J. Appl. Math. Comput. (2015) 48:349–363
DOI 10.1007/s12190-014-0806-1

ORIGINAL RESEARCH

Implicit factorization of unbalanced RSA moduli

Abderrahmane Nitaj · Muhammad Rezal Kamel Ariffin

Received: 30 March 2014 / Published online: 26 July 2014

Abstract Let $N_1 = p_1q_1$ and $N_2 = p_2q_2$ be two RSA moduli, not necessarily of the same bit-size. In 2009, May and Ritzenhofen proposed a method to factor N_1 and N_2 given the implicit information that p_1 and p_2 share an amount of least significant bits. In this paper, we propose a generalization of their attack as follows: suppose that some unknown multiples a_1p_1 and a_2p_2 of the prime factors p_1 and p_2 share an amount of their Most Significant Bits (MSBs) or an amount of their Least Significant Bits (LSBs). Using a method based on the continued fraction algorithm, we propose a method that leads to the factorization of N_1 and N_2. Using simultaneous diophantine approximations and lattice reduction, we extend the method to factor $k \geq 3$ RSA moduli $N_i = p_iq_i, i = 1, \ldots, k$ given the implicit information that there exist unknown multiples $a_1p_1, \ldots, a_kp_k$ sharing an amount of their MSBs or their LSBs. Also, this paper extends many previous works where similar results were obtained when the p_i's share their MSBs or their LSBs.

Mathematics Subject Classification 94A60 · 11Y05

1 Introduction

Research in determining pre-requisites for strong primes for the integer factorization problem (IFP) of a product of two primes $N = pq$ has been intriguing and have

A. Nitaj (✉)
Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Basse Normandie, Caen, France
e-mail: abderrahmane.nitaj@unicaen.fr

M. R. K. Ariffin
Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Department of Mathematics, Faculty of Science, Universiti Putra Malaysia (UPM), Selangor, Malaysia
e-mail: rezal@upm.edu.my

captured the attention of researchers since IFP came into prominence via the RSA algorithm. The simplicity of the problem statement raised interest on whether such a simple problem statement describing the IFP could only be solved in exponential time for all cases, i.e. all types of primes. As can be found in the literature, this is not the case. So-called weak primes were identified by researchers and this caused an avalanche of research output on this matter. In this paper, we focus on IFP when $N = pq$ is unbalanced, that is when q is much smaller than p.

In PKC 2009, May and Ritzenhofen [5] presented a method for factoring large integers with some implicit hints. More precisely, let $N_1 = p_1q_1$ and $N_2 = p_2q_2$ be two RSA moduli of the same bit-size such that q_1 and q_2 are α-bit primes and p_1 and p_2 share at least t least significant bits (LSBs). The method of May and Ritzenhofen is a lattice based method that allows to find the factorization of N_1 and N_2 when $t \geq 2\alpha + 3$. May and Ritzenhofen's method heuristically generalizes to a lattice based method to simultaneously factor k RSA moduli $N_1 = p_1q_1, \ldots, N_k = p_kq_k$ when the p_i's share $t \geq \frac{k}{k-1}\alpha$ many LSBs.

In [8], Sarkar and Maitra reconsidered the method of May and Ritzenhofen for two RSA moduli. Sarkar and Maitra's method works when $N_1 = p_1q_1$ and $N_2 = p_2q_2$ are such that p_1 and p_2 share their LSBs or most significant bits (MSBs) as well as a contiguous portion of bits at the middle.

In PKC 2010, Faugère, Marinier and Renault [1] presented a new and rigorous lattice-based method that addresses the implicit factoring problem when p_1 and p_2 share t MSBs. Moreover, when $N_1 = p_1q_1$ and $N_2 = p_2q_2$ are two RSA moduli of the same bit-size and the prime factors q_i are α-bit primes, the method of Faugère et al. provably factors N_1 and N_2 as soon as p_1 and p_2 share $t \geq 2\alpha + 3$ MSBs. The method heuristically generalizes to the case when p_1 and p_2 share an amount of bits in the middle. It also heuristically generalizes to k RSA moduli $N_1 = p_1q_1, \ldots, N_k = p_kq_k$ when the p_i's share $t \geq \frac{k}{k-1}\alpha + 6$ of MSBs.

In IWSEC 2013, Kurosawa and Ueda [3] presented a lattice-based method to factor two RSA moduli $N_1 = p_1q_1$ and $N_2 = p_2q_2$ of the same bit size when p_1 and p_2 share t LSBs with $t \geq 2\alpha + 1$ where $q_1 \approx q_2 \approx 2^\alpha$. Their method takes advantage on using Gaussian reduction techniques. It slightly improves the bound $t \geq 2\alpha + 3$ of May and Ritzenhofen. We notice that Kurosawa and Ueda did not study a number of possible extensions of their method, namely, when p_1 and p_2 share t MSBs and also when the multiple of the primes share LSB's and MSB's.

All the former attacks apply when the RSA moduli $N_1 = p_1q_1, \ldots, N_k = p_kq_k$ are of the same bit-size and the p_i's share an amount of MSBs, LSBs or bits in the middle. In this paper, we present novel approaches of implicit factoring that generalize the former attacks and apply when some unknown multiples a_ip_i of the prime factors p_i share an amount of MSBs or of LSBs.

