APPLICATION OF ECM TO A CLASS OF RSA KEYS

Abderrahmane Nitaj

Département de mathématiques
Université de Caen
Boulevard Maréchal Juin
14032 Caen Cedex, France
E-mail address: nitaj@math.unicaen.fr

July 07, 2006

Abstract. Let $N = pq$ be an RSA modulus where p, q are large primes of the same bitsize and $\phi(N) = (p-1)(q-1)$. We study the class of the public exponents e for which there exist integers X, Y, Z satisfying

$$eX + \phi(N)Y = NZ,$$

with $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ and all prime factors of $|Y|$ are less than 10^{40}. We show that these exponents are of improper use in RSA cryptosystems and that their number is at least $O\left(N^{\frac{1}{2}-\varepsilon}\right)$ where ε is a small positive constant. Our method combines continued fractions, Coppersmith's lattice-based technique for finding small roots of bivariate polynomials and H. W. Lenstra's elliptic curve method (ECM) for factoring.

1 Introduction

The RSA cryptosystem was invented by Rivest, Shamir and Adleman [15] in 1978 and is currently the most widely known and widely used public key cryptosystem. Let p, q be large distinct primes of the same size. A typical size for p and q is 512 bits, i.e., 155 decimal digits. Define $N = pq$ and let e and d be two integers satisfying $ed \equiv 1 \pmod{\phi(N)}$ where $\phi(N) = (p-1)(q-1)$ is the Euler totient function at N. The integers N, e and d are commonly called the *modulus*, the *public exponent* and the *private exponent* respectively. The public key is the pair (N, e) and the secret key is d.

2000 *Mathematics Subject Classification.* 94A60, 11Y05.

Key words and phrases. RSA cryptosystem, Continued fractions, Coppersmith's algorithm, Elliptic curve method, LLL algorithm.

Typeset by $\mathcal{AMS}$-TeX

In order to speed up the decryption time, one may be tempted to use a small secret key. Unfortunately, based on the convergents of the continued fraction expansion of $\frac{e}{N}$, Wiener showed that RSA is insecure if $d < \frac{1}{3}N^{\frac{1}{4}}$. Verheul and van Tilborg [16] and Dujella [6] proposed extensions of Wiener's attack to $d < N^{\frac{1}{4}+\gamma}$ by doing an exhaustive research of about $2\gamma + 8$ bits. In 1999, Boneh and Durfe proposed an attack on RSA with secret exponents $d < N^{0.292}$. This attack is based on Coppersmith's lattice-based technique [5] for finding small roots of bivariate modular polynomials. A similar attack was proposed by Blömer and May [2] and works if $d < N^{0.290}$. In 2004, Blömer and May [3] proposed a generalization of Wiener's attack to the public exponents satisfying $ex + y = k\phi(n)$ with $1 \leq x < \frac{1}{3}N^{\frac{1}{4}}$ and $|y| < N^{-\frac{3}{4}}ex$.

The previous attacks exploit arithmetical properties relating the public key e and $\phi(N)$. Alternative attacks were recently introduced by the author in [12] and [13]. They concern classes of RSA keys with arithmetical properties in connection with special functions $F(p, q)$. In this paper, we exploit another arithmetical property satisfied by the public exponent. Since N, $\phi(N)$ and e are pairwize relatively prime, then the diophantine equation

$$eX + \phi(N)Y = NZ, \tag{1}$$

has infinitely many solutions (see Section 2). First, we will show that it is possible to find X and $Z - Y$ by the continued fraction algorithm if $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$. Next, we will show how to find Y and Z by applying Coppersmith's technique provided $|p - q| < N^{\beta}$ with $\beta < \frac{3}{8}$. More generally, we show how to find Y and Z if all the prime factors of $|Y|$ are less than 10^{40} by factoring $M = |eX - N(Z - Y)|$ using the elliptic curve method (ECM). Our method is based on the combinaison of essential algorithms in computational number theory, namely the continued fraction algorithm, LLL [11] (or PSLQ [1]) and ECM [10]. LLL is the Lenstra, Lenstra, Lovasz lattice basis reduction algorithm, PSLQ is Bailey and Fegusson's partial sum of least squares algorithm and ECM is H. W. Lenstra's elliptic curve method for factoring.

ECM was invented in 1985 by H. W. Lenstra Jr. [10] and is suited to find small prime factors -say up to 40 decimal digits- of large numbers. For notational convenience, let

$$B = 10^{40}, \tag{2}$$

be the ECM bound.

The rest of this paper is organized as follows. In Section 2 we review some well-known facts about the diophantine equation (1), the continued fraction expansion of rational numbers, the lattice based technique of Coppersmith, and finally the elliptic curve method for factoring. In Section 3 we present the general attack combining diophantine approximations and Coppersmith's technique or ECM. Section 4 gives an estimation of the number of the public exponents $e < N$ satisfying (1) with $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$. We briefly conclude the paper in Section 4.

2 Preliminaries

In this section, some facts that will be used troughout this paper are briefly introduced concerning the diophantine equation (1), the continued fraction expansion of a rational number, Coppermith's technique and the elliptic curve method for factoring.

