/2}$ is larger than e^{mw}, where $w \approx (1 - \delta)m^2$ is the dimension of the lattice used in the basis reduction procedure.*

Proof. For simplicity, we carry out the proof with "Bit" operation $B(x)$ as defined above. Obviously,

$$B(e^{mw}) \approx B(e^{m\cdot(1-\delta)m^2}) \approx (1 - \delta)m^3 B(e) \tag{25}$$

And on the other hand, we have

$$\begin{aligned} B(\gamma) = B((w2^w)^{w/2}) &\approx \frac{w}{2}\cdot(B(w) + w) \\ &\approx \frac{(1-\delta)m^2}{2}\cdot(2B(m) + (1-\delta)m^2) \\ &> \frac{(1-\delta)^2 m^4}{2} \end{aligned} \tag{26}$$

If we take a large $m > 2B(e)/(1 - \delta)$, then

$$B(\gamma) > \frac{(1-\delta)^2 m^4}{2} = (1-\delta)m^3 B(e)\cdot\frac{m(1-\delta)}{2B(e)} > B(e^{mw}) \tag{27}$$

Thus, we must have $\gamma > e^{mw}$, since γ has more bits than e^{mw}. □

Theorem 1 indicates that, if we choose a large parameter m as required, the "small" constant $\gamma = (w2^w)^{w/2}$ is actually larger than e^{mw}; thus, it is not negligible. On the other hand, if we take a smaller m, then the lower order parts $o(m^3)$ cannot be ignored. This leads to a modification on the upper bound on δ in Boneh–Durfee's algorithm.

More particularly, we find that Boneh–Durfee's algorithm succeeds to break RSA with $\delta < 0.292$ only with sufficiently large N and e (say, 100,000 bits), such that γ and $o(m^3)$ are both negligible. While for practical applications, considering cryptanalysis of RSA with $B(N) \approx 1024$ is more significant. In the next sections, we are to build a more precise analytic model for Boneh–Durfee's algorithm and modify the upper bounds on δ in RSA systems with ordinary public key (N, e).

4. Modified analytic model

For more precise analysis of the upper bound on δ, we take into account both of the effects due to γ and the lower order components in polynomials of m and t. According to Boneh–Durfee's algorithm, take $t = [(1 - 2\delta)m]$, then the dimension of the lattice L_1 used is

$$w \approx -m(m+1)\delta + m^2 \tag{28}$$

To compute the determinant of L_1, divide the lattice into two parts Δ and M'_y corresponding to the x-shifts and y-shifts, respectively (see details in Section 5 in Ref. [4]). The determinant for the Δ component is 2

$$\det(\Delta) = e^{\frac{m(m+1)(m+2)}{3}\delta + \frac{5m(m+1)(m+2)}{12}} \tag{29}$$

and that for M'_y is

$$\det(M'_y) \approx e^{-\frac{m(m+1)(2m+1)}{6}\delta^2 - \frac{m(m+1)(8m+1)}{12}\delta + \frac{m(m+1)(5m+1)}{12}} \tag{30}$$

Then the determinant of the whole lattice L_1 is defined by

$$\det(L_1) = \det(\Delta)\cdot\det(M'_y). \tag{31}$$

Use the denotation $B(n)$ as in Section 3, we can compute

$$B(\det(L_1)) = B(\det(\Delta)) + B(\det(M'_y))$$
$$\approx B(e)\left\{-\frac{m(m+1)(2m+1)}{6}\delta^2 - \frac{m(m+1)(4m-7)}{12}\delta + \frac{m(m+1)(10m+11)}{12}\right\} \quad (32)$$

On the other hand, we have

$$B(e^{mw}) \approx B(e)\{-m^2(m+1)\delta + m^3\} \quad (33)$$

and

$$B(\gamma) \approx \frac{(1-\delta)m^2}{2}\cdot(2B(m) + (1-\delta)m^2) = \frac{m^4}{2}\delta^2 - (m^4 + m^2B(m))\delta + \frac{m^4}{2} + m^2B(m) \quad (34)$$

To satisfy the inequality $\det(L_1) < e^{mw}/\gamma$ so as to apply the cryptanalytic algorithm, we must have

$$B(\det(L_1)) - B(e^{mw}) + B(\gamma) < 0 \quad (35)$$

Plugging in all the values in Eqs. (31)–(33) implies that

$$a\delta^2 + b\delta + c < 0 \quad (36)$$

where a, b, c are variables depending on $B(e)$ and m:

$$a = -\frac{m(m+1)(2m+1)}{6}B(e) + \frac{m^4}{2}$$
$$b = -\frac{m(m+1)(4m-7)}{12}B(e) + m^2(m+1)B(e) - (m^4 + m^2B(m))$$
$$c = \frac{m(m+1)(10m+11)}{12}B(e) - m^3B(e) + \frac{m^4}{2} + m^2B(m) \quad (37)$$

