A New Vulnerable Class of Exponents in RSA

Abderrahmane Nitaj Laboratoire de Mathmatiques Nicolas Oresme
Universit de Caen, France

nitaj@math.unicaen.fr

http://www.math.unicaen.fr/~nitaj

Abstract

Let $N = pq$ be an RSA modulus, i.e. the product of two large unknown primes of equal bit-size. We consider the class of the public exponents satisfying an equation $eX - NY = (ap+bq)Z$ with $0 < a < q$, $b = \left[\frac{ap}{q}\right]$ (here $[x]$ denotes the nearest integer to x) and

$$|XZ| < \frac{N}{2(ap+bq)},$$

and all prime factors of $|Z|$ are less than 10^{50}. Using the continued fraction algorithm and the Elliptic Curve Method of factorization, we show that such exponents yield the factorization of the RSA modulus. Further, we show that the number of such weak keys is at least $N^{\frac{1}{2}-\varepsilon} \log N$. Thus, our attack applies to a relatively large class of weak keys in RSA.

KEYWORDS: RSA, Cryptanalysis, Factorization, Continued Fraction

1 Introduction

Invented by Rivest, Shamir and Adleman in 1977 [15], the RSA cryptosystem is most commonly used for providing privacy and ensuring authenticity of digital data. It is one of the most popular systems in use today. Let $N = pq$ be the product of two large primes of the same bit-size. Let e and d be two positive integers satisfying $ed \equiv 1 \pmod{\phi(N)}$ where $\phi(N) = (p-1)(q-1)$ is Euler's totient function. Commonly, N is called the RSA modulus, e the

encryption exponent and d the decryption exponent. The modular equation $ed \equiv 1 \pmod{\phi(N)}$ is sometimes used as an equation $ed - k\phi(N) = 1$, where k is some positive integer and is called the RSA key equation.

Since its publication, RSA has been analyzed for vulnerability by various methods (see [3] and [9]). The security of RSA is based on the well known problem of factoring large integers, which is widely believed to be intractable. If one can factor $N = pq$, then one can calculate $\phi(N) = (p-1)(q-1)$ and the private exponent d by solving the congruence $ed \equiv 1 \pmod{\phi(N)}$. Conversely, the recovery of the private exponent d is equivalent to factoring N (see [2]). Many attacks are based on trying to solve the key equation $ed - k\phi(N) = 1$ for particular exponents. In 1990, Wiener [16] showed that using continued fractions, one can efficiently recover the secret-exponent d from the public key (N, e) as long as $d < \frac{1}{3}N^{\frac{1}{4}}$. The number of such exponents can be estimated as $N^{\frac{1}{4}-\varepsilon}$ where $\varepsilon > 0$ is arbitrarily small for large N. The $N^{-\varepsilon}$ term corresponds to the number of exponents d such that $\gcd(d, \phi(N)) \neq 1$. The lattice-based Boneh-Durfee attack [4] and its variant given by Blömer and May [1] exploit the non-linear equation satisfied by the secret key

$$k\left(\frac{N+1}{2} + s\right) - 1 \equiv 0 \pmod{e},$$

where k and $s = -\frac{p+q}{2}$ are unknown integers. This gives an attack that heuristically succeeds in polynomial-time when $d < N^{0.292}$ and the number of the exponents for which this attack works can be estimated as $N^{0.292-\varepsilon}$. Based on the continued fraction algorithm and a theorem due to Coppersmith [7], Blömer and May [2] proposed in 2004 an attack on RSA using the variant equation $ex + y = k\phi(N)$ with $x < \frac{1}{3}N^{\frac{1}{4}}$ and $|y| = \mathcal{O}\left(N^{-\frac{3}{4}}ex\right)$. They showed that such exponents are vulnerable and that their number is at least $N^{\frac{3}{4}-\varepsilon}$. At Africacrypt 2008, Nitaj presented an attack on the class of the exponents satisfying an equation $eX - (p-u)(q-v)Y = 1$ with

$$1 \leq Y \leq X < 2^{-\frac{1}{4}}N^{\frac{1}{4}}, \quad |u| < N^{\frac{1}{4}}, \quad v = \left[-\frac{qu}{p-u}\right],$$

where $[x]$ denotes the nearest integer to the real number x. Combining the continued fraction algorithm, Coppersmith's method [7] and the Elliptic Curve Method of factorization [10], he showed that p, q can be found if all the prime factors of $p-u$ or $q-v$ are less than 10^{50}. The number of such exponents are estimated as $N^{\frac{1}{2}-\varepsilon}$. In a similar direction, Maitra and Sarkar [11]

presented in 2008 an attack using the equation $eX-(N-pu-v)Y=1$ with suitably small parameters X, Y, u and v. In contrast of the previous attacks, this attack uses only the technique of Coppersmith [7]. The number of such exponents are estimated as $N^{\frac{3}{4}-\varepsilon}$. In 2009, Nitaj [14] studied the class of the exponents e satisfying $eX-(N-(ap+bq))Y=Z$ where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ and X, Y, Z are suitably small integers. He showed that such exponents form a weak class of size $N^{\frac{3}{4}-\varepsilon}$.