Our first method concerns two RSA moduli $N_1 = p_1q_1$, $N_2 = p_2q_2$ of arbitrarily sizes in the situation that there exist two integers a_1, a_2 such that a_1p_1 and a_2p_2 share t many MSBs. We show that, using the continued fraction expansion of $\frac{N_2}{N_1}$, one can factor simultaneously N_1 and N_2 whenever $|a_1p_1 - a_2p_2| < \frac{p_1}{2a_2^2q_1q_2}$. In particular, when N_1 and N_2 are of the same bit size and q_1, q_2 are α-bit primes, then one can factor N_1 and N_2 whenever $a_i \leq 2^\beta$ for $i = 1, 2$ and $t \geq 2\alpha + 2\beta + 1$. When $\beta = 0$,

Table 1 Applicability of the methods for k RSA moduli

Method	MSBs		LSBs	
	$k=2$	$k\geq 3$	$k=2$	$k\geq 3$
May, Ritzenhofen [5]	No	No	Yes	Yes
Sarkar, Maitra [8]	Yes	No	Yes	No
Faugère et al. [1]	Yes	Yes	No	No
Kurosawa, Ueda [3]	No	No	Yes	No
Our methods	Yes	Yes	Yes	Yes

that is $a_1 = a_2 = 1$, our result becomes $t \geq 2\alpha+1$ and improves the bound $t \geq 2\alpha+3$ presented in [8] and [1] where the methods are based on lattice reduction techniques.

Our second method is a heuristic generalization of the first method to an arbitrary number $k \geq 3$ of RSA moduli $N_i = p_i q_i$, $i = 1, \ldots, k$ in the situation that there exist k integers a_i such that the $a_i p_i$'s share t many MSBs. When the RSA moduli are of the same bit size and the factors q_i, $i = 1, \ldots, k$, are α-bit primes, the method allows us to factor the RSA moduli as soon as

$$t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right), \tag{1}$$

where β is such that $a_i \leq 2^\beta$. Once again, with $\beta = 0$, we improve the bound presented in the attack of [1].

Our third method addresses the implicit factoring problem when two unbalanced RSA moduli $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ of arbitrarily sizes are such that there exist two integers a_1 and a_2 such that $a_1 p_1$ and $a_2 p_2$ share t many LSBs. We show that it is possible to factor both N_1 and N_2 if $a_1 a_2 q_1 q_2 < 2^{t-1}$. This method is also based on the continued fraction algorithm, applied to $\frac{T}{2^t}$ where $T \equiv N_2 N_1^{-1} \pmod{2^t}$. We notice that, when $a_1 = a_2 = 1$ and q_1, q_2 are α-bit primes, the former condition on t transforms to $t \geq 2\alpha + 1$ which improves the bound on t for LSBs in [5] and [8] and retrieves the bound of [3].

Our fourth method is a generalization of the third method to $k \geq 3$ RSA moduli $N_i = p_i q_i$, $i = 1, \ldots, k$. Assume that there exist k integers a_i such that the $a_i p_i$'s share t many LSBs. If the RSA moduli are of the same bit size and the q_i's are α-bit primes, our method allows us to address the implicit factoring problem whenever t satisfies (1) where β is such that $a_i \leq 2^\beta$.

In fact our findings under the four scenarios, further discus possible malicious key generation of RSA moduli by observing not only the difference between primes, but also the differences of the multiple of primes. At the same time it generalizes the previous works by [1,5,8] and [3]. Contrarily to the previous works, we study all the possible situations involving $k = 2$ as well as $k \geq 3$ in both cases of MSBs and LSBs. In Table 1, we compare the applicability of our methods against the previous methods for the different scenarios.

Also, we notice that not only the new bounds improve the previous ones, but also that the rank of the new underlying lattices are often lower than the ranks of the lattices used in the former methods. In Tables 2 and 3, we compare our results against the former results with k RSA moduli in terms of bounds and dimension of the lattices.

Table 2 Comparison of the bounds on t for k RSA moduli in the MSB case

Method for MSBs	Number of RSA moduli $k = 2$	Number of RSA moduli $k \geq 3$
May, Ritzenhofen [5]	Not studied	Not studied
Sarkar, Maitra [8]	For $q_1 \approx q_2 \approx 2^{\alpha}$ and $\lvert p_1 - p_2 \rvert < 2^t$, the bound is heuristically better than $t \geq 2\alpha + 3$ and the dimension of the lattice is at least 9 ($m = t = 1$).	Can not be applied
Faugère et al. [1]	For $q_1 \approx q_2 \approx 2^{\alpha}$ and $\lvert p_1 - p_2 \rvert < 2^t$, the rigorous bound is $t \geq 2\alpha + 3$ using 2-dimensional lattices of $\mathbb{Z}^3$.	For $q_1 \approx \cdots \approx q_k \approx 2^{\alpha}$ and $\lvert p_i - p_j \rvert < 2^t$, the heuristic bound is $t > \frac{k}{k-1}\alpha + 1 + \frac{k}{2(k-1)}\left(2 + \frac{\log_2(k)}{2} + \log_2(\pi e)\right)$ using k-dimensional lattices of $\mathbb{Z}^{\frac{k(k+1)}{2}}$.
Kurosawa, Ueda [3]	Not studied.	Can not be applied
Our results	For $q_1 \approx q_2 \approx 2^{\alpha}$ and $\lvert a_1 p_1 - a_2 p_2 \rvert < 2^t$ for some unknown integers $a_1, a_2 \leq 2^{\beta}$, the rigorous bound is $t \geq 2\alpha + 2\beta + 1$ using the continued fraction algorithm. For $a_1 = a_2 = 1$, $\beta = 0$ and the the rigorous bound is $t \geq 2\alpha + 1$.	For $q_1 \approx \cdots \approx q_k \approx 2^{\alpha}$ and $\lvert a_i p_i - a_j p_j \rvert < 2^t$ for some unknown integers $a_1, \ldots, a_k$, the heuristic bound is $t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right)$ using k-dimensional lattices of $\mathbb{Z}^k$. For $a_1 = \cdots = a_k = 1$, $\beta = 0$ and the the heuristic bound is $t > \frac{k}{k-1}\alpha + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right)$.

