Cryptanalysis of Public Key Cryptosystems

Abderrahmane Nitaj

Normandie Université

MEMOIRE D'HABILITATION A DIRIGER DES RECHERCHES

Spécialité : Mathématiques

Préparé au sein de l'Université de Caen Normandie

Cryptanalysis of Public Key Cryptosystems

Présenté et soutenu par

Abderrahmane NITAJ

Soutenu publiquement le 1er décembre 2016
devant le jury composé de

Mr Thierry BERGER	Professeur, Université de Limoges	Rapporteur
Mr Marc GIRAULT	Chercheur HDR, Orange Labs, Caen	Examinateur
Mr Marc JOYE	Chercheur HDR, NXP Semiconductors	Rapporteur
Mr Fabien LAGUILLAUMIE	Professeur, Université de Lyon 1	Rapporteur
Mr Denis SIMON	Professeur, Université de Caen	Examinateur
Mme Brigitte VALLEE	Directrice de Recherche, CNRS	Examinatrice

Mémoire préparé au Laboratoire de Mathématiques Nicolas Oresme (LMNO)

UNICAEN
UNIVERSITÉ
CAEN
NORMANDIE

LMNO

Contents

Remerciements

Je tiens à remercier ici toutes les personnes qui m'ont aidé de près ou de loin dans l'accomplissement de ce travail.

- Je suis très reconnaissant à Denis Simon qui a bien voulu prendre la responsablité de parrainer mon habilitation à diriger des recherches et qui a toujours été attentif à mes nombreuses sollicitudes.
- Je remercie très chaleureusement Brigitte Vallée qui me fait l'honneur d'être examinatrice de ce mémoire. Je suis fier de faire partie de la longue liste des personnes ayant bénéficié de sa présence dans le jury de leur thèse ou habilitation à diriger des recherches.
- Je suis très reconnaissant a Thierry Berger qui a bien voulu rapporter sur ce mémoire et qui me fait l'honneur de faire partie de ce jury.
- Je suis également très reconnaissant à Fabien Laguillaumie qui a accepté d'être rapporteur de ce mémoire et de faire partie du jury.
- Je remercie très chaleureusement Marc Joye pour l'honneur qu'il me fait en rapportant sur ce mémoire malgré son éloignement géographique.
- Je suis aussi très honoré par la présence de Marc Girault dans le jury de cette thèse et qui a accepté d'en être examinateur.
- Je tiens à remercier plusieurs collègues au sein du Laboratoire de Mathématiques Nicolas Oresme pour les nombreux services qu'ils m'ont rendus, notamment Francesco Amoroso, John Guaschi, John Boxall, Bruno Anglès, Bernard Leclerc, Gilles Damamme et Patrick Dehornoy.
- Je tiens aussi à remercier tous les co-auteurs de certaines de mes publications pour leurs efforts et leur soutien.
- Je remercie également le personnel administratif du Laboratoire de Mathématiques Nicolas Oresme pour leur aide et leur disponiblité. Merci donc à Sonia, Anita et Axelle.
- Une petite pensée pour mon professeur de toujours, Yves Hellegouarch et tous ceux que j'ai connus à mes débuts.
- Finalement, je remercie tous les membres de ma famille en France et au Maroc pour leurs soutiens.

List of Publications

The publications are sorted by type (book, book chapter, article, submitted, survey, etc) in anti-chronological order.

Books

1. David Pointcheval, Abderrahmane Nitaj, Tajjeeddine Rachidi (Eds): Progress in Cryptology - AFRICACRYPT 2016, Lecture Notes in Computer Science, Volume 9646 2016.
2. Said El Hajji, Abderrahmane Nitaj, Claude Carlet, El Mamoun Souidi (Eds.): Codes, Cryptology, and Information Security, First International Conference, C2SI 2015, Rabat, Morocco, May 26-28, 2015, Proceedings - In Honor of Thierry Berger Lecture Notes in Computer Science, Vol. 9084.
3. Amr Youssef, Abderrahmane Nitaj, Aboul Ella Hassanien (Eds.): Progress in Cryptology - AFRICACRYPT 2013, 6th International Conference on Cryptology in Africa, Cairo, Egypt, June 22-24, 2013. Proceedings. Lecture Notes in Computer Science 7918, Springer 2013.
4. Abderrahmane Nitaj, David Pointcheval (Eds.) Progress in Cryptology-Africacrypt 2011, 4th International Conference on Cryptology in Africa, Dakar, Senegal, July 5-7, 2011, LNCS, Springer 2011.

Book Chapters

1. A. Nitaj, The Mathematics of the NTRU Cryptosystem, In Addepalli VN Krishna (Eds.), Emerging Security Solutions Using Public and Private Key Cryptography: Mathematical Concepts, IGI Global, June, 2015.

2. A. Nitaj, Diophantine and Lattice Cryptanalysis of the RSA Cryptosystem. Artificial Intelligence, Evolutionary Computation and Metaheuristics (AIECM) –In the footsteps of Alan Turing (Turing 2012), AIECM2012, p. 139-168 (2012).

3. A. Nitaj, The Mathematical Cryptography of the RSA Cryptosystem. In Cryptography: Protocols, Design and Applications, Kamol Lek and Naruemol Rajapakse (Eds.), 2012.

Submitted papers

1. A. Nitaj, T. Rachidi: Lattice Attacks on the Homomorphic DGHV Scheme, Submitted to Discrete Applied Mathematics.

2. A. Nitaj, E. Fouotsa: A New Attack on RSA and Demytko's Elliptic Curve Cryptosystem, Submitted to Mathematics in Computer Science.

3. A. Nitaj, M.R.K. Ariffin: Generalizations of Former Attacks on RSA, Submitted to Journal of Mathematical Cryptology.

4. A. Nitaj, W. Susilo, J. Tonien: Factorization of RSA Modulus $N = p^r q$ with Small Prime Difference, Submitted to Theoretical Computer Science.

Refereed journal papers

1. A. Nitaj, D.I. Nassr, H.M. Bahig, A. Bhery: Another Look at Private Exponent Attack on RSA using Lattices, To appear in International Journal of Applied and Computational Mathematics.

2. M. Bunder, A. Nitaj, W. Susilo, J. Tonien: A generalized attack on RSA type cryptosystems, To appear in Theoretical Computer Science.

3. A. Nitaj, W. Susilo, J. Tonien: Dirichlet Product for Boolean Functions, To appear in Journal of Applied Mathematics and Computing, 2016.

4. A. Nitaj, M.R.K. Ariffin: Implicit factorization of unbalanced RSA moduli. J. Appl. Math. Comput. 48 (2015), no. 1-2, pp. 349–363 (2015)

5. A. Nitaj: A new attack on the KMOV cryptosystem. Bull. Korean Math. Soc. 51 (2014), no. 5, pp. 1347–1356.

6. A. Nitaj: Cryptanalysis of NTRU with two Public Keys, International Journal of Network Security, 16(2), pp. 112-117 (2014)

7. A. Nitaj, M. Ould Douh: A new attack on RSA with a composed decryption exponent (with M.O. Douh), International Journal on Cryptography and Information Security (IJCIS), Vol.3, No. 4, December 2013

8. A. Nitaj: A new attack on RSA with two or three decryption exponents. J. Appl. Math. Comput. 42 (2013), no. 1-2, pp. 309–319 (2013)

9. A. Nitaj: New weak RSA keys , JP Journal of Algebra, Number Theory and Applications. Volume 23, Number 2, 2011, pp. 131–148 (2011)

10. A. Nitaj: A new vulnerable class of exponents in RSA. JP J. Algebra Number Theory Appl. 21 (2011), no. 2, pp. 203–220 (2011).

11. A. Nitaj: Application of ECM to a class of RSA keys, Journal of Discrete Mathematical Sciences & Cryptography , 12, No. 2, pp. 121–137 (2009)

12. A. Nitaj: Cryptanalysis of RSA with constrained keys. Int. J. Number Theory 5 (2009), no. 2, pp. 311–325 (2009)

13. A. Nitaj: Isogenous of the elliptic curves over the rationals. J. Comput. Math. 20 (2002), no. 4, 337-348.

14. A. Nitaj: Invariants des courbes de Frey-Hellegouarch et grands groupes de Tate-Shafarevich. Acta Arith. 93 (2000), no. 4, 303-327.

15. A. Nitaj: Détermination de courbes elliptiques pour la conjecture de Szpiro. Acta Arith. 85 (1998), no. 4, 351-376.

16. A. Nitaj: Aspects expérimentaux de la conjecture abc. Number theory (Paris, 1993-1994), 145-156, London Math. Soc. Lecture Note Ser., 235, Cambridge Univ. Press, Cambridge, 1996.

17. A. Nitaj: La conjecture abc. Enseign. Math. (2) 42 (1996), no. 1-2, 3-24.

18. A. Nitaj: On a conjecture of Erdős on 3-powerful numbers, Bulletin of the London Mathematical Society 27 (1995), pp. 317-318.

19. A. Nitaj: L'algorithme de Cornacchia, Expositiones Math. 13 (1995), pp. 358-365 (1995)

20. A. Nitaj: An algorithm for finding good *abc*-examples, Comptes Rendus de l'Académie des Sciences de Paris, 317 (1993), pp. 811-815, (1993)

Refereed international conference papers

1. M. Bunder, A. Nitaj, W. Susilo, J. Tonien: A new attack on three variants of the RSA cryptosystem, Proceedings of ACISP, the 21st Australasian Conference on Information Security and Privacy, Lecture Notes in Computer Science 9723, 2016, 258–268.

2. A. Nitaj, T. Rachidi: Factoring RSA moduli with weak prime factors. Codes, cryptology, and information security, 361-374, Lecture Notes in Comput. Sci., 9084, Springer, Cham, (2015)

3. A. Nitaj, T. Rachidi: New attacks on RSA with moduli $N = p^r q$. Codes, Cryptology, and Information Security, pp. 352–360, Lecture Notes in Comput. Sci., 9084, Springer, Cham, (2015)

4. A. Nitaj, M.R.K. Ariffin, D.I. Nassr, H.M. Bahig: New attacks on the RSA cryptosystem. Progress in cryptology-AFRICACRYPT 2014, pp. 178–198, Lecture Notes in Comput. Sci., 8469, Springer, Cham, (2014)

5. A. Nitaj: An attack on RSA using LSBs of multiples of the prime factors. Progress in cryptology-AFRICACRYPT 2013, pp. 297–310, Lecture Notes in Comput. Sci., 7918, Springer, Heidelberg, 2013.

6. A. Nitaj: A new attack on RSA and CRT-RSA. Progress in cryptology-AFRICACRYPT 2012, pp. 221–233, Lecture Notes in Comput. Sci., 7374, Springer, Heidelberg, (2012)

7. A. Nitaj: Cryptanalysis of RSA using the ratio of the primes. Progress in cryptology-AFRICACRYPT 2009, pp. 98–115, Lecture Notes in Comput. Sci., 5580, Springer, Berlin, (2009)

8. A. Nitaj: Another generalization of Wiener's attack on RSA, in Vaudenay, S. (ed.) Africacrypt 2008. Lecture Notes in Computer Science, Springer-Verlag Vol. 5023, pp. 174–190 (2008)

9. Greaves, George, Nitaj, Abderrahmane: Some polynomial identities related to the abc-conjecture. Number theory in progress, Vol. 1 (Zakopane-Koscielisko, 1997), 229-236, de Gruyter, Berlin, 1999.

Technical Reports

1. A. Nitaj : La cryptographie et la confiance numérique.

2. A. Nitaj : Quantum and post quantum cryptography.

3. A. Nitaj : La cryptographie du futur.

4. A. Nitaj : Applications de l'algorithme LLL en cryptographie.

5. A. Nitaj : Cryptanalyse de RSA.

6. A. Nitaj : Cryptanalyse de RSA [Maple 12 Worksheet].

7. A. Nitaj : NTRU et ses variantes, sécurité et applications.

8. A. Nitaj : RSA and a higher degree diophantine equation.

9. A. Nitaj : A Maple Worksheet for elliptic curves.

10. A. Nitaj : Le problème du logarithme discret elliptique : Index et Xedni.

11. A. Nitaj : Le cryptosystème NTRU.

12. A. Nitaj : Table of all good abc examples.

13. A. Nitaj : Table of all good abc-Szpiro examples.

Chapter 1

Introduction

This document presents several results I obtained in cryptography, especially in the area of cryptanalysis of a few public key cryptosystems, mainly systems that are related to the RSA cryptosystem. The techniques used for the cryptanalytic attacks are based on adapting computational and algorithmic tools from Number Theory. The attacks concern the following cryptosystems and items

- The RSA cryptosystem.
- The CRT-RSA cryptosystem.
- The Prime Power RSA cryptosystem.
- The NTRU cryptosystem.
- The KMOV elliptic curve cryptosystem.
- The DGHV homomorphic cryptosystem.
- The Demytko elliptic curve cryptosystem.
- The RSA-type schemes based on singular cubic curves.
- The Dirichlet product for boolean functions.

I will describe the former systems and some of the most known cryptanalytic attacks on them. I will then describe my research results, mainly from the cryptanalytic point of view without proofs which can be found in the articles.

Notice that some attacks we conducted on RSA or on its variants are briefly presented in this work while they are included as appendices.

In my Ph.D. thesis, I studied many diophantine problems related to the *abc* conjecture and elliptic curves. The main tools used there were originated from algorithmic and computational number theory such as continued fractions, diophantine approximations, lattice reduction, diophantine equations and elliptic curves. Amazingly, the same techniques are used in cryptography and especially in cryptanalysis. Indeed, the main hard problems behind the security of most of the widely used public key cryptosystems are number theory problems, such as factorization, discrete logarithm and lattice problems. Often, to attack a public key cryptosystem, the starting point is to find a diophantine problem that can be transformed into a computational problem. In RSA, the diophantine problem is mainly the key equation $ed - k(p-1)(q-1) = 1$. In the key exchange protocol of Diffie-Hellman and in ElGamal cryptosystem, the diophantine problem is the discrete logarithm problem $b \equiv g^x \pmod{p}$. In the elliptic curve cryptography ECC, the diophantine problem is the elliptic discrete logarithm problem $Q = nP$. In NTRU, the diophantine problem is $h = \frac{g}{f} \pmod{q}$ where h, f and g are polynomials in the ring $\mathbb{Z}_q[X]/(X^N - 1)$.

The RSA cryptosystem [131], invented by Rivest, Shamir and Adleman in 1977 is the most widely used asymmetric cryptographic scheme. The RSA public-key cryptosystem is used for securing web traffic, e-mails, remote login sessions, and electronic credit card payment systems. The underlying one-way function in RSA is the integer factorization problem:
Multiplying two large primes is computationally easy, but factoring the resulting product is very hard.
Another hard problem in RSA is the difficulty of solving the so-called RSA problem:
Given an RSA public key (e, N) **and a ciphertext** $c \equiv m^e \pmod{N}$**, compute the plaintext** m**.**
Also, the security of RSA can be reduced to solving the key equation $ed - k(p-1)(q-1) = 1$ where $N = pq$ is the modulus, e is the public exponent and d is the private exponent. The prominent attacks on small private exponents in RSA are Wiener's continued fraction based attack [147] and the lattice

reduction attack of Boneh and Durfee [17].

There are many different ways to attack RSA by studying diophantine equations related to the encrypted message or the key equation. Many RSA variant equations, satisfied by the exponents, can potentially be used to factor the RSA modulus. This remark was used by many researchers such as Blömer and May [13]. In my research on RSA, I tried to exhibit variant key equations satisfied by the exponents and tried to solve them by number theoretical tools in order to factor the RSA modulus. For example, we studied in [104] the situation where the RSA public exponent satisfies an equation of the form $eX - (p-u)(q-v)Y = 1$,. We showed that if the parameters X, Y, u and v are suitably small, then one can solve the equation and break the system. The number of the exponents e satisfying the former equation is not negligible since this number can be lower bounded by $N^{\frac{1}{2}-\varepsilon}$ for a small positive constant ε. The method used for solving the equation combines the continued fraction algorithm and Coppersmith's technique. In [107], we studied another example of an RSA variant equation, namely $eX - (N - (ap + bq))Y = Z$ with sufficiently small parameters X, Y and Z where a and b are unknown positive integers such that $\frac{a}{b}$ is close to $\frac{q}{p}$. This equation is related to the requirement in section 4.1.2 of the ANSI X9.31:1998 standard for public key cryptography [1] to ovoid prime factors p and q in the RSA modulus $N = pq$ with a ratio close to a rational number $\frac{a}{b}$ with small a and b.

In some instances of RSA, using a system of diophantine equations ease the resolution and improve the bounds. For example, we considered in [111] the situation with the presence of two or three public exponents e_i, $i = 1, 2, 3$, satisfying the equations $e_i x_i - (p-1)(q-1)y_i = z_i$ with small parameters. We showed that the bounds are better than the situation where only one exponent is available. Similarly, in the presence of several RSA instances $N_i = p_i q_i$ with the same decryption exponent d, we have a system of diophantine equations $e_i d - k_i(p_i - 1)(q_i - 1) = 1$. In [114], we generalized this situation with the equations $e_i x - y_i(p_i - 1)(q_i - 1) = z_i$ and solved them by applying lattice reduction techniques to the simultaneous diophantine approximations.

There is another kind of attacks on RSA, called partial key exposure attacks, in which a fraction of the bits of a private parameter, such as p, q or d is known. Such attacks have been intensively studied by many researchers

(see e.g. [61, 71, 91]). Another example is when two RSA moduli $N_1 = p_1q_1$ and $N_2 = p_2q_2$ are such that the difference $|p_1 - p_2|$ is sufficiently small [45]. In [115], we generalized this situation by considering the difference of multiples of the primes in the form $|ap_1 - bp_2|$. When $|ap_1 - bp_2|$ is suitably small or in the form $|ap_1 - bp_2| = 2^m x$ for some positive integer m, we showed that one can apply the continued fraction algorithm and Coppersmith's method to factor both RSA moduli. Notice here that no known bits are involved. Moreover, the method can be generalized to k RSA moduli $N_i = p_iq_i$ such that the differences $|a_ip_i - a_ip_j|$ are simultaneously small by applying lattice reduction techniques [115].

In addition to the standard RSA cryptosystem where the modulus is $N = pq$ and the encryption of a message m is computed as $m^e \pmod N$, there are some variants that are of interest for efficiency reasons such as the CRT-RSA cryptosystem and the prime power RSA cryptosystem.

In CRT-RSA, the public exponent e and the private CRT-exponents d_p and d_q satisfy $ed_p \equiv 1 \pmod{(p-1)}$ and $ed_q \equiv 1 \pmod{(q-1)}$. One can further reduce the decryption time by carefully choosing d so that both d_p and d_q are small. Many known attacks such as [12] on CRT-RSA work by combining the key equations of d_p and d_q. In [110], we showed that applying the ideas of [59], one can reduce the situation by using solely one of the equations, which leads to the factorization of the RSA modulus. This enables us to attack more CRT-instances when only one of the decryption exponents d_p or d_q is small.

In the prime power RSA variant, the modulus is in the form $N = p^rq$ with $r \geq 2$. Such moduli are used in cryptography to speed up the decryption process in RSA [145]. Similarly to the standard RSA, there are various attacks that can be launched on the prime power RSA variant. In 2014, Sarkar [132] presented an attack using the key equation $ed - kp^{r-1}(p-1)(q-1) = 1$. The attack uses Coppersmith's technique and works for small values of d. In [116], we showed that the key equation can be generalized by using the more general equation $ex - p^{r-1}(p-1)(q-1)y = z$ in which the unknown parameters are suitably small. Then, using Coppersmith's technique, especially the ideas of [88], we showed that the prime power RSA modulus $N = p^rq$ can be factored in polynomial time.

RSA and many cryptographic schemes are vulnerable to quantum computers running Shor's algorithm. In the contrary, the NTRU cryptosystem [63], invented in 1996 by Hoffstein, Pipher, and Silverman, is still resistant to quantum attacks. This makes NTRU one of the post quantum candidates. In NTRU, the public key is a polynomial $h \in \mathbb{Z}_q[X]/(X^N - 1)$ with $h \equiv \frac{g}{f} \pmod{q, X^n - 1}$ where f and g are polynomials with small and sparse coefficients. In 1998, Coppersmith and Shamir [35] transformed the key equation into a linear one $f * h \equiv g \pmod{q}$ and exhibit a lattice $\mathcal{L}$ with dimension $2N$. By studying the short vectors of the lattice, they showed that one can retrieve the private keys f and g if some condition is satisfied. In [118], we extended the ideas of Coppersmith and Shamir in the presence of an instance of NTRU with two public keys, $h \equiv \frac{g}{f} \pmod{q, X^n - 1}$ and $h' \equiv \frac{G'}{F'} \pmod{q, X^n - 1}$ which can be rewritten as $h' \equiv \frac{g}{f} \pmod{q, X^n - 1}$ where $g \equiv \frac{G' * f}{F'} \pmod{q, X^n - 1}$. This means that any couple of public keys use the polynomial f. We combined the equations of h and h' to build a lattice $\mathcal{L}$. By studying the properties and short vectors of this lattice, we showed that the private polynomials f, g, g' can be found by lattice reduction techniques under the condition that g and g' share a certain amount of their coefficients.

Homomorphic encryption is a new research topic in cryptography. It aims to make cloud computing completely secure. Homomorphic encryption allows complex computation on encrypted data without decrypting it. By using properties of ideal lattices, Gentry [49] presented the first fully homomorphic scheme. Since then, several homomorphic schemes have been proposed, such as DGHV [41]. In DGHV, m integers $c_i = pq_i + r_i$, $i = 1, \cdots, m$, are public while the integers p, q_i and r_i are secret. In [36, 41, 87], the security of DGHV has been studied against several attacks. In [122], we showed that the parameters in DGHV always satisfy a linear equation of the form $a_2c_2 + \ldots + a_mc_m = a_1q_1$. Then, using a result of Herrmann and May [59] on solving multivariate linear equations, we launched an attack on DGHV to retrieve the secret parameters p, q_i and r_i simultaneously. This shows once again that Coppersmith's method can amazingly be applied in various situations.

In 1985, Koblitz [75] and Miller [95] independently proposed the idea

of using elliptic curves for cryptographic applications. The security of the elliptic curve schemes are based on the difficulty of solving the elliptic discrete logarithm. Since then, many cryptosystems based on the same problem have been proposed, and some of them include the difficulty of factoring large numbers.

In 1991, Koyama, Maurer, Okamoto and Vanstone proposed a scheme, called KMOV [79] using an RSA modulus $N = pq$, an elliptic curve over $\mathbb{Z}/N\mathbb{Z}$ with a public key e satisfying the key equation $ed - k(p+1)(q+1) = 1$. In [119], we studied the more general equation $ex - (p+1)(q+1)y = z$ and launched an attack on KMOV based on continued fractions and Coppersmith's technique to factor the modulus under some conditions on the size of the unknown parameters.

In 1994, Demytko [39] developed a cryptosystem using an elliptic curve $E_N(a,b)$ with equation $y^2 = x^3 + ax + b$ over the ring $\mathbb{Z}/N\mathbb{Z}$ where $N = pq$ is an RSA modulus. In this system, the public parameters are N, a, b and e satisfying $\gcd\left(e, \left(p^2 - t_p^2\right)\left(q^2 - t_q^2\right)\right) = 1$. The decryption exponent is an integer d satisfying an equation $ed - k\left(p + 1 - t_p\right)\left(q + 1 - t_q\right) = 1$ where $t_p = p + 1 - \#E_p(a,b)$ and $t_q = q + 1 - \#E_q(a,b)$. In [121], we considered a more general equation, namely $eu - (p-s)(q-r)v = w$, and used lattice reduction methods to solve the equation which leads to the factorization of the modulus.

Similarly to KMOV and Demytko's scheme, a few cryptosystems have been proposed using an RSA modulus $N = pq$. The following systems are more or less variants of RSA and use a public exponent e satisfying the key equation $ed - k\left(p^2 - 1\right)\left(q^2 - 1\right) = 1$,

- LUC : proposed in 1993 by Smith and Lennon [143]. LUC is based on a Lucas functions,
- Castagnos cryptosystem : proposed in 2007 by Castagnos [29]. This system is directly related to LUC and RSA,
- RSA with Gaussian primes : proposed in 2002 by Elkamchouchi, Elshenawy and Shaban [43]. This scheme is the RSA variant for Gaussian primes,
- RSA type schemes based on singular cubic curves : proposed in 1995 by

Kuwakado, Koyama, and Tsuruoka [81]. The scheme uses the singular cubic curve $E_N(b)$ with equation $y^2 = x^3 + bx^2 \mod N$.

In [25], we launched an attack that factors the modulus $N = pq$ in the former schemes by using the generalized equation $ex - \left(p^2 - 1\right)\left(q^2 - 1\right) y = z$ and by combining the continued fraction algorithm and Coppersmith's method.

The Boolean functions play an important role in cryptography, especially in symmetric encryption algorithms, pseudorandom number generators and hash functions. Consequently, a lot of attention has been given to constructing Boolean functions satisfying good cryptographic criteria such as high algebraic degree, balancedness and high nonlinearity [28]. A Boolean function with n variables is a map from the space $GF(2)^n = \{0,1\}^n$ into $GF(2) = \{0,1\}$. It can be uniquely represented by a truth table or by an algebraic normal form (ANF). The theory of Boolean function is full of remarkable properties that can be used to construct Boolean functions with good cryptographic criteria [136]. In this direction, we introduced the notion of Dirichlet product for Boolean functions in [120]. For two boolean functions f and g, we defined the concept of Dirichlet product by setting $(f * g)(x) = \sum_{u \preceq x} f(u)g(x-u)$ for all $x \in GF(2)^n$ where, for $u = (u_1, \ldots, u_n) \in GF(2)^n$ and $x = (x_1, \ldots, x_n) \in GF(2)^n$, $u \preceq x$ if and only if for each $i \in \{1, \ldots, n\}$, $u_i \leq x_i$. Many properties of Boolean functions can be then reformulated in terms of the Dirichlet product.

Chapter 2

Cryptanalysis of RSA

2.1 Introduction

The RSA cryptosystem is the first and most widely used cryptosystem. It was developed by Rivest, Shamir and Adleman in 1978 [131]. RSA is used in many industrial systems such as web servers, online payment systems and other systems requiring privacy and authenticity.

The main parameters of the RSA cryptosystem are p,q, N, $\phi(N)$ and d where

- p and q are two private large prime numbers of the same bit-size,
- $N = pq$ is the public modulus,
- $\phi(N) = (p-1)(q-1)$ is the private Euler totient function,
- e is a public positive integer such that $\gcd(e, \phi(N)) = 1$,
- $d \equiv e^{-1} \pmod{\phi(N)}$ is the private exponent.

In textbook RSA, to encrypt a message $m \in \{2, N-1\}$ with the public key (N, e), one computes $c \equiv m^e \pmod{N}$ and to decrypt c with the private key (N, d) one simply computes $m \equiv c^d \pmod{N}$. The complete way to encrypt and decrypt with RSA needs different technique as recommended in RSA OAEP [103]

Since its invention, RSA has been analyzed for vulnerabilities by applying many kinds of attacks such as factorization, algebraic attacks and side channel attacks. Algebraic attacks are mainly based on diophantine approximations that could be solved using continued fractions or lattice reduction.

In 1990, Wiener came up with an attack on RSA based on the continued fraction algorithm. The starting point of the attack is the RSA key equation $ed - k\phi(N) = 1$ which leads to

$$\frac{k}{d} \approx \frac{e}{\phi(N)}.$$

When the prime factors p and q are of the same bit-size, $\phi(N) \approx N$ and

$$\frac{k}{d} \approx \frac{e}{N}.$$

Using this approximation, Wiener showed that if d is small enough, namely $d < \frac{1}{3}N^{\frac{1}{4}}$, then $\frac{k}{d}$ can be found among the convergents of the continued fraction expansion of $\frac{e}{N}$ which leads to the factorization of N.

In 1996, Coppersmith [34] described two rigorous methods to find small modular roots of univariate polynomials and small integer roots of bivariate polynomials. As an application, Coppersmith showed how to factor an RSA modulus $N = pq$ if half of the bits of p are known. Coppersmith's method is based on lattice reduction techniques and has many application in cryptanalysis, especially for attacking the RSA cryptosystem [15, 61, 71, 91, 101].

In this chapter, we give an overview of the techniques used for the cryptanalysis of RSA using the key equation or variants of it. This includes the continued fraction algorithm, lattice reduction and Coppersmith's technique. Then we describe some of our attacks on specific variants of the key equation which lead to the factorization of the underlying RSA moduli.

2.2 Continued Fractions

The theory of diophantine approximations, named after Diophantus of Alexandria, deals with the approximation of real numbers by rational numbers. This can be achieved by continued fractions. Continued fractions have many properties and applications in Number Theory and cryptographic problems. They

are used to find good diophantine approximations to rational and irrational numbers, to solve diophantine equations and to build attacks on some instances of RSA. In this section, we examine the basic properties of continued fractions.

Definition 2.2.1 (Continued Fraction Expansion)**.** A continued fraction is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_m + \ddots}}} = [a_0, a_1, \ldots, a_m, \ldots],$$

where a_0 is an integer and for $n \geq 1$, a_n is a positive integer. The integers a_n are called the partial quotients of the continued fraction.

It is clear that every finite continued fraction defines a rational number. Conversely, every real number $x \neq 0$ can be expanded as a finite or infinite continued fraction by the continued fraction algorithm as follows.

Let $\lfloor x \rfloor$ denote the greatest integer less than or equal to x. Let $x_0 = x$ and $a_0 = \lfloor x_0 \rfloor$. Then, for $i \geq 0$, define

$$x_{i+1} = \frac{1}{x_i - a_i}, \quad a_{i+1} = \lfloor x_{i+1} \rfloor.$$

The procedure terminates only if $a_i = x_i$ for some $i \geq 0$, that is if x is a rational number.