2.1 The diophantine equation $eX + \phi(N)Y = NZ$.

A solution $(X, Y, Z) \in \mathbb{Z}^3$ to the diophantine equation (1) is said to be *proper* if $\gcd(X, Y, Z) = 1$. Since N, $\phi(N)$ and e are pairwise relatively prime, then (1) has infinitely many parametrized proper solutions. To see this, define X_0 to be the unique positive integer satisfying

$$X_0 \equiv -\phi(N)e^{-1} \pmod{N} \qquad \text{with} \qquad 1 \leq X_0 \leq N-1.$$

Then there exists a positive integer Z_0 such that

$$eX_0 + \phi(N) = NZ_0. \tag{3}$$

Next, for any $(a, Y) \in \mathbb{Z}^2$, let

$$X = X_0Y - aN, \qquad Z = Z_0Y - ae.$$

Using (3), we get

$$\begin{aligned} eX + \phi(N)Y &= e(X_0Y - aN) + \phi(N)Y \\ &= (eX_0 + \phi(N))Y - eaN \\ &= NZ_0Y - aeN \\ &= N(Z_0Y - ae) \\ &= NZ. \end{aligned}$$

Hence (X, Y, Z) is a solution to the equation (1). On the other hand, we have

$$\gcd(X, Y) = \gcd(X_0Y - aN, Y) = \gcd(aN, Y) = \gcd(a, Y),$$

unless $\gcd(Y, N)$ is non trivial which is unlikely for an RSA modulus. This means that the proper solutions are obtained by choosing a and Y without common factors.

2.2 The Continued fraction expansion and the Euclidean Algorithm.

Let a and b be relatively prime positive integers. It is well known that the process of finding the continued fraction expansion of the rational number $\frac{a}{b}$ is similar to the application of the Euclidean Algorithm to the pair a and b. The Euclidean Algorithm starts with $r_{-2} = a$, $r_{-1} = b$ and for $i \geq 0$,

$$a_i = \left\lfloor \frac{r_{i-2}}{r_{i-1}} \right\rfloor, \qquad r_i = r_{i-2} - a_i r_{i-1},$$

where $\lfloor x \rfloor$ is the integral part of x. This procedure stops when $r_s = 0$ for some positive integer $s \geq 1$. Using the successive quotients a_i $(i \geq 0)$, we form the sequences $\{p_i\}$ and $\{q_i\}$ according to the double recursions

$$\begin{array}{llll} p_i = a_i p_{i-1} + p_{i-2}, & i \geq 0, & p_{-2} = 0, & p_{-1} = 1, \\ q_i = a_i q_{i-1} + q_{i-2}, & i \geq 0, & q_{-2} = 1, & q_{-1} = 0. \end{array}$$

The rational numbers $\dfrac{p_i}{q_i}$ $(i \geq 0)$ are the convergents of $\dfrac{a}{b}$ and satisfy

$$\gcd(p_i, q_i) = 1, \qquad \text{and} \qquad aq_i - bp_i = (-1)^i r_i.$$

The running time of the Euclidean Algorithm for the positive integers a and b is

$$\mathcal{O}\left(\max\left(\log a, \log b\right)\right).$$

Conversely, to chek whether a rational number $\dfrac{x}{y}$ is a convergent of $\dfrac{a}{b}$ we will use the following classical theorem on diophantine approximations (see Corollary 2, [1, § 2] in [9]).

Theorem 2.1. (Legendre). *Let ξ be a real number. If the coprime integers x and y satisfy*

$$\left|\xi - \frac{X}{Y}\right| < \frac{1}{2Y^2},$$

then $\frac{X}{Y}$ is a convergent of ξ.

2.3 Coppersmith's technique.

Our attack makes use of Coppersmith's method for finding small roots of bivariate polynomials over $\mathbb{Z}$ [5].

Theorem 2.2. (Coppersmith). *Let $f(x,y) \in \mathbb{Z}[x,y]$ which is of maximum degree δ in x and y separately. Suppose that $f(x_0,y_0) = 0$ where $|x_0| \leq \widetilde{X}$, $|y_0| \leq \widetilde{Y}$. Let $W = \left|\left|f\left(\widetilde{X}x, \widetilde{Y}y\right)\right|\right|_\infty$, i.e. the absolute value of the largest coefficient of $f\left(\widetilde{X}x, \widetilde{X}y\right)$. If*

$$\widetilde{X}\widetilde{Y} \leq W^{\frac{2}{3\delta}},$$

then x_0 and y_0 can be found in polynomial time in $\left(\log W, 2^\delta\right)$.

2.4 The elliptic curve method for factoring (ECM).

The principle of ECM is based on Pollard's $(p-1)$-method [14]. We briefly descibe ECM.