We apply our results to the implicit factorization of $k \geq 2$ RSA for Paranoids [7] $N_i = p_i q_i, i = 1, \ldots, k$, where $p_i \approx 2^{4500}$ and $q_i \approx 2^{500}$. For example, we show that we can easily factor two RSA for Paranoids moduli $N_1 = p_1 q_1, N_2 = p_2 q_2$ if there exist two integers a_1 and a_2 such that $a_1 p_1$ and $a_2 p_2$ share t MSBs or t LSBs with $t \geq 1001 + 2\beta$ where β is such that $a_i \leq 2^{\beta}$ for $i = 1, 2$.

The rest of this paper is organized as follows. In Sect. 2, we introduce some useful background on continued fractions and lattice basis reduction. In Sect. 3, we present our first method to address the problem of implicit factoring of two RSA moduli $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ when $a_1 p_1$ and $a_2 p_2$ share t MSBs. In Sect. 4, we present a generalization to $k \geq 3$ RSA moduli $N_i = p_i q_i, i = 1, \ldots, k$, in the situation that the $a_i p_i$'s share t MSBs. In Sect. 5, we present an attack on two RSA moduli $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ when $a_1 p_1$ and $a_2 p_2$ share t LSBs and we generalize this attack to $k \geq 3$ RSA moduli in Sect. 6. In Sect. 7, we present our experiments and we conclude in Sect. 8.

2 Preliminaries

In this section, we review some knowledge background on continued fractions and lattice basis reduction.

Table 3 Comparison of the bounds on t for k RSA moduli in the LSB case

Method for LSBs	Number of RSA moduli $k = 2$	Number of RSA moduli $k \geq 3$
May, Ritzenhofen [5]	For $q_1 \approx q_2 \approx 2^\alpha$ and $p_1 \equiv p_2 \pmod{2^t}$, the rigorous bound is $t \geq 2\alpha + 3$ using 2-dimensional lattices of $\mathbb{Z}^2$	For $q_1 \approx \cdots \approx q_k \approx 2^\alpha$ and $p_i \equiv p_j \pmod{2^t}$, the heuristic bound is $t \geq \frac{k}{k-1}\alpha$ using k-dimensional lattices of $\mathbb{Z}^k$
Sarkar, Maitra [8]	For $q_1 \approx q_2 \approx 2^\alpha$ and $p_1 \equiv p_2 \pmod{2^t}$, the bound is heuristically better than $t \geq 2\alpha + 3$ and the dimension of the lattice is at least 9 ($m = t = 1$)	Can not be applied
Faugère et al. [1]	Not studied	Not studied
Kurosawa, Ueda [3]	For $q_1 \approx q_2 \approx 2^\alpha$ and $p_1 \equiv p_2 \pmod{2^t}$, the rigorous bound is $t \geq 3\alpha + 1$ using 2-dimensional lattices of $\mathbb{Z}^2$	Can not be applied
Our results	For $q_1 \approx q_2 \approx 2^\alpha$ and $\lvert a_1 p_1 - a_2 p_2 \rvert < 2^t$ for some unknown integers $a_1, a_2 \leq 2^\beta$, the rigorous bound is $t \geq 2\alpha + 2\beta + 1$ using the continued fraction algorithm. For $a_1 = a_2 = 1$, $\beta = 0$ and the the rigorous bound is $t \geq 2\alpha + 1$	For $q_1 \approx \cdots \approx q_k \approx 2^\alpha$ and $a_i p_i \equiv a_j p_j \pmod{2^t}$ for some unknown integers $a_1, \ldots, a_k$, the heuristic bound is $t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right)$ using k-dimensional lattices of $\mathbb{Z}^k$. For $a_1 = \cdots = a_k = 1$, $\beta = 0$ and the the heuristic bound is $t > \frac{k}{k-1}\alpha + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right)$

2.1 Continued fractions

First we give the definition of continued fractions and state a related theorem. The details can be referenced in [2]. For any positive real number ξ, define $\xi_0 = \xi$ and for $i = 0, 1, \ldots, n$, $a_i = \lfloor \xi_i \rfloor$, $\xi_{i+1} = 1/(\xi_i - a_i)$ unless ξ_n is an integer. Then ξ can be expanded as a continued fraction in the following form

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{\ddots}}}},$$

which, for simplicity, can be rewritten as $\xi = [a_0, a_1, \ldots, a_n, \ldots]$. If ξ is a rational number, then the process of calculating the continued fraction expansion would be finished in some finite index n and then $\xi = [a_0, a_1, \ldots, a_n]$. The convergents $\frac{a}{b}$ of ξ are the fractions defined by $\frac{a}{b} = [a_0, \ldots, a_i]$ for $i \geq 0$. We note that, if $\xi = \frac{a}{b}$ is

a rational number, then the continued fraction expansion of ξ is finite with the total number of convergents being polynomial in $\log(b)$.

Another important result on continued fractions that will be used throughout this paper is the following (Theorem 184 of [2]).

Theorem 1 (Legendre) *Let ξ be a positive number. Suppose* $\gcd(a, b) = 1$ *and*

$$\left|\xi - \frac{a}{b}\right| < \frac{1}{2b^2}.$$

Then $\frac{a}{b}$ is one of the convergents of the continued fraction expansion of ξ.