The continued fraction of a rational number $x = \frac{a}{b}$ with $\gcd(a, b) = 1$ can be computed by the Euclidean Algorithm in time $\mathcal{O}(\log b)$. Set $r_0 = a$ and $r_1 = b$. For $i \geq 0$, divide r_i by r_{i+1}:

$$r_i = a_i r_{i+1} + r_{i+2}, \quad 0 \leq r_{i+2} < r_{i+1}.$$

This process stops when $r_{m+2} = 0$ for some $m \geq 0$.

In 1990, Wiener [147] proposed an attack on RSA with modulus N and small private exponent d. The attack is based on the convergents of the continued fraction expansion of $\frac{e}{N}$.

Definition 2.2.2 (Convergent)**.** For $0 \leq n \leq m$, the nth convergent of the continued fraction $[a_0, a_1, \cdots, a_m]$ is $[a_0, a_1, \cdots, a_n]$.

For each $n \geq 0$, we define

$$\begin{aligned} p_{-2} &= 0, \quad p_{-1} = 1, \quad p_n = a_n p_{n-1} + p_{n-2}, \\ q_{-2} &= 1, \quad q_{-1} = 0, \quad q_n = a_n q_{n-1} + q_{n-2}. \end{aligned}$$

It is well known that the nth convergent of the continued fraction expansion satisfies $[a_0, a_1, \cdots, a_n] = \frac{p_n}{q_n}$. More generally, there are various results satisfied by the convergents of a continued fraction. We need only the following result on diophantine approximations (for more general information see [57] and [32]).

Theorem 2.2.3. *Let x be a real positive number. If a and b are positive integers such that* $\gcd(a, b) = 1$ *and*

$$\left| x - \frac{a}{b} \right| < \frac{1}{2b^2},$$

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

2.3 Lattice Reduction

The most powerful attacks on RSA are based on techniques that use lattice basis reduction algorithms, such as the LLL algorithm. Invented by Lenstra, Lenstra and Lovász [86] in 1982, LLL is a polynomial time algorithm for lattice basis reduction with many applications in cryptography [101]. A typical example of the powers of the LLL algorithm is the following problem.

Small roots of a modular polynomial problem: Given a composite N with unknown factorization and a polynomial $f(x)$ of degree d, find all small solutions x_0 to the polynomial equation $f(x) \equiv 0 \pmod{N}$.

In his seminal work, Coppersmith [34] solved this problem in 1996 for solutions x_0 satisfying $|x_0| < N^{\frac{1}{d}}$ using the LLL algorithm.

In this section, we give the mathematical background on lattices and the LLL algorithm for basis reduction. We start by giving a formal definition of a lattice.

Definition 2.3.1 (Lattice)**.** Let $n \leq m$ be two positive integers and $b_1, \cdots, b_n \in \mathbb{R}^m$ be n linearly independent vectors. A lattice $\mathcal{L}$ spanned by $\{b_1, \cdots, b_n\}$ is the set of all integer linear combinations of $b_1, \cdots, b_n$, that is

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

The set $\langle b_1 \ldots, b_n \rangle$ is called a lattice basis for $\mathcal{L}$. The lattice dimension is $\dim(\mathcal{L}) = n$.

In general, a basis for $\mathcal{L}$ is any set of independent vectors that generates $\mathcal{L}$. Any two bases for a lattice $\mathcal{L}$ are related by a matrix having integer coefficients and determinant equal to ± 1. Hence, all the bases have the same Gramian determinant $\det_{1 \leq i,j \leq n} \langle b_i, b_j \rangle$ where $\langle b_i, b_j \rangle$ denotes the scalar product of vectors b_i, b_j. The determinant of the lattice is then

$$\det(\mathcal{L}) = \left(\det_{1 \leq i,j \leq n} \langle b_i, b_j \rangle \right)^{\frac{1}{2}}.$$

Let $v = \sum_{i=1}^{n} x_i b_i$ be a vector of $\mathcal{L}$. We define the Euclidean norm of v as

$$\|v\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

Given a basis $\langle b_1 \ldots, b_n \rangle$ of the lattice $\mathcal{L}$, the Gram-Schmidt process gives an orthogonal set $\langle b_1^* \ldots, b_n^* \rangle$. The determinant of the lattice is then $\det(\mathcal{L}) = \prod_{i=1}^{n} \|b_i^*\|$. The Gram-Schmidt procedure starts with $b_1^* = b_1$, and then for $i \geq 2$,

$$i \geq 2, \quad b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{i,j} b_j^*, \quad \text{where} \quad \mu_{i,j} = \frac{\langle b_i, b_j^* \rangle}{\langle b_j^*, b_j^* \rangle} \quad \text{for} \quad 1 \leq j < i.$$

Note that $\langle b_1^* \ldots, b_n^* \rangle$ is not a basis of the lattice $\mathcal{L}$. Since every nontrivial lattice has infinitely many bases, some bases are better than others. The most important quality measure is the length of the basis vectors. For arbitrary lattices, the problem of computing a shortest vector is known to be NP-hard under randomized reductions [3]. However, in many applications, the LLL algorithm computes in polynomial time a reduced basis with nice properties.

Definition 2.3.2 (LLL Reduction)**.** Let $B = \langle b_1, \ldots, b_n \rangle$ be a basis for a lattice $\mathcal{L}$ and let $B^* = \langle b_1^*, \ldots, b_n^* \rangle$ be the associated Gram-Schmidt orthogonal basis. Let

$$\mu_{i,j} = \frac{\langle b_i, b_j^* \rangle}{\langle b_j^*, b_j^* \rangle} \quad \text{for} \quad 1 \leq j < i.$$

The basis B is said to be LLL reduced if it satisfies the following two conditions:

$$|\mu_{i,j}| \leq \frac{1}{2}, \quad \text{for} \quad 1 \leq j < i \leq n,$$
$$\frac{3}{4}\|b_{i-1}^*\|^2 \leq \|b_i^* + \mu_{i,i-1} b_{i-1}^*\|^2 \quad \text{for} \quad 1 < i \leq n.$$

Below we give useful inequalities satisfied by an LLL reduced basis derived from the LLL reduction definition (for a proof see e.g. [86], [32], [91]).

Theorem 2.3.3. *Let $\mathcal{L}$ be a lattice of dimension n. Let $B = \langle b_1, \ldots, b_n \rangle$ be an LLL reduced basis and let $B^* = \{b_1^*, \ldots, b_n^*\}$ be the associated Gram-Schmidt orthogonal basis. Then*

$$\|b_1\| \leq \|b_2\| \leq \ldots \leq \|b_i\| \leq 2^{\frac{n(n-i)}{4(n+1-i)}} (\det(\mathcal{L}))^{\frac{1}{n+1-i}} \quad \textit{for} \quad 1 \leq i \leq n.$$

2.4 Coppersmith's Method

In 1988, Hastad [55] and Toffin, Girault and Vallée [51] used lattice reduction techniques to find very small solutions of modular polynomial equations of the form $f(x) \equiv 0 \pmod{N}$ where $f(x) \in \mathbb{Z}[x]$. They showed that this technique can be applied for cryptanalytic purposes. In 1996, Coppersmith [34] further improved the former bounds and developed two rigorous methods for finding small solutions of polynomial equations, the first for the univariate modular case and the second one for the bivariate case.

For a polynomial $f(x) \in \mathbb{Z}[x]$ of degree δ and a known positive integer N, the univariate modular case is the equation $f(x) \equiv 0 \pmod{N}$.
The bivariate case is the equation $g(x, y) = 0$ over the integers where $g(x, y) = \sum_{i,j} g_{i,j} x^i y^j \in \mathbb{Z}[x, y]$.

Using lattice reduction techniques such as the LLL algorithm [86], Coppersmith showed that one can find all solutions x_0 of the modular equation $f(x) \equiv 0 \pmod{N}$ with $|x_0| \leq N^{\delta}$ and all solutions (x_0, y_0) of the equation $g(x, y) = 0$ with $|x_0| < X$ and $|y_0| < Y$ with $XY < W^{\frac{2}{3\delta}}$ where $W = \max_{i,j} |g_{i,j}| X^i Y^j$.

Since 1996, many cryptanalytic applications have been based on Coppersmith's method, for example the factorization of $N = pq$ knowing a fraction of the most significant bits on each factor. Another well-known application of Coppersmith's method is the cryptanalysis of RSA with small private key. In 1999, based on the seminal work of Coppersmith, Boneh and Durfee [17] presented an attack on RSA which recovers the factors p and q of an RSA modulus $N = pq$ if $d < N^{0.292}$. This result improves the well known bound $d < \frac{1}{3} N^{\frac{1}{4}}$ of Wiener [147]. To simplify Coppersmith's methods, Howgrave-Graham [65] and Coron [36] revisited the problem of finding small solutions of polynomial equations. Later, the ideas of Coppersmith have been generalized to multivariate polynomials. The generalizations use the Euclidean norm of a multivariate polynomial and Howgrave-Graham's Theorem.

Definition 2.4.1. Let $f(x_1, \ldots, x_n) = \sum_{i_1, \ldots, i_n} a_{i_1, \ldots, i_n} x^{i_1} \cdots x^{i_n}$ be a polynomial and $X_1, \ldots, X_n$ be n real numbers. The Euclidean norm of the polynomial $f(X_1 x_1, \ldots, X_n x_n)$ is defined as

$$\|f(X_1 x_1, \ldots, X_n x_n)\| = \left(\sum_{i_1, \ldots, i_n} \left(a_{i_1, \ldots, i_n} X_1^{i_1} \cdots X_n^{i_n} \right)^2 \right)^{\frac{1}{2}}.$$

Theorem 2.4.2 (Howgrave-Graham)**.** *Let $h(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ be a polynomial with at most ω monomials. Suppose that*

$$\begin{aligned} &h(x_1^{(0)}, \ldots, x_n^{(0)}) \equiv 0 \pmod{B}, \\ &|x_1^{(0)}| < X_1, \ldots, |x_n^{(0)}| < X_n, \\ &\|h(X_1 x_1, \ldots, X_n x_n)\| < \frac{B}{\sqrt{\omega}}. \end{aligned}$$

Then $h(x_1^{(0)}, \ldots, x_n^{(0)}) = 0$ holds over the integers.

2.5 Attacks on RSA Using a Variant of the Key Equation

This section concerns the cryptanalysis of RSA using a variant of the key equation. Many attacks on RSA exploit the RSA key equation

$$ed - k(p-1)(q-1) = 1.$$

Indeed, the starting point in Wiener's attack [147] is based on the RSA key equation and uses the approximation $(p-1)(q-1) \approx N$. Similarly, in the attack of Boneh and Durfee [17], the RSA key equation is transformed into a modular polynomial equation of the form

$$x\left(\frac{N+1}{2} - y\right) = 1 \pmod{e}.$$

While Wiener's attack is based on the continued fraction algorithm, the attack of Boneh and Durfee is entirely based on Coppersmith's method.

Some attacks on RSA combine both techniques. Blömer and May [13] presented an attack on RSA using the RSA variant equation $ex + y = k(p-1)(q-1)$ and combined the continued fraction algorithm and Coppersmith's method to solve it when $x < \frac{1}{3}N^{1/4}$ and $|y| = O\left(N^{-3/4}ex\right)$.

Such variant of the key equation can be extended in various ways. In [107], we considered a typical example where the set of the public exponents e satisfy the equation

$$eX - (N - (ap + bq))Y = Z, \tag{2.1}$$

with small parameters X, Y and Z where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ with $a \geq 1$. Here, p and q are of the same bit size and are ordered such that $q < p < 2q$. This situation is related to the requirement of the ANSI X9.31:1998 standard for public key cryptography [1] where it is advised to ovoid prime numbers p and q for RSA with a ratio close to a rational number $\frac{a}{b}$ with small a and b.

The study of equation (2.1) uses two techniques. First, we apply the continued fraction algorithm to determine $\frac{Y}{X}$ among the convergents of the continued fraction of $\frac{e}{N}$. Second, we compute p and q by applying Coppersmith's method.

First we present the result based on continued fractions.

Lemma 2.5.1. *[107] Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let a, b be coprime positive integers such that $ap + bq = N^{\frac{1}{2}+\alpha}$ with $\alpha < \frac{1}{2}$. Let e be a public exponent satisfying the equation $eX - (N - (ap + bq))Y = Z$ with $\gcd(X, Y) = 1$. If $|Z| < N^{\frac{1}{2}+\alpha}X$ and $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$, then $\frac{Y}{X}$ is a convergent of $\frac{e}{N}$.*

The proof of Lemma 2.5.1 is entirely based on Legendre's Theorem 2.2.3. Once X and Y are found, we set $M = N - \frac{eX}{Y}$ and then we compute

$$|ap + bq - M| = \frac{|Z|}{Y}.$$

To study the equation (2.1), we will consider three cases according to the size of the difference $|ap - bq|$:

1. case 1 : $|ap - bq|$ is small, i.e. $|ap - bq| < (abN)^{\frac{1}{4}}$, which corresponds approximately to $b > 2^{\frac{1}{2}}N^{\frac{1}{6}}$
2. case 2 : $|ap - bq|$ is medium, i.e. $(abN)^{\frac{1}{4}} < |ap - bq| < aN^{\frac{1}{4}}$, which corresponds approximately to $2^{\frac{1}{2}}N^{\frac{1}{6}} > b > 2^{\frac{1}{4}}N^{\frac{1}{8}}$,
3. case 3 : $|ap - bq|$ is large, i.e. $|ap - bq| > aN^{\frac{1}{4}}$, which corresponds approximately to $b < 2^{\frac{1}{4}}N^{\frac{1}{8}}$.

In the three cases, we will use the following results. The first result shows that one can find ab if an approximation of $ap+bq$ is known. Here the integer closest to x is denoted $[x]$.

Lemma 2.5.2. *[107] Let $N = pq$ be an RSA modulus with $q < p < 2q$ and $\frac{a}{b}$ a convergent of the continued fraction expansion of $\frac{q}{p}$ with $a \geq 1$. Let $ap + bq = N^{\frac{1}{2}+\alpha}$ with $\alpha < \frac{1}{2}$. If $|ap + bq - M| < \frac{1}{2}N^{\frac{1}{2}-\alpha}$, then*

$$ab = \left[\frac{M^2}{4N}\right].$$

The second result is related to the seminal work of Coppersmith [34].

Theorem 2.5.3. *Let $n = pq$ be the product of two unknown integers such that $q < p < 2q$. Given an approximation of p with additive error at most $n^{\frac{1}{4}}$, one can find p and q in polynomial time.*

2.5.1 An Attack for Small Difference $|ap - bq|$

The small difference corresponds to case 1.

Theorem 2.5.4. *[107] Let $N = pq$ be an RSA modulus with unknown factors p, q such that $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued fraction expansion of $\frac{q}{p}$ with $a \geq 1$ and $|ap - bq| < (abN)^{\frac{1}{4}}$. Let e be a public exponent satisfying an equation $eX - (N - ap - bq)Y = Z$ with $\gcd(X, Y) = 1$. Set $ap + bq = N^{\frac{1}{2}+\alpha}$ with $0 < \alpha < \frac{1}{2}$. If $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ and $|Z| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$, then N can be factored in polynomial time.*

To prove this result, one first use the continued fraction algorithm to find $\frac{Y}{X}$ among the convergents of $\frac{e}{N}$. If we set $M = N - \frac{eX}{Y}$, then one find that M is an approximation of $ap + bq$ with error term less than $(abN)^{\frac{1}{4}}$. Using Lemma 2.5.2, this allows us to find the exact value of the product ab as

$$ab = \left\lfloor \frac{M^2}{4N} \right\rfloor.$$

If we assume that

$$|Z| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y,$$

then one finds that the term $\frac{M}{2}$ is an approximation of the factor ap of $n = abN$ with additive error at most $n^{\frac{1}{4}}$. Hence, using Theorem 2.5.3 with n and $\frac{M}{2}$, we find ap, and since $a < q$, we get $p = \gcd(N, ap)$.

2.5.2 An Attack for Medium Difference $|ap - bq|$

The case with medium difference $|ap - bq|$ corresponds to the case when $2^{\frac{1}{2}}N^{\frac{1}{6}} > b > 2^{\frac{1}{4}}N^{\frac{1}{8}}$. The method here uses the Elliptic Curve Method (ECM) [84] which can find factors of about 52-digits. Assuming the efficiency of ECM, every step in this attack can be done in polynomial time and the number of convergents is bounded by $\mathcal{O}(\log N)$.

Theorem 2.5.5. *[107] Let $N = pq$ be an RSA modulus with unknown factors p, q such that $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued fraction expansion of $\frac{q}{p}$ such that $a \geq 1$, $(abN)^{\frac{1}{4}} < |ap - bq| < aN^{\frac{1}{4}}$ and*

$b \leq 10^{52}$. *Let e be a public exponent satisfying an equation* $eX - (N - ap - bq)Y = Z$ *with* $\gcd(X, Y) = 1$. *Set* $M = N - \frac{eX}{Y}$ *and* $ap + bq = N^{\frac{1}{2}+\alpha}$ *with* $0 < \alpha < \frac{1}{2}$. *If* $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ *and* $|Z| < \min\left(aN^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$, *then, under ECM, N can be factored efficiently.*

First, using the conditions of the theorem, one can apply Lemma 2.5.1 to find $\frac{Y}{X}$ among the convergents of $\frac{e}{N}$. Using X and Y, we compute $M = N - \frac{eX}{Y}$, and, using Lemma 2.5.2, we find $ab = \left[\frac{M^2}{4N}\right]$. If moreover, $b \leq 10^{52}$, then, applying the Elliptic Curve Method with $\left[\frac{M^2}{4N}\right]$, we can efficiently find a and b. Then, using the assumption $|ap - bq| < aN^{\frac{1}{4}}$, we find that $\frac{M}{2a}$ is an approximation of p with error term at most $N^{\frac{1}{4}}$. Then, using Theorem 2.5.3, one can find p and then q.

2.5.3 An Attack for Large Difference $|ap - bq|$

In the case 3, we assume that the difference $|ap - bq|$ is large, which corresponds to $b < 2^{\frac{1}{4}}N^{\frac{1}{8}}$.

Theorem 2.5.6. *[107] Let $N = pq$ be an RSA modulus with unknown factors p, q such that $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued fraction expansion of $\frac{q}{p}$ such that $a \geq 1$ and $b \leq 10^{52}$. Let e be a public exponent satisfying an equation* $eX - (N - (ap + bq))Y = Z$ *with* $\gcd(X, Y) = 1$. *Let $M = N - \frac{eX}{Y}$. Set $D = \sqrt{|M^2 - 4abN|}$ and $ap + bq = N^{\frac{1}{2}+\alpha}$ with $0 < \alpha < \frac{1}{2}$. If $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ and $|Z| < \frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}Y$ then, under ECM, N can be factored efficiently.*

Similarly to case 1 and case 2, we find $\frac{eX}{Y}$ among the convergents of $\frac{e}{N}$, and by setting $M = N - \frac{eX}{Y}$, we find that M is an approximation $ap + bq$ with an error term of at most $aN^{\frac{1}{4}}$ and that $ab = \left[\frac{M^2}{4N}\right]$. Then, if $b \leq 10^{52}$, then applying the Elliptic Curve Method with $\left[\frac{M^2}{4N}\right]$, we can find a and b. Next, let $D = \sqrt{|M^2 - 4abN|}$. Then $\pm D$ is an approximation of $ap - bq$ with an error term of at most $aN^{\frac{1}{4}}$. Combining the approximations of $ap + bq$ and $ap - bq$, we find the approximation $\frac{M \pm D}{2a}$ is an approximation of p with

additive error at most $N^{\frac{1}{4}}$. We can then apply Theorem 2.5.3 to the values $\frac{M \pm D}{2a}$. The correct term will lead to the factorization of N.

For a given RSA modulus $N = pq$, the number of exponents e satisfying the equation $eX - (N - (ap + bq))Y = Z$ with suitably small parameters and $e < \phi(N) = (p-1)(q-1)$ is at least $N^{\frac{3}{4}-\varepsilon}$ where ε is arbitrarily small for suitably large N. This is much larger than the number ($\approx N^{\frac{1}{4}-\varepsilon}$) of exponents that are weak for Wiener's attack [147]. It is also much larger than the number ($\approx N^{0.292-\varepsilon}$) of exponents that are weak for the attack of Boneh and Durfee [17].

We note that the attack described in this section answers positively the requirement of the ANSI X9.31:1998 standard for public key cryptography [1] where it is advised to choose the prime numbers p and q for RSA with a ratio not too close to a rational number $\frac{a}{b}$ with small a and b.

2.6 An Attack on RSA Unbalanced Moduli

In PKC 2009, May and Ritzenhofen [94] presented a method for factoring two RSA moduli $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ 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 is a lattice based method that allows to find the factorization of N_1 and N_2 when $t \geq 2\alpha + 3$. The method can be heuristically generalized to a lattice based method to factor k RSA moduli $N_1 = p_1 q_1, \ldots, N_k = p_k q_k$ when the p_i's share $t \geq \frac{k}{k-1}\alpha$ many LSBs. The method of May and Ritzenhofen was reconsidered in [135] by Sarkar and Maitra. Their method works when $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ are such that p_1 and p_2 share their least significant bits (LSBs) or most significant bits (MSBs) as well as a contiguous portion of bits at the middle. In PKC 2010, Faugère, Marinier and Renault [45] presented a new and rigorous lattice-based method that addresses the implicit factoring problem when p_1 and p_2 share $t \geq 2\alpha + 3$ MSBs where the prime numbers q_i are α-bit primes. The method heuristically generalizes to k RSA moduli $N_1 = p_1 q_1, \ldots, N_k = p_k q_k$ when the prime numbers p_i share $t \geq \frac{k}{k-1}\alpha + 6$ of MSBs. For two RSA moduli $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ of the same bit size, Kurosawa and Ueda [82] presented a lattice-based method to factor the two

moduli when p_1 and p_2 share t LSBs with $t \geq 2\alpha + 1$ where $q_1 \approx q_2 \approx 2^\alpha$.

In the rest of this section, we show that it is possible to factor k RSA moduli $N_1 = p_1q_1, \ldots, N_k = p_kq_k$ when some unknown multiples a_ip_i of the prime factors p_i share an amount of MSBs or of LSBs. This method generalizes all the former attacks. This is a brief description of the the attack presented in [117].

2.6.1 Implicit factorization of two RSA Moduli

We study here the problem of factoring two RSA moduli $N_1 = p_1q_1$ and $N_2 = p_2q_2$ where some unknown multiples a_1p_1 and a_2p_2 coincide on the t most significant bits (MSBs), that is when $|a_2p_2 - a_1p_1|$ is sufficiently small.

Theorem 2.6.1. *[117] 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.*

The proof is based in the continued fraction algorithm. Indeed, if $a_1 < p_2$, $a_2 < p_1$ and $|a_1p_1 - a_2p_2| < \frac{p_1}{2a_2q_1q_2}$, then we have

$$\left|\frac{N_2}{N_1} - \frac{a_1q_2}{a_2q_1}\right| = \frac{|a_1p_1 - a_2p_2|q_2}{a_2p_1q_1} < \frac{p_1}{2a_2q_1q_2} \times \frac{q_2}{a_2p_1q_1} = \frac{1}{2(a_2q_1)^2}.$$

This implies by Theorem 2.2.3 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 $q_2 = \gcd(N_2, a_1q_2)$ and $p_2 = \frac{N_2}{q_2}$.

Notice that the result of Theorem 2.6.1 is valid even when the RSA moduli are not of the same size. Comparatively, the attacks presented by Sarkar and Maitra in [135] and Faugère et al. in [45] are valid only if $N_1 \approx N_2$ and $q_1 \approx q_2$.

As an application of Theorem 2.6.1 to factor two unbalanced RSA moduli of the same size, we have the following result

Corollary 2.6.2. *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 $a_1 p_1$ and $a_2 p_2$ 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.

Notice that, with $\beta = 0$ in Corollary 2.6.2, 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 [45]. This shows that the bound found by Faugère et al. with lattice reduction techniques can be achieved using the continued fraction algorithm instead.

2.6.2 Implicit factorization of k RSA Moduli

Let $N_i = p_i q_i$, $i = 1 \ldots, k$, be $k \geq 3$ RSA moduli of the same bit size. We show that if k multiples $a_i p_i$, $i = 1 \ldots, k$ share t most significant bits, then one can factor the k moduli. The method here is based on lattice reduction, especially the LLL algorithm [86].

Theorem 2.6.3. *[117] 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 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 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.

The idea to prove Theorem 2.6.3 is to transform the k differences $x_i = a_i p_i - a_1 p_1$ into k simultaneous equations

$$a q_1 N_i - \frac{a a_1 q_i}{a_i} N_1 = \frac{a q_1 q_i x_i}{a_i},$$

where $a = a_1 a_2 \cdots a_k$. To solve the equations, we apply the LLL algorithm

to the lattice $\mathcal{L}$ spanned by the rows of the matrix

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

where C is a positive integer to be optimized later. Consider the vector

$$v = \left(Caq_1, \frac{aq_1q_2x_2}{a_2}, \dots, \frac{aq_1q_kx_k}{a_k} \right) \in \mathbb{Z}^k.$$

Since

$$v = \left(aq_1, \frac{aa_1q_2}{a_2} \dots, \frac{aa_1q_k}{a_k} \right) \times M,$$

then, by reducing the lattice $\mathcal{L}$, one can find v among the smallest vectors. Indeed, the Gaussian Heuristic for $\mathcal{L}$ asserts that the length of its shortest non-zero vector is usually $\sigma(\mathcal{L})$ where

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

If we choose C such that $\sigma(\mathcal{L}) > \|v\|$, then v can be found among the shortest non-zero vectors of the lattice $\mathcal{L}$. This leads to the factorization of the k RSA moduli by computing $q_1 = \gcd(N_1, aq_1)$ and for $i \geq 2$, $q_i = \gcd\left(N_i, \frac{aa_1q_i}{a_i}\right)$. Any value for C with $C \geq 2^{n-t}$ will be optimal for the reduction and will lead to the condition

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

where $q_i \approx 2^\alpha$ and $a_i \leq 2^\beta$ for $i = 1, \dots, k$.

It is also possible to consider the same situation as before where the least significant bits (LSBs) are shared instead of the most significant bits (MSBs). The techniques and the result are similar.

Chapter 3

Cryptanalysis of Variants of RSA

In this chapter, we consider three variants of the RSA cryptosystem and show how to adapt the cryptanalytic attacks on standard RSA to break the underlying systems. The variant RSA cryptosystems studied in this chapter are KMOV, Demytko's scheme and four variants of the RSA cryptosystem with the same key equations, namely the Kuwakado-Koyama-Tsuruoka cryptosystem, a cryptosystem of Castagnos, a cryptosystem based on Gaussian integers and LUC cryptosystem.

3.1 Cryptanalysis of KMOV

In this section, we describe the attack presented in [119] on the KMOV cryptosystem.

In 1991, Koyama, Maurer, Okamoto and Vanstone [79] introduced a new public key cryptosystem, called KMOV. The KMOV cryptosystem is based on elliptic curves over the ring $\mathbb{Z}_n$ where $n = pq$ is an RSA modulus, that is, the product of two large unknown primes of equal bit-size. The KMOV public key is denoted by (n, e) where $n = pq$ and e is an integer satisfying $\gcd(e, (p+1)(q+1)) = 1$. The corresponding private exponent d is an integer satisfying $ed \equiv 1 \pmod{(p+1)(q+1)}$ which can be reformulated as an equation

$$ed - k(p+1)(q+1) = 1.$$

In this section, we consider KMOV with a public exponent e satisfying the

generalized equation

$$ex - (p+1)(q+1)y = z,$$

where x and y are co-prime positive integers.

In 1995, Pinch [128] extended Wiener's attack [147] to KMOV using similar techniques applied with the key equation $ed - k(p+1)(q+1) = 1$, that is when $z = 1$ in our generalization. Similarly, Ibrahimpašić [69], studied the security of KMOV with short secret exponents using the equation $ed - k(p+1)(q+1) = 1$.

The KMOV cryptosystem uses the arithmetic of elliptic curves over the ring $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ where $n = pq$ is the product of two large distinct primes p and q. An elliptic curve $E_n(a,b)$ over $\mathbb{Z}_n$ is the set of points $(x,y) \in \mathbb{Z}_n^2$ satisfying

$$y^2 = x^3 + ax + b \pmod{n}$$

together with the point at infinity $\mathcal{O}$. The addition law can be extended for points in a curve $E_n(a,b)$ over $\mathbb{Z}_n$ using the same rules than the addition over a finite field. $E_n(a,b)(\mathbb{Z}_n)$ is not a group. By the Chinese Remainder Theorem, the mapping

$$E_n(a,b) \to E_p(a,b) \times E_q(a,b)$$

defined by the natural projections is a bijection. Thus, a point (x,y) of the elliptic curve $E_n(a,b)$ is associated to the point

$$((x \pmod{p}, y \pmod{p}), (x \pmod{q}, y \pmod{q})) \in E_p(a,b) \times E_q(a,b).$$

The points $(\mathcal{O}, P)$ and $(P, \mathcal{O})$ can not be represented like this. Finding such a point is, however, very unlikely and would lead to the factorization of n (see [84]).

Let $\#E_p(a,b)$ denote the number of distinct points of the elliptic curve $E_p(a,b)$. We have the following result which is a consequence of the Chinese Remainder Theorem .

Lemma 3.1.1. *Let $n = pq$ be an RSA modulus and $E_n(a,b)$ an elliptic curve over $\mathbb{Z}_n$ with $\gcd\left(4a^3 + 27b^2, n\right) = 1$. Then for any $P \in E_n(a,b)$ and any integer k, we have*

$$(1 + k\#E_p(a,b)\#E_q(a,b))P = P.$$

The algorithms in KMOV work as follows.

- **Key Generation**
  INPUT: The bit-length k of the RSA modulus.
  OUTPUT: The public key (n, e) and the private key (n, d).