Let M be the integer to factor which is divisble by at least two different primes p, q with $p < q$. Let $\mathbb{P}^2(\mathbb{Z}/M\mathbb{Z})$ be the projective plane over $\mathbb{Z}/M\mathbb{Z}$ and $E/\mathbb{Q}$ be an elliptic curve with the homogeonous Weierstrass equation

$$E(\mathbb{Z}/M\mathbb{Z}) = \{(x:y:z) \in \mathbb{P}^2(\mathbb{Z}/M\mathbb{Z}),\ y^2z \equiv x^3 + axz^2 + bz^3 \pmod{M}\},$$

where $a, b \in \mathbb{Z}/M\mathbb{Z}$ with $4a^3 + 27b^2 \neq 0$. The point at infinity is $\mathcal{O} = (0:1:0)$. Let

$$\begin{aligned} E(\mathbb{Z}/M\mathbb{Z}) &\longrightarrow E(\mathbb{Z}/p\mathbb{Z}) \\ P &\longmapsto \widetilde{P} \end{aligned}$$

be the reduction modulo p. Pick at random a point $P \in E(\mathbb{Z}/M\mathbb{Z})$ and fixe two bounds B_1, B_2 with $0 < B_1 < B_2$. The first phase of ECM works as follows. Calculate $Q = kP$ where

$$k = \prod_{\substack{p \leq B_1 \\ p \text{ prime}}} p^{e_p} \qquad \text{with} \qquad e_p = \left\lfloor \frac{\log B_1}{\log p} \right\rfloor.$$

If $\widetilde{Q} = \widetilde{\mathcal{O}}$ and $Q \neq \mathcal{O}$, then the z-coordinate of Q is a multiple of p, hence $\gcd(z, M)$ will reveal the factor p (or another factor of M). If the first phase has not been successful in finding a factor of M, we may continue with a second phase. For each prime p' satisfying $B_1 < p' < B_2$, compute $p'Q = (x' : y' : z')$ and check if $\widetilde{p'Q} = \widetilde{\mathcal{O}}$. This can be done by testing whether $\gcd(z', M) > 1$.

An advantage of ECM over Pollard's $(p-1)$-method is the possibility to choose another elliptic curve if no factor was found. The expected running time of ECM is

$$\mathcal{O}\left(e^{\left(\sqrt{2}+o(1)\right)\sqrt{\log p \log\log p}}\alpha(\log M)\right),$$

to find the smallest factor p of M where $\alpha(\log M)$ is the time required to multiply numbers modulo M and the $o(1)$ term tends to 0 as $p \to +\infty$.

The ECMNET project [7] is devoted to find large factors by ECM. The largest prime factor found thus far by ECM is a 66-digit (220-bit) factor of the 180-digit (598-bit) number $3^{466} + 1$. The computation was carried out by B. Dodson and reported on April 2005.

2.5 A useful lemma.

We state a simple lemma that will be used troughout the paper. We give bounds for $p+q$ when p and q are primes of the same bitsize.

Lemma 2.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then*

$$2\sqrt{N} < p+q < \frac{3\sqrt{2}}{2}\sqrt{N}.$$

Proof. Assume that $q < p < 2q$. Then multiplying by q, we get $q^2 < N < 2q^2$. Similarly, multiplying by p we get $N < p^2 < 2N$. This gives

$$\frac{\sqrt{2}}{2}N^{\frac{1}{2}} < q < N^{\frac{1}{2}} < p < \sqrt{2}N^{\frac{1}{2}}.$$

Rewrite

$$p+q = p+\frac{N}{p}.$$

An easy computation shows that $p+\frac{N}{p}$ is minimized at $p = N^{\frac{1}{2}}$ and maximized at $p = \sqrt{2}N^{\frac{1}{2}}$. This leads to

$$2N^{\frac{1}{2}} < p+q < \frac{3\sqrt{2}}{2}N^{\frac{1}{2}},$$

and terminates the proof. ■

3 The new attack

3.1 Application of Legendre's theorem.

Theorem 3.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Suppose that e satisfies an equation $eX + \phi(N)Y = NZ$ with*

$$|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}.$$

Then $\dfrac{Z-Y}{X}$ is a convergent of $\dfrac{e}{N}$.

Proof. Using $\phi(N) = N - (p+q-1)$, we rewrite the equation $eX + \phi(N)Y = NZ$ as

$$eX - N(Z-Y) = (p+q-1)Y.$$

This gives us

$$\left|\frac{e}{N}-\frac{Z-Y}{X}\right|=\frac{(p+q-1)|Y|}{N|X|}.$$

By Lemma 2.3, we have $p+q-1<\frac{3\sqrt{2}}{2}N^{\frac{1}{2}}$. Then under the assumption $|XY|<\frac{\sqrt{2}}{6}N^{\frac{1}{2}}$, we get

$$\left|\frac{e}{N}-\frac{Z-Y}{X}\right|\leq\frac{3\sqrt{2}N^{\frac{1}{2}}|Y|}{2N|X|}<\frac{1}{2X^2},$$

so that, by Legendre's theorem 2.1, $\frac{Z-Y}{X}$ is a convergent of $\frac{e}{N}$ which proves the lemma. ■

3.2 Application of Coppersmith's technique.

In this section, we present an algorithm to factor N if the public exponent e satisfies (1) with $|XY|<\frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ and the prime difference $p-q$ satisfies

$$p-q\leq N^{\beta}\qquad\text{with}\qquad\beta<\frac{3}{8},$$

Our approach makes use of Coppersmith's technique (Theorem 2.2). Notice that if $p-q\leq N^{\frac{1}{4}}$, an algorithm of Fermat finds the factorization of N in polynomial time (see [17]). Our approach makes use of the following lemma.