2.2 Lattice reduction

Let us present some basics on lattice reduction techniques. Let $b_1 \ldots, b_d$ be d linearly independent vectors of $\mathbb{R}^n$ with $d \leq n$. The set of all integer linear combinations of the b_i forms a lattice $\mathcal{L}$. Namely,

$$\mathcal{L} = \left\{ \sum_{i=1}^{d} x_i b_i \mid x_i \in \mathbb{Z} \right\}.$$

The integer n is the rank of the lattice $\mathcal{L}$ and d is its dimension. The set $(b_1, \ldots, b_d)$ is called a basis of $\mathcal{L}$. The determinant of $\mathcal{L}$ is defined as $\det(\mathcal{L}) = \sqrt{B^t B}$ where B is the basis matrix, i.e., the matrix of the b_i's in the canonical basis of $\mathbb{R}^n$. The determinant is invariant under unimodular basis transformations of B and reduces to $\det(\mathcal{L}) = |\det(B)|$ when $d = n$. Let us denote by $\|v\|$ the Euclidean norm of a vector $v \in \mathcal{L}$. A central problem in lattice reduction is to find short non-zero vectors in $\mathcal{L}$. Vectors with short norm can be computed by the LLL algorithm of Lenstra, Lenstra, and Lovász [4].

Theorem 2 (LLL) *Let $\mathcal{L}$ be a lattice spanned by a basis $(u_1, \ldots, u_d)$. Then the LLL algorithm produces a new basis $(b_1, \ldots, b_d)$ of $\mathcal{L}$ satisfying*

$$\|b_1\| \leq 2^{\frac{d-1}{4}} \det(\mathcal{L})^{\frac{1}{d}}.$$

On the other hand, for comparison, the Gaussian Heuristic says that the length of the shortest non-zero vector of a lattice $\mathcal{L}$ is usually approximately $\sigma(\mathcal{L})$ where

$$\sigma(\mathcal{L}) \approx \sqrt{\frac{d}{2\pi e}} \det(\mathcal{L})^{\frac{1}{d}}.$$

3 Factoring two RSA Moduli in the MSB Case

In this section, we study the problem of factoring two RSA moduli $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ where $a_1 p_1$ and $a_2 p_2$ coincide on the t most significant bits (MSBs), that is when $|a_2 p_2 - a_1 p_1|$ is sufficiently small.

3.1 The general attack for two RSA Moduli in the MSB Case

We begin by the following result which applies to two RSA moduli not necessarily of the same bit size.

Theorem 3 *Let $N_1 = p_1q_1$, $N_2 = p_2q_2$ be two RSA moduli. If there exist two integers a_1, a_2 such that $a_1 < p_2, a_2 < p_1$ and $|a_1p_1 - a_2p_2| < \frac{p_1}{2a_2q_1q_2}$, then one can factor N_1 and N_2 in polynomial time.*

Proof For $N_1 = p_1q_1$ and $N_2 = p_2q_2$, let $x = a_1p_1 - a_2p_2$. Multiplying x by q_2, we get $a_1p_1q_2 - a_2N_2 = xq_2$. Suppose that $|x| < \frac{p_1}{2a_2q_1q_2}$. Then, dividing by $a_2N_1 = a_2p_1q_1$, we get

$$\left|\frac{N_2}{N_1} - \frac{a_1q_2}{a_2q_1}\right| = \frac{|x|q_2}{a_2p_1q_1} < \frac{p_1}{2a_2q_1q_2} \times \frac{q_2}{a_2p_1q_1} = \frac{1}{2(a_2q_1)^2}.$$

Hence, from Theorem 1, it follows that $\frac{a_1q_2}{a_2q_1}$, in lowest term is one of the convergents in the continued fraction expansion of $\frac{N_2}{N_1}$. If we assume $a_1 < p_2, a_2 < p_1$, then using $\frac{a_1q_2}{a_2q_1}$, we get $q_1 = \gcd(N_1, a_2q_1)$ and therefore $p_1 = \frac{N_1}{q_1}$. Similarly, we get $q_1 = \gcd(N_2, a_1q_2)$ and $p_2 = \frac{N_2}{q_2}$. □

Remark 1 The result of Theorem 3 is valid even when the RSA moduli are not of the same size. Comparatively, the attacks presented by Sarkar and Maitra in [8] and Faugère et al. in [1] are valid only if $N_1 \approx N_2$ and $q_1 \approx q_2$.

Example 1 Consider the following RSA moduli

$$N_1 = 6343178298641262531091215558254707197227984863447 9,$$
$$N_2 = 9946006657067710178027582903059286609914354223.$$

The first partial quotients of $\frac{N_2}{N_1}$ are

$$[0, 6377, 1, 1, 1, 1, 2, 2, 3, 1, 1, 3, 9, 1, 1, 1, 1, 7, 1, 19, 1, 1, 11,$$
$$1, 1, 23, 1, 1, 3, 2, 3, 2, 3, 4, 2, 1, 1, 1, 8, 1, 322, 3, 4, 1, 1, 2, \ldots]$$

Each convergent $\frac{a}{b}$ of $\frac{N_2}{N_1}$ is a candidate for $\frac{a_1q_2}{a_2q_1}$ and the good one will reveal q_1 and q_2 if the conditions of Theorem 3 are fulfilled. Indeed, the 40th convergent is $\frac{a}{b} = \frac{1351300027964332}{8618068847003717463}$ and gives

$$q_1 = \gcd(N_1, b) = 2125300178867,$$
$$p_1 = \frac{N_1}{q_1} = 29846034747067203786403150576377329237,$$
$$q_2 = \gcd(N_2, a) = 9531501481,$$
$$p_2 = \frac{N_2}{q_2} = 10434879202289356679403932941653273 83.$$

We notice that p_1 and p_2 do not share any amount of LSBst nor MSBs nor bits in the middle. This shows that the attacks presented in [8] and [1] will not give a result in this situation.

3.2 Application to unbalanced RSA and RSA for Paranoids

As an application of Theorem 3 to factor two unbalanced RSA moduli of the same bit-size, we get the following result.