In this paper, we introduce a new attack on RSA. The attack works for all public keys (N,e) satisfying an equation

$$eX-NY=(ap+bq)Z,$$

with

$$b=\left[\frac{ap}{q}\right], \quad |XZ|<\frac{N}{2(ap+bq)},$$

where a is an unknown positive integer satisfying $a<q$ ($[x]$ means the nearest integer of the real number x) and all prime factors of Z are less than the Elliptic Curve Method of factorization [10] bound $B_{ECM}=10^{50}$. The new attack is based on the continued fraction algorithm and the Elliptic Curve Method (ECM). In contrast to the previous attacks, it does not make use of Coppersmith's technique. We show that for integers X, a, b and Z within the given bounds, this attack yields the factorization of the RSA modulus $N=pq$ for the class of the public exponents e with the structure

$$e\equiv (ap+bq)ZX^{-1} \pmod{N},$$

where $\gcd(X,(ap+bq)Z)=1$. Observe that this condition implies $\gcd(X,Y)=1$ where Y satisfies $eX-NY=(ap+bq)Z$. We show that the number of the exponents e with $e<N$ for which this method works can be estimated as $N^{\frac{1}{2}-\varepsilon}$.

The new attack works as follows. We use the continued fraction algorithm to recover X and Y among the convergents of $\frac{e}{N}$. Then we use the Elliptic Curve Method to recover Z among the divisors of $eX-NY$. Afterwards, we find p and q using the equation $ap+bq=\frac{|eX-NY|}{|Z|}$ and the inequality $|ap-bq|<2\sqrt{N}$. This yields the factorization of N.

The remainder of this paper is organized as follows. In Section 2, we begin with some notations and a brief review of basic facts about the continued fraction algorithm and the Elliptic Curve Method of factorization. In Section

3, we present some useful lemmas needed for the attack. In Section 4 we present our attack on RSA. In section 5, we estimate the size of the exponents that are weak with this attack. We conclude in Section 6.

2 Preliminaries

Let x be a real number. In this paper, $[x]$ denotes the nearest integer to x and $\lfloor x \rfloor$ the largest integer less than or equal to x.

2.1 The Continued Fraction Algorithm

Here we recall some facts from the theory of continued fractions. Let $x \neq 0$ be a real number. Put $x_0 = x$ and $a_0 = \lfloor x_0 \rfloor$. Recursively, for $n \geq 1$, if $x_{n-1} \neq a_{n-1}$, define

$$x_n = \frac{1}{x_{n-1} - a_{n-1}}, \quad a_n = \lfloor x_n \rfloor.$$

For $n \geq 1$, since $0 < x_{n-1} - a_{n-1} < 1$, then $x_n > 1$ and $a_n \geq 1$. This process is called the continued fraction algorithm and yields an expression of the form

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots}}},$$

which is called the continued fraction expansion of x. This expression is often used in the form $x = [a_0, a_1, a_2, \cdots]$. Any rational number $\frac{a}{b}$ can be expressed as a finite continued fraction $x = [a_0, a_1, a_2, \cdots, a_m]$. For $i \geq 0$ ($0 \leq i \leq m$ in the finite case), we define the i^{th} convergent of the continued fraction $[a_0, a_1, a_2, \cdots]$ to be $[a_0, a_1, a_2, \cdots, a_i]$. Each convergent is a rational number. The main results from the theory of continued fractions that we use in this paper are the following two theorems (see, e.g. Theorem 164 and Theorem 184 of [8]).