Lemma 3.2. *Let β be a real value. Let $N=pq$, where p and q are two prime integers such that $q<p<2q$ and $p-q\leq N^{\beta}$. Let $S=\left\lfloor 2N^{\frac{1}{2}}\right\rfloor+1$. Then*

$$0<p+q-1-S<N^{\beta}.$$

Proof. Assume that $q<p<2q$ and $p-q\leq N^{\beta}$. Let

$$S=\left\lfloor 2N^{\frac{1}{2}}\right\rfloor+1.$$

By the definition of the intgeral part, we have

$$S\leq 2N^{\frac{1}{2}}+1<S+1.$$

Combining this with Lemma 2.3, we get $S+1\leq 2N^{\frac{1}{2}}+2<p+q$ and $p+q-1-S>0$. Again applying Lemma 2.3, we get

$$2q<2N^{\frac{1}{2}}<2N^{\frac{1}{2}}+1<S+1.$$

It follows that $p+q-1-S<p-q<N^{\beta}$ which terminates the proof. ■

Theorem 3.3. *Let β and δ be real values such that $\delta + 4\beta \leq \frac{3}{2}$. Let $N = pq$, where p and q are two prime integers such that $q < p < 2q$ and $p - q \leq N^{\beta}$. Let e be a public exponent and X, Y, Z be unknown integers satisfying $eX + \phi(N)Y = NZ$ with $|Y| \leq N^{\delta}$ and $|X| \leq \frac{\sqrt{2}}{6}N^{\frac{1}{2}-\delta}$. Then given (N, e) one can factor N in polynomial time.*

Proof. Assume that $|Y| \leq N^{\delta}$ and $|X| \leq \frac{\sqrt{2}}{6}N^{\frac{1}{2}-\delta}$. Then $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$, and, by Theorem 3.1, we can find X and $Z - Y$ among the convergents of $\frac{e}{N}$. Rewrite (1) as $(p+q-1)Y = eX - N(Z-Y)$. Let $M = |eX - N(Z-Y)|$. We want to solve the equation

$$(p + q - 1)|Y| = M, \tag{4}$$

with the unknowns p, q and Y. Assume that $p - q \leq N^{\beta}$. Let

$$S = \left\lfloor 2N^{\frac{1}{2}} \right\rfloor + 1.$$

Then, by Lemma 3.2, we have $|p+q-1-S| < N^{\beta}$, which means that S is an approximation of $p + q - 1$ with an error term of at most N^{β}. Put

$$\widetilde{X} = N^{\beta}, \tag{5}$$

and define x_0 by

$$x_0 = p + q - 1 - S.$$

Then $|x_0| < \widetilde{X}$. Next, let

$$T = \left\lfloor \frac{M}{S} \right\rfloor.$$

Since $p + q - 1 > S > 2N^{\frac{1}{2}}$, then

$$|Y| = \frac{M}{p + q - 1} < \frac{M}{S}.$$

Hence $|Y| \leq \left\lfloor \frac{M}{S} \right\rfloor = T$. Combining this and $p + q - 1 > 2N^{\frac{1}{2}}$, we get

$$\begin{aligned} 0 < T - |Y| &< \frac{M}{S} - \frac{M}{p+q-1} \\ &= \frac{M(p+q-1-S)}{(p+q-1)S} \\ &< \frac{MN^{\beta-\frac{1}{2}}}{2S}. \end{aligned}$$

Hence T is an approximation of $|Y|$ with an error term bounded by

$$\widetilde{Y} = \frac{MN^{\beta-\frac{1}{2}}}{2S}. \tag{6}$$

Define y_0 by

$$y_0 = |Y| - T.$$

Then $|y_0| < \widetilde{Y}$. Using $p+q-1 = S + x_0$ and $|Y| = T + y_0$ in (4), we get

$$(S + x_0)(T + y_0) = x_0y_0 + Tx_0 + Sy_0 + ST = M.$$

We can define the following polynomial

$$f(x, y) = xy + Tx + Sy + ST - M,$$

with a root $(x_0, y_0) = (p + q - 1 - S, |Y| - T)$ satisfying

$$|x_0| < \widetilde{X}, \qquad |y_0| < \widetilde{Y}.$$

Let W denote the largest absolute value of the coefficients of $f\left(\widetilde{X}x, \widetilde{Y}y\right)$ with

$$f(\widetilde{X}x, \widetilde{Y}y) = \widetilde{X}\widetilde{Y}xy + T\widetilde{X}x + S\widetilde{Y}y + ST - M.$$

Using (6), we get

$$W = \max\left(\widetilde{X}\widetilde{Y}, T\widetilde{X}, S\widetilde{Y}, |ST - M|\right) \geq S\widetilde{Y} = \frac{1}{2}MN^{\beta-\frac{1}{2}}.$$