  1. Find two primes, p and q, of length $k/2$ bits satisfying $p \equiv q \equiv 2 \pmod 3$.
  2. Compute the RSA modulus $n = pq$.
  3. Choose a public key e co-prime to $(p+1)(q+1)$.
  4. Compute the inverse d of e mod $((p+1)(q+1))$.
  5. Return the public key (n, e) and the private key (n, d).

- **KMOV Encryption**
  INPUT: The public key (n, e) and the plaintext message m.
  OUTPUT: The cyphertext (c_1, c_2).

  1. Represent the message m as a couple $(m_1, m_2) \in \mathbb{Z}_n^2$.
  2. Compute $b = m_2^2 - m_1^3 \pmod n$.
  3. Compute the point $(c_1, c_2) = e(m_1, m_2)$ on the elliptic curve $y^2 = x^3 + b \pmod n$.
  4. Return (c_1, c_2).

- **KMOV Decryption**
  INPUT: The private key (n, d) and the cyphertext (c_1, c_2).
  OUTPUT: The plaintext message (m_1, m_2).

  1. Compute $b = c_2^2 - c_1^3 \pmod n$. Note that the receiver of a message never need to compute b, but he can compute it.
  2. Compute the point $(m_1, m_2) = d(c_1, c_2)$ on the elliptic curve $y^2 = x^3 + b \pmod n$.
  3. Return (m_1, m_2).

The choice of the prime numbers p and q with $p \equiv q \equiv 2 \pmod 3$ is motivated by the following lemma.

Lemma 3.1.2. *Let $p > 3$ be a prime satisfying $p \equiv 2 \pmod 3$ and $0 < b < p$. Then*

$$\#E_p(0, b) = p + 1.$$

The starting point for our attack on KMOV is the equation

$$ex - (p + 1)(q + 1)y = z.$$

When the unknown parameters x, y and z are suitably small, then by applying the continued fraction algorithm and Coppersmith's method, one can solve the equation and then factor the RSA modulus $n = pq$.

Theorem 3.1.3. *[119] Let $n = pq$ be an RSA modulus with $q < p < 2q$. Suppose that the public exponent e satisfies an equation $ex-(p+1)(q+1)y = z$ where x and y are positive integers with $\gcd(x, y) = 1$ and*

$$|z| < n^{\frac{1}{4}}y, \quad xy < \frac{\sqrt{2}\sqrt{n}}{12}.$$

Then $\frac{y}{x}$ is one of the convergents of the continued fraction expansion of $\frac{e}{n}$.

The proof begins by transforming the equation $ex - (p+1)(q+1)y = z$ into $ex - ny = (p+q+1)y + z$. Using some properties on the size of p and q, namely $p + q < \frac{3\sqrt{2}}{2}\sqrt{n}$ and assuming that $|z| < n^{\frac{1}{4}}y$ and $xy < \frac{\sqrt{2}\sqrt{n}}{12}$, we get

$$\left|\frac{e}{n} - \frac{y}{x}\right| = \frac{|(p + q + 1)y + z|}{nx} < \frac{3\sqrt{2}\sqrt{n}y}{nx} < \frac{1}{2x^2}.$$

This implies, by Theorem 2.2.3, that $\frac{y}{x}$ is a convergent of the continued fraction expansion of $\frac{e}{n}$. Moreover, under an extra condition, one can use the fraction $\frac{y}{x}$ to factor the modulus $n = pq$.

Theorem 3.1.4. *[119] Let $n = pq$ be an RSA modulus with $q < p < 2q$. Suppose that e is an exponent satisfying an equation $ex - (p+1)(q+1)y = z$ with $\gcd(x, y) = 1$ and*

$$|z| < \frac{(p - q)n^{\frac{1}{4}}y}{3(p + q)}, \quad xy < \frac{\sqrt{2}\sqrt{n}}{12}.$$

Then n can be factored in polynomial time.

If we set

$$U = \frac{ex}{y} - n - 1, \qquad V = \sqrt{|U^2 - 4n|},$$

then using the equation $ex - (p+1)(q+1)y = z$, we get

$$|U - p - q| = \left|\frac{ex}{y} - n - 1 - p - q\right| = \frac{|z|}{y} < \frac{(p-q)n^{\frac{1}{4}}}{3(p+q)}.$$

Hence, if $|z| < \frac{(p-q)n^{\frac{1}{4}}y}{3(p+q)}$, then $|U - p - q| < n^{\frac{1}{4}}$. This leads to $|p - q - V| < n^{\frac{1}{4}}$. Then, combining the approximations of $p+q$ and $p-q$, we get $\left|p - \frac{U+V}{2}\right| < n^{\frac{1}{4}}$. Hence, using Coppersmith's Theorem 2.5.3, we get p and then the factorization of n.

3.2 Cryptanalysis of Demytko's cryptosystem

In this section, we describe the attack presented in [121] on the Demytko cryptosystem.

Let p be a prime number and a and b be two positive integers such that $\gcd(4a^3 + 27b^2, p) = 1$. Consider the elliptic curve $E_p(a, b)$ over the field $\mathbb{F}_p$. It is the set of points $P = (x, y) \in \mathbb{F}_p^2$ such that $y^2 \equiv x^3 + ax + b \pmod{p}$ together with the point at infinity. The number of such points is denoted $\#E_p(a, b)$ and satisfies $\#E_p(a, b) = p + 1 - t_p$ where, according to Hasse Theorem [140] satisfies $|t_p| \leq 2\sqrt{p}$.

Let $N = pq$ be an RSA modulus and let a and b be two integers such that $\gcd(4a^3 + 27b^2, N) = 1$. An elliptic curve $E_N(a, b)$ is the set of points (x, y) such that

$$y^2 \equiv x^3 + ax + b \pmod{N},$$

together with the point at infinity $\mathcal{O}$. It is well known that chord-and-tangent method in the case of elliptic curves $E_p(a, b)$ defined over the finite filed $\mathbb{F}_p$ still hold for $E_N(a, b)$ unless the inversion of a non-zero number Q does not exist modulo N. This case would lead to find a factor of N by computing $\gcd(Q, N)$. When the prime factors p, q in $N = pq$ are large, then with overwhelming probability the inversion of a non-zero number will exist modulo N.

In 1994, Demytko [39] developed a cryptosystem using an elliptic curve $E_N(a, b)$ over the ring $\mathbb{Z}/N\mathbb{Z}$ where $N = pq$ is an RSA modulus which can be summarized as follows.

1. **Key Generation:**
   - Choose two distinct prime numbers p and q of similar bit-length.
   - Compute $N = pq$.
   - Select two integers $a, b < p$ such that $\gcd\left(N, 4a^3 + 27b^2\right) = 1$.
   - Choose e such that $\gcd\left(e, \left(p^2 - t_p^2\right)\left(q^2 - t_q^2\right)\right) = 1$.
   - Keep p, q secret and publish N, e, a, b.
2. **Encryption:**
   - Transform the message m as the x-coordinate of a point $P = (m_x, m_y)$ on the elliptic curve $E_N(a, b)$.
   - Compute the ciphertext point $C = eP = (c_x, c_y) = e(m_x, m_y)$ on the elliptic curve $y^2 = x^3 + ax + b \pmod N$.
3. **Decryption:**
   - Compute $u = c_x^3 + ac_x + b \pmod N$.
   - Compute the Legendre symbols $u_p = \left(\frac{u}{p}\right)$ and $u_q = \left(\frac{u}{q}\right)$.
   - If $(u_p, u_q) = (1, 1)$, then compute $d \equiv e^{-1} \pmod{\text{lcm}(p + 1 - t_p, q + 1 - t_q)}$.
   - If $(u_p, u_q) = (1, -1)$, then compute $d \equiv e^{-1} \pmod{\text{lcm}(p + 1 - t_p, q + 1 + t_q)}$.
   - If $(u_p, u_q) = (-1, 1)$, then compute $d \equiv e^{-1} \pmod{\text{lcm}(p + 1 + t_p, q + 1 - t_q)}$.
   - If $(u_p, u_q) = (-1, -1)$, then compute $d \equiv e^{-1} \pmod{\text{lcm}(p+1+t_p, q+1 + t_q)}$.
   - Compute m as the x-coordinate of $dC = deP = P = (m_x, m_y)$ on the elliptic curve $y^2 = x^3 + ax + b \pmod N$.

In the following, we consider $d \equiv e^{-1} \pmod{(p+1 \pm t_p, q+1 \pm t_q)}$ instead of modulo $\mathrm{lcm}(p+1 \pm t_p, q+1 \pm t_q)$. This implies that e and d satisfy an equation of the form

$$ed - k\,(p-s)\,(q-r) = 1, \quad s = \mp t_p - 1, \quad r = \mp t_q - 1.$$

Then, a generalization of this equation is of the form $eu-(p-s)(q-r)v = w$.

Under some conditions on the size of the unknown parameters, one can apply Coppersmith's technique [34] and the elliptic curve method for factoring [84] to solve the equation and then find p and q which factors N and breaks the system. This will be achieved in two steps. The first step enables us to find the product $(p-s)(q-r)$.

Theorem 3.2.1 ([121], Theorem 3.1)**.** *Let $N = pq$ be an RSA modulus and $e = N^\beta$ be a public exponent. Suppose that e satisfies the equation $eu-(p-s)(q-r)v = w$ with $|r|, |s| < N^\alpha$, $u < N^\delta$ and $|w| < N^\gamma$. If*

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)} - \varepsilon,$$

then one can find the product $(p-s)(q-r)$ in polynomial time.

The general idea in the proof is to use Coppersmith's technique combined with Jochemz-May strategy [72], Howgrave-Graham's Theorem [65] and lattice reduction.

Next, assume that $(p-s)(q-r)$ is computed and that one of the factors $p-s$ or $q-r$ is B-smooth, that is all prime factors of $(p-s)(q-r)$ are less than B where B is a parameter bound for the elliptic curve method for factoring ECM. Then, one can use ECM to factor $(p-s)(q-r)$ and to find all divisors. As $p-s$ is one of these factors and assuming $s < N^{\frac{1}{4}}$, then one can apply Coppersmith's Theorem 2.5.3 to find p and then q.

Theorem 3.2.2 ([121], Theorem 3.2)**.** *Let $N = pq$ be an RSA modulus and $e = N^\beta$ be a public exponent. Suppose that e satisfies the equation $eu-(p-s)(q-r)v = w$ with $|r|, |s| < N^\alpha < N^{\frac{1}{4}}$, $u < N^\delta$ and $|w| < N^\gamma$. Let B be an ECM-efficiency bound for the Elliptic Curve Method. If $(p-s)$ or $(q-r)$ is B-smooth and*

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)},$$

then one can find p and q in polynomial time.

We can now apply the former result to Demytko's scheme. In this scheme, the RSA modulus is $N = pq$ and the elliptic curve $E_N(a, b)$ is such that $\#E_p(a, b) = p + 1 - t_p$ and $\#E_q(a, b) = q + 1 - t_q$ where, according to Hasse Theorem, $|t_p| < 2\sqrt{p}$ and $|t_q| < 2\sqrt{q}$. The public exponent e and the private exponent d satisfy one of the four equations

$$eu - (p + 1 \pm t_p)(q + 1 \pm t_q)v = w.$$

These equations can be transformed into one of the form $eu-(p-s)(q-r)v = w$ where $s = \mp t_p - 1$ and $t = \mp t_q - 1$. Consequently, when t_p and t_q are suitably small and satisfy some specific conditions, then one can solve the equation, find p and q and then break the system.

3.3 Cryptanalysis of Some RSA Type cryptosystems

In this section, we describe the attack presented in [25] on the Demytko cryptosystem.

RSA is the first and widely most used public key cryptosystem. In order to improve the implementation of the RSA cryptosystem, many schemes have been presented giving rise to some RSA type cryptosystems. A way to extend RSA is to consider the modulus $N = pq$ and the exponent e with specific arithmetical operations such as elliptic curves [79, 81] and Gaussian domains [43].

In 1993, Smith and Lennon presented a cryptosystem called LUC [143]. LUC is based on Lucas sequences. The public exponent e and the private exponent d are integers such that $ed \equiv 1 \pmod{\left(p^2 - 1\right)\left(q^2 - 1\right)}$ which can be transformed into the equation $ed - k\left(p^2 - 1\right)\left(q^2 - 1\right) = 1$.

In 1995, Kuwakado, Koyama and Tsuruoka [81] presented a scheme with an RSA modulus $N = pq$ and a singular cubic curve with equation $y^2 = x^3 + bx^2 \mod N$. The addition of two points in the singular cubic curve is similar than the addition for elliptic curves. In some cases, the addition is not possible if the inversion modulo N is not possible. In this situation, one can certainly find a factor of N. When N is an RSA modulus with

large prime factors, this will happen very rarely. Note that the elliptic curve method factoring method of Lenstra [84] is based on this idea. In the scheme of Kuwakado-Koyama-Tsuruoka, a message is then transformed into a point $M = (m_x, m_y)$ on the singular cubic equation. In the scheme, the public exponent e and the private exponent d satisfy an equation of the form $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$.

In 2002, Elkamchouchi, Elshenawy and Shaban [43] presented an extension of RSA to the Gaussian domain. The modulus is of the form $N = PQ$ where P and Q are two Gaussian primes. The public exponent e and the private exponent d are two positif integers satisfying $ed \equiv 1 \pmod{(|P|-1)(|Q|-1)}$. When $P = p$ and $Q = q$ are integer prime numbers, the modular equations can be transformed into the equation $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$.

In 2007 Castagnos [29] presented a scheme with an RSA modulus $N = pq$ and a public exponent e such that $\gcd\left(e, \left(p^2-1\right)\left(q^2-1\right)\right) = 1$. This condition is equivalent to the equation $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$ where d and k are positive integers.

As we can see, the moduli and the public exponents of the former schemes satisfy the same equation, namely $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$. In the rest of this section, we show how to apply the continued fraction algorithm and Coppersmith's method to solve the generalized equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$. We consider that the prime factors in $N = pq$ are of the same bit size and ordered such that $q < p < 2q$. Then, we can find some useful results on the size of the primes.

Lemma 3.3.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then*

$$2\sqrt{N} < p + q < \frac{3\sqrt{2}}{2}\sqrt{N} \quad \textit{and} \quad 2N < p^2 + q^2 < \frac{5}{2}N.$$

The proof of the lemma starts by applying the increasing function $f(x) = x + \frac{1}{x}$ on $1 < \frac{p}{q} < 2$. This leads to $2N < p^2 + q^2 < \frac{5}{2}N$. Similarly, applying the function f on $1 < \sqrt{\frac{p}{q}} < \sqrt{2}$, we get $2\sqrt{N} < p + q < \frac{3\sqrt{2}}{2}\sqrt{N}$.

To solve the equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$, we need some extra conditions as in the following result.

Theorem 3.3.2 ([25], Theorem 3)**.** *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let e be a public exponent satisfying an equation $ex - \left(p^2 - 1\right)\left(q^2 - 1\right) y = z$ with coprime positive integers x and y. If*

$$xy < 2N - 4\sqrt{2}N^{\frac{3}{4}} \quad \textit{and} \quad |z| < (p - q)N^{\frac{1}{4}}y,$$

then one can find p and q in polynomial time in $\log(N)$.

The proof transforms the equation $ex - \left(p^2 - 1\right)\left(q^2 - 1\right) y = z$ into

$$ex - \left(N^2 + 1 - \frac{9}{4}N\right) y = z - \left(p^2 + q^2 - \frac{9}{4}N\right) y.$$

Using various assumptions and proved results, we find the inequality

$$\left|\frac{e}{N^2 + 1 - \frac{9}{4}N} - \frac{y}{x}\right| < \frac{1}{2x^2}.$$

It follows that $\frac{y}{x}$ is a convergent of the continued fraction expansion of $\frac{e}{N^2+1-\frac{9}{4}N}$. Then using x and y, we compute an approximation $\widetilde{p}$ of p with an error term of at most $N^{\frac{1}{4}}$, with

$$\widetilde{p} = \frac{1}{2}\left(\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|} + \sqrt{\left|(N-1)^2 - \frac{ex}{y}\right|}\right).$$

Then applying Coppersmith's Theorem 2.5.3, we find p in a polynomial time which leads to the factorization of N.

We note that the attacks on the equation $ed - k\left(p^2 - 1\right)\left(q^2 - 1\right) = 1$ are very rare and a lot of the attacks on RSA can be adapted to solve the equation.

Chapter 4

Cryptanalysis of NTRU

4.1 Introduction

The NTRU public key cryptosystem was proposed by J. Hoffstein, J. Pipher and J. H. Silverman [63] in 1996. It is one of the fastest known public key cryptosystems. It offers both encryption (NTRUencrypt) and digital signature (NTRUSign). In comparison with RSA [131] and ECC [75, 95], NTRU is faster and has smaller keys. Moreover, NTRU is still resistant to quantum attacks because its security is conjectured to rely on the hardness of certain lattice problems. Since its presentation, NTRU has been scrutinized for weaknesses and was standardized in 2009 by IEEE Std 1363.1-2008 [70] and in 2010 by ANSI X9.98-2010 [2].

In 1998, Coppersmith and Shamir [35] presented a lattice based attack on NTRU. This attack exploits the structure of the public key in NTRU. Using the key equation, the method builds a lattice and applies lattice reduction techniques to find short vectors in the lattice that could exhibit the private key. However, the method is not practicable in large dimensions. The updated parameters in [64] make NTRU secure against lattice attacks.

Other types of attacks on NTRU have been considered in the last decade. In 2007, Howgrave-Graham [67] presented a hybrid attack of lattice reduction and meet-in-the-middle attack on NTRU. In 2015, Kirchner and Fouque [73, 74] presented a heuristic subexponential-time algorithm on NTRU and in

2016, independently, Cheon, Jeong and Lee [31] and Albrecht, Bai and Ducas [4] described similar attacks on NTRU. The two attacks are based on the fact that for any cyclotomic number field, there exists a subfield that allows to reduce the dimension of the lattice. To avoid the attacks that exploit special structures of the rings used in NTRU, Bernstein et al. [11] presented recently a variant of NTRU, called NTRU Prime where the underlying ring is in the form $\mathbb{Z}_q[X]/(X^p - X - 1)$ where $\mathbb{Z}_q = \mathbb{Z}/qZ$ and p is a prime number.

In this chapter, we show that the attack of Coppersmith and Shamir can be improved for NTRU with two public keys if the private keys share an amount of the coefficients.

4.2 Description of NTRU

The main objects in NTRU are polynomials from the ring of truncated polynomials $\mathcal{R} = \mathbb{Z}_q[X]/\left(X^N - 1\right)$ where $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$. NTRU uses four public parameter sets, $\mathcal{L}_f$, $\mathcal{L}_g$, $\mathcal{L}_m$ and $\mathcal{L}_r$. These sets and the other parameters can be categorized as follows where, for a positive integers d, $\mathcal{B}(d)$) is the set of polynomials of $\mathcal{R}$ with d coefficients equal to 1 and all the other coefficients equal to 0:

- N is a public prime and is sufficiently large.
- p is a small public modulus, typically $p = 3$.
- q is a large public modulus with $\gcd(p, q) = 1$.
- $\mathcal{L}_f = \mathcal{B}(d_f)$ is a set of small polynomials from which the private keys are selected.
- $\mathcal{L}_g = \mathcal{B}(d_g)$ is a similar set of small polynomials from which other private keys are selected.
- $\mathcal{L}_m = \mathbb{Z}_p[X]/\left(X^N - 1\right)$ is the plaintext space. It is a set of polynomials $m \in \mathbb{Z}_p[X]/(X^N - 1)$ that represent encryptable messages.
- $\mathcal{L}_r = \mathcal{B}(d_r)$ is a set of polynomials from which the blinding value used during encryption is selected.

The key generation, encryption and decryption primitives are as follows:

1. **Key generation**
   - Randomly choose a polynomial $f \in \mathcal{L}_f$ such that f is invertible in $\mathcal{R}$ modulo p and modulo q.
   - Compute $f_p \equiv f^{-1} \pmod{p}$ and $f_q \equiv f^{-1} \pmod{q}$.
   - Randomly choose a polynomial $g \in \mathcal{L}_g$.
   - Compute $h \equiv g * f_q \pmod{q}$.
   - Publish the public key (N, h) and the set of parameters p, q, $\mathcal{L}_f$, $\mathcal{L}_g$, $\mathcal{L}_r$ and $\mathcal{L}_m$.
   - Keep the private key (f, f_p).
2. **Encryption**
   - Represent the message as a polynomial $m \in \mathcal{L}_m$.
   - Randomly choose a polynomial $r \in \mathcal{L}_r$.
   - Encrypt m with the public key (N, h) using the rule $e \equiv p*r*h+m \pmod{q}$.
3. **Decryption**
   - The receiver computes $a \equiv f * e \pmod{q}$.
   - Using a centering procedure, transform a to a polynomial with coefficients in the interval $\left[-\frac{q}{2}, \frac{q}{2}\right[$.
   - Compute $m \equiv f_p * a \pmod{p}$.

If one of the coefficients of the polynomial $p * r * g + f * m$ do not lie in the interval $\left[-\frac{q}{2}, \frac{q}{2}\right[$, then the original message m can not be recovered. This situation is called decryption failure and when the parameters are suitably chosen, the decryption is always correct [141].

4.3 The attack of Coppersmith and Shamir on NTRU

The recovery of the NTRU private key f from public key $h \equiv g * f_q \pmod{q}$ can be transformed into as a lattice problem as was presented by Coppersmith

and Shamir [35]. Let λ be a positive value to be optimized later. The equation $h \equiv g * f_q \pmod{q}$ is equivalent to $h * f - q * u = g$ where $u \in \mathcal{R}$ and can be rewritten as

$$\begin{bmatrix} \lambda & 0 \\ h & q \end{bmatrix} \begin{bmatrix} f \\ -u \end{bmatrix} = \begin{bmatrix} \lambda f \\ g \end{bmatrix}.$$

Using the coordinates of f and g

$$\begin{aligned} f &= (f_0, f_1, \ldots, f_{N-1}), \quad g = (g_0, g_1, \ldots, g_{N-1}), \\ h &= (h_0, h_1, \ldots, h_{N-1}), \quad u = (u_0, u_1, \ldots, u_{N-1}), \end{aligned}$$

we get

$$\left[\begin{array}{cccc||cccc} \lambda & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 0 & 0 & \cdots & 0 \\ \hline\hline h_0 & h_{N-1} & \cdots & h_1 & q & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & h_2 & 0 & q & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & \cdots & h_0 & 0 & 0 & \cdots & q \end{array}\right] * \left[\begin{array}{c} f_0 \\ f_1 \\ \vdots \\ f_{N-1} \\ \hline\hline -u_0 \\ -u_1 \\ \vdots \\ -u_{N-1} \end{array}\right] = \left[\begin{array}{c} \lambda f_0 \\ \lambda f_1 \\ \vdots \\ \lambda f_{N-1} \\ \hline\hline g_0 \\ g_1 \\ \vdots \\ g_{N-1} \end{array}\right].$$

Consider the lattice $\mathcal{L}$ spanned by the rows of the underlying matrix. Then $(\lambda f, g) \in \mathcal{L}$ and $\mathcal{L} \subset \mathbb{Z}^{2N}$ is a lattice with dimension $2N$ and determinant $\det(\mathcal{L}) = \lambda^N q^N$. The Euclidean norm of the vector $(\lambda f, g)$ is $\sqrt{\lambda^2 d_f + d_g}$. The Gaussian heuristic asserts that $(\lambda f, g)$ is the shortest vectors of the lattice overwhelming probability and so lattice reduction might be used to find it. More precisely, the Gaussian heuristic says that the length of the shortest non-zero vector is usually approximately $\sigma(\mathcal{L})$ where

$$\begin{aligned} \sigma(\mathcal{L}) &= \sqrt{\frac{\dim(\mathcal{L})}{2\pi e}} \left(\det \mathcal{L}\right)^{1/\dim(\mathcal{L})} \\ &= \sqrt{\frac{2N}{2\pi e}} (\lambda q)^{\frac{N}{2N}} \\ &= \sqrt{\frac{\lambda q N}{\pi e}}. \end{aligned}$$

We need $\sigma(\mathcal{L}) \approx \|(\lambda f, g)\|$. This can be achieved by taking $\lambda = \|g\|/\|f\|$ which leads to a ratio

$$c = \frac{\|(\lambda f, g)\|}{\sigma(L)} = \sqrt{\frac{2\pi e \|g\| \|f\|}{qN}}.$$

When the ratio c is small, then lattice reduction is expected to find the vector $(\lambda f, g)$ and then the private key f.

4.4 An attack of NTRU with two public keys: Case 1

In this section, we describe the attack presented in [118] on the Demytko cryptosystem.

Consider NTRU with two public keys h and h' defined by the same parameters (N, p, q) and

$$\begin{aligned} h &= f_q^{-1} * g \pmod q, \\ h' &= F_q'^{-1} * G' \pmod q. \end{aligned}$$

We can rewrite h' using f_q^{-1} and f as

$$h' = f_q^{-1} \left(f * F_q'^{-1} * G' \right) = f_q^{-1} * g' \pmod q,$$

where $g' = f * F_q'^{-1} * G' \pmod q$. This means that all the public keys in NTRU can be expressed with the same polynomials $f \in \mathcal{R}$ and f_q^{-1}. Combining h and h', we get

$$(h - h') * f = g - g' \pmod q.$$

Consider the lattice $\mathcal{L}$ spanned by the rows of the matrix

$$M(h, h') = \begin{bmatrix} \lambda I_N & H - H' \\ 0 & qI_N \end{bmatrix},$$

where $H - H'$ is the circulant matrix

$$\begin{bmatrix} h_0 - h'_0 & h_1 - h'_1 & \cdots & h_{N-1} - h'_{N-1} \\ h_{N-1} - h'_{N-1} & h_0 - h'_0 & \cdots & h_{N-2} - h'_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ h_1 - h'_1 & h_2 - h'_2 & \cdots & h_0 - h'_0 \end{bmatrix}.$$

Then one can observe that $(\lambda f, g - g')$ is in the lattice $\mathcal{L}$. Using the Gaussian heuristic, one can expect to find $(\lambda f, g - g')$ as the shortest vector of the lattice. The Gaussian heuristic says that the length of the shortest non-zero vector is approximately

$$\sigma(L) = \sqrt{\frac{\dim(L)}{2\pi e}} \det L^{1/\dim(L)} = \sqrt{\frac{\lambda N q}{\pi e}}.$$

On the other hand, the Euclidean norm of $(\lambda f, g - g')$ is $\|(\lambda f, g - g')\| = \sqrt{\lambda^2 \|f\|^2 + \|g - g'\|^2}$ and the ratio

$$c_1 = \frac{\|(\lambda f, g - g')\|}{\sigma(\mathcal{L}(h, h'))},$$

is minimized for

$$\lambda = \frac{\|g - g'\|}{\|f\|},$$

which gives

$$c_1 = \sqrt{\frac{2\pi e \|g - g'\| \|f\|}{qN}}.$$

Recall that in the attack of Coppersmith and Shamir, the ratio is

$$c = \sqrt{\frac{2\pi e \|g\| \|f\|}{qN}}.$$

The new attack is more efficient when $c_1 < c$ which is satisfied when

$$\|g - g'\| < \min(\|g\|, \|g'\|).$$

This means that, whenever g and g' share many coefficients, then the new attack is more efficient to find the private key f than the attack of Coppersmith and Shamir.

4.5 An attack of NTRU with two public keys: Case 2

In this section, we describe the second attack on NTRU with two public keys as presented in [118].

Let h and h' be two public keys with the same parameters

$$\begin{aligned} h &= f_q^{-1} * g \pmod q, \\ h' &= F_q'^{-1} * G' \pmod q. \end{aligned}$$

We suppose that g and G' are invertible modulo q. We set $h_q = h^{-1} \pmod q$ and $h'_q = h'^{-1} \pmod q$. Then

$$\begin{aligned} h_q &= g^{-1} * f \pmod q, \\ h'_q &= G'^{-1} * F' = g^{-1} * \left(g * G'^{-1} * F'\right) = g^{-1} * f' \pmod q. \end{aligned}$$

where $f' = g * G'^{-1} * F'$. Then, we get

$$g * h_q = f \pmod q, \qquad g * h'_q = f' \pmod q.$$

Let

$$h_q(X) = \sum_{i=0}^{N-1} h_{q,i} X^i, \qquad h'_q(X) = \sum_{i=0}^{N-1} h'_{q,i} X^i.$$

For a positive constant λ, we define the $2N$-dimension lattice $\mathcal{L}'$

$$\mathcal{L}' = \left\{(\lambda v, w) \in \mathcal{R}^2 \ : \ w = v * \left(h_q - h'_q\right) \pmod q\right\}.$$

We can see that the lattice is generated by the row vectors of the matrix M' with

$$M' = \begin{bmatrix} \lambda I_N & H_q - H'_q \\ 0 & qI_N \end{bmatrix},$$

where $H_q - H'_q$ is the circulant matrix

$$\begin{bmatrix} h_{q,0} - h'_{q,0} & \cdots & h_{q,N-1} - h'_{q,N-1} \\ h_{q,N-1} - h'_{q,N-1} & \cdots & h_{q,N-2} - h'_{q,N-2} \\ \vdots & \ddots & \vdots \\ h_{q,1} - h'_{q,1} & \cdots & h_{q,0} - h'_{q,0} \end{bmatrix}.$$

Assume that $g * h_q = f + qv$ and $g * h'_q = f' + qv'$. Then $(g, -v + v') * M' = (\lambda g, f - f')$, and $(\lambda g, f - f')$ is a vector of $\mathcal{L}'$. Hence, by reducing the lattice, one can find $(\lambda g, f - f')$ among the shortest vectors if a certain condition is

satisfied. Indeed, by the Gaussian heuristic, the shortest non-zero vector is approximately

$$\sigma(\mathcal{L}') = \sqrt{\frac{\dim(\mathcal{L}')}{2\pi e}} \det \mathcal{L}'^{1/\dim(\mathcal{L}')} = \sqrt{\frac{\lambda N q}{\pi e}}.$$

To compare $\sigma(\mathcal{L}')$ and the length $\|(\lambda g, f - f')\| = \sqrt{\lambda^2 \|g\|^2 + \|f - f'\|^2}$ of the target vector $(\lambda g, f - f')$, we introduce the ratio

$$c_2 = \frac{\|(\lambda g, f - f')\|}{\sigma(\mathcal{L}')}.$$

To increase the chances of lattice reductions to find the vector $(\lambda g, f - f')$, the ratio c_2 should be as small as possible. This is achieved by taking

$$\lambda = \frac{\|f - f'\|}{\|g\|},$$

which leads to

$$c_2 = \sqrt{\frac{2\pi e \|f - f'\| \|g\|}{qN}}.$$

In comparison, in the attack of Coppersmith and Shamir, the ratio is

$$c = \sqrt{\frac{2\pi e \|g\| \|f\|}{qN}}.$$

Then, when $c_2 < c$, the new method is be more efficient. A sufficient condition for this is about the shortness of the length of the difference $f - f'$:

$$\|f - f'\| < \min\left(\|f\|, \|f'\|\right).$$

Hence, if f and f' share a certain amount of their coefficients, the new method will find the private keys f and f' more efficiently than the attack of Coppersmith and Shamir.