Theorem 2.1. *Let x be a real number. If $\frac{Y}{X}$ is a convergent of x, then*

$$\left| x - \frac{Y}{X} \right| < \frac{1}{X^2}.$$

Theorem 2.2. *Let x be a real number. If X and Y are coprime integers such that*

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

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

2.2 The Elliptic Curve Factorization Method (ECM)

The Elliptic Curve Method of factoring (ECM) was originally proposed by H.W. Lenstra [10] and subsequently extended by Brent [5], [6] and Montgomery [12]. The first part of the method proposed by Lenstra is called Phase 1, and the extension by Brent and Montgomery is called Phase 2. ECM is suited to find small prime factors of large numbers. Let n be an integer with a prime factor p. Let E be a random elliptic curve defined over $\mathbb{Z}/n\mathbb{Z}$ by a projective equation

$$y^2 z \equiv x^3 + axz^2 + bz^3 \pmod{n},$$

where $(x : y : z) \in \mathbb{P}^2(\mathbb{Z}/n\mathbb{Z})$, the projective plane over $\mathbb{Z}/n\mathbb{Z}$ and $a, b \in \mathbb{Z}/n\mathbb{Z}$. The point at infinity is $O = (0 : 1 : 0)$. Let $P \in E(\mathbb{Z}/n\mathbb{Z})$. Let B_1 and B_2 be two bounds for prime numbers with $B_1 < B_2$. Phase 1 and phase 2 of ECM work as follows:
Phase 1: Calculate $Q = (x_Q : y_Q : z_Q) = kP$ where

$$k = \prod_{\substack{q=2 \\ q \text{ prime}}}^{B_1} q^{e_q}, \qquad e_q = \left\lfloor \frac{\log B_1}{\log q} \right\rfloor.$$

If the order of E over $\mathbb{Z}/p\mathbb{Z}$ divides k, then Q could be the point at infinity of $E(\mathbb{Z}/p\mathbb{Z})$, which means that z_Q is a multiple of p. Thus $p = \gcd(z_Q, n)$.
Phase 2: For each prime number k such that $B_1 < k < B_2$, calculate $Q = (x_Q : y_Q : z_Q) = kP$ and test if $1 < \gcd(z_Q, n) < n$. For similar reasons as in Phase 1, this can reveals a prime factor of n.

Let $M(n)$ be the cost of one multiplication (mod n). Then ECM finds a factor p of n with the sub-exponentiel run time

$$\mathcal{O}\left(\exp\left\{c\sqrt{\log p \log\log p}\right\} M(n)\right),$$

where $c \approx 2$ is a constant. According to Brent (see [6]), the evolution of the ECM record satisfies the equation

$$\sqrt{D} = \frac{Y - 1932.3}{9.3},$$

where D is the decimal digits in the largest factor found by ECM up to the date Y. Extrapolating, a 69-digit factor could be found in 2010. In([17]), it is announced that a 73-digit prime factor of the special number $2^{1181} - 1$ was found by J. Bos, T. Kleinjung, A. Lenstra and P. Montgomery in 2010. Consequently, in this paper, we consider that ECM is efficient to find prime factors up to the the bound $B_{ECM} = 10^{50}$.

3 Useful Lemmas

In this section, we prove three useful lemmas. We begin by a very simple lemma on the size of the primes of the RSA modulus $N = pq$.

Lemma 3.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then*

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

Proof. Multiplying $q < p < 2q$ by p we get $N < p^2 < 2N$, which gives $N^{\frac{1}{2}} < p < 2^{\frac{1}{2}} N^{\frac{1}{2}}$. Similarly, multiplying by q we get $q^2 < N < 2q^2$ which gives in turn $2^{-\frac{1}{2}} N^{\frac{1}{2}} < q < N^{\frac{1}{2}}$. This terminates the proof. □

A key role in all our arguments is played by the following lemma.

Lemma 3.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let a be an integer and $b = \left\lceil \frac{ap}{q} \right\rfloor$. Set $S = ap + bq$. Then*

$$ab = \left\lfloor \frac{S^2}{4N} \right\rfloor \quad \text{and} \quad |ap - bq| = \sqrt{S^2 - 4 \left\lfloor \frac{S^2}{4N} \right\rfloor N}.$$

Proof. Let a be an integer and let $b = \left\lceil \frac{ap}{q} \right\rfloor$. Set $S = ap + bq$. Since

$$\left| \frac{ap}{q} - b \right| \leq \frac{1}{2},$$

then using Lemma 3.1, we get

$$|ap - bq| \leq \frac{q}{2} < \frac{1}{2}N^{\frac{1}{2}}.$$

On the other hand, $S^2 - 4abN = (ap + bq)^2 - 4abN = (ap - bq)^2 > 0$. Hence

$$0 < \frac{S^2}{4N} - ab = \frac{S^2 - 4abN}{4N} = \frac{(ap - bq)^2}{4N} < \frac{\left(\frac{1}{2}N^{\frac{1}{2}}\right)^2}{4N} = \frac{1}{16} < 1,$$

from which we deduce $ab = \left\lfloor \frac{S^2}{4N} \right\rfloor$ according to the definition of the floor function. On the other hand, using again $S^2 - 4abN = (ap - bq)^2 > 0$, we get

$$|ap - bq| = \sqrt{S^2 - 4abN} = \sqrt{S^2 - 4\left\lfloor \frac{S^2}{4N} \right\rfloor N}.$$

This terminates the proof. □

We terminate with the following lemma which will be used for counting the number of exponents that are vulnerable to our attack.