On the other hand, combining (5) and (6), we have

$$\widetilde{X}\widetilde{Y} = \frac{MN^{2\beta-\frac{1}{2}}}{2S} < \frac{MN^{2\beta-\frac{1}{2}}}{4N^{\frac{1}{2}}} = \frac{1}{4}MN^{2\beta-1}.$$

In order to apply Coppersmith's theorem (Theorem 2.2), we have to satisfy the condition

$$\left(\widetilde{X}\widetilde{Y}\right)^3 \leq W^2.$$

This leads to

$$\frac{1}{64}M^3N^{6\beta-3} \leq \frac{1}{4}M^2N^{2\beta-1},$$

which in turn gives

$$M \leq 16N^{2-4\beta}.$$

By assumption $|Y| < N^{\delta}$. Then applying Lemma 2.3 we get

$$M = (p + q - 1)|Y| < \frac{3\sqrt{2}}{2}N^{\frac{1}{2}}N^{\delta} = \frac{3\sqrt{2}}{2}N^{\frac{1}{2}+\delta}.$$

Hence, it suffices that

$$\frac{3\sqrt{2}}{2}N^{\frac{1}{2}+\delta} \leq 16N^{2-4\beta}.$$

From this we get

$$N^{4\beta+\delta-\frac{3}{2}} \leq \frac{16\sqrt{2}}{3},$$

which is satisfied if $4\beta + \delta - \frac{3}{2} \leq 0$. With this condition, Coppersmith's technique finds the solution (x_0, y_0). Finaly, using $x_0 = p + q - 1 - S$, we can find the factorization of N. This terminates the proof. ■

3.3 Application of ECM.

In this section, we present an algorithm that on input (N, e) outputs the factors p and q of N if e satisfies the equation (1) with $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ and if all prime factors of $|Y|$ are less than the ECM bound $B = 10^{40}$.

Let $N = pq$ be an RSA modulus with $q < p < 2q$ and e a public exponent satisfying

$$eX + \phi(N)Y = NZ,$$

for some unknown X, Y and Z such that $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$. Remember that X and $Z - Y$ could be recovered using Theorem 3.1. Let $M = |eX - N(Z - Y)|$. Then (1) transforms to a factorization problem, namely

$$M = (p + q - 1)|Y|,$$

with the unknown factors $p + q - 1$ and $|Y|$. A crucial step in our attack consists in computing a set of divisors of M by extracting the primes $p_1, p_2, \cdots, p_s$ that divide M and are less than B. Since only partial factorization of M is required, ECM is a good candidate for this task. Write

$$M = p_1^{r_1} p_2^{r_2} \cdots p_s^{r_s} M', \qquad \text{with} \qquad 1 < p_1 < p_2 < \cdots < p_s \leq B,$$

to be the factorizaion of M as a product of powers of distinct primes less than B where $M' = 1$ or M' has no prime divisor less than B. Then the B-smooth divisors of M are the numbers

$$d = p_1^{x_1} p_2^{x_2} \cdots p_s^{x_s}, \qquad \text{with} \qquad 0 \leq x_i \leq r_i.$$

Define $\tau_B(M)$ to be the number of the B-smooth divisors of M. Then

$$\tau_B(M) = \prod_{i=1}^{s}(x_i + 1).$$

By the prime number theorem

$$p_1 p_2 \cdots p_s = e^{(1+o(1))p_s} = e^{(1+o(1))s \log s} \leq M.$$

From this we find the inequality $s \log s < \log M$ and so

$$s < C\frac{\log M}{\log\log M}, \tag{7}$$

where C is an absolute constante (see [8]). This gives an upper bound for the number of the prime factors of the B-smooth part of M. For every i, $1 \leq i \leq s$, we have $2^{x_i} \leq p_i^{x_i} \leq M$. Then

$$x_1 \leq \frac{\log M}{\log 2}. \tag{8}$$

Let

$$\varepsilon_0 = \frac{\log 2}{\log\left(2\sqrt{N}\right)}.$$

By Lemma 2.3, we have $M \geq p+q-1 > 2N^{\frac{1}{2}}$. then

$$\varepsilon_0 \geq \frac{\log 2}{\log(M)}. \tag{9}$$

Combining (8) and (9), we derive an upper bound for $\tau_B(M)$

$$\tau_B(M) = \prod_{i=1}^{s}(x_i+1) \leq \prod_{i=1}^{s}(1+\varepsilon_0)\frac{\log M}{\log 2} = (1+\varepsilon_0)^s\left(\frac{\log M}{\log 2}\right)^s.$$

In practice, from Hardy and Ramanujan, we know that the average number $\tau(M)$ of divisors of M is

$$\tau(M) \sim \log M \leq \log N,$$

and $\tau_B(M) \leq \tau(M)$. Morever, let

$$D_1 = \left\lfloor \frac{M}{\frac{3\sqrt{2}}{2}\sqrt{N}} \right\rfloor \qquad \text{and} \qquad D_2 = \left\lceil \frac{M}{2\sqrt{N}} \right\rceil.$$