Chapter 5

Cryptanalysis of the DGHV Cryptosystem

5.1 Introduction

The purpose of homomorphic encryption is to allow computation on encrypted data without decrypting it. The idea of computing on encrypted data was first proposed by Rivest, Adleman and Dertouzos [130] in 1978. In 1991, Feigenbaum and Merritt [46] posed the problem if it is possible to design an encryption function E such that both $E(x+y)$ and $E(x.y)$ are easy to compute from $E(x)$ and $E(y)$. In 2009, Gentry [49] proposed a positive answer and theoretically demonstrated the possibility of construction such an encryption function. In the last few years, homomorphic encryption schemes have become of great interest in many different cryptographic protocols such as Medical Applications [9], cloud computing, multiparty computation, election and voting protocols (see [6, 98] for more applications). For this reason, homomorphic encryption schemes have been studied extensively to improve implementations and applications. Many classical systems are partially homomorphic. For example, RSA [131] is multiplicatively homomorphic while the Goldwasser-Micali scheme [52] and ElGamal [42] are additively homomorphic. Boneh, Goh and Nissim [19] were the first to construct a scheme capable of performing an arbitrary number of additions but one multiplication. In 2009, Gentry [49] constructed a fully homomorphic encryption

scheme (FHE) with the possibility of evaluating an arbitrary number of additions and multiplications. The security of Gentry's proposal relies on hard assumptions in lattices such as the Sparse Subset Sum Problem (SSSP).

In 2010, van Dijk, Gentry, Halevi and Vaikuntanathan [41] presented an efficient and simple scheme, called DGHV. In DGHV, the private key is an odd integer p. A bit m is then encrypted as $c = pq + 2r + m$ where q and r are secret random integers. To decrypt from c, one just computes $m = (c \pmod p) \pmod 2$. The security of DGHV relies on the hardness of SSSP as well as that of the Approximate Greatest Common Divisor problem (AGCD) as introduced by Howgrave-Graham [66]. The AGCD problem is to recover a prime number p when many approximate multiples $q_i p + r_i$ of p are given with small r_i.

Many attacks on DGHV have been proposed [30,36,41,87] and served to fix the parameters. In this chapter, we describe two new attacks. The first attack concerns the instance where $c_1 = pq_1$ and $c_i = pq_i + r_i$ with $r_i \neq 0$. We apply Coppersmith's method to solve the equation $a_2c_2 + \ldots + a_mc_m = a_1q_1$ under suitable conditions. The second attack works for the general instance $c_i = pq_i + r_i$ for $i = 1, \ldots, m$. We study the linear equation $a_1q_1 + \ldots + a_mq_m = 0$ and apply lattice reduction techniques to solve it under some conditions. In both attacks, we retrieve the secret parameters p and q_i.

5.2 Description of the Parameters in DGHV

DGHV is described in [41] as a somewhat homomorphic scheme with a single private key p, which is a prime number. To encrypt $m \in \{0, 1\}$, one compute $c = pq + 2r + m$ where q is a large random integer and r is a small random integer. To decrypt c, one computes ($c \bmod p$) $\bmod 2 = m$. If $c_1 = pq_1 + 2r_1 + m_1$ and $c_2 = pq_2 + 2r_2 + m_2$ are two ciphertexts, then

$$\begin{aligned} c_1 + c_2 &= p(q_1 + q_2) + 2(r_1 + r_2) + m_1 + m_2, \\ c_1c_2 &= p(pq_1q_2 + 2q_1r_2 + q_1m_2 + 2q_2r_1 \\ &\quad + q_2m_1) + 2(2r_1r_2 + r_1m_2 + m_1r_2) + m_1m_2, \end{aligned}$$

which implies that DGHV is a homomorphic scheme.

To ease the notation, the ciphertext $c = pq + 2r + m$ can be rewritten as $c = pq + r$. In DGHV, the public parameters are the integers $c_i = pq_i + r_i$ where the parameters p, q_i and r_i are as follows.

- For $i = 1, \ldots, m$, c_i is a public integer of bit-length γ.
- p is a private prime number of bit-length η.
- For $i = 1, \ldots, m$, q_i is a private integer of bit-length $\gamma - \eta$.
- For $i = 1, \ldots, m$, r_i is a private random integer with $|r_i| < 2^\rho$.

In [41], it is shown that the scheme is semantically secure under the Approximate-GCD assumption which states the following:

Definition 5.2.1 (Approximate-GCD assumption)**.** Let γ, η, ρ be positive integers. For any η-bit prime number p, given m many positive integers $c_i = pq_i + r_i$ with m many $(\gamma - \eta)$-bit integers q_i and m many integers r_i satisfying $|r_i| < 2^\rho$, it is hard to find p.

There exists a variant of DGHV with $c_1 = pq_1$ where q_1 is a large integer such that it is hard to find any prime factor that divides c_1.

In [36, 41, 87], the security of DGHV has been studied against several attacks. These attacks served to propose optimal parameters for η, ρ, and γ in order to improve the security of DGHV. These attacks can be categorized according to their underling techniques:

- Brute force search [30, 41]: When $c_1 = pq_1$, this technique consists in removing the noise, say r_2 from c_2 by trying all possibilities for $r_2 \in (-2^\rho, 2^\rho)$ and computing $\gcd(c_1, c_2 - r_2)$ which gives p with overwhelming probability.
- Continued fractions [41, 87]: This consists on recovering q_i/q_j from c_i/c_j using continued fractions, which yields immediate calculation of $p = \lfloor c_i/q_i \rceil$.
- Attacks on the Approximate-GCD assumption [41, 87]: The recovery of p through the recovery of r_i or q_i, $i = 1, \ldots, m$, using a combination of

lattice reduction and other techniques. These attacks include Coppersmith's technique [34], the method for solving simultaneous diophantine equations [86] and the orthogonal lattice attacks [41, 87].

In the case $c_1 = pq_1$, the parameters in DGHV are chosen such that direct factorization is not possible [87].

5.3 The First Proposed attack on DGHV

In this section, we briefly describe the attack on DGHV as presented in [122].

In this section, we assume that $c_1 = pq_1$ and $c_i = pq_i + r_i$ for $i = 2, \ldots, m$. Then one can show that there exist infinitely many solutions of the linear equation $a_2c_2 + \ldots + a_mc_m = a_1q_1$, in integers $a_1, \ldots, a_m$. We derive a condition on the size of each $|a_i|$ under which the above equation can be solved leading to the cryptanalysis of the scheme. To solve the linear equation, we use Coppersmith's technique for linear equations with unknown modulus, as presented by Herrmann and May in [59].

Theorem 5.3.1 (Herrmann-May)**.** *Let N be a composite integer of unknown factorization with a divisor $p \geq N^\beta$. Let $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ be a linear polynomial in n variables. One can find in polynomial time all solutions $\left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ of the equation $f(x_1, \ldots, x_n) \equiv 0 \pmod{p}$ with $\left|x_1^{(0)}\right| < N^{\lambda_1}, \ldots, \left|x_n^{(0)}\right| < N^{\lambda_n}$ if*

$$\sum_{i=1}^{n} \lambda_i < 1 - (1-\beta)^{\frac{n+1}{n}} - (n+1)\left(1 - \sqrt[n]{1-\beta}\right)(1-\beta).$$

Since q_1 is an unknown factor of c_1, we apply the former result and get the following result.

Theorem 5.3.2 ([122], Theorem 4.1.)**.** *Let $c_1 = pq_1$ and $c_i = pq_i + r_i$, $i = 2, \ldots, m$, be m positive integers with $2^{\eta-1} < p < 2^\eta$, $2^{\gamma-1} < c_i < 2^\gamma$ and $|r_i| < p$ for $i = 2, \ldots, m$. Let $a_1, \ldots, a_m$ be m integers satisfying $|a_i| < 2^{\alpha_i}$*

for $i = 2, \ldots, m$ *and* $a_2c_2 + \ldots + a_mc_m = a_1q_1$. *Define* $\beta = \frac{\gamma-\eta-1}{\gamma}$. *If*

$$\sum_{i=2}^{m} \alpha_i < \left(1 - (1-\beta)^{\frac{m}{m-1}} - m\left(1 - \sqrt[m-1]{1-\beta}\right)(1-\beta)\right)(\gamma - 1),$$

then, one can find p, $q_1, \ldots, q_m$, $r_2, \ldots, r_m$ *in polynomial time.*

The proof of this result is a straightforward application of Herrmann-May's Theorem with the linear modular equation $a_2c_2 + \ldots + a_mc_m \equiv 0 \pmod{q_1}$ with the specific bounds of the unknown parameters. Once a solution $(a_2, \ldots, a_m)$ is found, we compute

$$q_1 = \gcd(c_1, a_2c_2 + \ldots + a_mc_m), \quad p = \frac{c_1}{q_1}.$$

Then for $i = 2, \ldots, m$, we get $r_i \equiv c_i \pmod{p}$ and $q_i = \frac{c_i - r_i}{p}$. In [122], a comparison with the former attacks shows that the new attack is significantly more efficient.

5.4 The Second Proposed attack on DGHV

Next, we describe the second attack on DGHV as presented in [122].

In this section, we consider the situation where the DGHV public values are of the general form $c_i = pq_i + r_i$, $i = 1, \ldots, m$. Since there are infinitely many linear integer relations between the q_i of the form $a_1q_1 + \ldots + a_mq_m = 0$, we show that one can find such unknown integers under some conditions.

Theorem 5.4.1 ([122], Theorem 5.1.)**.** *Let* $c_i = pq_i + r_i$, $i = 1, \ldots, m$, *be* m *positive integers with* $c_1 < \ldots < c_m$ *and* $|r_i| < 2^{\rho}$ *for* $i = 1, \ldots, m$. *Let* $a_1, \ldots, a_m$ *be* m *integers satisfying* $|a_i| < 2^{\alpha}$ *for* $i = 1, \ldots, m$ *and* $a_1q_1 + \ldots + a_mq_m = 0$. *If*

$$\alpha < \frac{1}{m}\log_2(c_m) + \log_2\left(\frac{\sqrt{m}}{m+1}\right) - \rho,$$

then, one can find p, $q_1, \ldots, q_m$, $r_1, \ldots, r_m$ *in polynomial time.*

The proof of this theorem is based on the LLL algorithm [86] for lattice reduction. Using $c_i = pq_i + r_i$ and the relation $a_1q_1 + \ldots + a_mq_m = 0$, we get $a_1c_1 + \ldots + a_mc_m = a_1r_1 + \ldots + a_mr_m$. The idea is that, if $a_1r_1 + \ldots + a_mr_m$ is close to 0, then so is $a_1c_1 + \ldots + a_mc_m$. Lattice reduction are then used to solve this equation to find $a_1, \ldots, a_m$. Then using these values, we need $a_1r_1 + \ldots + a_mr_m$ to be close to 0. Again, a second lattice reduction will find the values $r_1, \ldots, r_m$. Hence, $p = \gcd(c_1 - r_1, c_2 - r_2)$ and for $i = 1, \ldots, m$, we get $q_i = \frac{c_i - r_i}{p}$.

Appendices

Appendix A:

Another Generalization of Wiener's Attack on RSA

AFRICACRYPT 2008

This paper presents an attack on the RSA cryptosystem using a generalization of the key equation. The variant equation is in the form $eX - (p - u)(q - v)Y = 1$. For suitably small parameters, the continued fraction algorithm and Coppersmith's technique are used to solve the equation and factor the RSA modulus $N = pq$.

Appendix B:

Cryptanalysis of RSA Using the Ratio of the Primes

AFRICACRYPT 2009

In this paper, we consider the RSA cryptosystem with a modulus $N = pq$ and a public exponent e satisfying an equation of the form $eX-(N-(ap+bq))Y = Z$ with suitably small integers X, Y, Z, where $\frac{a}{b}$ is an unknown convergent of the continued fraction expansion of $\frac{q}{p}$. We combine the continued fraction algorithm, Coppersmith's technique and the elliptic curve method for factorization (ECM) to solve the equation and factor the modulus.

Appendix C:

A New Attack on RSA with Two or Three Decryption Exponents

Journal of Applied Mathematics and Computing 2013

In this paper, we consider two or three instances of RSA with the same modulus $N = pq$ and two or three different public exponents e_i satisfying equations of the form $e_i x_i - (p-1)(q-1)y_i = z_i$. We show how, under some specific conditions, one can solve the system of equations and then factor the RSA modulus.

Appendix D:

An Attack on RSA Using LSBs of Multiples of the Prime Factors

`AFRICACRYPT 2013`

One of the recommendations of the ANSI Standard X9.31-1998 is to choose an RSA modulus $N = pq$ such that the ratio of the prime factors is not too close to a rational fraction with small integers. In this paper, we consider an instance of RSA where the public exponent satisfies an equation of the form $ed - k(N + 1 - ap - bq) = 1$ where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$. We show that such instance is insecure when the integers d, a and b are suitably small.

Appendix E:

Implicit Factorization of Unbalanced RSA Moduli

with Muhammad Rezal Kamel Ariffin

`Journal of Applied Mathematics and Computing 2015`

In this paper, we consider $k \geq 2$ RSA muduli $N_i = p_i q_i$ such that some unknown multiples $a_i p_i$ share an amount of least or most significant bits. We show that using continued fraction and lattice reduction techniques, one can factor the moduli under some suitable conditions. This paper generalizes many results on the problem of implicit factorization.

Appendix F:

Factoring RSA Moduli with Weak Prime Factors

with Tajjeeddine Rachidi
C2SI 2015
The paper presents an attack to factor a RSA modulus $N = pq$ with k constants M_i if the factor p satisfies a linear equation $u_0+M_1u_1+\ldots+M_ku_k = ap$ where the unknown parameters u_i and a are small for given constants M_i. The method is based on Coppersmith's technique and lattice reduction.

Appendix G:
New attacks on RSA with Moduli $N = p^rq$
with Tajjeeddine Rachidi
C2SI 2015
In this paper, we consider the prime power RSA variant of RSA with modulus $N = p^rq$. We study the generalized equation $ex - \phi(N)y = z$ where $\phi(N) = p^{r-1}(p-1)(q-1)$ and show that Coppersmith's method can be applied to solve it under suitable conditions on the unknown parameters. The attack of this paper generalizes the former attacks on the prime power RSA with the key equation $es - \phi(N)k = 1$.

Appendix H:
A New Attack on the KMOV Cryptosystem
Bulletin of the Korean Mathematical Society 2014
KMOV is a cryptosystem introduced by Koyama, Maurer, Okamoto and Vanstone in 1991. It is based on an elliptic curve with equation $y^2 = x^3 + b \pmod n$ where $n = pq$ is an RSA modulus. The public key is an integer e satisfying an equation $ed - k(p+1)(q+1) = 1$. In this paper, we show how to attack KMOV by solving the general equation $ex-(p+1)(q+1)y = z$ by applying lattice reduction techniques.

Appendix I:

A Generalized Attack on RSA Type Cryptosystems
with Martin Bunder, Willy Susilo and Joseph Tonien
`Theoretical Computer Science 2016`
In this paper, we consider three RSA variants of the RSA cryptosystem, one based on singular cubic curves, one based in Lucas sequences and one based on Gaussian integers. The three schemes use an RSA modulus $N = pq$ and a public key e satisfying an equation $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$. We apply continued fractions and Coppersmith's technique to solve the general equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ with suitably small parameters. Any solution of the equation will be used to break the scheme.

Appendix J:
Cryptanalysis of NTRU with two Public Keys
`International Journal of Network Security 2014`
We consider an instance of NTRU with two public keys. We show that, if the private keys share an amount of coefficients, then one can apply lattice reduction to break the system. The attack on NTRU with two public keys extends the former attack of Coppersmith and Shamir when only one public key is available.

Appendix K:
Dirichlet Product for Boolean Functions
`Journal of Applied Mathematics and Computing 2016`
Boolean functions play a central role in symmetric cryptography. For security reasons, boolean functions are required to satisfy a certain number of cryptographic properties. In this paper, we introduce a new notion, called Dirichlet product and study various properties and applications.

Appendix L:

New Attack on RSA and Demytko's Elliptic Curve Cryptosystem
with Emmanuel Fouotsa
Submitted to `Mathematics in Computer Science`
In this paper, we study the elliptic curve scheme of Demytko. This system uses an RSA modulus N and an elliptic curve on the ring $\mathbb{Z}/N\mathbb{Z}$. The public key satisfies a variant equation of the RSA equation. We show that, under some conditions on the unknown parameters of the equation, on can solve the equation and break the system. The attack is based on Coppersmith's method and the elliptic method for factoring ECM.

Appendix M:
Lattice Attacks on the DGHV Homomorphic Encryption Scheme
with Tajjeeddine Rachidi
Submitted to `Discrete Applied Mathematics`
In this paper, we study the somewhat homomorphic encryption scheme of van Dijk, Gentry, Halevi, and Vaikuntanathan. We consider a set of m public ciphers $c_i = pq_i + r_i$, $i = 1, \cdots, m$ where p is a private prime number and q_i, r_i are private integers. We propose two lattice based attacks to retrieve all the parameters when some suitable conditions are fulfilled.

Appendix A

Another Generalization of Wiener's Attack on RSA

AFRICACRYPT 2008
[104]

Abstract :
A well-known attack on RSA with low secret-exponent d was given by Wiener in 1990. Wiener showed that using the equation $ed-(p-1)(q-1)k=1$ and continued fractions, one can efficiently recover the secret-exponent d and factor $N=pq$ from the public key (N,e) as long as $d<\frac{1}{3}N^{\frac{1}{4}}$. In this paper, we present a generalization of Wiener's attack. We show that every public exponent e that satisfies $eX-(p-u)(q-v)Y=1$ with

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

and all prime factors of $p-u$ or $q-v$ are less than 10^{50} yields the factorization of $N=pq$. We show that the number of these exponents is at least $N^{\frac{1}{2}-\varepsilon}$.

A.1 Introduction

The RSA cryptosystem invented by Rivest, Shamir and Adleman [131] in 1978 is today's most important public-key cryptosystem. The security of RSA depends on mainly two primes p, q of the same bit-size and two integers e, d satisfying $ed \equiv 1 \pmod{(p-1)(q-1)}$. Throughout this paper, we label the primes so that $q < p < 2q$. The RSA modulus is given by $N = pq$ and Euler's totient function is $\phi(N) = (p-1)(q-1)$. The integer e is called the public (or encrypting) exponent and d is called the private (or decrypting) exponent.

To reduce the decryption time or the signature-generation time, one may wish to use a short secret exponent d. This was cryptanalysed by Wiener [147] in 1990 who showed that RSA is insecure if $d < \frac{1}{3}N^{0.25}$. Wiener's method is based on continued fractions. These results were extended by Boneh and Durfee [17] in 1999 to $d < N^{0.292}$. The method of Boneh and Durfee is based on Coppersmith's results for finding small solutions of modular polynomial equations [34]. In 2004, Blömer and May [13] presented a generalization of Wiener's attack by combining continued fractions and Coppersmith's method. They showed that RSA is insecure for every (N, e) satisfying $ex + y \equiv 0 \pmod{\phi(N)}$ with $x < \frac{1}{3}N^{1/4}$ and $|y| = O\left(N^{-3/4}ex\right)$.

In this paper, we present another generalization of Wiener's attack. Our method combines continued fractions, integer partial factorization, integer relation detection algorithms and Coppersmith's method. Let us introduce the polynomial

$$\psi(u, v) = (p - u)(q - v).$$

Observe that $\psi(1, 1) = (p-1)(q-1) = \phi(N)$, so ψ could be seen as a generalization of Euler's function. We describe an attack on RSA that works for all public exponents e satisfying

$$eX - \psi(u, v)Y = 1, \tag{A.1}$$

with integers X, Y, u, v such that

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

with the extra condition that all prime factors of $p-u$ or $q-v$ are less than the Elliptic Curve Method of Factoring smoothness bound $B_{\text{ecm}} = 10^{50}$. Here and throughout this paper, we let $[x]$ and $\lfloor x \rfloor$ denote the nearest integer to the real number x and the fractional part of x.

Observe that when $u = 1$, we get $v = -1$ and rewriting (A.1) as

$$eX - (p-1)(q+1)Y = 1,$$

a variant of Wiener's attack enables us to compute p and q without assuming any additional condition on the prime divisors of $p-1$ nor $q+1$.

Our new method works as follows: We use the Continued Fraction Algorithm (see e.g. [57], p. 134) to find the unknowns X and Y among the convergents of $\frac{e}{N}$. Then we use Lenstra's Elliptic Curve Factorization Method (ECM) [84] to partially factor $\frac{eX-1}{Y}$. Afterwards, we use an integer relation detection algorithm (notably LLL [86] or PSLQ [47]) to find the divisors of the B_{ecm}-smooth part of $\frac{eX-1}{Y}$ in a short interval. Finally, we show that a method due to Coppersmith [34] can be applied. Moreover, we show that the number of keys (N, e) for which our method works is at least $N^{\frac{1}{2}-\varepsilon}$.

Organization of the paper. Section 2 presents well known results from number theory that we use. After presenting some useful lemmas in Section 3, and some properties of ψ in Section 4, we present our attack in Section 5 and in Section 6, we show that the number of keys (N, e) for which our method works is lower bounded by $N^{\frac{1}{2}-\varepsilon}$. We briefly conclude the paper in Section 7.

A.2 Preliminaries

A.2.1 Continued fractions and Wiener's attack

The continued fraction expansion of a real number ξ is an expression of the form

$$\xi = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

where $a_0 \in \mathbb{Z}$ and $a_i \in \mathbb{N}-\{0\}$ for $i \geq 1$. The numbers $a_0, a_1, a_2, \cdots$ are called the partial quotients. As usual, we adopt the notation $\xi = [a_0, a_1, a_2, \cdots]$. For $i \geq 0$, the rationals $\frac{r_i}{s_i} = [a_0, a_1, a_2, \cdots, a_i]$ are called the convergents of the continued fraction expansion of ξ. If $\xi = \frac{a}{b}$ is rational with $\gcd(a, b) = 1$, then the continued fraction expansion is finite and the Continued Fraction Algorithm (see [57], p. 134) finds the convergents in time $O((\log b)^2)$. We recall a result on diophantine approximations (see Theorem 184 of [57]).

Theorem A.2.1. *Suppose* $\gcd(a, b) = \gcd(x, y) = 1$ *and*

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

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

Let us recall Wiener's famous attack on RSA with $N = pq$ and $q < p < 2q$. The idea behind Wiener's attack on RSA [147] with small secret exponent d is that for $d < \frac{1}{3}N^{1/4}$, the fraction e/N is an approximation to k/d and hence, using Theorem A.2.1, k/d can be found from the convergents of the continued fraction expansion of e/N. Wiener's attack works as follows. Since $ed - k\phi(N) = 1$ with $\phi(N) = N - (p + q - 1)$ and $p + q - 1 < 3\sqrt{N}$ then $kN - ed = k(p + q - 1) - 1$. Therefore,

$$\left|\frac{k}{d} - \frac{e}{N}\right| = \frac{|k(p + q - 1) - 1|}{Nd} < \frac{3k\sqrt{N}}{Nd}.$$

Now, assume that $d < \frac{1}{3}N^{1/4}$. Since $k\phi(N) = ed - 1 < ed$ and $e < \phi(N)$, then $k < d < \frac{1}{3}N^{1/4}$. Hence

$$\left|\frac{k}{d} - \frac{e}{N}\right| < \frac{N^{3/4}}{Nd} = \frac{1}{dN^{1/4}} < \frac{1}{2d^2}.$$

From Theorem A.2.1, we know that k/d is one of the convergents of the continued fraction expansion of e/N.

A.2.2 Coppersmith's method

The problem of finding small modular roots of a univariate polynomial has been extensively studied by Coppersmith [34], Howgrave-Graham [65], May [91]

and others. Let $f(x)$ be a monic univariate polynomial with integer coefficients of degree δ. Let N be an integer of unknown factorization and $B = N^{1/\delta}$. The problem is to find all integers x_0 such that $|x_0| < B$ and $f(x_0) \equiv 0 \pmod{N}$. In 1997, Coppersmith presented a deterministic algorithm using $\left(2^\delta \log N\right)^{O(1)}$ bit operations to solve this problem. The algorithm uses lattice reduction techniques, and as an application, the following theorem was proved (see also [91], Theorem 11).

Theorem A.2.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Given an approximation $\tilde{p}$ of p with $|p - \tilde{p}| < N^{\frac{1}{4}}$, N can be factored in time polynomial in $\log N$.*

A.2.3 Smooth numbers

A few words about notation: let f and g be functions of x. The notation $f \asymp g$ denotes that $f(x)/g(x)$ is bounded above and below by positive numbers for large values of x. The notation $f = O(g)$ denotes that $\exists c$ such that $f(x) \leq cg(x)$. The notation $f \sim g$ denotes that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

Let y be a positive constant. A positive number n is y-smooth if all prime factors of n are less than y. As usual, we use the notation $\Psi(x, y)$ for the counting function of the y-smooth numbers in the interval $[1, x]$, that is,

$$\Psi(x, y) = \#\left\{n : 1 \leq n \leq x,\ n \text{ is } y\text{-smooth}\right\}.$$

The ratio $\Psi(x, y)/[x]$ may be interpreted as the probability that a randomly chosen number n in the interval $[1, x]$ has all its prime factors less than y. The function $\Psi(x, y)$ plays a central role in the running times of many integer factoring and discrete logarithm algorithms, including the Elliptic Curve Method (ECM) [84] and the number field sieve method (NFS) [85]. Let $\rho(u)$ be the Dickman-de Bruijn function (see [53]). In 1986, Hildebrand [60] showed that

$$\Psi(x, y) = x\rho(u)\left\{1 + O\left(\frac{\log(u+1)}{\log y}\right)\right\} \quad \text{where} \quad x = y^u \tag{A.2}$$

holds uniformly in the range $y > \exp\left\{(\log\log x)^{5/3+\varepsilon}\right\}$. Studying the distribution in short intervals of integers without large prime factors, Friedlander

and Granville [48] showed that

$$\Psi(x+z,y)-\Psi(x,y)\geq c\frac{z}{x}\Psi(x,y), \tag{A.3}$$

in the range $x\geq z\geq x^{\frac{1}{2}+\delta}$, $x\geq y\geq x^{1/\gamma}$ and x is sufficiently large where δ and γ are positive constants and $c=c(\delta,\gamma)>0$.

In order to study the distribution of divisors of a positive integer n, Hall and Tenenbaum [54] studied the counting function

$$U(n,\alpha)=\#\left\{(d,d'):d|n,d'|n,\gcd(d,d')=1,\left|\log\frac{d}{d'}\right|<(\log n)^{\alpha}\right\}, \tag{A.4}$$

where α is a real number. They proved that for any fixed $\alpha<1$ and almost all n,

$$U(n,\alpha)\leq(\log n)^{\log 3-1+\alpha+o(1)}, \tag{A.5}$$

where the $o(1)$ term tends to 0 as n tends to $+\infty$.

A.2.4 ECM

The Elliptic Curve Method (ECM) was originally proposed by Lenstra [84] in 1984 and then extended by Brent [21] and Montgomery [96]. It is suited to find small prime factors of large numbers. The original part of the algorithm proposed by Lenstra is referred to as Phase 1, and the extension by Brent and Montgomery is called Phase 2. ECM relies on Hasse's theorem: if p is a prime factor of a large number M, then an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ has group order $p+1-t$ with $|t|<2\sqrt{p}$, where t depends on the curve. If $p+1-t$ is a smooth number, then ECM will probably succeed and reveal the unknown factor p. ECM is a sub-exponential factoring algorithm, with expected run time of

$$O\left(\exp\left\{\sqrt{(2+o(1))\log p\log\log p}\right\}Mult(M)\right)$$

where the $o(1)$ term tends to 0 as p tends to $+\infty$ and $Mult(M)$ denotes the cost of multiplication mod M. The largest factor known to have been found by ECM is a 67-digit factor of the number $10^{381}+1$, found by B. Dodson with P. Zimmerman's GMP-ECM program in August 2006 (see [150]). According

Table A.1: Running times for factoring $N = pq$ with $q < p < 2q$

n = Number of bits of q	60	70	80	90	100	110	120	130
T = Time in seconds	0.282	0.844	3.266	13.453	57.500	194.578	921.453	3375.719

to Brent's formula [22] $\sqrt{D} = (Y - 1932.3)/9.3$ where D is the number of decimal digits in the largest factor found by ECM up to a given date Y, a 70-digit factor could be found by ECM around 2010.