Lemma 3.3. *Let m and n be positive integers. Then*

$$m\frac{\phi(n)}{n} - 2^{\omega(n)} < \sum_{\substack{k=1 \\ \gcd(k,n)=1}}^{m} 1 < m\frac{\phi(n)}{n} + 2^{\omega(n)},$$

where $\omega(n)$ is the number of distinct prime factors of n.

Proof. Let $\mu(d)$ be the Möbius function which is defined by $\mu(1) = 1$, $\mu(d) = 0$ if d is not square-free and $\mu(d) = (-1)^{\omega(d)}$ if d is square-free where $\omega(d)$ is the number of distinct prime factors of d. Then the Lengedre Formula gives

$$\sum_{\substack{k=1 \\ \gcd(k,n)=1}}^{m} 1 = \sum_{d|n} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor.$$

Since $\left\lfloor \frac{m}{d} \right\rfloor \leq \frac{m}{d} < \left\lfloor \frac{m}{d} \right\rfloor + 1$, then

$$\begin{aligned}
\sum_{d|n} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor &= \sum_{\substack{d|n \\ \mu(d)=1}} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor + \sum_{\substack{d|n \\ \mu(d)=-1}} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor \\
&> \sum_{\substack{d|n \\ \mu(d)=1}} \mu(d) \left(\frac{m}{d} - 1 \right) + \sum_{\substack{d|n \\ \mu(d)=-1}} \mu(d) \frac{m}{d} \\
&= \sum_{d|n} \mu(d) \frac{m}{d} - \sum_{\substack{d|n \\ \mu(d)=1}} 1 \\
&> m \sum_{d|n} \frac{\mu(d)}{d} - \sum_{d|n} |\mu(d)| \\
&= m \frac{\phi(n)}{n} - 2^{\omega(n)}.
\end{aligned}$$

The Möbius function satisfies (see 16.3.1 of [8])

$$\sum_{d|n} \frac{\mu(d)}{d} = \frac{\phi(n)}{n}.$$

It also satisfies $\sum_{d|n} |\mu(d)| = 2^{\omega(n)}$ (Theorem 264 of [8]). It follows that

$$\sum_{\substack{k=1 \\ \gcd(k,n)=1}}^{m} 1 > m \frac{\phi(n)}{n} - 2^{\omega(n)}.$$

On the other hand, using $\left\lfloor \frac{m}{d} \right\rfloor \leq \frac{m}{d} < \left\lfloor \frac{m}{d} \right\rfloor + 1$, we get

$$\begin{aligned}
\sum_{d|n} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor &= \sum_{\substack{d|n \\ \mu(d)=1}} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor + \sum_{\substack{d|n \\ \mu(d)=-1}} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor \\
&< \sum_{\substack{d|n \\ \mu(d)=1}} \mu(d) \frac{m}{d} + \sum_{\substack{d|n \\ \mu(d)=-1}} \mu(d) \left(\frac{m}{d} - 1 \right) \\
&= \sum_{d|n} \mu(d) \frac{m}{d} + \sum_{\substack{d|n \\ \mu(d)=-1}} 1 \\
&< m \sum_{d|n} \frac{\mu(d)}{d} + \sum_{d|n} |\mu(d)| \\
&= m \frac{\phi(n)}{n} + 2^{\omega(n)}.
\end{aligned}$$

where we used similar results. This concludes the proof. □

4 The New Class of Weak Keys in RSA

In this section, we present a strategy to find the prime factors p and q of the modulus $N = pq$ of an RSA instance with a public exponent e satisfying an equation $eX - NY = (ap + bq)Z$ for some a, b and Z satisfying

$$0 < a < q, \quad b = \left\lceil \frac{ap}{q} \right\rceil, \quad |XZ| < \frac{N}{2(ap + bq)}.$$

Notice that if $X < 0$ then $e(-X) - N(-Y) = (ap + bq)(-Z)$ where $-X > 0$. Thus, for symmetrical reason, we will consider only on the scenario when $X > 0$.