Corollary 1 *Let $N_1 = p_1q_1$, $N_2 = p_2q_2$ be two unbalanced RSA moduli of the same bit-size n. Suppose that $q_i \approx 2^{\alpha}$, $p_i \approx 2^{n-\alpha}$ for $i = 1, 2$. Let a_1, a_2 be two integers such that $a_i \leq 2^{\beta}, i = 1, 2$. If a_1p_1 and a_2p_2 share t most significant bits with $t \geq 2\alpha + 2\beta + 1$, then one can factor N_1 and N_2 in polynomial time.*

Proof Let $N_1 = p_1q_1$, $N_2 = p_2q_2$ be two RSA moduli with $N_1 \approx N_2 \approx 2^n$ and $q_1 \approx q_2 \approx 2^{\alpha}$. Suppose that a multiple a_1p_1 and a multiple a_2p_2 share the t most significant bits, that is $a_1p_1 - a_2p_2 = x$ with $|x| \leq 2^{n-\alpha+\beta-t}$. Assume that $t \geq 2\alpha + 2\beta + 1$. Then

$$2a_2q_1q_2|x| < 2^{1+\beta+2\alpha+n-\alpha+\beta-t} \leq 2^{n-\alpha} \approx p_1,$$

which can be transformed into the inequality $|x| < \frac{p_1}{2a_2q_1q_2}$. Hence, as in Theorem 3, it follows that $\frac{a_1q_2}{a_2q_1}$ is a convergent of the continued fraction of $\frac{N_2}{N_1}$ which leads to the factorization of N_1 and N_2. □

Remark 2 If we consider $\beta = 0$ in Corollary 1, that is, if $a_1 = a_2 = 1$, a sufficient condition to factor the two RSA moduli is $t \geq 2\alpha + 1$ which slightly improves the bound $t \geq 2\alpha + 3$ found by Faugère et al. in [1]. This shows that the bound found by Faugère et al. with lattice reduction techniques can be achieved using the continued fraction algorithm instead.

Consider two RSA for Paranoids moduli $N_i = p_iq_i$ with $N_i \approx 2^{5000}$, $q_i \approx 2^{500}$ and $p_i \approx 2^{4500}$ for $i = 1, 2$. Then $\alpha = 500$ and by Corollary 1, it is possible to factor N_1 and N_2 if a multiple a_1p_1 and a multiple a_2p_2 share the t MSBs whenever $t \geq 2\alpha + 2\beta + 1$, that is whenever $t \geq 1001 + 2\beta$.

4 Factoring k RSA Moduli in the MSB Case

The attack mounted for two RSA moduli can be generalized to an arbitrary number $k \geq 3$ of moduli $N_i = p_iq_i, i = 1 \ldots, k$ where the q_i's are α-bit primes and the a_ip_i's share t MSBs. Instead of using the continued fraction algorithm, we use a lattice based method to find simultaneous diophantine approximations.

Theorem 4 *Let $N_i = p_iq_i, i = 1 \ldots, k$, be $k \geq 3$ n-bit RSA moduli where the q_i's are α-bit primes. Suppose that there exist k integers $a_1, \ldots, a_k$ with $a_i \leq 2^{\beta}, i = 1, \ldots, k$, such that the a_ip_i's share all t most significant bits. If*

$$t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right),$$

then, under the Gaussian Heuristic assumption, one can factor the k RSA moduli $N_1, \cdots, N_k$ *in polynomial time.*

Proof For $2 \le i \le k$, we set $x_i = a_i p_i - a_1 p_1$. Then, multiplying by $q_1 q_i$, we get $a_i q_1 N_i - a_1 q_i N_1 = q_1 q_i x_i$. Define $a = \prod_{j=1}^{k} a_j$. Multiplying by $\frac{a}{a_i}$, we get

$$aq_1 N_i - \frac{aa_1 q_i}{a_i} N_1 = \frac{aq_1 q_i x_i}{a_i}.$$

Let C be a number to be fixed later. Consider the vector

$$v = \left(Caq_1, \frac{aq_1 q_2 x_2}{a_2}, \ldots, \frac{aq_1 q_k x_k}{a_k}\right) \in \mathbb{Z}^k. \tag{2}$$

Then $v = \left(aq_1, \frac{aa_1 q_2}{a_2} \ldots, \frac{aa_1 q_k}{a_k}\right) \times M$, where M is the $k \times k$-matrix

$$M = \begin{bmatrix} C & N_2 & N_3 & \ldots & N_{k-1} & N_k \\ 0 & -N_1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & -N_1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -N_1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & -N_1 \end{bmatrix}.$$

Let $\mathcal{L}$ be the lattice defined by the rows of M. The dimension of $\mathcal{L}$ is k and the determinant is $\det(\mathcal{L}) = C N_1^{k-1}$. The Gaussian Heuristics for $\mathcal{L}$ asserts that the length of its shortest non-zero vector is usually $\sigma(\mathcal{L})$ where

$$\sigma(\mathcal{L}) \approx \sqrt{\frac{k}{2\pi e}} \det(\mathcal{L})^{\frac{1}{k}} = \sqrt{\frac{k}{2\pi e}} C^{\frac{1}{k}} N_1^{\frac{k-1}{k}}. \tag{3}$$

If we choose C such that $\sigma(\mathcal{L}) > \|v\|$, then v can be found among the shortest non-zero vectors of the lattice $\mathcal{L}$. Using (2), we get

$$\|v\|^2 = C^2 a^2 q_1^2 + \sum_{i=2}^{k} \frac{a^2 q_1^2 q_i^2 x_i^2}{a_i^2}. \tag{4}$$

Suppose that for $i = 1, \ldots, k$, we have

$$N_i \approx 2^n, \quad q_i \approx 2^\alpha, \quad p_i \approx 2^{n-\alpha}, \quad a_i \le 2^\beta.$$