In Table 1, we give the running times obtained on a Intel(R) Pentium(R) 4 CPU 3.00 GHz to factor an RSA modulus $N = pq$ of size $2n$ bits with $q < p < 2q$ with ECM, using the algebra system Pari-GP [125].

Extrapolating Table 1, we find the formula

$$\log T = 2.609\sqrt{n} - 21.914 \quad \text{or equivalently} \quad T = \exp\left\{2.609\sqrt{n} - 21.914\right\},$$

where T denotes the running time to factor an RSA modulus $N = pq$ with $2n$ bits. Extrapolating, we can extract a prime factor of 50 digits (≈ 166 bits) in 1 day, 9 hours and 31 minutes. Throughout this paper, we then assume that ECM is efficient to extract prime factors up to the bound $B_{\text{ecm}} = 10^{50}$.

A.3 Useful lemmas

In this section we prove three useful lemmas. We begin with a simple lemma fixing the sizes of the prime factors of the RSA modulus.

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

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

Proof. Assume $q < p < 2q$. Multiplying by p, we get $N < p^2 < 2N$ or equivalently $N^{\frac{1}{2}} < p < \sqrt{2}N^{\frac{1}{2}}$. Since $q = \frac{N}{p}$, we obtain $2^{-\frac{1}{2}} N^{\frac{1}{2}} < q < N^{\frac{1}{2}}$ and the lemma follows. □

Our second lemma is a consequence of Theorem A.2.2 and Lemma A.3.1.

Lemma A.3.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Suppose $|u| < N^{\frac{1}{4}}$. If $p - u < N^{\frac{1}{2}}$ or $p - u > \sqrt{2}N^{\frac{1}{2}}$, then the factorization of N can be found in polynomial time.*

Proof. Assume $q < p < 2q$ and $|u| < N^{\frac{1}{4}}$. If $p - u < N^{\frac{1}{2}}$, then $p < N^{\frac{1}{2}} + u < N^{\frac{1}{2}} + N^{\frac{1}{4}}$. Combining this with Lemma A.3.1, we obtain

$$N^{\frac{1}{2}} < p < N^{\frac{1}{2}} + N^{\frac{1}{4}}.$$

It follows that $\tilde{p} = N^{\frac{1}{2}}$ is an approximation of p with $0 < p - \tilde{p} < N^{\frac{1}{4}}$. By Theorem A.2.2, we deduce that the factorization of N can be found in polynomial time.

Similarly, if $p - u > \sqrt{2}N^{\frac{1}{2}}$, then $p > \sqrt{2}N^{\frac{1}{2}} + u > \sqrt{2}N^{\frac{1}{2}} - N^{\frac{1}{4}}$ and using Lemma A.3.1, we get

$$\sqrt{2}N^{\frac{1}{2}} > p > \sqrt{2}N^{\frac{1}{2}} - N^{\frac{1}{4}}.$$

It follows that $\tilde{p} = \sqrt{2}N^{\frac{1}{2}}$ satisfies $0 > p - \tilde{p} > -N^{\frac{1}{4}}$. Again, by Theorem A.2.2, we conclude that the factorization of N can be found in polynomial time. □

Our third lemma is a consequence of the Fermat Factoring Method (see e.g. [146]). We show here that it is an easy consequence of Theorem A.2.2 and Lemma A.3.1.

Lemma A.3.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. If $p - q < N^{\frac{1}{4}}$, then the factorization of N can be found in polynomial time.*

Proof. Assume $q < p < 2q$ and $p - q < N^{\frac{1}{4}}$. Combining with Lemma A.3.1, we get

$$N^{\frac{1}{2}} < p < q + N^{\frac{1}{4}} < N^{\frac{1}{2}} + N^{\frac{1}{4}}.$$

It follows that $\tilde{p} = N^{\frac{1}{2}}$ is an approximation of p with $0 < p - \tilde{p} < N^{\frac{1}{4}}$. By Theorem A.2.2, we conclude that the factorization of N can be found in polynomial time. □

A.4 Properties of $\psi(u, v)$

Let $N = pq$ be an RSA modulus with $q < p < 2q$. The principal object of investigation of this section is the polynomial $\psi(u, v) = (p - u)(q - v)$ when p and q are fixed.

Lemma A.4.1. *Let u be an integer with $|u| < N^{\frac{1}{4}}$. Put $v = \left[-\frac{qu}{p-u}\right]$. Then*

$$|\psi(u, v) - N| < 2^{-\frac{1}{2}} N^{\frac{1}{2}}.$$

Proof. Since v is the nearest integral value to $-\frac{qu}{p-u}$, then

$$-\frac{1}{2} \leq -\frac{qu}{p-u} - v < \frac{1}{2}.$$

Hence

$$q + \frac{qu}{p-u} - \frac{1}{2} \leq q - v < q + \frac{qu}{p-u} + \frac{1}{2}.$$

Multiplying by $p - u$, we get

$$N - \frac{1}{2}(p-u) \leq (p-u)(q-v) < N + \frac{1}{2}(p-u).$$

It follows that

$$|(p-u)(q-v) - N| \leq \frac{1}{2}(p-u).$$

Since $|u| < N^{\frac{1}{4}}$, then by Lemma A.3.2, we can assume $p - u < \sqrt{2} N^{\frac{1}{2}}$ and we obtain

$$|(p-u)(q-v) - N| \leq 2^{-\frac{1}{2}} N^{\frac{1}{2}}.$$

This completes the proof. □

Lemma A.4.2. *Let u be an integer with $|u| < N^{\frac{1}{4}}$. Set $v = \left[-\frac{qu}{p-u}\right]$. Then $|v| \leq |u|$.*

Proof. Assume $q < p < 2q$ and $|u| < N^{\frac{1}{4}}$. By Lemma A.3.3, we can assume that $p - q > N^{\frac{1}{4}}$. Then

$$u < N^{\frac{1}{4}} < p - q,$$

and $q < p - u$. Hence

$$|v| = \left\lfloor \frac{q|u|}{p-u} \right\rfloor \leq \frac{q|u|}{p-u} + \frac{1}{2} < |u| + \frac{1}{2}.$$

Since u and v are integers, then $|v| \leq |u|$ and the lemma follows. □

Lemma A.4.3. *Let u, u', be integers with $|u|, |u'| < N^{\frac{1}{4}}$. Define*

$$v = \left\lfloor -\frac{qu}{p-u} \right\rfloor \quad \textit{and} \quad v' = \left\lfloor -\frac{qu'}{p-u'} \right\rfloor.$$

If $v = v'$, then $|u' - u| \leq 1$.

Proof. Suppose $v' = v$. Then, from the definitions of v and v', we obtain

$$\left| \frac{qu'}{p-u'} - \frac{qu}{p-u} \right| < 1,$$

Transforming this, we get

$$|u' - u| < \frac{(p-u)(p-u')}{N}.$$

By Lemma A.3.3 we can assume that $p - u < \sqrt{2}N^{\frac{1}{2}}$ and $p - u' < \sqrt{2}N^{\frac{1}{2}}$. Then

$$|u' - u| < \frac{\left(\sqrt{2}N^{\frac{1}{2}}\right)^2}{N} = 2.$$

Since u and u' are integers, the lemma follows. □

Lemma A.4.4. *Let u, u', be integers with $|u|, |u'| < N^{\frac{1}{4}}$. Define*

$$v = \left\lfloor -\frac{qu}{p-u} \right\rfloor \quad \textit{and} \quad v' = \left\lfloor -\frac{qu'}{p-u'} \right\rfloor.$$

If $\psi(u, v) = \psi(u', v')$, then $u = u'$.

Proof. Assume that $\psi(u, v) = \psi(u', v')$, that is $(p-u)(q-v) = (p-u')(q-v')$. If $v = v'$, then $p - u = p - u'$ and $u = u'$. Next, assume for contradiction that

$v \neq v'$. Without loss of generality, assume that $u > u'$. Put $\psi = \psi(u,v) = \psi(u',v')$ and let $U(\psi,\alpha)$ as defined by (A.4), i.e.

$$U(\psi,\alpha) = \#\left\{(d,d') : d|\psi, d'|\psi, \gcd(d,d') = 1, \left|\log\frac{d}{d'}\right| < (\log\psi)^{\alpha}\right\}.$$

Let $g = \gcd(p-u, p-u')$, $d = \frac{p-u}{g}$ and $d' = \frac{p-u'}{g}$. Hence $\gcd(d,d') = 1$. We have

$$\frac{d}{d'} = \frac{p-u}{p-u'} = 1 - \frac{u-u'}{p-u'}.$$

By Lemma A.3.2, we can assume that $p - u > N^{\frac{1}{4}}$. For $N > 2^8$ we have

$$0 < \frac{u-u'}{p-u'} < \frac{2N^{\frac{1}{4}}}{N^{\frac{1}{2}}} = 2N^{-\frac{1}{4}} < \frac{1}{2}.$$

Using that $|\log(1-x)| < 2x$ holds for $0 < x < \frac{1}{2}$ this yields

$$\left|\log\frac{d}{d'}\right| = \left|\log\left(1 - \frac{u-u'}{p-u'}\right)\right| < 2 \times \frac{u-u'}{p-u'} < 2\sqrt{2}N^{-\frac{1}{4}} = (\log\psi)^{\alpha},$$

where

$$\alpha = \frac{\log\left(2\sqrt{2}N^{-\frac{1}{4}}\right)}{\log(\log(\psi))}.$$

It follows that $U(\psi,\alpha) \geq 1$. On the other hand, we have

$$\alpha = \frac{\log\left(2\sqrt{2}N^{-\frac{1}{4}}\right)}{\log(\log(\psi))} \leq \frac{\log\left(2\sqrt{2}N^{-\frac{1}{4}}\right)}{\log\left(\log\left(N - 2^{-\frac{1}{2}}N^{\frac{1}{2}}\right)\right)} < 1 - \log 3,$$

where we used Lemma A.4.1 in the medium step and $N > 2^7$ in the final step. Using the bound (A.5), we have actually

$$U(\psi,\alpha) \leq (\log\psi)^{\log 3 - 1 + \alpha + o(1)} \leq (\log N)^{\delta + o(1)},$$

where $\delta = \log 3 - 1 + \alpha < 0$ and we deduce $U(\psi,\alpha) = 0$, a contradiction. Hence $v = v'$, $u = u'$ and the lemma follows. □

Lemma A.4.5. *Let u, u' be integers with $|u|, |u'| < N^{\frac{1}{4}}$. Define*

$$v = \left[-\frac{qu}{p-u}\right] \quad \textit{and} \quad v' = \left[-\frac{qu'}{p-u'}\right].$$

Assume that $\psi(u,v) < \psi(u',v')$. *Let* $[a_0, a_1, a_2, \cdots]$ *be the continued fraction expansion of* $\frac{\psi(u,v)}{\psi(u',v')}$. *Then* $a_0 = 0$, $a_1 = 1$ *and* $a_2 > 2^{-\frac{1}{2}}N^{\frac{1}{2}} - \frac{1}{2}$.

Proof. Let us apply the continued fraction algorithm (see e.g. of [57], p. 134). Assuming $\psi(u,u) < \psi(u',v')$, we get

$$a_0 = \left\lfloor \frac{\psi(u,v)}{\psi(u',v')} \right\rfloor = 0.$$

Next, we have

$$a_1 = \left\lfloor \frac{1}{\frac{\psi(u,v)}{\psi(u',v')} - a_0} \right\rfloor = \left\lfloor \frac{\psi(u',v')}{\psi(u,v)} \right\rfloor.$$

By Lemma A.4.1, we have

$$0 < \psi(u',v') - \psi(u,v) \leq |\psi(u,v) - N| + |\psi(u',v') - N| < \sqrt{2}N^{\frac{1}{2}}. \quad \text{(A.6)}$$

Combining this with Lemma A.4.1, we get

$$0 < \frac{\psi(u',v')}{\psi(u,v)} - 1 = \frac{\psi(u',v') - \psi(u,v)}{\psi(u,v)} < \frac{\sqrt{2}N^{\frac{1}{2}}}{\psi(u,v)} < \frac{\sqrt{2}N^{\frac{1}{2}}}{N - 2^{-\frac{1}{2}}N^{\frac{1}{2}}} < 1.$$

From this, we deduce $a_1 = 1$. Finally, combining (A.6) and Lemma A.4.1, we get

$$a_2 = \left\lfloor \frac{1}{\frac{\psi(u',v')}{\psi(u,v)} - a_1} \right\rfloor = \left\lfloor \frac{\psi(u,v)}{\psi(u',v') - \psi(u,v)} \right\rfloor > \frac{N - 2^{-\frac{1}{2}}N^{\frac{1}{2}}}{\sqrt{2}N^{\frac{1}{2}}} = 2^{-\frac{1}{2}}N^{\frac{1}{2}} - \frac{1}{2}.$$

This completes the proof. □

A.5 The new attack

In this section we state our new attack. Let $N = pq$ be an RSA modulus with $q < p < 2p$. Let e be a public exponent satisfying an equation $eX - \psi(u,v)Y = 1$ with integers X, Y, u, v such that

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

and with the condition that all prime factors of $p-u$ or $q-v$ are $\leq B_{\text{ecm}} = 10^{50}$. Our goal is to solve this equation. As in Wiener's approach, we use the continued fraction algorithm to recover the unknown values X and Y.

Theorem A.5.1. *Let $N = pq$ be an RSA modulus with $q < p < 2p$. Suppose that the public exponent e satisfies an equation $eX - \psi(u,v)Y = 1$ with*

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

Then $\frac{Y}{X}$ is one of the convergents of the continued fraction expansion of $\frac{e}{N}$.

Proof. Starting with the equation $eX - \psi(u,v)Y = 1$, we get

$$eX - NY = 1 - (N - \psi(u,v))Y.$$

Together with Lemma A.4.1, this implies

$$\begin{aligned} \left| \frac{e}{N} - \frac{Y}{X} \right| &= \frac{|1 - (N - \psi(u,v))Y|}{NX} \\ &\leq \frac{1 + |(N - \psi(u,v))|\, Y}{NX} \\ &\leq \frac{1 + 2^{-\frac{1}{2}} N^{\frac{1}{2}} Y}{NX} \\ &\leq \frac{2 + \sqrt{2} N^{\frac{1}{2}} (X-1)}{2NX}. \end{aligned}$$

Suppose we can upperbound the right-hand side term by $\frac{1}{2X^2}$, that is

$$\frac{2 + \sqrt{2} N^{\frac{1}{2}} (X-1)}{2NX} < \frac{1}{2X^2},$$

then, applying Theorem A.2.1 the claim follows. Rearranging to isolate X, this leaves us with the condition

$$\sqrt{2} N^{\frac{1}{2}} X^2 - \left(\sqrt{2} N^{\frac{1}{2}} - 2 \right) X - N < 0.$$

It is not hard to see that the condition is satisfied if $X < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$. This gives us the theorem. □

Afterwards, we combine ECM, integer relation detection algorithms and Coppersmith's method to factor $N = pq$.

Theorem A.5.2. *Let $N = pq$ be an RSA modulus with $q < p < 2p$. Let B_{ecm} be the ECM-bound. Suppose that the public exponent $e < N$ satisfies an equation $eX - \psi(u, v)Y = 1$ with*

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

If $p - u$ or $q - v$ is B_{ecm}-smooth, then we can efficiently factor N.

Proof. By Theorem A.5.1 we know that X and Y can be found among the convergents of the continued expansion of $\frac{e}{N}$. From X and Y, we get

$$\psi(u, v) = (p - u)(q - v) = \frac{eX - 1}{Y}.$$

Without loss of generality, suppose that $p - u$ is B_{ecm}-smooth. Using ECM, write $\frac{eX-1}{Y} = M_1 M_2$ where M_1 is B_{ecm}-smooth. Let $\omega(M_1)$ denote the number of distinct prime factors of M_1. Then the prime factorization of M_1 is of the form

$$M_1 = \prod_{i=1}^{\omega(M_1)} p_i^{a_i},$$

where the $a_i \geq 1$ are integers and the p_i are distinct primes $\leq B_{\text{ecm}}$. Since $p - u$ is B_{ecm}-smooth, then $p - u$ a divisor of M_1, so that

$$p - u = \prod_{i=1}^{\omega(M_1)} p_i^{x_i}, \tag{A.7}$$

where the x_i are integers satisfying $0 \leq x_i \leq a_1$. By Lemma A.3.2, we can assume that $N^{\frac{1}{2}} < p - u < \sqrt{2} N^{\frac{1}{2}}$. Combining this with (A.7) and taking logarithms, we get

$$0 < \sum_{i=1}^{\omega(M_1)} x_i \log p_i - \frac{1}{2} \log N < \frac{1}{2} \log 2. \tag{A.8}$$

These inequalities are related to Baker's famous theory of linear forms in logarithms [7] and can be formulated as a nearly closest lattice problem in the 1-norm. They can be solved using the LLL [86] or the PSLQ algorithm [47]. The complexity of LLL and PSLQ depends on $\omega(M_1)$. Since Hardy and Ramanujan (see e.g.Theorem 431 of [57]), we know that, in average, $\omega(M_1) \sim \log\log M_1$ if M_1 is uniformly distributed. Since $X < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$, we have for $e < N$

$$M_1 \leq \frac{eX - 1}{Y} < \frac{eX}{Y} \leq eX < N^{\frac{5}{4}},$$

This implies that the number of primes dividing M_1 satisfies

$$\omega(M_1) \sim \log\log M_1 \sim \log\log N.$$

Next, let us investigate the number of solutions of (A.8) which is related to the number of divisors of M_1. Let $\tau(M_1)$ denote the number of positive divisors of M_1. The prime decomposition of M_1 gives the exact value

$$\tau(M_1) = \prod_{i=1}^{\omega(M_1)} (1 + a_i).$$

By Dirichlet's Theorem, we know that if M_1 is uniformly distributed, then the average order of $\tau(M_1)$ is $\log M_1$ (see Theorem 319 of [57]). It follows that the average number of divisors of M_1 is

$$\tau(M_1) \sim \log(M_1) \sim \log(N).$$

This gives in average the number of solutions to the inequalities (A.8).

Next, let D be a divisor of M_1 satisfying (A.8). If D is a good candidate for $p - u$ with $|u| < N^{\frac{1}{4}}$, then applying Theorem A.2.2, we get the desired factor p. This concludes the theorem. □

Notice that the running time is dominated by ECM since every step in our attack can be done in polynomial time and the number of convergents and divisors are bounded by $O(\log N)$.

A.6 The number of exponents for the new method

In this section, we estimate the number of exponents for which our method works. Let $N = pq$ be an RSA modulus with $q < p < 2q$. The principal object of investigation of this section is the set

$$H(N) = \left\{ e : e < N,\ \exists u \in V(p),\ \exists X < 2^{-\frac{1}{4}} N^{\frac{1}{4}},\ e \equiv X^{-1} \pmod{\psi(u,v)} \right\},$$

where

$$V(p) = \left\{ u :\ |u| < p^{\frac{1}{2}},\ p - u \text{ is } B_{\text{ecm}}\text{-smooth} \right\}, \tag{A.9}$$

and $v = \left\lfloor -\frac{qu}{p-u} \right\rfloor$.

We will first show that every public exponent $e \in H(N)$ is uniquely defined by a tuple (u, X). We first deal with the situation when an exponent e is defined by different tuples (u, X) and (u, X').

Lemma A.6.1. *Let $N = pq$ be an RSA modulus with $q < p < 2p$. Let u, v, X, X' be integers with $1 \le X, X' < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$ and $\gcd(XX', \psi(u,v)) = 1$ where $v = \left\lfloor -\frac{qu}{p-u} \right\rfloor$. Define*

$$e \equiv X^{-1} \pmod{\psi(u,v)} \quad \textit{and} \quad e' \equiv X'^{-1} \pmod{\psi(u,v)}.$$

If $e = e'$, then $X = X'$.

Proof. Since $e \equiv X^{-1} \pmod{\psi(u,v)}$, there exists a positive integer Y such that $eX - \psi(u,v)Y = 1$ with $\gcd(X, Y) = 1$. Similarly, e' satisfies $e'X' - \psi(u,v)Y' = 1$ with $\gcd(X', Y') = 1$. Assume that that $e = e'$. Then

$$\frac{1 + \psi(u,v)Y}{X} = \frac{1 + \psi(u,v)Y'}{X'}.$$

Combining this with Lemma A.4.1, we get

$$|XY' - X'Y| = \frac{|X' - X|}{\psi(u,v)} < \frac{2^{-\frac{1}{4}} N^{\frac{1}{4}}}{N - 2^{-\frac{1}{2}} N^{\frac{1}{2}}} < 1.$$

Hence $XY' = X'Y$ and since $\gcd(X, Y) = 1$, we get $X' = X$ and the lemma follows. □

Next, we deal with the situation when an exponent e is defined by different tuples (u, X) and (u', X') with $u \neq u'$ and $v = v'$.

Lemma A.6.2. *Let $N = pq$ be an RSA modulus with $q < p < 2p$. Let u, u' be integers with $|u|, |u'| < N^{\frac{1}{4}}$. Let X, X' be integers with $1 \leq X, X' < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$, $\gcd(X, \psi(u, v)) = 1$, $\gcd(X', \psi(u', v')) = 1$ where $v = \left\lfloor -\frac{qu}{p-u} \right\rceil$ and $v' = \left\lfloor -\frac{qu'}{p-u'} \right\rceil$. Define*

$$e \equiv X^{-1} \pmod{\psi(u, v)} \quad \text{and} \quad e' \equiv X'^{-1} \pmod{\psi(u', v')}.$$

If $v = v'$ and $e = e'$, then $X = X'$ and $u = u'$.

Proof. As in the proof of Lemma A.6.1, rewrite e and e' as

$$e = \frac{1 + \psi(u, v)Y}{X} \quad \text{and} \quad e' = \frac{1 + \psi(u', v')Y'}{X'}.$$

Suppose $e = e'$. Then

$$|\psi(u', v')XY' - \psi(u, v)X'Y| = |X' - X|. \tag{A.10}$$

Assuming $v = v'$ and using $\psi(u, v) = (p-u)(q-v)$, $\psi(u', v') = (p-u')(q-v)$ in (A.10), we get

$$(q - v)\,|(p - u')XY' - (p - u)X'Y| = |X' - X|.$$

By Lemma A.2.2, we have $q - v > 2^{-\frac{1}{2}} N^{\frac{1}{2}} - N^{\frac{1}{4}} > N^{\frac{1}{4}}$ and since $|X' - X| < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$, we get

$$\begin{cases} X' - X & = 0, \\ (p - u')XY' - (p - u)X'Y & = 0. \end{cases}$$

Hence $X = X'$ and $(p - u')Y' = (p - u)Y$. Suppose for contradiction that $u' \neq u$. Put $g = \gcd(p - u', p - u)$. Then g divides $(p - u) - (p - u') = u' - u$. Since $v = v'$, by Lemma A.4.3 we have $|u' - u| \leq 1$, so $g = 1$. Hence $\gcd(p - u', p - u) = 1$ and $p - u$ divides Y'. Since $p - u > N^{\frac{1}{2}}$ and $Y' < X' < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$, this leads to a contradiction, so we deduce that $u' = u$. This terminates the proof. □

Using the methods used to prove Lemma A.6.1 and Lemma A.6.2 plus some additional arguments, we shall prove the following stronger result.

Theorem A.6.3. *Let $N = pq$ be an RSA modulus with $q < p < 2p$. Let u, u' be integers with $|u|, |u'| < N^{\frac{1}{4}}$. Let X, X' be integers with $1 \leq X, X' < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$, $\gcd(X, \psi(u,v)) = 1$, $\gcd(X', \psi(u',v')) = 1$ where $v = \left\lfloor -\frac{qu}{p-u} \right\rfloor$ and $v' = \left\lfloor -\frac{qu'}{p-u'} \right\rfloor$. Define*

$$e \equiv X^{-1} \pmod{\psi(u,v)} \quad \text{and} \quad e' \equiv X'^{-1} \pmod{\psi(u',v')}.$$

If $e = e'$, then $u = u'$, $v = v'$ and $X = X'$.

Proof. Assume that $e = e'$. Then, as in the proof of Lemma A.6.2, e and e' satisfy (A.10). We first take care of some easy cases.

If $u = u'$, then $v = v'$ and by Lemma A.6.1, we get $X = X'$.

If $v = v'$, then by Lemma A.6.2, we get $u = u'$ and $X = X'$.

Without loss of generality, suppose that $\psi(u,v) < \psi(u',v')$. Transforming (A.10), we get

$$\left| \frac{XY'}{X'Y} - \frac{\psi(u,v)}{\psi(u',v')} \right| = \frac{|X' - X|}{X'Y\psi(u',v')} \leq \frac{\max(X', X)}{X'Y\psi(u',v')} < \frac{1}{2(X'Y)^2},$$

where the final step is trivial since, for $N \geq 2^{10}$

$$2\max(X', X)X'Y < 2 \times \left(2^{-\frac{1}{4}} N^{\frac{1}{4}}\right)^3 < N - 2^{-\frac{1}{2}} N^{\frac{1}{2}} < \psi(u',v').$$

Combined with Theorem A.2.1, this implies that $\frac{XY'}{X'Y}$ is one of the convergents of the continued fraction expansion of $\frac{\psi(u,v)}{\psi(u',v')}$. By Lemma A.4.5, the first non trivial convergents are $\frac{1}{1}$ and $\frac{a_2}{a_2+1}$ where $a_2 > 2^{-\frac{1}{2}} N^{\frac{1}{2}} - \frac{1}{2}$. Observe that

$$a_2 + 1 > 2^{-\frac{1}{2}} N^{\frac{1}{2}} - \frac{1}{2} + 1 = 2^{-\frac{1}{2}} N^{\frac{1}{2}} + \frac{1}{2} > 2^{-\frac{1}{2}} N^{\frac{1}{2}} = \left(2^{-\frac{1}{4}} N^{\frac{1}{4}}\right)^2 > X'Y.$$

This implies that the only possibility for $\frac{XY'}{X'Y}$ to be a convergent of $\frac{\psi(u,v)}{\psi(u',v')}$ is $\frac{1}{1}$. This gives $XY' = X'Y$. Since $\gcd(X, Y) = \gcd(X', Y') = 1$ then $X = X'$ and $Y = Y'$. Replacing in (A.10), we get $\psi(u',v') = \psi(u,v)$ and by Lemma A.4.4, we deduce $u = u'$. This completes the proof. □

We now determine the order of the cardinality of the set $H(N)$. Recall that the elements of $H(N)$ are uniquely defined by the congruence

$$e \equiv X^{-1} \pmod{\psi(u,v)},$$

where $|u| < N^{\frac{1}{4}}$, $v = \left\lfloor -\frac{qu}{p-u} \right\rceil$, $1 \leq X < 2^{-\frac{1}{4}} N^{\frac{1}{4}}$ and $\gcd(X, \psi(u,v)) = 1$. In addition, $p-u$ is B_{ecm}-smooth.

Theorem A.6.4. *Let $N = pq$ be an RSA modulus with $q < p < 2p$. We have*

$$\#H(N) \geq N^{\frac{1}{2}-\varepsilon},$$

where ε is a small positive constant.

Proof. Assume $B_{\text{ecm}} < p - p^{\frac{1}{2}}$. Let us consider the set $V(p)$ as defined by (A.9). Put $x = p - p^{\frac{1}{2}}$, $z = 2p^{\frac{1}{2}}$ and $y = B_{\text{ecm}}$. Define $\delta > 0$ and $\gamma > 0$ such that

$$x^{\frac{1}{2}+\delta} \leq z, \qquad y = x^{1/\gamma}.$$

Then $x \geq z \geq x^{\frac{1}{2}+\delta}$, $x \geq y \geq x^{1/\gamma}$ and the conditions to apply (A.3) are fulfilled. On the other hand, we have $y > \exp\left\{(\log\log x)^{5/3+\varepsilon}\right\}$ for $x < \exp\left\{10^{7-\varepsilon}\right\}$ and the condition to apply (A.2) is fulfilled. Combining (A.3) and (A.2), we get

$$\#V(p) = \Psi(x+z,y) - \Psi(x,y) \geq c\frac{z}{x}\Psi(x,y) = cz\rho(\gamma)\left\{1 + O\left(\frac{\log(\gamma+1)}{\log(y)}\right)\right\},$$

where $c = c(\delta,\gamma) > 0$ and $\rho(\gamma)$ is the Dickman-de Bruijn ρ-function (see Table 2). Hence

$$\#V(p) \geq c\rho(\gamma)\, z = 2c\rho(\gamma)\, p^{\frac{1}{2}}.$$

Since trivially $\#V(p) < z = 2p^{\frac{1}{2}}$, we get $\#V(p) \asymp p^{\frac{1}{2}}$. Combining this with Table 2, we conclude that $\#V(p)$ is lower bounded as follows

$$\#V(p) \geq p^{\frac{1}{2}-\varepsilon'} = N^{\frac{1}{4}-\varepsilon_1},$$

with small constants $\varepsilon' > 0$ and $\varepsilon_1 > 0$.