4.1 The New Attack

We begin by linking the solutions of the equation $eX - NY = (ap + bq)Z$ with the convergents of $\frac{e}{N}$.

Theorem 4.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let a, b be integers such that $0 < a < q$ and $b = \left\lceil \frac{ap}{q} \right\rceil$. Let e be an exponent satisfying*

the equation $eX - NY = (ap + bq)Z$. *If* $\gcd(X, Y) = 1$ *and* $X|Z| < \frac{N}{2(ap+bq)}$, *then* $\frac{Y}{X}$ *is a convergent of* $\frac{e}{N}$.

Proof. Assume $0 < a < q$ and $b = \left[\frac{ap}{q}\right]$. Suppose $eX - NY = (ap + bq)Z$ with $X > 0$. Then

$$\left|\frac{e}{N} - \frac{Y}{X}\right| = \frac{|(ap + bq)Z|}{NX}.$$

Assume $X|Z| < \frac{N}{2(ap+bq)}$. Then $\frac{|(ap+bq)Z|}{NX} < \frac{1}{2X^2}$. From this we deduce

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

and by Theorem 2.2, we conclude that $\frac{Y}{X}$ is one of the convergents of the continued fraction expansion of $\frac{e}{N}$. □

Notice that the continued fraction algorithm is a polynomial time algorithm and the number of convergents of $\frac{e}{N}$ is bounded by $\mathcal{O}(\log N)$. This results in a very efficient method for finding the solution (X, Y) in the equation $eX - NY = (ap + bq)Z$. The following lemma shows how to find the parameter Z assuming the efficiency of the Elliptic Curve method.

Lemma 4.2. *Let* $N = pq$ *be an RSA modulus with* $q < p < 2q$. *Let* e *be an exponent satisfying the equation* $eX - NY = (ap + bq)Z$ *with positive integers* a, b *where* $\frac{Y}{X}$ *is a convergent of* $\frac{e}{N}$. *If all prime factors of* Z *are less than the ECM-bound* B_{ECM}, *then* Z *can be found efficiently.*

Proof. Suppose e satisfies $eX - NY = (ap + bq)Z$ where $\frac{Y}{X}$ is a convergent of $\frac{e}{N}$ and all prime factors of Z are less than B_{ECM}. Set $M = |eX - NY|$. Let $p_1, p_2, \cdots, p_k$ denote the distinct prime factors of M that are less than B_{ECM}. Such primes can be efficiently determined by applying ECM to M, namely, let

$$M = M' \prod_{i=1}^{k} p_i^{\alpha_i},$$

be the factorization of M where $M' = 1$ or M' has no prime factor less than B_{ECM}. By results of Hardy and Ramanujan (see, e.g., Theorem 430 and Theorem 431 of [8]), we know that, in average, the number of prime divisors

of M is $\mathcal{O}(\log\log M)$ if M is uniformly distributed. Since $\frac{Y}{X}$ is a convergent of $\frac{e}{N}$, then by Theorem 2.1, we have

$$M = (ap + bq)|Z| = |eX - NY| < \frac{N}{X} \leq N.$$

It follows that the average number of prime divisors of M is bounded by $\log\log N$. On the other hand, we know that $M = (ap + bq)|Z|$ where $|Z|$ is B_{ECM}-smooth. Hence $|Z|$ is a divisor of $\prod_{i=1}^{k} p_i^{\alpha_i}$. Thus

$$|Z| = \prod_{i=1}^{k} p_i^{x_i}, \quad \text{with} \quad 0 \leq x_i \leq \alpha_i.$$

The number of divisors of $\prod_{i=1}^{k} p_i^{\alpha_i}$ is $\prod_{i=1}^{k}(1 + \alpha_i)$. Nevertheless, by results of Dirichlet on the distribution of divisors of a random integer (see, e.g., Theorem 432 of [8]), the number of divisors of M is $\mathcal{O}(\log M)$ where $M < N$. This shows that the average number of candidates for Z is polynomially bounded by $\mathcal{O}(\log N)$. This results in an efficient way to find Z depending the efficiency of ECM. □

Given X, Y and Z satisfying $eX - NY = (ap + bq)Z$. The following theorem shows how to find the remainder parameters, namely a, b, p and q.