Since $|Y| = \frac{M}{p+q-1}$ is a divisor of M, then by Lemma 2.3, we have

$$D_1 \leq |Y| \leq D_2. \tag{10}$$

Writing $|Y| = p_1^{x_1}p_2^{x_2}\cdots p_s^{x_s}$ and taking logarithms in (10), we obtain

$$\log D_1 \leq x_1 \log p_1 + x_2 \log p_2 + \cdots + x_s \log p_s \leq \log D_2,$$

in the unknowns $0 \leq x_1 \leq r_1$, $0 \leq x_2 \leq r_2$, $\ldots$, $0 \leq x_s \leq r_s$. This is a subset problem which can be solved with polynomial time algorithms such LLL [11] (the exact integral method of de Weger [18]) or PSLQ [1]. On the other hand, the estimation (7) gives an upper bound for the number of the prime factors of $|Y|$, and consequently, the number of the entries in PSLQ algorithm and the dimension of the lattice used by LLL to solve the subset problem (e.g. $s = 81$ for a 1024 bit integer N).

Finally, let d be a divisor of M so that $w = \frac{M}{d}$ is a candidate for $p+q-1$. Then, using $N = pq$, we get the quadratic equation

$$p^2 - (w+1)\,p + N = 0,$$

which is solvable for $p \in \mathbb{N}$ if $\Delta = (w+1)^2 - 4N$ is a perfect square.

Summarising, we describe the algorithm attack as follows:

Algorithm :

```
INPUT: $(N, e)$, where $N = pq$ and $eX + \phi(N)Y = NZ$ for some unknown $X$, $Y$, $Z$
        with $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ and all prime factors of $|Y|$ are less than $B = 10^{40}$.
1. Compute the continued fraction expansion of $\frac{e}{N}$.
2. For every convergent $\frac{u}{v}$ with $v < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ :
    (i) Compute $M = |ev - Nu|$.
    (ii) Compute $D_1 = \left\lfloor \frac{M}{\frac{3\sqrt{2}}{2}\sqrt{N}} \right\rfloor$ and $D_2 = \left\lceil \frac{M}{2\sqrt{N}} \right\rceil$.
    (iii) Find the $B$-smooth part $M_1$ of $M$ by ECM.
    (iv) Find the divisors $d$ of $M_1$ in the interval $[D_1, D_2]$ by LLL or PSLQ.
    (v) For every such divisor $d$ :
        (a) Compute $w = \frac{M}{d}$.
        (b) Compute $\Delta = (w+1)^2 - 4N$.
        (c)) If $\Delta$ is a perfect square then compute $\tilde{p} = \frac{(w+1)+\sqrt{\Delta}}{2}$.
        (d) If $\tilde{p}|N$, then stop.
3. Output $p = \tilde{p}$, $q = \frac{N}{\tilde{p}}$.
```

It is well known that the continued fraction algorithm has polynomial time complexity. It follows that Step 1 of the algorithm outputs at most $\mathcal{O}(\log N)$ convergents of $\frac{e}{N}$. Let $p_s < B$ be the largest prime factor of $M = |eX - N(Z-Y)| = (p+q-1)|Y|$ with unknown factors $p+q-1$ and $|Y|$. ECM will find p_s in

$$\exp\left(\left(\sqrt{2}+o(1)\right)\sqrt{\log p_s \log\log p_s}\right) \approx \exp\left(\left(\sqrt{2}+o(1)\right)\sqrt{\log B \log\log B}\right),$$

expected running time. On the other hand, LLL performs Step 2, (iv) on lattices with dimension at most $\frac{\log M}{\log\log M}$ and entries smaller than $\log p_s \leq \log B$. Hence Step (2), (iv) terminates in

$$\mathcal{O}\left(\left(\frac{\log M}{\log\log M}\right)^6 (\log B)^3\right) = \mathcal{O}\left(\left(\frac{\log N}{\log\log N}\right)^6 (\log B)^3\right).$$

Finally, Step 2, (v) concerns $\mathcal{O}(\log N)$ divisors.

Summarising, our attack requires

$$\mathcal{O}\left(\left(\frac{\log N}{\log\log N}\right)^6 (\log B)^3 (\log N)^2 \exp\left(\left(\sqrt{2}+o(1)\right)\sqrt{\log B \log\log B}\right)\right)$$

operations.

4. The number of the exponents satisfying (1)

In this section we give a lower bound of the size of the public exponents satisfying the equation (1) with $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$. To do this, we define a subclass of public exponents by

$$e \equiv -\frac{\phi(N)Y}{X} \pmod N \qquad \text{with} \qquad 1 \leq e < N,$$

where X and Y are two integers satisfying $\gcd(X, NY) = 1$, $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ and all prime factors of $|Y|$ are less than the ECM bound $B = 10^{40}$. The following lemma shows that different tuples (X, Y) lead to different public keys.