Moreover, suppose that the $a_i p_i$'s share all t MSBs. Then, for $i \geq 2$, we have

$$|x_i| = |a_i p_i - a_1 p_1| \leq 2^{n-\alpha+\beta-t}.$$

Hence (4) leads to

$$\begin{aligned}\|v\|^2 &< C^2 \times 2^{2k\beta+2\alpha} + (k-1)2^{2k\beta+4\alpha+2(n+\beta-\alpha-t)-2\beta} \\ &= C^2 \times 2^{2k\beta+2\alpha} + (k-1) \times 2^{2k\beta+2\alpha+2n-2t}.\end{aligned}$$

Define C such that $C^2 \times 2^{2k\beta+2\alpha} \geq 2^{2k\beta+2\alpha+2n-2t}$, that is $C \geq 2^{n-t}$. Then $\|v\|^2 < kC^2 \times 2^{2k\beta+2\alpha}$. On the other hand, using $N_i \approx 2^n$ in (3), we get

$$\sigma(\mathcal{L})^2 \approx \frac{k}{2\pi e} C^{\frac{2}{k}} \times 2^{\frac{2n(k-1)}{k}}.$$

Suppose $\sigma(L) > \|v\|$. Then $\sigma(L)^2 > \|v\|^2$, that is

$$\frac{k}{2\pi e} C^{\frac{2}{k}} 2^{\frac{2n(k-1)}{k}} > kC^2 \times 2^{2k\beta+2\alpha}.$$

Hence

$$C^{\frac{2(k-1)}{k}} < \frac{1}{\pi e} 2^{\frac{2n(k-1)}{k}-2k\beta-2\alpha-1}.$$

Plugging $C \geq 2^{n-t}$ and extracting t, we get

$$t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right).$$

Using (2), we get $q_1 = \gcd(Caq_1, N_1)$ and for $i = 2, \ldots, k$, $q_i = \gcd(\frac{aa_1q_i}{a_i}, N_i)$. This terminates the proof. □

We notice that with $\beta = 0$, that is $a_i = 1$ for $i = 1, \ldots, k$, we get

$$t > \frac{k}{k-1}\alpha + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right),$$

which slightly improves the bound obtained by Faugère et al. in [1]. This shows that our result extends the result of Faugère et al. where they considered only the case when the p_i's share t MSBs.

5 Factoring two RSA Moduli in the LSB Case

The study of implicit factorization when p_1, p_2 share some LSBs has been considered in [1,5,8] and [3]. In this section, we extend the former attacks to the case where an unknown multiple $a_1 p_1$ of p_1 and an unknown multiple $a_2 p_2$ of p_2 share their t LSBs.

5.1 The general attack

Theorem 5 *Let $N_1 = p_1q_1, N_2 = p_2q_2$ be two RSA moduli. Assume that there exist two integers a_1, a_2 with $a_1 < p_2, a_2 < p_1$ such that a_1p_1 and a_2p_2 share t many LSBs. If $a_1a_2q_1q_2 < 2^{t-1}$, then one can factor N_1 and N_2 in polynomial time.*

Proof Let $N_1 = p_1q_1$ and $N_2 = p_2q_2$. Assume that a_1p_1 and a_2p_2 share t many LSBs. Then $a_1p_1 - a_2p_2 = 2^t x$ for some integer x and we have

$$q_1q_2(a_1p_1 - a_2p_2) = N_1a_1q_2 - N_2a_2q_1 = 2^t xq_1q_2.$$

Then $N_1a_1q_2 - N_2a_2q_1 \equiv 0 \pmod{2^t}$. Since $gcd(N_1, 2) = 1$, then $N_1^{-1} \pmod{2^t}$ exists and $a_1q_2 - a_2q_1N_2N_1^{-1} \equiv 0 \pmod{2^t}$. Define $T \equiv N_2N_1^{-1} \pmod{2^t}$. Then $a_1q_2 - a_2q_1T \equiv 0 \pmod{2^t}$ and there exists an integer y such that

$$a_1q_2 = a_2q_1T - 2^t y. \tag{5}$$

Suppose that $a_1a_2q_1q_2 < 2^{t-1}$. Then dividing by $2^ta_2q_1$, we get

$$\left|\frac{T}{2^t} - \frac{y}{a_2q_1}\right| = \frac{\left|a_2q_1T - 2^t y\right|}{2^ta_2q_1} = \frac{a_1q_2}{2^ta_2q_1} < \frac{a_1q_2}{2a_1a_2q_1q_2a_2q_1} = \frac{1}{2(a_2q_1)^2}.$$

Therefore from Theorem 1, it follows that $\frac{y}{a_2q_1}$ is one of the convergents in the continued fraction expansion of $\frac{T}{2^t}$. Since $a_2 < p_1$, we get $q_1 = \gcd(N1, a_2q_1)$ and $p_1 = \frac{N_1}{q_1}$. Using (5), we get $a_1q_2 = a_2q_1T - 2^t y$. Similarly, since $a_1 < p_2$, we get $q_2 = \gcd(N_2, a_1q_2)$ and $p_2 = \frac{N_2}{q_2}$. This terminates the proof. □

5.2 Application to unbalanced RSA and RSA for Paranoids

Here we apply Theorem 5 in the situation that the two RSA moduli $N_1 = p_1q_1, N_2 = p_2q_2$ are of the same shape, that is N_1 and N_2 are of the same bit-size and the q_i's are α-bit primes.