Theorem 4.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let e be a public exponent satisfying the equation $eX - NY = (ap+bq)Z$ with $0 < a < q$ and $b = \left\lceil \frac{ap}{q} \right\rceil$. If $\frac{Y}{X}$ is a convergent of $\frac{e}{N}$ and Z is a B_{ECM}-smooth integer, then p and q can be found in polynomial time.*

Proof. Suppose that e satisfies the equation $eX - NY = (ap + bq)Z$ with known parameters X, Y and Z where p, q, a and b are the unknown parameters. Assume in addition that $0 < a < q$ and $b = \left\lceil \frac{ap}{q} \right\rceil$. Define

$$S = ap + bq = \frac{eX - NY}{Z}$$

By Lemma 3.2 we get

$$ab = \left\lfloor \frac{S^2}{4N} \right\rfloor \quad \text{and} \quad |ap - bq| = \sqrt{S^2 - 4\left\lfloor \frac{S^2}{4N} \right\rfloor N}.$$

Combining with $ap + bq = S$, we find

$$ap = \frac{S \pm \sqrt{S^2 - 4\left\lfloor \frac{S^2}{4N} \right\rfloor N}}{2}.$$

Since $0 < a < q$, we recover p using $p = \gcd(ap, N)$. Hence $q = \frac{N}{p}$. Since every calculation can be done in polynomial time, this concludes the proof. □

Now, we give the factorization algorithm.

Algorithm 1 The new attack

Input: a public key (N, e) satisfying $N = pq$, $q < p < 2q$ and $eX - NY = (ap+bq)Z$ for small parameters X, Y, Z where $\gcd(X, Y) = 1$, $0 < a < q$ and $b = \left\lfloor \frac{ap}{q} \right\rfloor$.

Output: the prime factors p and q.

Compute the continued fraction expansion of $\frac{e}{N}$.
For every convergent $\frac{Y}{X}$ of $\frac{e}{N}$ with $X < \frac{1}{2}N^{\frac{1}{2}}$ **do**
  Compute $M = |eX - NY|$ and apply ECM to find the B_{ECM}-smooth part M_0 of M.
  Compute the divisors of M_0.
  For every divisor Z of M_0 with $Z < \frac{\sqrt{N}}{2X}$ **do**
    Compute $S = \frac{M}{Z}$, $N_0 = \left\lfloor \frac{S^2}{4N} \right\rfloor$ and $D = \sqrt{|S^2 - 4N_0N|}$.
    If $D \in \mathbb{N}$ **then**
      Compute $p = gcd\left(N, \frac{S+D}{2}\right)$.
      If $1 < p < N$ **then**
        Output p, $q = \frac{N}{p}$, Stop
      End if
    End if
  End for
End for

5 Estimation of Weak Keys

In this section, we give a very conservative estimation of the number of exponents for which our attack works. To be more precise, let a be an

integer with $0 < a < q$ and let $b = \left\lfloor \frac{ap}{q} \right\rfloor$. Define α such that $ap + bq = N^{\frac{1}{2}+\alpha}$. Let us consider the number of exponents e satisfying an equation

$$eX - NY = ap + bq,$$

with $\gcd(X, ap + bq) = 1$, $e < N$ and $X < \frac{N}{2(ap+bq)}$. Observe that since $\gcd(X, N) = 1$, then reducing the equation $eX - NY = ap + bq$ modulo N yields

$$e \equiv (ap + bq)X^{-1} \pmod{N}.$$

We begin by the following result. It shows that for fixed a and b, different parameters X_1, X_2 with $X_1, X_2 < X < \frac{N}{2(ap+bq)}$ define different exponents e_1, e_2 of the desired form.

Lemma 5.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let a and b be positive integers such that $a < q$ and $b < p$. For $i = 1, 2$, let e_i be exponents satisfying $e_i < N$ and $e_i X_i - NY_i = ap + bq$ with $\gcd(X_i, ap + bq) = 1$ and $X_i < \frac{N}{2(ap+bq)}$. If $X_1 \neq X_2$ then $e_1 \neq e_2$.*

Proof. Suppose that for $i = 1, 2$, we have $e_i X_i - NY_i = ap + bq$. Assume for contradiction that $e_1 = e_2$. Then

$$(ap + bq)X_1^{-1} \equiv (ap + bq)X_2^{-1} \pmod{N},$$

which can be rewritten as

$$(ap + bq)\left(X_1^{-1} - X_2^{-1}\right) \equiv 0 \pmod{N}.$$

Notice that, since $\gcd(ap + bq, N) = 1$, then $X_2^{-1} - X_1^{-1} \equiv 0 \pmod{N}$ and $X_2 - X_1 \equiv 0 \pmod{N}$. On the other hand, we have

$$|X_2 - X_1| \leq X_2 + X_1 < 2\left(\frac{N}{2(ap + bq)}\right) < N.$$

Hence $X_2 - X_1 = 0$ and $X_1 = X_2$. This terminates the proof. □

Now we present a result which will be required for counting the weak exponents e with the structure $e \equiv (ap+bq)X^{-1} \pmod{N}$ where $\gcd(X, (ap+bq)) = 1$ and $X < \frac{N}{2(ap+bq)}$.