Next, for every integer u with $|u| < N^{\frac{1}{4}}$ put

$$W(u) = \left\{X :\ 1 \leq X < 2^{-\frac{1}{4}} N^{\frac{1}{4}},\ (X, \psi(u,v)) = 1\right\},$$

where $v = \left\lfloor -\frac{qu}{p-u} \right\rceil$. Setting $m = \left\lfloor 2^{-\frac{1}{4}} N^{\frac{1}{4}} \right\rfloor$, we have

$$\#W(u) = \sum_{\substack{X=1 \\ (X,\psi(u,v))=1}}^{m} 1 = \sum_{d \mid \psi(u,v)} \mu(d) \left\lfloor \frac{m}{d} \right\rfloor \geq m \sum_{d \mid \psi(u,v)} \frac{\mu(d)}{d} = \frac{m\phi(\psi(u,v))}{\psi(u,v)}$$

Table A.2: Table of values of $\rho(\gamma)$ with $\left(p-\sqrt{p}\right)^{\frac{1}{\gamma}} = B_{\text{ecm}} = 10^{50}$

Number of bits of p	256	512	1024	2048
$\gamma = \dfrac{\log\left(p-\sqrt{p}\right)}{\log B_{\text{ecm}}} \approx$	1.5	3	6.25	12.50
$\rho(\gamma) \approx$(see [53])	5.945×10^{-1}	4.861×10^{-2}	9.199×10^{-6}	1.993×10^{-15}

where $\mu(.)$ is the Möbius function and $\phi(.)$ is the Euler totient function. We shall need the well known result (see Theorem 328 of [57]),

$$\frac{\phi(n)}{n} \geq \frac{C}{\log\log n},$$

where C is a positive constant. Applying this with $n = \psi(u,v)$ and using Lemma A.4.1, we get

$$\#W(u) \geq \frac{Cm}{\log\log\psi(u,v)} \geq \frac{2^{-\frac{1}{4}}CN^{\frac{1}{4}}}{\log\log\left(N+2^{-\frac{1}{2}}N^{\frac{1}{2}}\right)} = N^{\frac{1}{4}-\varepsilon_2},$$

with a small constant $\varepsilon_2 > 0$.

It remains to show that $\#H(n) \geq N^{\frac{1}{4}-\varepsilon}$ where ε is a positive constant. Indeed, for every $u \in V(p)$ there are at least $N^{\frac{1}{4}-\varepsilon_2}$ integers $X \in W(u)$. Hence

$$\#H(n) \geq \#V(p)\#W(u) \geq N^{\frac{1}{2}-\varepsilon_1-\varepsilon_2}.$$

Setting $\varepsilon = \varepsilon_1 + \varepsilon_2$, this completes the proof of the theorem. □

A.7 Conclusion

Wiener's famous attack on RSA with $d < \frac{1}{3}N^{0.25}$ shows that using the equation $ed - k(p-1)(q-1) = 1$ and a small d makes RSA insecure. In this paper, we performed a generalization of this attack. We showed that we can find any X and Y with $1 \leq Y < X < 2^{-0.25}N^{0.25}$ from the continued fraction expansion of e/N when they satisfy an equation

$$eX - Y(p-u)\left(q + \left[\frac{qu}{p-u}\right]\right) = 1,$$

and if $p - u$ or $q + [qu/(p - u])$ is smooth enough to factor, then p and q can be found from X and Y. Our results illustrate that one should be very cautious when choosing some class of RSA exponent. Note that our attack, as well as all the attacks based on continued fractions do not apply to RSA with modulus N and small public exponents as the popular values $e = 3$ or $e = 2^{16} + 1$ because the non-trivial convergents of $\frac{e}{N}$ are large enough to use diophantine approximation techniques, namely Theorem A.2.1.

Appendix B

Cryptanalysis of RSA Using the Ratio of the Primes

AFRICACRYPT 2009
[107]

Abstract :
Let $N = pq$ be an RSA modulus, i.e. the product of two large unknown primes of equal bit-size. In the X9.31-1997 standard for public key cryptography, Section 4.1.2, there are a number of recommendations for the generation of the primes of an RSA modulus. Among them, the ratio of the primes shall not be close to the ratio of small integers. In this paper, we show that if the public exponent e satisfies an equation $eX - (N - (ap + bq))Y = Z$ with suitably small integers X, Y, Z, where $\frac{a}{b}$ is an unknown convergent of the continued fraction expansion of $\frac{q}{p}$, then N can be factored efficiently. In addition, we show that the number of such exponents is at least $N^{\frac{3}{4}-\varepsilon}$ where ε is arbitrarily small for large N.

B.1 Introduction

The RSA public-key cryptosystem [131] was invented by Rivest, Shamir, and Adleman in 1978. Since then, the RSA system has been the most widely accepted public key cryptosystem. In the RSA cryptosystem, the modulus $N = pq$ is a product of two primes of equal bit-size. Let e be an integer coprime with $\phi(N) = (p-1)(q-1)$, the Euler function of N. Let d be the integer solution of the equation $ed \equiv 1 \pmod{\phi(N)}$ with $d < \phi(N)$. We call e the public exponent and d the private exponent. The pair (N, e) is called the public key and the pair (N, d) is the corresponding private key.

RSA is computationally expensive as it requires exponentiations modulo the large RSA modulus N. For efficient modular exponentiation in the decryption/signing phase, one may be tempted to choose a small d. Unfortunately, Wiener [147] showed in 1990 that using continued fractions, one can efficiently recover the secret exponent d from the public key (N, e) as long as $d < \frac{1}{3}N^{\frac{1}{4}}$. Wiener's attack is based on solving the equation $ex - \phi(N)y = 1$ where $x < \frac{1}{3}N^{\frac{1}{4}}$. Since then, attacking RSA using information encoded in the public key (N, e) has been a stimulating area of research.

Based on the lattice basis reduction, Boneh and Durfee [17] proposed in 1999 a new attack on the use of short secret exponent d, namely, they improved the bound to $d < N^{0.292}$.

In 2004, Blömer and May [13] showed that N can be factored in polynomial time for every public key (N, e) satisfying an equation $ex - (N + 1 - (p + q))k = y$, with $x < \frac{1}{3}N^{\frac{1}{4}}$ and $|y| < N^{-\frac{3}{4}}ex$.

Another attack using information encoded in (N, e) was recently proposed by Nitaj in [104]. The idea of [104] is based on solving the equation satisfied by the public exponent e. Suppose e satisfies an equation $eX - (p-u)(q-v)Y = 1$ with $1 \leq Y \leq X < 2^{-\frac{1}{4}}N^{\frac{1}{4}}$, $|u| < N^{\frac{1}{4}}$ and $v = \left\lfloor -\frac{qu}{p-u} \right\rfloor$. If the prime factors of $p-u$ or $q-v$ are less than 10^{50}, then N can be factored efficiently.

In this paper, we propose new attacks on RSA. Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued

fraction expansion of $\frac{q}{p}$. Define α such that $ap + bq = N^{\frac{1}{2}+\alpha}$ with $0 < \alpha < \frac{1}{2}$. We focus on the class of the public exponents satisfying an equation

$$eX - (N - (ap + bq))Y = Z,$$

with small parameters X, Y, Z satisfying

$$1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}, \ \gcd(X, Y) = 1,$$

and Z depends on the size of $|ap-bq|$. We present three attacks according to the size of the difference $|ap-bq|$. The first attack concerns small difference, i.e. $|ap-bq| < (abN)^{\frac{1}{4}}$, the second attack will work for medium difference, i.e. $(abN)^{\frac{1}{4}} < |ap-bq| < aN^{\frac{1}{4}}$, and the third attack concerns large difference, i.e. $|ap - bq| > aN^{\frac{1}{4}}$. The first attack always lead to the factorization of N. The second and the third attacks work if, in addition, $b \leq 10^{52}$. This corresponds to the current limit of the Elliptic Curve Method [84] to find large factors of integers.

The attacks combine techniques from the theory of continued fractions, Coppersmith's method [34] for finding small roots of bivariate polynomial equations and the Elliptic Curve Method [84] for Integer Factorization. We also show that the set of exponents e for which our approach works is at least $N^{\frac{3}{4}-\varepsilon}$ where ε is a small positive constant depending only on N.

Our approach is more efficient if $\frac{q}{p}$ is close to $\frac{a}{b}$ with small integers a and b. This is a step in the direction of the recommendations of the $X9.31$-1997 standard for public key cryptography (Section 4.1.2) which requires that the ratio of the primes shall not be close to the ratio of small integers. It is important to notice that, since $q < p < 2q$, then $\frac{0}{1}$ and $\frac{1}{1}$ are among the convergents of the continued fraction expansion of $\frac{q}{p}$ (see Section 2). For $a = 0$, $b = 1$, the equation $eX - (N - (ap + bq))Y = Z$ becomes

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

and was studied by Nitaj [104] with suitably small parameters X, Y, Z. Consequently, in this paper, we focus on the convergents $\frac{a}{b}$ with $a \geq 1$. For $a = b = 1$, our third attack applies and matches the attack of Blömer and May [13].

The rest of the paper is organized as follows. In Section 2 we give a brief introduction to continued fractions, Coppersmith's lattice-based method for finding small roots of polynomials [34] and the Elliptic Curve Method of Factorization. In Section 3 we study the properties of the convergents of the continued fraction expansion of the ratio of the primes of $N = pq$. In Section 4 we present the new attacks. In Section 5, we give an estimate for the size of the set of the public exponents for which our attacks work. Section 6 concludes the paper.

B.2 Preliminaries on Continued Fractions, Coppersmith's Method and The Elliptic Curve Method (ECM)

We first introduce some notation. The integer closest to x is denoted $[x]$ and the largest integer less than or equal to x is denoted $\lfloor x \rfloor$.

B.2.1 Continued Fractions and the Euclidean Algorithm

We briefly recall some basic definitions and facts that we use about continued fractions and the Euclidean algorithm, which can be found in [57].

The process of finding the continued fraction expansion of a rational number $\frac{q}{p}$ involves the same series of long divisions that are used in the application of the Euclidean algorithm to the pair of integers (q, p). Starting with $r_{-2} = q$ and $r_{-1} = p$, define the recursions

$$a_i = \left\lfloor \frac{r_{i-2}}{r_{i-1}} \right\rfloor, \quad r_i = r_{i-2} - a_i r_{i-1}, \ i \geq 0, \tag{B.1}$$

where a_i is the integer quotient $\lfloor r_{i-2}/r_{i-1} \rfloor$ and r_i is the integer remainder that satisfies $0 \leq r_i < r_{i-1}$. The Euclidean algorithm terminates with a series of remainders satisfying

$$0 = r_m < r_{m-1} < \cdots < r_2 < r_1 < r_0 < r_{-1} = p.$$

The continued fraction expansion of $\frac{q}{p}$ is then

$$\frac{q}{p} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_m}}}},$$

or alternatively, $\frac{q}{p} = [a_0, a_1, \cdots, a_m]$. The rational number $[a_0, a_1, \cdots, a_i]$ with $0 \leq i \leq m$ is called the i-th convergent of $\frac{q}{p}$ and satisfies

$$[a_0, a_1, \cdots, a_i] = \frac{p_i}{q_i},$$

where the integers p_i and q_i are coprime positive integers. Note that the integers p_i and q_i are also defined by the double recursions

$$p_{-2} = 0, \quad p_{-1} = 1, \quad p_i = a_i p_{i-1} + p_{i-2}, \ i \geq 0, \tag{B.2}$$

$$q_{-2} = 1, \quad q_{-1} = 0, \quad q_i = a_i q_{i-1} + q_{i-2}, \ i \geq 0. \tag{B.3}$$

Since $q < p < 2q$, we have $\frac{q}{p} < 1$ and taking $i = 0$ in (B.1), (B.2) and (B.3), we get

$$a_0 = \left\lfloor \frac{r_{-2}}{r_{-1}} \right\rfloor = \left\lfloor \frac{q}{p} \right\rfloor = 0, \quad r_0 = q, \quad p_0 = 0, \quad q_0 = 1.$$

Similarly, we have $1 < \frac{p}{q} < 2$ and taking $i = 1$ in (B.1), (B.2) and (B.3), we get

$$a_1 = \left\lfloor \frac{r_{-1}}{r_0} \right\rfloor = \left\lfloor \frac{p}{q} \right\rfloor = 1, \quad p_1 = 1, \quad q_1 = 1.$$

From this we deduce that the first convergents of the continued fraction expansion of $\frac{q}{p}$ are $\frac{0}{1}$ and $\frac{1}{1}$.

Proposition B.2.1. *Let $\frac{q}{p} = [a_0, a_1, \cdots, a_m]$ be a continued fraction. For $0 \leq i < m$, we have*

$$\left| \frac{q}{p} - \frac{p_i}{q_i} \right| < \frac{1}{q_i^2}.$$

We terminate with a famous result on good rational approximations.

Theorem B.2.2. *Let* $\frac{q}{p} = [a_0, a_1, \cdots, a_m]$. *If* a *and* b *are coprime positive integers such that* $b < p$ *and*

$$\left|\frac{q}{p} - \frac{a}{b}\right| < \frac{1}{2b^2},$$

then $a = p_i$ *and* $b = q_i$ *for some* i *with* $0 \leq i \leq m$.

B.2.2 Coppersmith's Method

At Eurocrypt'96, Coppersmith [34] introduced two lattice reduction based techniques to find small roots of polynomial diophantine equations. The first technique works for modular univariate polynomials, the second for bivariate integer polynomial equations. Since then, Coppersmith's techniques have been used in a huge variety of cryptanalytic applications. Coppersmith illustrated his technique for solving bivariate integer polynomial equations with the problem of finding the factors of $n = xy$ if we are given the high order $\frac{1}{4}\log_2 n$ bits of y.

Theorem B.2.3. *Let* $n = xy$ *be the product of two unknown integers such that* $x < y < 2x$. *Given an approximation of* y *with additive error at most* $n^{\frac{1}{4}}$, *then* x *and* y *can be found in polynomial time.*

B.2.3 The Elliptic Curve Method of Factorization

The difficulty of factoring a large number is an element of the security of the RSA system. In the recent years, the limits of the best factorization algorithms have been extended greatly. There are two classes of algorithms for finding a nontrivial factor p of a composite integer n. The algorithms in which the run time depends on the size of n: Lehman's algorithm [83], the Continued Fraction algorithm [97], the Multiple Polynomial Quadratic Sieve algorithm [142], the Number Field Sieve [85]. And the algorithms in which the run time depends on the size of p: Trial Division, Pollard's "rho" algorithm [129], Lenstra's Elliptic Curve Method [84].

The Elliptic Curve Method (ECM for short) was originally proposed by H.W. Lenstra [84] and subsequently extended by Brent [21], [22], and

Montgomery [96]. The original part of the algorithm proposed by Lenstra is typically referred to as Phase 1, and the extension by Brent and Montgomery is called Phase 2. ECM is suited to find small factors p of large numbers n and has complexity

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

where $c > 0$ and $M(n)$ denotes the cost of multiplication $\pmod n$. R. Brent [22] extrapolated that the Elliptic Curve Method record will be a D-digit factor in year $Y(D) = 9.3\sqrt{D} + 1932.3$. According to this formula, $Y(50) \approx 1998$ and $Y(67) \approx 2008$. A table of the largest factors found using the ECM is maintained by Zimmermann [150]. The largest prime factor found using the ECM had 67 decimal digits and was found by B. Dodson on August 24, 2006.

B.3 Useful Lemmas and Properties

First we recall a very useful lemma (see [104]).

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

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

The following lemma shows that a and b are of the same bit-size.

Lemma B.3.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. If $\frac{a}{b}$ is a convergent of $\frac{q}{p}$ with $a \geq 1$, then $a \leq b \leq 2a$.*

Proof. If $b = 1$, then $a = 1$ and the inequalities $a \leq b \leq 2a$ are satisfied. Next, suppose $b \geq 2$. Observe that if $\frac{a}{b}$ is a convergent of $\frac{q}{p}$ then by Proposition B.2.1 we have $|ap - bq| \leq \frac{p}{b} \leq \frac{p}{2}$. Isolating bq and dividing by q, we get

$$a\frac{p}{q} - \frac{p}{2q} \leq b \leq a\frac{p}{q} + \frac{p}{2q}.$$

Combining this with $1 < \frac{p}{q} < 2$, we get

$$a - \frac{p}{2q} < a\frac{p}{q} - \frac{p}{2q} \leq b \leq a\frac{p}{q} + \frac{p}{2q} < 2a + \frac{p}{2q}.$$

Since $p < 2q$, then $0 < \frac{p}{2q} < 1$. Hence $a \leq b \leq 2a$ which completes the proof. □

The following lemma plays an important role in this paper. Recall that the integer closest to x is denoted $[x]$.

Lemma B.3.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$ and $\frac{a}{b}$ a convergent of the continued fraction expansion of $\frac{q}{p}$ with $a \geq 1$. Let $ap + bq = N^{\frac{1}{2}+\alpha}$ with $\alpha < \frac{1}{2}$. If $|ap + bq - M| < \frac{1}{2}N^{\frac{1}{2}-\alpha}$, then*

$$ab = \left[\frac{M^2}{4N}\right].$$

Proof. Set $M = ap + bq + x$. Using $(ap - bq)^2 = (ap + bq)^2 - 4abN$, we get, after rearrangement,

$$M^2 - 4abN = (ap + bq + x)^2 - 4abN = (ap - bq)^2 + 2(ap + bq)x + x^2. \quad \text{(B.4)}$$

Consider the term $(ap - bq)^2$ on the right side of (B.4). If $b = 1$, then by Lemma B.3.2, $a = 1$. Hence, since $q < p < 2q$, we have $|ap - bq| = |p - q| = p - q < \frac{p}{2}$. If $b \geq 2$, then by Proposition B.2.1, we have $|ap - bq| < \frac{p}{b} \leq \frac{p}{2}$. Combining with Lemma B.3.1, we get in both cases

$$(ap - bq)^2 < \left(\frac{p}{2}\right)^2 < \left(\frac{2^{\frac{1}{2}}N^{\frac{1}{2}}}{2}\right)^2 = \frac{N}{2}.$$

Hence, using $|x| < \frac{1}{2}N^{\frac{1}{2}-\alpha}$, the right side of (B.4) becomes

$$\begin{aligned}
\left|(ap - bq)^2 + 2(ap + bq)x + x^2\right| &\leq (ap - bq)^2 + 2(ap + bq)|x| + x^2 \\
&< \frac{N}{2} + 2N^{\frac{1}{2}+\alpha} \cdot \frac{1}{2}N^{\frac{1}{2}-\alpha} + \frac{1}{4}N^{1-2\alpha} \\
&= \left(\frac{1}{2} + 1 + \frac{1}{4}N^{-2\alpha}\right)N \\
&< 2N,
\end{aligned}$$

where we used $\alpha > 0$. Plugging this in (B.4) and dividing by $4N$, we get

$$\left|\frac{M^2}{4N} - ab\right| = \frac{\left|M^2 - 4abN\right|}{4N} = \frac{\left|(ap - bq)^2 + 2(ap + bq)x + x^2\right|}{4N} < \frac{2N}{4N} = \frac{1}{2}.$$

It follows that $ab = \left[\frac{M^2}{4N}\right]$ which terminates the proof. □

The following lemma indicates that ap and bq are of the same bit-size.

Lemma B.3.4. *Let $N = pq$ be an RSA modulus with $q < p < 2q$ and $\frac{a}{b}$ a convergent of the continued fraction expansion of $\frac{q}{p}$ with $a \geq 1$. Then*

$$ap < bq < 2ap \quad or \quad bq < ap < 2bq$$

Proof. First, assume $ap < bq$. By Lemma B.3.2, we have $b \leq 2a$. Combining this with $q < p$, we get $bq < 2ap$, and consequently $ap < bq < 2ap$.

Next, assume $bq < ap$. By Lemma B.3.2, we have $a \leq b$. Combining this with $p < 2q$, we get $ap < 2bq$ and finally $bq < ap < 2bq$. This terminates the proof. □

B.4 The New Attacks on RSA

In this section, we show how to factor the RSA modulus N if (N, e) is a public key satisfying an equation $eX - (N - (ap + bq))Y = Z$ with small parameters X, Y and Z where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ with $a \geq 1$. We shall consider separately the cases when the difference $|ap - bq|$ is small, i.e. $|ap - bq| < (abN)^{\frac{1}{4}}$, medium, i.e. $(abN)^{\frac{1}{4}} < |ap - bq| < aN^{\frac{1}{4}}$, and large, i.e. $|ap - bq| > aN^{\frac{1}{4}}$. This corresponds approximately to $b > 2^{\frac{1}{2}}N^{\frac{1}{6}}$, $2^{\frac{1}{2}}N^{\frac{1}{6}} > b > 2^{\frac{1}{4}}N^{\frac{1}{8}}$ and $b < 2^{\frac{1}{4}}N^{\frac{1}{8}}$ respectively.

First we present a result based on continued fractions.

Lemma B.4.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let a, b be coprime positive integers such that $ap + bq = N^{\frac{1}{2}+\alpha}$ with $\alpha < \frac{1}{2}$. Let e be a public exponent satisfying the equation $eX - (N - (ap + bq))Y = Z$ with $\gcd(X, Y) = 1$. If $|Z| < N^{\frac{1}{2}+\alpha}X$ and $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$, then $\frac{Y}{X}$ is a convergent of $\frac{e}{N}$.*

Proof. Set $ap + bq = N^{\frac{1}{2}+\alpha}$ with $\alpha < \frac{1}{2}$. Rewrite $eX - (N - ap - bq)Y = Z$ as $eX - NY = Z - (ap + bq)Y$. Now suppose $|Z| < N^{\frac{1}{2}+\alpha}X$, $1 \leq Y \leq X$ and

$\gcd(X, Y) = 1$. Then

$$\begin{aligned}\left|\frac{e}{N} - \frac{Y}{X}\right| &= \frac{|eX - NY|}{NX} \\ &= \frac{|Z - (ap + bq)Y|}{NX} \\ &\leq \frac{|Z|}{NX} + \frac{(ap + bq)Y}{NX} \\ &< \frac{N^{\frac{1}{2}+\alpha}}{N} + \frac{N^{\frac{1}{2}+\alpha}}{N} \\ &= 2N^{-\frac{1}{2}+\alpha}.\end{aligned}$$

Since $X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$, then $2N^{-\frac{1}{2}+\alpha} < \frac{1}{2X^2}$. Hence, by Theorem B.2.2, $\frac{Y}{X}$ is one of the convergents of the continued fraction expansion of $\frac{e}{N}$. □

B.4.1 An Attack for Small Difference $|ap - bq|$

We now present the first attack.

Theorem B.4.2. *Let $N = pq$ be an RSA modulus with unknown factors p, q such that $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued fraction expansion of $\frac{q}{p}$ with $a \geq 1$ and $|ap - bq| < (abN)^{\frac{1}{4}}$. Let e be a public exponent satisfying an equation $eX - (N - ap - bq)Y = Z$ with $\gcd(X, Y) = 1$. Set $ap + bq = N^{\frac{1}{2}+\alpha}$ with $0 < \alpha < \frac{1}{2}$. If $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ and $|Z| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$, then N can be factored in polynomial time.*

Proof. Assume $|Z| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$, $1 \leq Y \leq X$ with $\gcd(X, Y) = 1$. Then

$$|Z| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)X \leq \frac{1}{2}N^{\frac{1}{2}-\alpha}X < N^{\frac{1}{2}+\alpha}X.$$

Hence by Lemma B.4.1, $\frac{Y}{X}$ is one of the convergents of $\frac{e}{N}$. Set $M = N - \frac{eX}{Y}$. Starting with the equation $eX - (N - (ap + bq))Y = Z$, we get

$$|ap + bq - M| = \frac{|Z|}{Y} < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right) < \frac{1}{2}N^{\frac{1}{2}-\alpha}.$$

Hence, by Lemma B.3.3, we find $ab = \left\lfloor \frac{M^2}{4N} \right\rfloor$. On the other hand, we have

$$|ap + bq - M| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right) < (abN)^{\frac{1}{4}}.$$

Moreover, if $|ap - bq| < (abN)^{\frac{1}{4}}$, then

$$\left|ap - \frac{M}{2}\right| \leq \frac{1}{2}\,|ap + bq - M| + \frac{1}{2}|ap - bq| < \frac{1}{2}(abN)^{\frac{1}{4}} + \frac{1}{2}(abN)^{\frac{1}{4}} = (abN)^{\frac{1}{4}}.$$

It follows that the term $\frac{M}{2}$ is an approximation of the factor ap of $n = abN$ with additive error at most $n^{\frac{1}{4}}$. In addition, by Lemma B.3.4, the factors ap and bq of n are of the same bit-size. Hence, using Theorem B.2.3 with n and $\frac{M}{2}$, we find ap, and since $a < q$, we get $p = \gcd(N, ap)$ which terminates the proof. □

Let us summarize the first factorization algorithm.

Algorithm 1 Small $|ap - bq|$

Require: a public key (N, e) satisfying $N = pq$, $q < p < 2q$ and $eX - (N - (ap + bq))Y = Z$ for small parameters X, Y, Z where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ with $a \geq 1$.

Ensure: the prime factors p and q.

1: Compute the continued fraction expansion of $\frac{e}{N}$.
2: For every convergent $\frac{Y}{X}$ of $\frac{e}{N}$ with $X < \frac{1}{2}N^{\frac{1}{4}}$:
3: Compute $M = N - \frac{eX}{Y}$ and $N_0 = \left\lfloor \frac{M^2}{4N} \right\rfloor$.
4: Apply Coppersmith's algorithm (Theorem B.2.3) with $n = N_0 N$ and $\frac{M}{2}$ as an approximation of y.
5: Compute $g = \gcd(y, N)$. If $1 < g < N$, then stop.

B.4.2 An Attack for Medium Difference $|ap - bq|$

Here we present the second attack. It is based on the Elliptic Curve Method (ECM) which can find factors of about 52-digits. Assuming the efficiency of ECM, every step in this attack can be done in polynomial time and the number of convergents is bounded by $\mathcal{O}(\log N)$. To express this fact, the term *efficient* is used.

Theorem B.4.3. *Let $N = pq$ be an RSA modulus with unknown factors p, q such that $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued fraction expansion of $\frac{q}{p}$ such that $a \geq 1$, $(abN)^{\frac{1}{4}} < |ap - bq| < aN^{\frac{1}{4}}$ and $b \leq 10^{52}$. Let e be a public exponent satisfying an equation $eX - (N - ap - bq)Y = Z$ with $\gcd(X, Y) = 1$. Set $M = N - \frac{eX}{Y}$ and $ap + bq = N^{\frac{1}{2}+\alpha}$ with $0 < \alpha < \frac{1}{2}$. If $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ and $|Z| < \min\left(aN^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$, then, under ECM, N can be factored efficiently.*

Proof. Assume $|Z| < \min\left(aN^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$, $1 \leq Y \leq X$ and $\gcd(X, Y) = 1$. Then

$$|Z| < \min\left(aN^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)X \leq \frac{1}{2}N^{\frac{1}{2}-\alpha}X < N^{\frac{1}{2}+\alpha}X.$$

It follows, by Lemma B.4.1, that $\frac{Y}{X}$ is among the convergents of $\frac{e}{N}$.
Next, set $M = N - \frac{eX}{Y}$. Using the equation $eX - (N - (ap + bq))Y = Z$, we get

$$|ap + bq - M| = \frac{|Z|}{Y} < \min\left(aN^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right) \leq \frac{1}{2}N^{\frac{1}{2}-\alpha}.$$

Hence, by Lemma B.3.3, we find $ab = \left[\frac{M^2}{4N}\right]$ and by Lemma B.3.2, we know that a and b are of equal bit-size. Hence, applying the Elliptic Curve Method with $\left[\frac{M^2}{4N}\right]$, we can efficiently find a and b assuming $b \leq 10^{52}$.
From $|ap + bq - M| < aN^{\frac{1}{4}}$, we get

$$\left|p + \frac{bq}{a} - \frac{M}{a}\right| < \frac{aN^{\frac{1}{4}}}{a} = N^{\frac{1}{4}}. \tag{B.5}$$

On the other hand, by assumption, $|ap - bq| < aN^{\frac{1}{4}}$. Then $\left|p - \frac{bq}{a}\right| < N^{\frac{1}{4}}$,

and combining with (B.5), we get

$$\begin{aligned}\left|p-\frac{M}{2a}\right| &= \left|\frac{1}{2}\left(p+\frac{bq}{a}-\frac{M}{a}\right)+\frac{1}{2}\left(p-\frac{bq}{a}\right)\right| \\ &\leq \frac{1}{2}\left|p+\frac{bq}{a}-\frac{M}{a}\right|+\frac{1}{2}\left|p-\frac{bq}{a}\right| \\ &< \frac{1}{2}N^{\frac{1}{4}}+\frac{1}{2}N^{\frac{1}{4}} \\ &= N^{\frac{1}{4}}.\end{aligned}$$

This implies that $\frac{M}{2a}$ is an approximation of p with additive error at most $N^{\frac{1}{4}}$. Then, using Theorem B.2.3, this gives p which terminates the proof. □

Here we summarize the second factorization algorithm.

Algorithm 2 Medium $|ap - bq|$

Require: a public key (N, e) satisfying $N = pq$, $q < p < 2q$ and $eX - (N - (ap + bq))Y = Z$ for small parameters X, Y, Z where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ with $a \geq 1$.

Ensure: the prime factors p and q.

1: Compute the continued fraction expansion of $\frac{e}{N}$.
2: For every convergent $\frac{Y}{X}$ of $\frac{e}{N}$ with $X < \frac{1}{2}N^{\frac{1}{4}}$:
3: Compute $M = N - \frac{eX}{Y}$ and $N_0 = \left\lfloor \frac{M^2}{4N} \right\rfloor$.
4: **if** $N_0 < 10^{104}$ **then**
5: Apply ECM to find a and b such that $N_0 = ab$ and $a \leq b \leq 2a$.
6: Apply Coppersmith's algorithm (Theorem B.2.3) with $n = N$ and $\frac{M}{2a}$ as an approximation of y. If Coppersmith's algorithm outputs the factors p and q of N, then stop.
7: **end if**

B.4.3 An Attack for Large Difference $|ap - bq|$

Here we present the last attack. We suppose $|ap - bq| > aN^{\frac{1}{4}}$ so that the Small and the Medium difference attacks should not succeed. This attack depends on the efficiency of the Elliptic Curve Method (ECM) to find factors up to 10^{52}. Assuming the efficiency of ECM, the term *efficient* is also used to express the fact that every step in this attack can be done in polynomial time.