Lemma 4.1. *Let X, Y, X', Y' be integers with* $\gcd(X, NY) = \gcd(X', NY') = 1$, $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ *and* $|X'Y'| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$. *Let*

$$e \equiv -\frac{\phi(N)Y}{X} \pmod N, \qquad e' \equiv -\frac{\phi(N)Y'}{X'} \pmod N$$

with $1 \leq e, e' < N$. If $(X, Y) \neq (X', Y')$, then $e \neq e'$.

Proof. Suppose that $(X, Y) \neq (X', Y')$ and, for contradiction, that $e = e'$ where

$$e \equiv -\frac{\phi(N)Y}{X} \pmod N, \qquad e' \equiv -\frac{\phi(N)Y'}{X'} \pmod N.$$

Since $\gcd(X, N) = \gcd(X', N) = 1$, then $e = e'$ implies

$$\phi(N)YX' \equiv \phi(N)XY' \pmod N.$$

This means that there exists an integer k such that

$$\phi(N)(YX' - YX') = Nk.$$

Since $\gcd(\phi(N), N) = 1$, then N is a factor of $YX' - YX'$ which is impossible since

$$|YX' - YX'| \leq |YX'| + |YX'| < \frac{\sqrt{2}}{3}N^{\frac{1}{2}} < N.$$

Hence $(X, Y) = (X', Y')$ which terminates the proof. ■

Theorem 4.2. *The size of the set of the exponents e satisfying $e \equiv -\frac{\phi(N)Y}{X} \pmod N$ with* $\gcd(X, NY) = 1$*,* $|XY| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}}$ *and all prime factors of* $|Y|$ *are less than* $B = 10^{40}$ *is at least*

$$O\left(N^{\frac{1}{2}-\varepsilon}\right),$$

where ε is a small positive constant.

Proof. Define α_0 by

$$N^{\alpha_0} = B = 10^{40}.$$

Let Ω denote the number of the exponents satisfying

$$e \equiv -\frac{\phi(N)Y}{X} \pmod N,$$

with $\gcd(X, NY) = 1$, $|X| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}-\alpha_0}$ and $|Y| < N^{\alpha_0}$. Let $Y_0 = \lfloor N^{\alpha_0} \rfloor$ and $X_0 = \left\lfloor \frac{\sqrt{2}}{6}N^{\frac{1}{2}-\alpha_0} \right\rfloor$. Then

$$\begin{aligned}\Omega = \sum_{|Y|=1}^{Y_0} \sum_{\substack{|X|=1\\ \gcd(X,NY)=1}}^{X_0} 1 &= \sum_{|Y|=1}^{Y_0} \sum_{\substack{|X|=1\\ \gcd(X,Y)=1}}^{Y_0-1} 1 + \sum_{|Y|=1}^{Y_0} \sum_{\substack{|X|=Y_0\\ \gcd(X,Y)=1}}^{X_0} 1\\ &> \sum_{|Y|=1}^{Y_0} \sum_{\substack{|X|=1\\ \gcd(X,Y)=1}}^{|Y|-1} 1 + \sum_{|X|=Y_0}^{X_0} \sum_{\substack{|Y|=1\\ \gcd(X,Y)=1}}^{Y_0} 1.\end{aligned}$$

First consider the sum

$$S_1 = \sum_{|Y|=1}^{Y_0} \sum_{\substack{|X|=1\\ \gcd(X,Y)=1}}^{|Y|-1} 1.$$

Observe that

$$\sum_{\substack{|X|=1\\ \gcd(X,Y)=1}}^{|Y|-1} 1 = \phi(|Y|).$$

Recall that $\phi(\cdot)$ is the Euler totient function and satisfies (see [19])

$$\phi(x) \geq \frac{e^{-\gamma}x}{9\log\log(x)}, \tag{11}$$

where γ is the Euler-Mascheroni constant. Applying this with $|Y| < N^{\alpha_0}$, we get

$$\phi(|Y|) \geq \frac{e^{-\gamma}|Y|}{9\log\log|Y|} \geq \frac{e^{-\gamma}|Y|}{9\log\log(N^{\alpha_0})} = N^{-\varepsilon_1}|Y|,$$

where ε_1 is a small positive constant. Hence

$$S_1 = \sum_{|Y|=1}^{Y_0} \sum_{\substack{|X|=1\\ \gcd(X,Y)=1}}^{|Y|-1} 1 > \sum_{|Y|=1}^{Y_0} \phi(|Y|) \geq \sum_{|Y|=1}^{Y_0} N^{-\varepsilon_1}|Y| = Y_0(Y_0+1)N^{-\varepsilon_1}. \tag{12}$$

Next consider the sum

$$S_2 = \sum_{|X|=Y_0}^{X_0} \sum_{\substack{|Y|=1\\ \gcd(X,Y)=1}}^{Y_0} 1.$$

Using the Möbius function $\mu(\cdot)$, we have

$$\sum_{\substack{|Y|=1\\ \gcd(X,Y)=1}}^{Y_0} 1 \geq \ Y_0 \sum_{\substack{d||X|\\ \gcd(X,Y)=1}} \frac{\mu(d)}{d} \geq Y_0 \frac{\phi(|X|)}{|X|}.$$