Lemma 5.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let a and b be fixed positive integers such that $a < q$ and $\left|b - \left[\frac{ap}{q}\right]\right| \leq 1$. Define α by $ap+bq = N^{\frac{1}{2}+\alpha}$. The number of the exponents e of the form $e \equiv (ap+bq)X^{-1} \pmod N$ with $\gcd(X, ap+bq) = 1$ and $X < \frac{1}{2}N^{\frac{1}{2}-\alpha}$ is $\mathcal{O}\left(N^{\frac{1}{2}-\alpha-\varepsilon}\right)$ where $\varepsilon > 0$ is arbitrary small for suitably large N.*

Proof. Let a be a positive integer. Define α by $ap + bq = N^{\frac{1}{2}+\alpha}$ and let

$$X_0 = \left\lfloor \frac{1}{2} N^{\frac{1}{2}-\alpha} \right\rfloor.$$

Let $\mathcal{N}(a)$ denote the number of the exponents e satisfying $e \equiv (ap+bq)X^{-1} \pmod N$ with $\gcd(X, ap+bq) = 1$ and $X < \frac{1}{2}N^{\frac{1}{2}-\alpha}$. We have

$$\mathcal{N}(a) = \sum_{\substack{X=1 \\ \gcd(X, ap+bq)=1}}^{X_0} 1. \tag{1}$$

Using Lemma 3.3 with $n = ap + bq$ and $m = X_0$, we get

$$X_0 \frac{\phi(ap+bq)}{ap+bq} - 2^{\omega(ap+bq)} < \mathcal{N}(a) < X_0 \frac{\phi(ap+bq)}{ap+bq} + 2^{\omega(ap+bq)}, \tag{2}$$

Here, $2^{\omega(ap+bq)}$ is the number of square free divisors of $ap+bq$ which is upper bounded by the total number $\tau(ap+bq)$ of divisors of $ap+bq$. We recall that $\tau(n)$ satisfies $\tau(n) = \mathcal{O}(\log\log n)$ (Theorems 430-431 of [8]). It follows that the dominant term in (2) is $X_0 \frac{\phi(ap+bq)}{ap+bq}$. Using this with $n = ap + bq = N^{\frac{1}{2}+\alpha}$ and $X_0 = \left\lfloor \frac{1}{2} N^{\frac{1}{2}-\alpha} \right\rfloor$, this leads to

$$\mathcal{N}(a) = \mathcal{O}\left(N^{-2\alpha} \phi\left(ap+bq\right)\right).$$

On the other hand, for $n \geq 2$, we have (see Theorem 328 of [8])

$$\phi(n) > \frac{cn}{\log\log n},$$

where c is a positive constant. Taking $n = ap + bq = N^{\frac{1}{2}+\alpha}$, this implies

$$\mathcal{N}(a) = \mathcal{O}\left(\frac{N^{\frac{1}{2}-\alpha}}{\log\log N^{\frac{1}{2}+\alpha}}\right) = \mathcal{O}\left(N^{\frac{1}{2}-\alpha-\varepsilon}\right),$$

where ε satisfies $N^{\varepsilon} = \log\log N$ and depends only on N. This terminates the proof. □

Now we are able to prove an estimation for the number of exponents for which our attack works. Actually, we give a very conservative estimation with $a = 1$. Observe that since $q < p < 2q$, then $b = \left[\frac{p}{q}\right]$ satisfies

$$1 \leq \left[\frac{p}{q}\right] \leq 2,$$

so that $b = 1$ or $b = 2$.