Theorem B.4.4. *Let $N = pq$ be an RSA modulus with unknown factors p, q such that $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of the continued*

fraction expansion of $\frac{q}{p}$ *such that* $a \geq 1$ *and* $b \leq 10^{52}$. *Let* e *be a public exponent satisfying an equation* $eX-(N-(ap+bq))Y = Z$ *with* $\gcd(X,Y) = 1$. *Let* $M = N - \frac{eX}{Y}$. *Set* $D = \sqrt{|M^2 - 4abN|}$ *and* $ap + bq = N^{\frac{1}{2}+\alpha}$ *with* $0 < \alpha < \frac{1}{2}$. *If* $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ *and* $|Z| < \frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}Y$ *then, under ECM,* N *can be factored efficiently.*

Proof. Combining Proposition B.2.1 and Lemma B.3.1, we have

$$|ap - bq| < \frac{p}{b} < \frac{2^{\frac{1}{2}}N^{\frac{1}{2}}}{b}.$$

Hence, since $a \leq b$, this gives

$$\frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha} < \frac{1}{3}a \cdot \frac{2^{\frac{1}{2}}N^{\frac{1}{2}}}{b} \cdot N^{-\frac{1}{4}-\alpha} \leq \frac{2^{\frac{1}{2}}}{3}N^{\frac{1}{4}-\alpha}. \tag{B.6}$$

Now, suppose $|Z| < \frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}Y$, $1 \leq Y \leq X$ and $\gcd(X,Y) = 1$. Then using (B.6), we get

$$|Z| < \frac{2^{\frac{1}{2}}}{3}N^{\frac{1}{4}-\alpha}X < N^{\frac{1}{2}+\alpha}X.$$

Consequently, by Lemma B.4.1, $\frac{Y}{X}$ is a convergent of $\frac{e}{N}$. Next, set $M = N - \frac{eX}{Y}$. Using the equation $eX - (N - ap - bq)Y = Z$, we get

$$|ap + bq - M| = \frac{|Z|}{Y} < \frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}. \tag{B.7}$$

Then using (B.6), we get

$$|ap + bq - M| < \frac{2^{\frac{1}{2}}}{3}N^{\frac{1}{4}-\alpha} < \frac{1}{2}N^{\frac{1}{2}-\alpha}.$$

Hence, by Lemma B.3.3, $ab = \left\lceil \frac{M^2}{4N} \right\rceil$ and by Lemma B.3.2, we know that a and b are of the same bit-size. Hence, if $b \leq 10^{52}$, then applying the Elliptic Curve Method with $\left\lceil \frac{M^2}{4N} \right\rceil$, we can find a and b.

Next, using $|ap - bq| < 2N^{\frac{1}{2}}$, we can rewrite (B.7) as

$$|ap + bq - M| < \frac{1}{3}a \cdot 2N^{\frac{1}{2}} \cdot N^{-\frac{1}{4}-\alpha} = \frac{2}{3}aN^{\frac{1}{4}-\alpha} < aN^{\frac{1}{4}}. \tag{B.8}$$

Now, let $D = \sqrt{|M^2 - 4abN|}$. Then

$$\begin{aligned} \big||ap-bq|^2 - D^2\big| &= \big||ap-bq|^2 - \big|M^2 - 4abN\big|\big| \\ &\leq \big|(ap-bq)^2 - M^2 + 4abN\big| \\ &= \big|(ap+bq)^2 - M^2\big|. \end{aligned}$$

From this we deduce

$$\big||ap-bq| - D\big| \leq \frac{|ap+bq-M||ap+bq+M|}{|ap-bq|+D}.$$

Next, by (B.8), we have $|ap+bq-M| < aN^{\frac{1}{4}}$. Then $M < ap+bq+aN^{\frac{1}{4}}$ and

$$ap+bq+M < 2(ap+bq) + aN^{\frac{1}{4}} < 3(ap+bq) = 3N^{\frac{1}{2}+\alpha}.$$

Combining with (B.7), this leads to

$$\big||ap-bq| - D\big| < \frac{3 \cdot \frac{1}{3}a|ap-bq|N^{-\frac{1}{4}-\alpha}N^{\frac{1}{2}+\alpha}}{|ap-bq|} = aN^{\frac{1}{4}}.$$

If $ap - bq > 0$, then combining with (B.8), we get

$$\begin{aligned} |2ap - M - D| &= \big|ap+bq-M+|ap-bq|-D\big| \\ &\leq \big|ap+bq-M\big| + \big||ap-bq|-D\big| \\ &< 2aN^{\frac{1}{4}}. \end{aligned}$$

Dividing by $2a$, we find that $\frac{M+D}{2a}$ is an approximation of p with additive error at most $N^{\frac{1}{4}}$.

If $ap - bq < 0$, then combining with (B.8), we get

$$\begin{aligned} |2ap - M + D| &= |ap+bq-M-(bq-ap-D)| \\ &< |ap+bq-M| + \big||ap-bq|-D\big| \\ &< 2aN^{\frac{1}{4}}. \end{aligned}$$

Dividing again by $2a$, we find that $\frac{M-D}{2a}$ is an approximation of p with additive error at most $N^{\frac{1}{4}}$. We can then apply Theorem B.2.3 to the values $\frac{M\pm D}{2a}$. The correct term will lead to the factorization of N. □

Now we summarize the third factorization algorithm.

Algorithm 3 Large $|ap - bq|$

Require: a public key (N, e) satisfying $N = pq$, $q < p < 2q$ and $eX - (N - (ap + bq))Y = Z$ for small parameters X, Y, Z where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ with $a \geq 1$.

Ensure: the prime factors p and q.

1: Compute the continued fraction expansion of $\frac{e}{N}$.
2: For every convergent $\frac{Y}{X}$ of $\frac{e}{N}$ with $X < \frac{1}{2}N^{\frac{1}{4}}$:
3: Compute $M = N - \frac{eX}{Y}$ and $N_0 = \left\lfloor \frac{M^2}{4N} \right\rceil$.
4: **if** $N_0 < 10^{104}$ **then**
5: Apply ECM to find a and b such that $N_0 = ab$ and $a \leq b \leq 2a$.
6: Compute $D = \sqrt{|M^2 - 4N_0N|}$.
7: Compute $m_1 = \frac{M+D}{2a}$ and $m_2 = \frac{M-D}{2a}$.
8: Apply Coppersmith's algorithm (Theorem B.2.3) with $n = N$ and m_1 and m_2 as approximations of y. If Coppersmith's algorithm outputs the factors p and q, then stop.
9: **end if**

B.5 Estimation of the Public Exponents for which the Attacks Apply

In this Section, we will study the size of the class of the public keys for which our attacks can be applied. Let $\frac{a}{b}$ be a convergent of $\frac{q}{p}$ with $a \geq 1$. Define α by $ap + bq = N^{\frac{1}{2}+\alpha}$ with $0 < \alpha < \frac{1}{2}$ and let

$$\mathcal{P}(a,b) = \left\{ (X, Y, z) \;\middle|\; 1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}},\ \gcd(X,Y) = 1,\ |z| < N^{\frac{1}{4}-\frac{\alpha}{2}} \right\},$$

be the set of the parameters and

$$\mathcal{E}(a,b) = \left\{ e \;\middle|\; e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z,\ (X, Y, z) \in \mathcal{P}(a,b) \right\},$$

the set of the exponents. We will show that much of these exponents are vulnerable to our attacks. To find a lower bound for the size of the sets $\mathcal{E}(a,b)$, we show that different convergents $\frac{a}{b}$ of $\frac{q}{p}$ and different parameters in the set $\mathcal{P}(a,b)$ define different exponents in the sets $\mathcal{E}(a,b)$.

First, we show that our attacks will work for the exponents in $\mathcal{E}(a,b)$: given an exponent in $\mathcal{E}(a,b)$, it is possible to find the factorization of N according to Theorem B.4.2, Theorem B.4.3 or Theorem B.4.4. First, we start with a result for small difference $|ap - bq|$.

Corollary B.5.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of $\frac{q}{p}$ with $a \geq 1$ and $|ap - bq| < (abN)^{\frac{1}{4}}$. Let X, Y be unknown coprime positive integers with $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ where $ap + bq = N^{\frac{1}{2}+\alpha}$ and $0 < \alpha < \frac{1}{2}$. If $e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z$ is a public exponent with*

$$|z| < \inf\left((abN)^{\frac{1}{4}}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha}{2}}\right),$$

then N can be factored in polynomial time.

Proof. Set $e_0 = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor$, $e = e_0 + z$, $Z = eX - (N - (ap + bq))Y$. We want to show that the conditions of Theorem B.4.2 are satisfied. Assume that $|z| < \inf\left((abN)^{\frac{1}{4}}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha}{2}}\right)$. Then, since

$$\left|(N - (ap + bq))\frac{Y}{X} - e_0\right| < 1,$$

we get

$$\begin{aligned} |Z| = |eX - (N - (ap + bq))Y| &= |(e_0 + z)X - (N - (ap + bq))Y| \\ &\leq |e_0X - (N - (ap + bq))Y| + |z|X \\ &< (1 + |z|)X. \end{aligned}$$

Observe that $(1 + |z|)X < (abN)^{\frac{1}{4}}Y$ and, assuming $X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$, we find

$$(1 + |z|)X < N^{\frac{1}{4}-\frac{\alpha}{2}} \cdot \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}} \leq \frac{1}{2}N^{\frac{1}{2}-\alpha}Y.$$

From this, we deduce $|Z| < \inf\left((abN)^{\frac{1}{4}}, \frac{1}{2}N^{\frac{1}{2}-\alpha}\right)Y$. It follows that the conditions of Theorem B.4.2 are fulfilled which leads to the factorization of N. □

Next, we give a result for medium difference $|ap - bq|$.

Corollary B.5.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of $\frac{q}{p}$ with $a \geq 1$, $b \leq 10^{52}$ and $(abN)^{\frac{1}{4}} < |ap - bq| < aN^{\frac{1}{4}}$. Let X, Y be unknown coprime positive integers with $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ where $ap + bq = N^{\frac{1}{2}+\alpha}$ and $0 < \alpha < \frac{1}{2}$. If $e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z$ is a public exponent with*

$$|z| < \inf\left(aN^{\frac{1}{4}}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha}{2}}\right),$$

then, under ECM, N can be factored efficiently.

Proof. The proof is similar to that of Corollary B.5.1 and the parameters satisfy the condition of Theorem B.4.3. □

Finally, we give a result which concerns large difference $|ap - bq|$.

Corollary B.5.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\frac{a}{b}$ be an unknown convergent of $\frac{q}{p}$ with $a \geq 1$, $b \leq 10^{52}$ and $|ap - bq| > aN^{\frac{1}{4}}$. Let X, Y be unknown coprime positive integers with $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ where $ap + bq = N^{\frac{1}{2}+\alpha}$ and $0 < \alpha < \frac{1}{2}$. If $e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z$ is a public exponent with*

$$|z| < \min\left(\frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha}{2}}\right),$$

then, under ECM, N can be factored efficiently.

Proof. Let $Z = eX - (N - (ap + bq))Y$, $e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z$ with $|z| < \min\left(\frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha}{2}}\right)$ and $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$. Using the same arguments as in the proof of Corollary B.5.1, we get

$$|Z| < (1 + |z|)X < \frac{1}{3}a|ap - bq|N^{-\frac{1}{4}-\alpha}Y.$$

It follows that all the conditions of Theorem B.4.4 are fulfilled which leads to the factorization of N. □

The following result shows that distinct parameters from $\mathcal{P}(a, b)$ define different exponents in $\mathcal{E}(a, b)$.

Lemma B.5.4. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\frac{a}{b}$ be a convergent of $\frac{q}{p}$ with $a \geq 1$ and $ap + bq = N^{\frac{1}{2}+\alpha}$. Let $(X, Y, z), (X', Y', z') \in \mathcal{P}(a, b)$. Let*

$$e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z, \quad e' = \left\lfloor (N - (ap + bq))\frac{Y'}{X'} \right\rfloor + z'.$$

If $e = e'$ then $X = X'$, $Y = Y'$ and $z = z'$.

Proof. Let $e_0 = \left\lfloor (N-(ap+bq))\frac{Y}{X} \right\rfloor$, $e_0' = \left\lfloor (N-(ap+bq))\frac{Y'}{X'} \right\rfloor$. If $e = e_0 + z$ and $e' = e_0' + z'$ then

$$\begin{aligned}
&\left| (N-(ap+bq)) \left(\frac{Y'}{X'} - \frac{Y}{X} \right) - e' + e \right| \\
&\leq \left| (N-(ap+bq))\frac{Y'}{X'} - e_0' - z' \right| + \left| (N-(ap+bq))\frac{Y}{X} - e_0 - z \right| \\
&\leq \left| (N-(ap+bq))\frac{Y'}{X'} - e_0' \right| + |z'| + \left| (N-(ap+bq))\frac{Y}{X} - e_0 \right| + |z| \\
&< 2 + |z| + |z'|.
\end{aligned}$$

Suppose $e = e'$. Then, multiplying by XX', we get

$$(N-(ap+bq))\,|Y'X - YX'| < (2+|z|+|z'|)XX'. \tag{B.9}$$

We want to compare the sides of (B.9). Assume that $X, X' < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha}{2}}$ and $|z|, |z'| < N^{\frac{1}{4}-\frac{\alpha}{2}}$. Then

$$(2+|z|+|z'|)XX' < 2N^{\frac{1}{4}-\frac{\alpha}{2}} \cdot \frac{1}{4}N^{\frac{1}{2}-\alpha} = \frac{1}{2}N^{\frac{3}{4}-\frac{3\alpha}{2}}.$$

On the other hand, we have

$$N-(ap+bq) = N - N^{\frac{1}{2}+\alpha} = N^{\frac{3}{4}-\frac{3\alpha}{2}} \left(N^{\frac{1}{4}+\frac{3\alpha}{2}} - N^{-\frac{1}{4}+\frac{5\alpha}{2}} \right).$$

Since $0 < \alpha < \frac{1}{2}$, then $\frac{1}{4} + \frac{3\alpha}{2} > -\frac{1}{4} + \frac{5\alpha}{2}$ and $N^{\frac{1}{4}+\frac{3\alpha}{2}} > N^{-\frac{1}{4}+\frac{5\alpha}{2}} + 1$. Hence $N-(ap+bq) > N^{\frac{3}{4}-\frac{3\alpha}{2}}$. From our comparison of the sides of (B.9), we conclude that $Y'X - YX' = 0$. Since $\gcd(X,Y) = 1$ and $\gcd(X',Y') = 1$, we find $X = X'$ and $Y = Y'$ and consequently $z = z'$. This terminates the proof. □

Finally, the following result shows that different convergents of $\frac{q}{p}$ lead to different exponents in $\mathcal{E}(a,b)$.

Lemma B.5.5. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\frac{a}{b}$ and $\frac{a'}{b'}$ be convergents of $\frac{q}{p}$ with $a \geq 1$, $a' \geq 1$, $ap+bq = N^{\frac{1}{2}+\alpha}$ and $a'p+b'q = N^{\frac{1}{2}+\alpha'}$. Let $(X,Y,z) \in \mathcal{P}(a,b)$ and $(X',Y',z') \in \mathcal{P}(a',b')$. Let*

$$e = \left\lfloor (N-(ap+bq))\frac{Y}{X} \right\rfloor + z, \quad e' = \left\lfloor (N-(a'p+b'q))\frac{Y'}{X'} \right\rfloor + z'.$$

If $e = e'$ then $X = X'$, $Y = Y'$, $a = a'$, $b = b'$ and $z = z'$.

Proof. Assume for contradiction that $a \neq a'$, $a < a'$ say. Then $b < b'$. Hence $ap + bq < a'p + b'q$ and $\alpha < \alpha'$. Combining with Lemma B.3.1, we get

$$N - (ap + bq) - (N - (a'p + b'q)) = (a' - a)p + (b' - b)q > p + q > p > N^{\frac{1}{2}},$$

which leads to

$$N - (ap + bq) > N - (a'p + b'q) + N^{\frac{1}{2}} \tag{B.10}$$

Now, set $e = \left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z$, $e' = \left\lfloor (N - (a'p + b'q))\frac{Y'}{X'} \right\rfloor + z'$ and assume $e = e'$. Then, since $|z| < N^{\frac{1}{4} - \frac{\alpha}{2}}$ and $|z'| < N^{\frac{1}{4} - \frac{\alpha'}{2}} < N^{\frac{1}{4} - \frac{\alpha}{2}}$, we get

$$\left| (N - (a'p + b'q))\frac{Y'}{X'} - (N - (ap + bq))\frac{Y}{X} \right| < 2 + |z| + |z'| < 2N^{\frac{1}{4} - \frac{\alpha}{2}} \tag{B.11}$$

On the other hand, we know that $\frac{Y}{X}$ and $\frac{Y'}{X'}$ are convergents of the continued fraction expansion of $\frac{e}{N}$. Hence $\frac{Y}{X} \approx \frac{Y'}{X'}$ and, combining (B.10) with $X < \frac{1}{2}N^{\frac{1}{4} - \frac{\alpha}{2}}$, we get

$$\begin{aligned}
(N - (ap + bq))\frac{Y}{X} &> (N - (a'p + b'q))\frac{Y}{X} + N^{\frac{1}{2}}\frac{Y}{X} \\
&> (N - (a'p + b'q))\frac{Y}{X} + N^{\frac{1}{2}} \cdot \frac{1}{\frac{1}{2}N^{\frac{1}{4} - \frac{\alpha}{2}}} \\
&\approx (N - (a'p + b'q))\frac{Y'}{X'} + 2N^{\frac{1}{4} + \frac{\alpha}{2}}
\end{aligned}$$

It follows that

$$\left| (N - (a'p + b'q))\frac{Y'}{X'} - (N - (ap + bq))\frac{Y}{X} \right| > 2N^{\frac{1}{4} + \frac{\alpha}{2}}.$$

Comparing with (B.11), we get a contradiction. Hence $a = a'$ and $b = b'$. Now, we have $\left\lfloor (N - (ap + bq))\frac{Y}{X} \right\rfloor + z = \left\lfloor (N - (ap + bq))\frac{Y'}{X'} \right\rfloor + z'$. By Lemma B.5.4, we conclude that $X = X'$, $Y = Y'$ and $z = z'$. This terminates the proof. □

Let us now prove a lower bound for the size of the number of the exponents e that are vulnerable to our approach. Note that we do not require $\gcd(e, \phi(N)) = 1$ as usual.

Theorem B.5.6. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then the number of the exponents $e \in \mathcal{E}(a, b)$ that are vulnerable to the attacks for some convergent $\frac{a}{b} \neq \frac{0}{1}$ of $\frac{q}{p}$ is at least $N^{\frac{3}{4}-\varepsilon}$ where ε is arbitrarily small for suitably large N.*

Proof. We focus on $\mathcal{E}(1, 1)$ since the total number of exponents is much higher. Let α_0 such that $p + q = N^{\frac{1}{2}+\alpha_0}$. Since $q < p$, then $2q < p + q < 2p$ and by Lemma B.3.1, we get $2^{\frac{1}{2}} N^{\frac{1}{2}} < N^{\frac{1}{2}+\alpha_0} < 2^{\frac{3}{2}} N^{\frac{1}{2}}$. From this we deduce $\alpha_0 \approx 0$. On the other hand, by Corollary B.5.3, we need

$$|z| < \min\left(\frac{1}{3}|p - q|N^{-\frac{1}{4}-\alpha_0}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha_0}{2}}\right),$$

where $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha_0}{2}}$ and $\gcd(X, Y) = 1$. Observe that for the normal RSA, we have $p - q > cN^{\frac{1}{2}}$ with a constant $c > 0$. So let

$$|z| < \min\left(\frac{c}{3}N^{\frac{1}{4}-\alpha_0}\frac{Y}{X}, N^{\frac{1}{4}-\frac{\alpha_0}{2}}\right),$$

and put

$$X_0 = \left\lfloor \frac{1}{2}N^{\frac{1}{4}-\frac{\alpha_0}{2}} \right\rfloor.$$

We want to estimate

$$\#\mathcal{E}(1, 1) = \sum_{X=1}^{X_0} \sum_{\substack{Y=1 \\ \gcd(X,Y)=1}}^{X-1} |z|.$$

Taking $|z| < \frac{c}{3}N^{\frac{1}{4}-\alpha_0}\frac{Y}{X}$, we get

$$\#\mathcal{E}(1, 1) = \frac{c}{3}N^{\frac{1}{4}-\alpha_0} \sum_{X=1}^{X_0} \sum_{\substack{Y=1 \\ \gcd(X,Y)=1}}^{X-1} \frac{Y}{X} = \frac{c}{6}N^{\frac{1}{4}-\alpha_0} \sum_{X=1}^{X_0} \phi(X), \tag{B.12}$$

where we used the well known identity

$$\sum_{\substack{Y=1 \\ \gcd(X,Y)=1}}^{X-1} Y = \frac{1}{2}X\phi(X).$$

Similarly, taking $|z| < N^{\frac{1}{4}-\frac{\alpha_0}{2}}$, we get

$$\#\mathcal{E}(1,1) = N^{\frac{1}{4}-\frac{\alpha_0}{2}} \sum_{X=1}^{X_0} \sum_{\substack{Y=1 \\ \gcd(X,Y)=1}}^{X-1} 1 = N^{\frac{1}{4}-\frac{\alpha_0}{2}} \sum_{X=1}^{X_0} \phi(X). \tag{B.13}$$

We can rewrite (B.12) and (B.13) in a single expression

$$\#\mathcal{E}(1,1) = N^{\frac{1}{4}-\varepsilon_0} \sum_{X=1}^{X_0} \phi(X),$$

for a suitable $\varepsilon_0 > 0$. It is well known (see Theorem 328 of [57]), that

$$\phi(X) > \frac{CX}{\log\log X},$$

where C is a positive constant. Since $X < N$, then $\phi(X) > XN^{-\varepsilon_1}$ for a small positive constant ε_1. From this, we deduce

$$\#\mathcal{E}(1,1) > N^{\frac{1}{4}-\varepsilon_0-\varepsilon_1} \sum_{X=1}^{X_0} X > N^{\frac{1}{4}-\varepsilon_0-\varepsilon_1} \frac{X_0^2}{2} > \frac{1}{8} N^{\frac{3}{4}-\alpha_0-\varepsilon_0-\varepsilon_1},$$

where we used $X_0 \approx \frac{1}{2} N^{\frac{1}{4}-\frac{\alpha_0}{2}}$. We get finally $\#\mathcal{E}(1,1) > N^{\frac{3}{4}-\varepsilon}$, with a constant $\varepsilon \approx \alpha_0+\varepsilon_0+\varepsilon_1$ depending only on N. This terminates the proof. □

B.6 Conclusion

In this paper, we showed how to perform three attacks on RSA using the ratio of the primes. The attacks apply when the public key (N, e) satisfies an equation $eX - (N - (ap + bq))Y = Z$ with suitably small parameters X, Y and Z where $\frac{a}{b}$ is an unknown convergent of $\frac{q}{p}$ with $a \geq 1$. The attacks combine a variety of techniques, including continued fractions, Coppersmith's lattice based method and H.W. Lenstra's Elliptic Curve Method for Factoring (ECM). Our results illustrate once again the fact that we should be very cautious when using RSA with specific exponents. Moreover, we showed that the number of such exponents is at least $N^{\frac{3}{4}-\varepsilon}$. Using the notion of weak keys, as defined by Blömer and May [13], the results of this paper show that this set of RSA public keys is a class of weak keys.

Appendix C

A New Attack on RSA with Two or Three Decryption Exponents

Journal of Applied Mathematics and Computing 2013
[111]

Abstract :
Let $N = pq$ be an RSA modulus, i.e. the product of two large unknown primes of equal bit-size. In this paper, we describe an attack on RSA in the presence of two or three exponents e_i with the same modulus N and satisfying equations $e_i x_i - \phi(N) y_i = z_i$, where $\phi(N) = (p-1)(q-1)$ and x_i, y_i, z_i are unknown parameters. The new attack is an extension of Guo's continued fraction attack as well as the Blömer and May lattice-reduction basis attack.

C.1 Introduction

The RSA public-key cryptosystem was invented by Rivest, Shamir, and Adleman [131] in 1978. Since then, the RSA system has been the best known and most widely accepted public key cryptosystem. Encryption and decryption in RSA each requires an exponentiation modulo a large modulus N which is the product of two large primes, p and q. The exponents in the exponentiations are the public exponent e for encryption and the private exponent d for decryption. The exponents e and d are related by the equation $ed - k\phi(N) = 1$ for some positive integer k where $\phi(N) = (p-1)(q-1)$ is the Euler totient function of N. To reduce the decryption time or signature generation, it may be tempting to use a small private exponent d. Unfortunately, based on the convergents of the continued fraction expansion of $\frac{e}{N}$, Wiener [147] showed that the RSA system can be totally broken if $d < \frac{1}{3}N^{\frac{1}{4}}$. Then, in 1999, based on lattice basis reduction, Boneh and Durfee [17] proposed a new attack on the use of short secret exponents. They showed that the RSA system can be totally broken if $d < N^{0.292}$. In 1994, Blömer and May [13] proposed a different attack on RSA with a public exponent e satisfying the equation $ex + y = k\phi(N)$ with $x < \frac{1}{3}N^{\frac{1}{4}}$ and $|y| < \mathcal{O}\left(N^{-\frac{3}{4}}ex\right)$. This attack combines the convergents of the continued fraction expansion of $\frac{e}{N}$ and the seminal work of Coppersmith [34] for solving bivariate polynomial equations.

In 1999, Guo (see [68]) proposed an attack on RSA when there are two or more instances of RSA, having the same modulus, with public exponents e_i, $i = 1, 2, \ldots$. The attack is based on the continued fraction algorithm and can be used to factor the modulus if the private exponents d_i are each less than $N^{\frac{1}{3}-\varepsilon}$ for some $\varepsilon > 0$. In 1999, Howgrave-Graham and Seifert [68] proposed an extension of Guo's attack that allows the RSA system to be broken in the presence of two decryption exponents (d_1, d_2) with $d_1, d_2 < N^{\frac{5}{14}}$. In the presence of three decryption exponents, Howgrave-Graham and Seifert improved the bound to $N^{\frac{2}{5}}$. The attack of Howgrave-Graham and Seifert is based on lattice reduction methods. Very recently, Sarkar and Maitra [133] used a different lattice based technique and improved the bound $N^{\frac{5}{14}}$ for the case of two decryption exponents up to $N^{0.416}$. In [133], Sarkar and Maitra proposed a generalized attack when $n \geq 2$ many decryption exponents d_i are

used with the same RSA modulus N and $d_i < N^{\frac{3n-1}{4n+4}}$ for each i, $1 \leq i \leq n$.

In this paper, we combine the attack of Guo and the attack of Blömer and May to mount a new attack on RSA with two or three decryption exponents and a common modulus. Let $N = pq$ be an RSA modulus with $q < p < 2q$ and e_i, $i = 1, 2, \ldots$, be two or three public exponents. Assume that each exponent satisfies an equation $e_i x_i - \phi(N) y_i = z_i$. We show, that, depending on certain inequalities verified by the parameters x_i, y_i, z_i, one can find the factorization of the RSA modulus N. The new approach still uses the continued fraction algorithm and the lattice-reduction basis technique of Coppersmith [34].

The rest of this paper is organized as follows. In Section 2 we present the attack of Guo as well as the attack of Blömer and May. In Section 3, we prove three lemmas to be used in our new approach. We present the new approach for two exponents in Section 4 and for three exponents in Section 5. We conclude the paper in Section 6.

C.2 Former Attacks

Since the motivation for our new attack originates from Guo's continued fraction attack and the Blömer and May lattice attack, we revisit these attacks in this section.

C.2.1 Guo's attack for two exponents

Guo's attack was described in [68] by Howgrave-Graham and Seifert (see also [62]). It is based on the continued fraction algorithm and makes use of the following result (see [57], Theorem 184).

Theorem C.2.1 (Legendre)**.** *Let ξ be a real number. If a and b are coprime integers such that*

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

then $\frac{a}{b}$ is a convergent of the continued fraction expansion of ξ.

Guo's attack concerns at least two public exponents e_1, e_2 such that $e_1d_1 - k_1\phi(N) = 1$ and $e_2d_2 - k_2\phi(N) = 1$, where $\phi(N) = (p-1)(q-1)$. Eliminating $\phi(N)$, we find the equation $e_1d_1k_2 - e_2d_2k_1 = k_2 - k_1$. Dividing by $e_2d_1k_2$, we get

$$\left|\frac{e_1}{e_2} - \frac{d_2k_1}{d_1k_2}\right| = \frac{|k_2 - k_1|}{e_2d_1k_2}.$$

Hence, if $2|k_2 - k_1|d_1k_2 < e_2$, then

$$\frac{|k_2 - k_1|}{e_2d_1k_2} < \frac{1}{2(d_1k_2)^2}.$$

Thus, by Theorem C.2.1, $\frac{d_2k_1}{d_1k_2}$ must be one of the convergents of the continued fraction of $\frac{e_1}{e_2}$. Moreover, if d_1 and d_2 are bounded, $d_1 < N^\delta$, $d_2 < N^\delta$ say, then k_1 and k_2 are also bounded since for $e_i < \phi(N)$ we have

$$k_i = \frac{e_id_i - 1}{\phi(N)} < \frac{e_id_i}{\phi(N)} < d_i.$$

It follows that the condition $2|k_2 - k_1|d_1k_2 < e_2$ reduces to $2N^{3\delta} < N$, or equivalently $\delta < \frac{1}{3} - \varepsilon$, where ε is a small positive constant.

In practice, Guo's attack is effective if one can find d_1 or d_2 using the convergent $\frac{d_2k_1}{d_1k_2}$. This means that the quantities d_i, k_i, $i = 1, 2$, must satisfy $\gcd(d_1k_2, d_2k_1) = 1$. Moreover, it is necessary to factor d_1k_2 or d_2k_1. Since $k_i < d_i < N^\delta$, $i = 1, 2$, then $\max(d_1k_2, d_2k_1) < N^{2\delta} < N^{\frac{2}{3}}$. Depending on the structure of the quantities d_i and k_i, $i = 1, 2$, the numbers d_1k_2 and d_2k_1 are not expected to be of a difficult factorization shape and can be factored easily. Using the exact values of d_1 and k_1 in $e_1d_1 - k_1\phi(N) = 1$, this gives the factorization of N.