Now we get a modified theoretical upper bound on δ

$$\delta < \frac{-b - \sqrt{(b^2 - 4ac)}}{-2a}, \text{ where } \quad a < 0 \quad (38)$$

Remark. In inequality (36), cases with $a \geqslant 0$ are actually not satisfying. For simplicity, ignore the small constants and lower order parts, and we can simplify a, b and c to $a \approx -\frac{1}{3}B(e) + \frac{m^4}{2}$, $b \approx \frac{2}{3}B(e) - m^4 = -2a$ and $c \approx -\frac{1}{6}B(e) + \frac{m^4}{2}$. When $a = 0$, we have $b \approx 0$ and $c > 0$, so the left part in (36) is larger than 0. Furthermore, when $a > 0$, we have $c > a$ and $b^2 - 4ac < 0$; hence, there is no δ satisfying inequality (36).

5. More precise analysis

The above result is achieved by taking $\gamma = (w2^w)^{(w-1)/2}$, similar to Hinek's analysis [8]. We will show in Section 6 that in this case the theoretic bound is actually far from experimental results. For a more precise analysis, we introduce the following heuristic by Ngyuen, Stehlé and Gama [9,10]:

Heuristic 1. Given as input a random basis of almost any lattice L of sufficiently high dimension d, the LLL algorithm (and its improved version L^2 in [11]) outputs a basis whose first vector b_1 satisfies

$$\|b_1\| \approx 1.02^d \det(L)^{1/d} \quad (39)$$

The heuristic is supported by lots of experimental results, though not proved as theorem. Thereby, γ should be replaced by

$$\gamma \approx (w1.02^w)^{(w-1)/2} \quad (40)$$

Accordingly, the coefficients a, b and c in inequality (38) are also modified:

$$a \approx -\frac{m(m+1)(2m+1)}{6}B(e) + \frac{m^4}{2}\log_2 1.02$$
$$b \approx -\frac{m(m+1)(4m-7)}{12}B(e) + m^2(m+1)B(e) - (m^4\log_2 1.02 + m^2B(m)) \quad (41)$$
$$c \approx \frac{m(m+1)(10m+11)}{12}B(e) - m^3B(e) + \frac{m^4}{2}\log_2 1.02 + m^2B(m)$$

Remark. In Ref. [10], the authors give a more tight bound $\|b_1\| \approx 1.01^d \det(L)^{1/d}$, using the BKZ lattice reduction algorithm, also based on the mass of experimental results. This may still cause a small modification on δ. However, the BKZ algorithm is a blockwise generalization of LLL with potentially super-exponential complexity, and its performance may not satisfy with relatively large m and dimension w in Boneh–Durfee's algorithm.

6. Evaluation for the bounds on δ

According to the analysis above, we compare the theoretic bounds on δ, in the following different cases:

(a) Boneh and Durfee's original analysis, ignoring the "small" constant γ.
(b) Our modification, taking $\gamma = (w2^w)^{(w-1)/2}$.
(c) Our modification, taking $\gamma \approx (w1.02^w)^{(w-1)/2}$.

In Figs. 1 and 2 we show the bounds on δ with $B(e) = 1024$ and $B(e) = 10{,}240$, respectively. In both figures, δ for case (a) is approximating 0.293 with $m \to \infty$, just as Boneh and Durfee have claimed. For case (b), the theoretic bounds are actually far from experimental results; especially, when N and e are small; it is due to the fact that the LLL algorithm only gives a "supremum" for the norm of b_1, but not a tight bound. We care much more for case (c), the theoretic bound on δ is approximately 0.277 when $B(e) = 1024$, and approximately 0.288 when $B(e) = 10{,}240$.

Fig. 3 shows the upper bounds on δ in case (c), with different public key lengths. Notice that $\delta \to 0.292$ when $e \to \infty$.

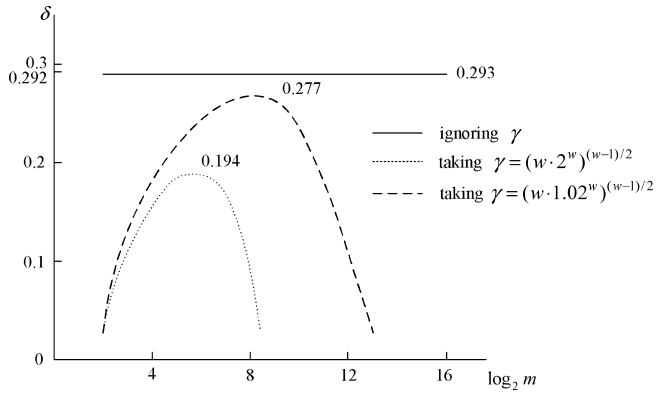

Fig. 1. Bounds evaluation of δ with 1024-bit public key (N, e).