Theorem 5.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. An estimation of the number of exponents $e < N$ such that $e \equiv (p+q)X^{-1} \pmod{N}$ is $\mathcal{O}\left(N^{\frac{1}{2}-\varepsilon}\right)$ where $\varepsilon > 0$ is arbitrarily small for suitably large N.*

Proof. Let $\mathcal{N}$ denote the number of exponents $e < N$ with the structure $e \equiv (p+q)X^{-1} \pmod{N}$ for some $X < \frac{N}{2(p+q)}$. Using Lemma 5.2, we get

$$\mathcal{N} = \mathcal{O}\left(N^{\frac{1}{2}-\alpha-\varepsilon_1}\right),$$

where α satisfies $p + q = N^{\frac{1}{2}+\alpha}$. Using Lemma 3.1, we get

$$\left(2^{-\frac{1}{2}}+1\right) N^{\frac{1}{2}} < p + q < \left(2^{\frac{1}{2}}+1\right) N^{\frac{1}{2}}.$$

It follows that α satisfies

$$\frac{\log\left(2^{-\frac{1}{2}}+1\right)}{\log N} < \alpha < \frac{\log\left(2^{\frac{1}{2}}+1\right)}{\log N}.$$

This implies that α is arbitrarily small for large N. Consequently, we get

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

where $\varepsilon = \varepsilon_1 + \alpha$ is arbitrarily small for large N. This terminates the proof. □

6 Conclusion

In this paper, we studied the class of exponents e satisfying an equation

$$eX - NY = (ap + bq)Z,$$

where X, Y, Z, a and b are integers satisfying

$$0 < a < q, \quad b = \left[\frac{ap}{q}\right], \quad X|Z| < \frac{N}{2(ap + bq)},$$

and all prime factors of Z are less than the Elliptic Curve Method of factorization bound $B_{ECM} = 10^{50}$. Using the continued fraction algorithm and the Elliptic Curve Method of factorization (ECM), we showed that such exponents are vulnerable and lead to the factorization of the RSA modulus $N = pq$. We also showed that the new class of weak exponents is sufficiently large since the size of this class can be estimated as $N^{\frac{1}{2}-\varepsilon}$ where $\varepsilon > 0$ is arbitrary small for suitably large N.

References

[1] Blömer, J., May, A.: Low secret exponent RSA revisited. In: Silverman, J.H. (ed.) CaLC 2001. LNCS, vol. 2146, pp. 419. Springer, Heidelberg (2001).

[2] Blömer, J., May, A.: A generalized Wiener attack on RSA. In Public Key Cryptography - PKC 2004, volume 2947 of Lecture Notes in Computer Science, Springer-Verlag (2004), 1-13.

[3] Boneh, D.: Twenty years of attacks on the RSA cryptosystem. Notices of the American Mathematical Society (AMS) 46 (2) (1999), 203-213 .

[4] Boneh, D., Durfee, G.: Cryptanalysis of RSA with private key d less than $N^{0.292}$, Advances in Cryptology Eurocrypt'99, Lecture Notes in Computer Science Vol. 1592, Springer-Verlag (1999), 1-11.

[5] Brent, R.P.: Some integer factorization algorithms using elliptic curves, Australian Computer Science Communications, vol. 8 (1986), 149-163.

[6] Brent, R.P.: Recent progress and prospects for integer factorisation algorithms, Springer-Verlag LNCS 1858 (2000), 3–22.

[7] Coppersmith, D.: Small solutions to polynomial equations, and low exponent RSA vulnerabilities. Journal of Cryptology, 10(4) (1997), 233-260.

[8] Hardy, G.H., Wright, E.M.: An Introduction to the Theory of Numbers. Oxford University Press, London (1975).

[9] Hinek, J.M.: Cryptanalysis of RSA and Its Variants, Chapman & Hall/CRC Press (2009).

[10] Lenstra, H.W.: Factoring integers with elliptic curves, Annals of Mathematics, vol. 126 (1987), 649-673.

[11] Maitra, S., Sarkar, S.: Revisiting Wiener's Attack - New Weak Keys in RSA. ISC 2008 (2008), 228–243.

[12] Montgomery, P.L.: Speeding the Pollard and elliptic curve methods of factorization, Mathematics of Computation, vol. 48 (1987), 243-264 .

[13] Nitaj, A.: Another generalization of Wiener's attack on RSA, In: Vaudenay, S. (ed.) Africacrypt 2008. LNCS, vol. 5023, Springer, Heidelberg (2008), 174–190.

[14] Nitaj, A.: Cryptanalysis of RSA using the ratio of the primes, In: B. Preneel (Ed.) Africacrypt 2009, LNCS 5580, 2009. Springer-Verlag, Berlin Heidelberg (2009), 98–115.

[15] Rivest, R., Shamir A., Adleman, L.: A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, Communications of the ACM, Vol. 21 (2) (1978), 120-126.

[16] Wiener, M.: Cryptanalysis of short RSA secret exponents, IEEE Transactions on Information Theory, Vol. 36 (1990), 553-558.

[17] Zimmermann, P.: 50 largest factors found by ECM `http://www.loria.fr/~zimmerma/records/top50.html`.