Corollary 2 *Let $N_1 = p_1q_1, N_2 = p_2q_2$ be two unbalanced n-bit size RSA moduli with $q_1 \approx q_2 \approx 2^\alpha$. Suppose that there exist two positive integers $a_1 \leq 2^\beta, a_2 \leq 2^\beta$ such that a_1p_1 and a_2p_2 share the t LSBs. If $t \geq 2\alpha + 2\beta + 1$, then one can factor N_1 and N_2 in polynomial time.*

Proof Let $N_1 = p_1q_1, N_2 = p_2q_2$ be two RSA moduli with $N_1 \approx N_2 \approx 2^n$ and, $q_1 \approx q_2 \approx 2^\alpha$. Suppose that a multiple a_1p_1 and a multiple a_2p_2 share the t least significant bits where $a_i \leq 2^\beta$ for $i = 1, 2$. Define $T \equiv N_2N_1^{-1} \pmod{2^t}$. As in the proof of Theorem 5, we have $a_1p_1 - a_2p_2 = 2^t x$ and $a_1q_2 = a_2q_1T - 2^t y$ for some integers x and y. Suppose that $t \geq 2\alpha + 2\beta + 1$. Then $a_1a_2q_1q_2 < 2^{2\beta+2\alpha} \leq$

2^{t-1}. Therefore, using the same arguments than Theorem 5, we conclude that $\frac{y}{a_2 q_1}$ is one of the convergents in the continued fraction expansion of $\frac{T}{2^t}$ which leads to the factorization of N_1 and N_2. □

Remark 3 Here again, if $\beta = 0$, then the condition of Corollary 2 becomes $t \geq 2\alpha + 1$ which improves the bounds found in the former approaches of [1,5,8] and retrieves the bound of [3].

As an application of Corollary 2, consider two 1024-bit RSA for Paranoids moduli $N_1 = p_1 q_1, N_2 = p_2 q_2$ where q_1, q_2 are 500-bit primes. Hence $\alpha = 500$ and using Corollary 2, one can factor N_1 and N_2 if there exist two integers $a_1 \leq 2^\beta$ and $a_2 \leq 2^\beta$ such that $a_1 p_1$ and $a_2 p_2$ share t LSBs with $t > 2001 + 2\beta$.

6 Factoring k RSA Moduli in the LSB Case

In this section, we assume that we are given $k \geq 3$ different RSA moduli $N_i = p_i q_i, i = 1, \ldots, k$ where some unknown multiples $a_i p_i$'s coincide on the t least significant bits. For suitably large t, we show that there is an efficient algorithm that recovers the factorization of the k RSA moduli. To this end, we use the lattice reduction techniques to solve a simultaneous diophantine approximations problem.

Theorem 6 *Let $N_i = p_i q_i, i = 1 \ldots, k$, be $k \geq 3$ n-bit RSA moduli where the q_i's are α-bit primes. Suppose that there exist k positive integers $a_1, \ldots, a_k$ with $a_i \leq 2^\beta, i = 1, \ldots, k$, such that the $a_i p_i$'s share all t least significant bits. If*

$$t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right),$$

then, under the Gaussian Heuristic assumption, one can factor the k RSA moduli $N_1, \cdots N_k$ in polynomial time.

Proof For $1 \leq i \leq k$, suppose that the $a_i p_i$'s share t least significant bits. Then, for $1 \leq i \leq k, a_i p_i - a_1 p_1 = 2^t x_i$. Multiplying by $q_1 q_i$, we get $a_i q_1 N_i - a_1 q_i N_1 = 2^t q_1 q_i x_i$. Define $a = \prod_{j=1}^{k} a_j$. Multiplying by $\frac{a}{a_i}$, we get

$$aq_1 N_i - \frac{aa_1 q_i}{a_i} N_1 = \frac{2^t a q_1 q_i x_i}{a_i}.$$

Transforming modulo 2^t, we get $aq_1 N_i N_1^{-t} - \frac{aa_1 q_i}{a_i} \equiv 0 \pmod{2^t}$. Define $T_i \equiv N_i N_1^{-1} \pmod{2^t}$. Then $aq_1 T_i - \frac{aa_1 q_i}{a_i} \equiv 0 \pmod{2^t}$ and there exists an integer y_i such that $aq_1 T_i - 2^t y_i = \frac{aa_1 q_i}{a_i}$. Consider the vector

$$v = \left(aq_1, \frac{aa_1 q_2}{a_2}, \ldots, \frac{aa_1 q_k}{a_k}\right) \in \mathbb{Z}^k. \tag{6}$$

Then $v = (aq_1, y_2 \ldots, y_k) \times M$, where M is the $k \times k$-matrix

$$M = \begin{bmatrix} 1 & T_2 & T_3 & \ldots & T_{k-1} & T_k \\ 0 & -2^t & 0 & \ldots & 0 & 0 \\ 0 & 0 & -2^t & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -2^t & 0 \\ 0 & 0 & 0 & \ldots & 0 & -2^t \end{bmatrix}.$$

Let $\mathcal{L}$ be the lattice defined by the rows of the matrix M. The dimension of $\mathcal{L}$ is k and the determinant is $\det(\mathcal{L}) = 2^{(k-1)t}$. The Gaussian Heuristics for $\mathcal{L}$ asserts that the length of its shortest non-zero vector is $\sigma(\mathcal{L})$ where

$$\sigma(\mathcal{L}) \approx \sqrt{\frac{k}{2\pi e}} \det(\mathcal{L})^{\frac{1}{k}} = \sqrt{\frac{k}{2\pi e}} 2^{\frac{(k-1)t}{k}}. \tag{7}$$

Observe that the norm of v satisfies

$$\|v\|^2 = a^2 q_1^2 + \sum_{i=2}^{k} \left(\frac{a a_1 q_k}{a_k} \right)^2.$$

If the $a_i p_i$'s share all t least significant bits, then, for $i = 1, \ldots, k$, we have

$$q_i \approx 2^{\alpha}, \quad a_i \leq 2^{\beta}, \quad |x_i| = \frac{|a_i p_i - a_1 p_1|}{2^t} < 2^{n-\alpha+\beta-t}.$$