Applying (11) with $|X| < \frac{\sqrt{2}}{6}N^{\frac{1}{2}-\alpha_0}$, we get

$$\phi(|X|) \geq \frac{e^{-\gamma}|X|}{9\log\log|X|} \geq \frac{e^{-\gamma}|X|}{9\log\log\left(\frac{\sqrt{2}}{6}N^{\frac{1}{2}-\alpha_0}\right)} \geq N^{-\varepsilon_2}|X|,$$

where ε_2 is a small positive constant. Hence

$$S_2 = \sum_{|X|=Y_0}^{X_0} \sum_{\substack{|Y|=1\\ \gcd(X,Y)=1}}^{Y_0} 1 \geq \sum_{|X|=Y_0}^{X_0} Y_0 \frac{\phi(|X|)}{|X|} \geq \sum_{|X|=Y_0}^{X_0} Y_0 N^{-\varepsilon_2} = 2Y_0(X_0-Y_0)N^{-\varepsilon_2}. \tag{13}$$

Plugging (12) and (13) in Ω and setting $\varepsilon = \max(\varepsilon_1, \varepsilon_2)$, we get

$$\Omega > S_1 + S_2 \geq Y_0(Y_0+1)N^{-\varepsilon} + 2Y_0(X_0-Y_0)N^{-\varepsilon} \geq X_0Y_0N^{-\varepsilon} \approx \frac{\sqrt{2}}{6}N^{\frac{1}{2}-\varepsilon}.$$

This gives the claimed result. ■

5. Conclusion

In this paper we investigated two attacks on the RSA keys e satisfying the linear diophantine equation $eX + \phi(N)Y = NZ$ with a small solution. The standard algorithms for solving this task are the continued fraction algorithm, the lattice based method of Coppersmith for solving bivariate polynomials, the elliptic curve method for factoring and the LLL (or the PSLQ) algorithm. For this problem, we were able to lower-bound the number of keys satisfying $e < N$ and $eX + \phi(N)Y = NZ$ with a small solution by $O\left(N^{\frac{1}{2}-\varepsilon}\right)$ where ε is a small positive constant.

References

1. D. H. Bailey and D. J. Broadhurst, *Parallel integer relation detection: techniques and applications*, Math. Comp. **70** (2000), 1719–1736.
2. J. Blömer, A. May, *Low secret exponent RSA revisited*, In Cryptography and Lattices - Proceedings of CALC '01, Lecture Notes in Computer Science, Springer-Verlag **2146** (2001), 4–19.
3. J. Blömer, A. May, *A generalized Wiener attack on RSA*, In Practice and Theory in Public Key Cryptography (PKC 2004), Lecture Notes in Computer Science, Springer-Verlag **2947** (2004), 1–13.
4. D. Boneh, G. Durfee, *Cryptanalysis of RSA with private key d less than $N^{0.292}$*, IEEE Transactions on Information Theory **46** (2000), 1339–1349.
5. D. Coppersmith, *Small solutions to polynomial equations, and low exponent RSA vulnerabilities*, Journal of Cryptology **10 (4)** (1997), 223–260.
6. A. Dujella, *Continued fractions and RSA with small secret exponent*, Tatra Mt. Math. Publ. **29** (2004), 101–112.
7. The ECMNET project, `http://www.loria.fr/ zimmerma/records/ecmnet.html`.
8. G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, 5th ed. Oxford University Press, 1981.
9. S. Lang, *Introduction to Diophantine Approximations*, Addison-Wesley Pub. Co, 1966.
10. H.W. Lenstra Jr., *Factoring integers with elliptic curves*, Annals of Mathematics, **126** (1987), 649–693.
11. A.K. Lenstra, H.W. Lenstra Jr., and L. Lovasz, *Factoring polynomials with rational coefficients*, Math. Annalen, **261** (1982), 515–534.
12. A. Nitaj, *Cryptanalysis of RSA with constrained keys*, Appl. Algebra Eng. Commun. Comput., Submitted.
13. A. Nitaj, *RSA and a higher degree diophantine equation*, Appl. Math. Comput., Submitted.
14. J. Pollard, *A Monte Carlo method for factorization*, Nordisk Tidskrift for Informationsbehandlung (BIT), **15** (1975), 331–334.
15. R. Rivest, A. Shamir, L. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Communication of ACM **21** (1978), 120–126.
16. E. R. Verheul, H. C. A. van Tilborg, *Cryptanalysis of 'less short' RSA secret exponents*, Appl. Algebra Eng. Commun. Comput. **8** (1997), 425–435.
17. B. de Weger, *Cryptanalysis of RSA with small prime difference*, Appl. Algebra Eng. Commun. Comput. **13** (2002), 17–28.
18. B. de Weger, *Algorithms for Diophantine Equations*, CWI Tract, vol. 65, (1989), Stichting Mathematisch Centrum, Amsterdam..
19. E. W. Weisstein, *Totient Function*, From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/TotientFunction.html`.
20. M. Wiener, *Cryptanalysis of short RSA secret exponents*, IEEE Transactions on Information Theory **36** (1990), 553-558.