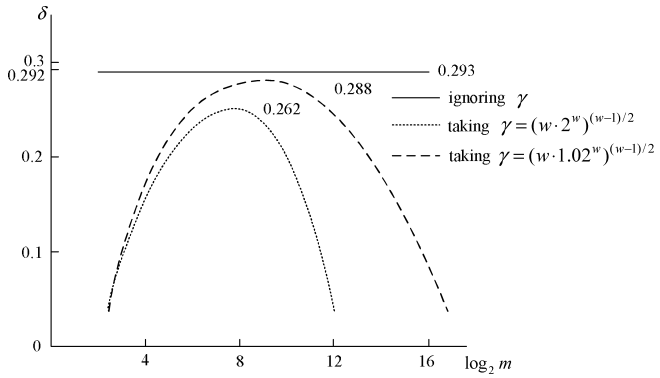

Fig. 2. Bounds evaluation of δ with 10,240-bit public key (N, e).

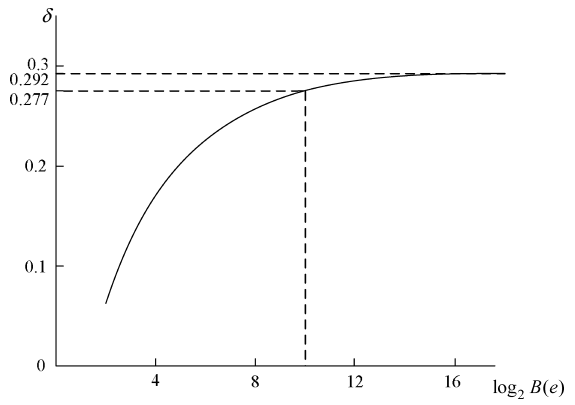

Fig. 3. Bounds evaluation of δ with various public key lengths in case (c).

Actually, the ideal result $d < N^{0.292}$ by Boneh and Durfee could be achieved with a 100,000-bit public key (N, e).

7. Conclusions

We conclude with the following facts for Boneh and Durfee's algorithm:

(1) The theoretic upper bound on δ is much tighter than Boneh and Durfee have claimed; especially, when the public key (N, e) are small.
(2) For RSA cryptosystems with a 1024-bit public key (N, e), the theoretic upper bound on δ is approximately 0.277.
(3) Boneh and Durfee's result, i.e. $\delta < 0.292$, can be achieved only in an ideal case with very large N and e.

Practical applications of the Boneh–Durfee algorithm may behave little better than the modified theoretical bounds in this paper, up to approximately $\delta < 0.280$ in experimental results. The difference is due to that we use Heuristic 1 in evaluation, which is based on the experiments. Actually, predicting and proving the precise output quality of lattice reduction algorithms is still an open problem to be solved.

Acknowledgements

This work is supported by the National Key Basic Research and Development (973) of China (Grant No. 2003 CB314805), 863 Project of China (Grant No. 2008 AA01Z419) and the National Natural Sciences Foundation of China (Grant No. 60873249 and 90304014).

References

[1] Rivest R, Shamir A, Adleman L. A method for obtaining digital signatures and public-key cryptosystems. Commun ACM 1978;21(2):120–6.
[2] Wiener M. Cryptanalysis of short RSA secret exponents. IEEE Trans Inform Theory 1990;36:553–8.
[3] Dujella A. Continued fractions and RSA with small secret exponent. Tatra Mt Math Publ 2004;29:101–12.
[4] Boneh D, Durfee G. Cryptanalysis of RSA with private key d less than N 0.292. IEEE Trans Inform Theory 2000;46(4):1339–49.
[5] Lenstra AK, Lenstra Jr HW, Lovász L. Factoring polynomials with rational coefficients. Math Ann 1982;261:513–34.
[6] Blömer J, May A. Low secret exponent RSA revisited. In: Proceedings of cryptography and lattice conference – CalC 2001. LNCS, vol. 2146. Berlin: Springer-Verlag; 2001. p. 4–19.
[7] Ernst M, Jochemsz E, May A, et al. Partial key exposure attacks on RSA up to full size exponents. In: Advances in cryptology – Eurocrypt 2005. LNCS, vol. 3494. Berlin: Springer-Verlag; 2005, p. 371–86.
[8] Hinek MJ, Low MK, Teske E. On some attacks on multi-prime RSA. In: Proceedings of SAC 2002. LNCS, vol. 2595. Berlin: Springer-Verlag; 2003. p. 385–404.
[9] Nguyen PQ, Stehlé D. LLL on the average. In: Proceedings of ANTS VII. LNCS, vol. 4076. Berlin: Springer-Verlag; 2006. p. 238–56.
[10] Gama N, Nguyen PQ. Predicting lattice reduction. In: Proceedings of Eurocrypt 2008; 2008. p. 31–51.
[11] Nguyen PQ, Stehlé D. Floating-point LLL revisited. In: Proceedings of Eurocrypt 2005. LNCS, vol. 3494. Berlin: Springer-Verlag; 2005, p. 215–33.