Hence

$$\|v\|^2 < 2^{2k\beta+2\alpha} + (k-1)2^{2k\beta+2\alpha} = k2^{2k\beta+2\alpha}. \tag{8}$$

Using (8) and (7) and transforming $\sigma(L)^2 > \|v\|^2$ into $\frac{k}{2\pi e} 2^{\frac{2(k-1)t}{k}} > k2^{2k\beta+2\alpha}$, we get

$$t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right).$$

Using (6), we get $q_1 = \gcd(aq_1, N_1)$ and for $i = 2, \ldots, k$, $q_i = \gcd(\frac{a a_1 q_i}{a_i}, N_i)$. This terminates the proof. □

Once again, if $\beta = 0$, then $a_i = 1$ and the bound of Theorem 6 transforms to $t > \frac{k}{k-1}\alpha + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right)$, which improves the bound of [1].

7 Experiments

In this section, we describe the experiments that we conducted for $k = 4, 10, 30$ and 50 RSA moduli, in connection with Theorem 4 and Theorem 6. We verified our assumptions by running experiments on a Core2 Duo 2GHz notebook. The lattice reduction basis technique was based on the LLL algorithm.

Assume that $a_1 p_1$ and the $a_i p_i$'s share t MSBs. Then since $a_i p_i \leq 2^{n-\alpha+\beta}$, we see that $|a_i p_i - a_1 p_1| \leq 2^{n-\alpha+\beta-t}$. Therefore, $t \leq n - \alpha + \beta$. Similarly, assume that $a_1 p_1$ and the $a_i p_i$'s share t LSBs. Then $|a_i p_i - a_1 p_1| = 2^t x_i$ with $t \leq n - \alpha + \beta$. In both cases, combining with the bound of t in Theorem 4 and Theorem 5, we get

$$n - \alpha + \beta \geq t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}\left(1 + \log_2(\pi e)\right),$$

which is satisfied if

$$\beta < \frac{n(k-1)}{k^2-k+1} - \frac{2k-1}{k^2-k+1}\alpha - \frac{k}{2(k^2-k+1)}\left(1 + \log_2(\pi e)\right). \tag{9}$$

Consequently, we only consider the situation where the bit-size β of the a_i's satisfies condition (9).

We generated many random 1024-bit RSA moduli for $k = 4, 10, 30, 50$ and various values of α and β according to the bound (9). All our experiments were successful

Table 4 Experiments for k RSA moduli in the MSB and the LSB cases

Number k of moduli	Bit-size α of the q_i's	Max bit-size β of the a_i's (9)	Used bit-size β of the a_i's	Minimal theoretical bound for t	Experimental bound for t in MSB case	Experimental bound for t in LSB case	Number of experiments
4	150	154	100	737	602	611	1000
4	250	100	80	763	655	662	1000
4	350	46	35	657	609	616	1,000
4	400	20	15	617	594	601	1,000
10	150	69	50	725	649	674	1,000
10	250	48	40	725	667	684	1,000
10	350	27	20	614	591	603	1,000
10	400	17	12	581	563	570	1,000
30	150	23	15	623	585	592	500
30	250	17	12	634	596	603	500
30	350	10	8	613	544	572	500
30	400	6	4	541	533	536	500
50	150	14	10	666	648	650	100
50	250	10	7	615	597	605	100
50	350	6	4	564	546	551	100
50	400	4	3	564	556	559	100

and the assumptions on the Gaussian Heuristics were verified. In Table 4, we notice the experimentally lowest values of t that have 100% success rate.

8 Conclusion

In this work we have designed a technique to factor $k \geq 2$ RSA moduli $N_i = p_i q_i$, $i = 1, \ldots, k$ when some unknown multiples $a_i p_i$ share t many Most Significant Bits (MSBs) or t many Least Significant Bits (LSBs). The new technique generalizes many previous results where the prime factors p_i share t many MSBs or t many LSBs. This provides practitioners tighter conditions for the primes that are generated for utilization with the RSA algorithm. On the other hand, our results also serve their purpose to provide a peace of mind for practitioners knowing that the generated RSA moduli does not fall into any of the categories mentioned in this work.

References

1. Faugère, J-C., Marinier, R., Renault, G.: Implicit factoring with shared most significant and middle bits. In P.Q. Nguyen and D. Pointcheval (Eds.): Public Key Cryptography, Lecture Notes in Computer Science, Springer 6056 (2010) 70–87.
2. Hardy, G.H., Wright, E.M.: An Introduction to the Theory of Numbers. Oxford University Press, London (1975)
3. Kurosawa, K., Ueda, T.: How to factor N_1 and N_2 when $p_1 = p_2 \pmod{2^t}$. In K. Sakiyama and M. Terada (Eds.): IWSEC 2013, Lecture Notes in Computer Science, Springer 8231 (2013) 217–225.
4. Lenstra, A.K., Lenstra, H.W., Lovász, L.: Factoring polynomials with rational coefficients. Math. Ann. **261**, 513–534 (1982)
5. May, A., M. Ritzenhofen, R.: Implicit factoring: On polynomial time factoring given only an implicit hint. In Stanislaw Jarecki and Gene Tsudik (Eds.): Public Key Cryptography, Lecture Notes in Computer Science, Springer 5443 (2009) 1–14.
6. Rivest, R., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**, 120–126 (1978)
7. Shamir, A.: RSA Laboratories CryptoBytes. RSA Paranoids **1**, 3–4 (1995)
8. Sarkar, S., Maitra, S.: Further results on implicit factoring in polynomial time. Adv. Math. Commun. **3**, 205–217 (2009)