Later, using lattice based techniques, Howgrave-Graham and Seifert [68] increased the bound up to $d_1, d_2 < N^{\frac{5}{14}}$. This bound was recently improved to $d_1, d_2 < N^{0.416}$ by Sarkar and Maitra [133].

C.2.2 Guo's attack for three exponents

To avoid the factorization problem, Guo proposed to use three exponents. Consider that three public exponents e_1, e_2, e_3 satisfying the key equations

$$e_1d_1 - k_1\phi(N) = 1, \qquad e_2d_2 - k_2\phi(N) = 1, \qquad e_3d_3 - k_3\phi(N) = 1,$$

satisfy also the inequalities $2|k_2 - k_1|d_1k_2 < e_2$ and $2|k_3 - k_1|d_1k_3 < e_3$. Combining the key equations, we get

$$e_1d_1k_2 - e_2d_2k_1 = k_2 - k_1, \qquad e_1d_1k_3 - e_3d_3k_1 = k_3 - k_1.$$

Proceeding as in Guo's first attack, we find the inequalities

$$\left|\frac{e_1}{e_2} - \frac{d_2k_1}{d_1k_2}\right| = \frac{|k_2 - k_1|}{e_2d_1k_2} < \frac{1}{2(d_1k_2)^2},$$
$$\left|\frac{e_1}{e_3} - \frac{d_3k_1}{d_1k_3}\right| = \frac{|k_3 - k_1|}{e_3d_1k_3} < \frac{1}{2(d_1k_3)^2}.$$

Using Theorem C.2.1, we see that $\frac{d_2k_1}{d_1k_2}$ is one of the convergents of the continued fraction of $\frac{e_1}{e_2}$ and similarly, $\frac{d_3k_1}{d_1k_3}$ is one of the convergents of the continued fraction of $\frac{e_1}{e_3}$. Suppose in addition that $\gcd(d_2k_1, d_1k_2) = 1$, $\gcd(d_3k_1, d_1k_3) = 1$, $\gcd(d_2, d_3) = 1$ and $\gcd(k_2, k_3) = 1$, then $\frac{d_2k_1}{d_1k_2}$ and $\frac{d_3k_1}{d_1k_3}$ are in lowest terms and

$$\gcd(d_1k_2, d_1k_3) = d_1, \qquad \gcd(d_2k_1, d_3k_1) = k_1.$$

With d_1 and k_1 known, the factorization of N becomes trivial using the equation $e_1d_1 - k_1\phi(N) = 1$.

In 1999, it was shown by Howgrave-Graham and Seifert [68] that the bound $d_i < N^{\frac{1}{3}}$ with $i = 1, 2, 3$ can be improved using lattice reduction techniques and very recently, Sarkar and Maitra [134] increased this bound up to $d_i < N^{\frac{1}{2}}$.

C.2.3 The Blömer and May attack

In 2004, Blömer and May [13] proposed an attack on RSA with a modulus $N = pq$ with $q < p < 2q$ and a public exponent e satisfying an equation $ex + y = k\phi(N)$. The attack is based on a combination of the continued fraction algorithm and Coppersmith's lattice-based technique for finding small roots of bivariate polynomial equation [34].

Theorem C.2.2 (Coppersmith)**.** *Let $N = pq$ be the product of two unknown primes such that $q < p < 2q$. Suppose we know an approximation $\widetilde{P}$ of p such that $\left|p - \widetilde{P}\right| < 2N^{\frac{1}{4}}$. Then N can be factored in polynomial time.*

Suppose that e satisfies an equation $ex + y = k\phi(N)$. Under the conditions

$$0 < x < \frac{1}{3}N^{\frac{1}{4}} \quad \text{and} \quad |y| = |ex - k\phi(N)| \leq \mathcal{O}\left(N^{-\frac{3}{4}}ex\right), \tag{C.1}$$

the fraction $\frac{k}{x}$ satisfies $\left|\frac{k}{x} - \frac{e}{N}\right| < \frac{1}{2x^2}$. By Theorem C.2.1, this shows that $\frac{k}{x}$ can be found among the convergents of the continued fraction expansion of $\frac{e}{N}$. Using $\phi(N) = (p-1)(q-1) = N+1-p-q$ in the equation $ex+y = k\phi(N)$, Blömer and May showed that $N+1-\frac{ex}{k}$ is an approximation of $p+q$ satisfying

$$\left|p + q - \left(N + 1 - \frac{ex}{k}\right)\right| = \frac{|y|}{k} < \frac{4}{3}cN^{\frac{1}{4}},$$

where $c < 1$ is a positive constant satisfying $p - q > cN^{\frac{1}{2}}$. Next, they derived that $\sqrt{\left(N + 1 - \frac{ex}{k}\right)^2 - 4N}$ is an approximation of $p - q$ up to an error term at most $9N^{\frac{1}{4}}$. Finally, combining the approximations of $p+q$ and $p-q$, they found an approximation of p up to an error term of at most $6N^{\frac{1}{4}}$ which leads to the exact value of p using Coppersmith's Theorem C.2.2.

C.3 Useful Lemmas

In this section we state and prove three lemmas needed for the new attack. The first is the following result.

Lemma C.3.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Suppose that S is an approximation of $p + q$ satisfying $S > 2\sqrt{N}$ and*

$$|p + q - S| < \frac{D}{S}N^{\frac{1}{4}}, \tag{C.2}$$

where $D = \sqrt{S^2 - 4N}$. Then $\widetilde{P} = \frac{1}{2}(S + D)$ is an approximation of p with $\left|p - \widetilde{P}\right| < 2N^{\frac{1}{4}}$.

Proof. Suppose that $S > 2\sqrt{N}$ is a positive integer satisfying (C.2) where

$D = \sqrt{S^2 - 4N}$. We have

$$\begin{aligned}
\left|D^2 - (p-q)^2\right| &= \left|\left|S^2 - 4N\right| - (p-q)^2\right| \\
&\leq \left|S^2 - 4N - (p-q)^2\right| \\
&= \left|S^2 - (p+q)^2\right| \\
&= (p+q+S)\,|p+q-S| \\
&\leq (p+q+S) \times \frac{DN^{\frac{1}{4}}}{S}.
\end{aligned}$$

Dividing by $p - q + D$, we get

$$|p-q-D| \leq \frac{p+q+S}{p-q+D} \times \frac{DN^{\frac{1}{4}}}{S}. \tag{C.3}$$

Let us find a bound for $\frac{p+q+S}{p-q+D}$. Since $D < S$, then from (C.2), we derive

$$p+q+S < 2S + \frac{DN^{\frac{1}{4}}}{S} < 2S + N^{\frac{1}{4}} < 3S.$$

On the other hand, we have $p - q + D > D$. Hence

$$\frac{p+q+S}{p-q+D} < \frac{3S}{D}$$

Plugging in (C.3), we deduce

$$|p-q-D| \leq \frac{3S}{D} \times \frac{DN^{\frac{1}{4}}}{S} = 3N^{\frac{1}{4}}. \tag{C.4}$$

Now, using (C.2) and (C.4), we get

$$\begin{aligned}
|2p - S - D| &= |p+q-S+(p-q-D)| \\
&\leq |p+q-S| + |p-q-D| \\
&< \frac{DN^{\frac{1}{4}}}{S} + 3N^{\frac{1}{4}} \\
&< 4N^{\frac{1}{4}}.
\end{aligned}$$

Dividing by 2, we find

$$\left|p - \frac{S+D}{2}\right| = \left|p - \tilde{P}\right| < 2N^{\frac{1}{4}},$$

which terminates the proof. □

Notice that when the primes p and q satisfy $q < p < 2q$, then $p + q > 2\sqrt{N}$ and if S is an approximation of $p + q$, then S also satisfies $S > 2\sqrt{N}$.

The second lemma is the following.

Lemma C.3.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let e_1, e_2 be integers satisfying the equations*

$$e_1x_1 - \phi(N)y_1 = z_1, \qquad e_2x_2 - \phi(N)y_2 = z_2.$$

If $2x_1y_2|z_1y_2 - z_2y_1| < e_2$ then $\frac{x_2y_1}{x_1y_2}$ is a convergent of $\frac{e_1}{e_2}$.

Proof. Suppose that e_1, e_2 satisfy the equations $e_1x_1 - \phi(N)y_1 = z_1$ and $e_2x_2 - \phi(N)y_2 = z_2$. Then eliminating $\phi(N)$, we get

$$e_1x_1y_2 - e_2x_2y_1 = z_1y_2 - z_2y_1.$$

Dividing both sides by $e_2x_1y_2$, we get

$$\left|\frac{e_1}{e_2} - \frac{x_2y_1}{x_1y_2}\right| = \frac{|z_1y_2 - z_2y_1|}{e_2x_1y_2}. \tag{C.5}$$

Suppose that the parameters satisfy the inequality $2x_1y_2|z_1y_2 - z_2y_1| < e_2$. Then (C.5) yields

$$\left|\frac{e_1}{e_2} - \frac{x_2y_1}{x_1y_2}\right| < \frac{1}{2(x_1y_2)^2}.$$

Combining with Theorem C.2.1, we see that $\frac{x_2y_1}{x_1y_2}$ is a convergent of $\frac{e_1}{e_2}$. □

Finally, we will use the following result.

Lemma C.3.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let e_1 be an integer satisfying the equation $e_1x_1 - \phi(N)y_1 = z_1$, with known positive parameters x_1, y_1. Let*

$$S = N + 1 - \frac{e_1x_1}{y_1} \quad \text{and} \quad D = \sqrt{|S^2 - 4N|}.$$

Then, under the conditions $S > 2\sqrt{N}$ and

$$|z_1| \ < \frac{D}{S}N^{\frac{1}{4}}y_1,$$

N can be factored in polynomial time.

Proof. Suppose that e_1 satisfies the conditions of the theorem where x_1, y_1 are known positive integers. Using $\phi(N) = N + 1 - p - q$ in the equation $e_1x_1 - \phi(N)y_1 = z_1$, we get

$$\left| p + q - \left(N + 1 - \frac{e_1x_1}{y_1} \right) \right| = \frac{|z_1|}{y_1} < \frac{D}{S} N^{\frac{1}{4}}.$$

This implies that $S = N+1-\frac{e_1x_1}{y_1}$ is an approximation of $p+q$ up to an error term satisfying the condition of Lemma C.3.1. Hence $\widetilde{P} = \frac{1}{2}\left(S + \sqrt{|S^2 - 4N|}\right)$ is an approximation of p up to an error term at most $2N^{\frac{1}{4}}$. Thus, using Coppersmith's Theorem C.2.2, one can find p in polynomial time.

□

C.4 The New Attack on RSA with Two Exponents

In this section, we investigate RSA with the same modulus and two public exponents e_1 and e_2 satisfying the equations $e_1x_1 - \phi(N)y_1 = z_1$, and $e_2x_2 - \phi(N)y_2 = z_2$, where the parameters satisfy

$$\gcd(x_2y_1, x_1y_2) = 1, \tag{C.6}$$

$$x_1y_2|z_1y_2 - z_2y_1| < \frac{e_2}{2}. \tag{C.7}$$

This means that the conditions of Lemma C.3.2 are satisfied which implies that $\frac{x_2y_1}{x_1y_2}$ can be found in the continued fraction expansion of $\frac{e_1}{e_2}$. The condition (C.6) implies that the convergent $\frac{x_2y_1}{x_1y_2}$ is in lowest terms which gives x_2y_1 and x_1y_2. Now, we wish to recover the values of the parameters x_1, y_1, x_2, y_2. Using the assumptions that x_1, y_2, $|z_1y_2 - z_2y_1|$ are at most N^{δ} and that $e_2 \approx N$, the condition (C.7) is satisfied whenever $N^{3\delta} < \frac{1}{2}N$, that is $\delta = \frac{1}{3} - \varepsilon$, for some small $\varepsilon > 0$. Moreover, if x_2y_1 and x_1y_2 are not of a difficult factorization shape, then their factorization is feasible for instances of RSA with a 1024-bit modulus. Thus, factoring x_2y_1 and x_1y_2 will reveal the parameters x_1, y_1. To find the prime factors p and q of the RSA modulus $N = pq$, we must make the assumption that the parameter z_1 satisfies

$$|z_1| \ < \frac{D}{S} N^{\frac{1}{4}} y_1,$$

where $S = N + 1 - \frac{e_1 x_1}{y_1}$ and $D = \sqrt{|S^2 - 4N|}$. Thus, the conditions of Lemma C.3.3 are satisfied which leads to the factorization of N.

We summarize the first attack in Algorithm 1.

Algorithm 4 Two exponents

Require: $N = pq$ with $q < p < 2q$, two public exponents e_i, $i = 1, 2$ satisfying $e_i x_i - \phi(N) y_i = z_i$ with unknown parameters x_i, y_i, z_i.
Ensure: The prime factors p and q.
  Compute the continued fraction expansion of $\frac{e_1}{e_2}$.
  for every convergent $\frac{p_k}{q_k}$ of $\frac{e_1}{e_2}$ with $\max(p_k, q_k) < N^{\frac{2}{3}}$ **do**
    Factor p_k and q_k.
    for every divisor y_1 of p_k **do**
      for every divisor x_1 of q_k **do**
        Compute $S = N + 1 - \frac{e_1 x_1}{y_1}$ and $\widetilde{P} = \frac{1}{2}\left(S + \sqrt{|S^2 - 4N|}\right)$.
        Apply Coppersmith's algorithm (Theorem C.2.2) with $\widetilde{P}$ as an approximation of p.

        if Coppersmith's algorithm outputs the factorization of N, **then**
          stop.
        end if
      end for
    end for
  end for

An Example for the New Attack with Two Exponents

As an example, let us take

$$\begin{aligned} N &= 78783023222142579402299, \\ e_1 &= 20339472065400293617, \\ e_2 &= 16071808231974749459. \end{aligned}$$

The first 30 partial quotients of $\frac{e_1}{e_2}$ are

$$\begin{aligned} [1,\ 3,\ 1,\ 3,\ 3,\ 1,\ 2,\ 59,\ 1,\ 2,\ 2,\ 2,\ 1,\ 1,\ 3,\ 1,\ 1, \\ 4,\ 3,\ 1,\ 7,\ 18,\ 10,\ 1,\ 13,\ 1,\ 1,\ 316,\ 4,\ 1,\ \cdots] \end{aligned}$$

Each convergent $\frac{a}{b}$ is a candidate for $\frac{x_2 y_1}{x_1 y_2}$. The 27th convergent is

$$\frac{a}{b} = \frac{3889559329731}{3073445144167}$$

We see that $a \approx N^{0.550}$, $b \approx N^{0.546}$ are not of difficult factorization shape. We get easily $a = 3^3 \cdot 229 \cdot 6079 \cdot 103483$ and $b = 41 \cdot 43 \cdot 71 \cdot 1693 \cdot 14503$ and we see that the largest prime factor is $103483 \approx N^{0.22}$. Next, the decomposition

$$a = x_2 y_1 = 71092821 \cdot 54711, \qquad b = x_1 y_2 = 211917889 \cdot 14503,$$

gives $x_1 = 211917889$, $y_1 = 54711$, $x_2 = 71092821$ and $y_2 = 14503$. Using the equation $e_1 x_1 - \phi(N) y_1 = z_1$, we get $p + q = N + 1 - \frac{e_1 x_1}{y_1} + \frac{z_1}{y_1}$. We then make the assumption that

$$p + q \approx S = N + 1 - \frac{e_1 x_1}{y_1} \approx 594807230437.$$

From this, we get the approximation

$$p - q \approx D = \sqrt{|S^2 - 4N|} \approx 196630487186.$$

Combining the approximations of $p + q$ and $p - q$, we get

$$p \approx \frac{S + D}{2} \approx 395718858812.$$

Then Coppersmith's algorithm C.2.2 gives $p = 395718860549$ and then we get $q = 199088370751$.

Note that, in this example, the private exponents $d_i \equiv e_i^{-1} \pmod{\phi(N)}$, $i = 1, 2$, are

$$\begin{aligned} d_1 &= 63426822067770650216953 \approx N^{0.996}, \\ d_2 &= 68134122111136587656939 \approx N^{0.997}, \end{aligned}$$

so that $d_1, d_2 > N^{\frac{1}{2}}$, which explains why Guo's attack would fail in this case. On the other hand, in connection with the attack of Blömer and May as described in Subsection C.2.3, the fraction $\frac{y_1}{x_1}$ is not among the convergents of the continued fraction of $\frac{e_1}{N}$. Similarly, $\frac{y_2}{x_2}$ is not among the convergents of the continued fraction of $\frac{e_2}{N}$. Moreover, all the convergents $\frac{x}{k}$ of $\frac{e_1}{N}$ with $x < \frac{1}{3} N^{\frac{1}{4}}$ are such that $|e_1 x - k\phi(N)| > N^{-\frac{3}{4}} e_1 x$ so that the condition (C.1) is never satisfied. We have a similar result with the convergents of $\frac{e_2}{N}$. This explains why Blömer and May's attack would also fail in this case.

C.5 The New Attack on RSA with Three Exponents

To avoid factoring integers of size $N^{\frac{2}{3}}$, we consider in this section that a third instance of RSA with the same modulus is available. Suppose we have three public exponents e_1, e_2, e_3 satisfying the equations

$$e_1x_1 - \phi(N)y_1 = z_1, \qquad e_2x_2 - \phi(N)y_2 = z_2, \qquad e_3x_3 - \phi(N)y_3 = z_3,$$

where the parameters satisfy $\gcd(x_2y_1, x_1y_2) = 1$, $\gcd(x_3y_1, x_1y_3) = 1$ and

$$\begin{aligned} x_1y_2 \quad & |z_1y_2 - z_2y_1| < \frac{e_2}{2}, \\ x_1y_3 \quad & |z_1y_3 - z_3y_1| < \frac{e_3}{2}. \end{aligned}$$

This immediately shows that the conditions of Lemma C.3.2 are satisfied for (e_1, e_2) and for (e_1, e_3). Hence $\frac{x_2y_1}{x_1y_2}$ is in lowest terms and is a convergent of $\frac{e_1}{e_2}$. Similarly, $\frac{x_3y_1}{x_1y_3}$ is in lowest terms and is a convergent of $\frac{e_1}{e_3}$. This gives

$$\gcd(x_1y_2, x_1y_3) = x_1, \qquad \gcd(x_2y_1, x_3y_1) = y_1.$$

Now, if the condition

$$|z_1| \ < \frac{D}{S}N^{\frac{1}{4}}y_1,$$

is satisfied, where $S = N + 1 - \frac{e_1x_1}{y_1}$ and $D = \sqrt{|S^2 - 4N|}$, then by Lemma C.3.3 one can find p using Coppersmith's Theorem with the approximation

$$\widetilde{P} = \frac{1}{2}\left(S + \sqrt{|S^2 - 4N|}\right).$$

of p.

We summarize the attack in Algorithm 2

An Example for the New Attack with Three Exponents

Here we take an RSA modulus $N = pq$ and three public exponents e_1, e_2 and

Algorithm 5 Three Exponents

Require: $N = pq$ with $q < p < 2q$, three public exponents e_i, $i = 1, 2, 3$, satisfying $e_i x_i - \phi(N) y_i = z_i$ with unknown parameters x_i, y_i, z_i.
Ensure: The prime factors p and q.
  Compute the continued fraction expansion of $\frac{e_1}{e_2}$.
  Compute the continued fraction expansion of $\frac{e_1}{e_3}$.
  for every convergent $\frac{a}{b}$ of $\frac{e_1}{e_2}$ **do**
    for every convergent $\frac{c}{d}$ of $\frac{e_1}{e_3}$ **do**
      Compute $x_1 = \gcd(b, d)$, $y_1 = \gcd(a, c)$.
      Compute $S = N + 1 - \frac{e_1 x_1}{y_1}$ and $\widetilde{P} = \frac{1}{2}\left(S + \sqrt{|S^2 - 4N|}\right)$.
      Apply Coppersmith's algorithm (Theorem C.2.2) with $\widetilde{P}$ as an approximation of p.
      if Coppersmith's algorithm outputs the factorization of N, **then**
        stop.
      end if
    end for
  end for

e_3 as

$$\begin{aligned} N &= 95026423511070214659367, \\ e_1 &= 988283832402044225959, \\ e_2 &= 35887685050144510339, \\ e_3 &= 4465685820126103902929. \end{aligned}$$

The candidates for $\frac{x_2 y_1}{x_1 y_2}$ are the convergents of $\frac{e_1}{e_2}$. Indeed, the 25th convergent of $\frac{e_1}{e_2}$ is $\frac{44398785042941}{1612259112200}$. Similarly, the candidates for $\frac{x_3 y_1}{x_1 y_3}$ are the convergents of $\frac{e_1}{e_3}$. The 35th convergent of $\frac{e_1}{e_3}$ is $\frac{6433869008153}{29072252986700}$. From the two convergents, we get

$$\begin{aligned} x_1 &= \gcd(1612259112200, 29072252986700) = 59365900, \\ y_1 &= \gcd(44398785042941, 6433869008153) = 617411. \end{aligned}$$

Using the equation $e_1 x_1 - \phi(N) y_1 = z1$, we get $p + q = N + 1 - \frac{e_1 x_1}{y_1} + \frac{z_1}{y_1}$, and neglecting $\frac{z_1}{y_1}$, we get

$$p + q \approx S = N + 1 - \frac{e_1 x_1}{y_1} \approx 642772787002.$$

From this, we get the approximation

$$p - q \approx D = \sqrt{|S^2 - 4N|} \approx 181799784560.$$

Using the approximations S and D, we get $p \approx \frac{S+D}{2} \approx 412286285781$. Finally, applying Coppersmith's Theorem C.2.2, we get

$$p = 412286285849, \qquad q = \frac{N}{p} = 230486501183.$$

We notice that, for $i = 1, 2, 3$, the integers d_i related to e_i by the relations $e_i d_i \equiv 1 \pmod{\phi(N)}$ satisfy $d_i > N^{0.98}$ which is far from Guo's upper bound $N^{\frac{1}{3}}$ as described in Subsection C.2.2. This shows that Guo's method would fail here. On the other hand, for $i = 1, 2, 3$, the convergents of the rational numbers $\frac{e_i}{N}$ are all different from the expected convergents $\frac{y_i}{x_i}$. Moreover, for $i = 1, 2, 3$, the conditions (C.1) are not satisfied by the convergents of $\frac{e_i}{N}$. This shows that the method of Blömer and May would also fail in this case.

C.6 Conclusion

In this paper, we have presented a new attack on RSA with the same modulus $N = pq$ and two or three exponents satisfying equations $e_i x_i - \phi(N) y_i = z_i$ with specific unknown parameters x_i, y_i, z_i. Our attack is an extension of Guo's attack as well as an extension of the Blömer and May attack. The new attack enables us to find p and q efficiently with two exponents and in polynomial time with three exponents. This proves once again that, under some conditions, RSA is insecure even when the private exponents are sufficiently large.

Appendix D

An Attack on RSA Using LSBs of Multiples of the Prime Factors

AFRICACRYPT 2013
[112]

Abstract :
Let $N = pq$ be an RSA modulus with a public exponent e and a private exponent d. Wiener's famous attack on RSA with $d < N^{0.25}$ and its extension by Boneh and Durfee to $d < N^{0.292}$ show that using a small d makes RSA completely insecure. However, for larger d, it is known that RSA can be broken in polynomial time under special conditions. For example, various partial key exposure attacks on RSA and some attacks using additional information encoded in the public exponent e are efficient to factor the RSA modulus. These attacks were later improved and extended in various ways. In this paper, we present a new attack on RSA with a public exponent e satisfying an equation $ed - k(N + 1 - ap - bq) = 1$ where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$. We show that RSA is insecure when certain amount of the Least Significant Bits (LSBs) of ap and bq are known. Further, we show that the existence of good approximations $\frac{a}{b}$ of $\frac{q}{p}$ with small a and b

substantially reduces the requirement of LSBs of ap and bq.

D.1 Introduction

The RSA cryptosystem was invented by Rivest, Shamir and Adleman [131] in 1977 and is today's most important public-key cryptosystem. The standard notations in RSA are as follows:

- p and q are two large primes of the same bit size.
- $N = pq$ is the RSA modulus and $\phi(N) = (p-1)(q-1)$ is Euler's totient function.
- e and d are respectively the public and the private exponents and satisfy $ed - k\phi(N) = 1$ for some positive integer k.

There have been a large number of attacks on RSA. Some attacks, called small private key attacks can break RSA in polynomial time when the private key is small. For example, Wiener [147] showed that if the private key satisfies $d < \frac{1}{3}N^{\frac{1}{4}}$, then N can be factored and Boneh and Durfee [17] showed that RSA is insecure if $d < N^{0.292}$. Some attacks, called partial key exposure attacks exploit the knowledge of a portion of the private exponent or of one of the prime factors. Partial key exposure attacks are mainly motivated by using side-channel attacks, such as fault attacks, power analysis and timing attacks ([76], [77]). Using a side-channel, an attacker can expose a part of one of the modulus prime factors p or q or of the private key d. In 1998, Boneh, Durfee and Frankel [18] presented several partial key exposure attacks on RSA with a public key $e < N^{1/2}$ where the attacker requires knowledge of most significant bits (MSBs) or least significant bits (LSBs) of the private exponent d. In [13], Ernest et al. [44] proposed several partial key exposure attacks that work for $e > N^{1/2}$. Notice that Wiener's attack [147] and the attack of Boneh and Durfee [17] can be seen as partial key exposure attacks because the most significant bits of the private exponent are known and are equal to zero. Sometimes, it is possible to factor the RSA modulus even if the private key is large and no bits are exposed. Such attacks exploit the

knowledge of special conditions verified by the modulus prime factors or by the exponents. In 2004, Blömer and May [13] showed that RSA can be broken if the public exponent e satisfies an equation $ex = y + k\phi(N)$ with $x < \frac{1}{3}N^{\frac{1}{4}}$ and $|y| < N^{-\frac{3}{4}}ex$. At Africacrypt 2009, Nitaj [107] presented an attack when the exponent e satisfies an equation $eX - (N - (ap + bq))Y = Z$ with the constraints that $\frac{a}{b}$ is an unknown convergent of the continued fraction expansion of $\frac{q}{p}$, $1 \leq Y \leq X < \frac{1}{2}\frac{N^{\frac{1}{4}}}{\sqrt{a}}$, $\gcd(X, Y) = 1$, and Z depends on the size of $|ap - bq|$. Nitaj's attack combines techniques from the theory of continued fractions, Coppersmith's method [34] for finding small roots of bivariate polynomial equations and the Elliptic Curve Method [84] for integer factorization.

In this paper we revisit Nitaj's attack by studying the generalized RSA equation $ed - k(N + 1 - ap - bq) = 1$ with different constraints using Coppersmith's method [34] only. We consider the situation when an amount of LSBs of ap and bq are exposed where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$, that is when $a = \left[\frac{bq}{p}\right]$. More precisely, assume that $ap = 2^{m_0}p_1 + p_0$ and $bq = 2^{m_0}q_1 + q_0$ where m_0, p_0 and q_0 are known to the attacker. We show that one can factor the RSA modulus if the public key e satisfies an equation $ed_1 - k_1(N + 1 - ap - bq) = 1$ where $e = N^{\gamma}$, $d_1 < N^{\delta}$, $2^{m_0} = N^{\beta}$ and $a < b < N^{\alpha}$ satisfy

$$\delta \leq \begin{cases} \delta_1 & \text{if} \quad \gamma \geq \frac{1}{2}(1 + 2\alpha - 2\beta), \\ \delta_2 & \text{if} \quad \gamma < \frac{1}{2}(1 + 2\alpha - 2\beta). \end{cases}$$

with

$$\begin{aligned} \delta_1 &= \frac{7}{6} + \frac{1}{3}(\alpha - \beta) - \frac{1}{3}\sqrt{4(\alpha - \beta)^2 + 4(3\gamma + 1)(\alpha - \beta) + 6\gamma + 1}, \\ \delta_2 &= \frac{1}{4}(3 - 2(\alpha - \beta) - 2\gamma). \end{aligned}$$

We notice the following facts

- When $a = b = 1$, the equation becomes $ed_1 - k_1(N + 1 - p - q) = 1$ as in standard RSA.
- When $\gamma = 1$ and $\alpha = \beta$, the RSA instance is insecure if $d < \frac{7}{6} - \frac{\sqrt{7}}{3} \approx$

0.284. This is a well known boundary in the cryptanalysis of RSA (see e.g. [17]).

- When $\gamma = 1$ and $\beta = 0$, that is no LSBs of ap nor of bq are known, the RSA instance is insecure if $\delta < \frac{7}{6} + \frac{1}{3}\alpha - \frac{1}{3}\sqrt{\alpha^2 + 16\alpha + 7}$. This considerably improve the bound $\delta < \frac{1}{4}(1 - 2\alpha)$ of [107].
- The ANSI X9.31 standard [1] requires that the prime factors p and q shall not be near the ratio of two small integers. Our new attack shows that this requirement is necessary and can be easily checked once one has generated two primes simply by computing the convergents of the continued fraction expansion of $\frac{q}{p}$.

The rest of the paper is organized as follows. In Section 2 we review some basic results from lattice theory and their application to solve modular equations as well as two useful lemmas. In Section 3 we describe the new attack on RSA. In Section 4, we present various numerical experiments. Finally, we conclude in Section 5.

D.2 Preliminaries

D.2.1 Lattices

Let ω and n be two positive integers with $\omega \leq n$. Let $b_1, \cdots, b_\omega \in \mathbb{R}^n$ be ω linearly independent vectors. A lattice $\mathcal{L}$ spanned by $\{b_1, \cdots, b_\omega\}$ is the set of all integer linear combinations of $b_1, \cdots, b_\omega$, that is

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

The set $\langle b_1 \ldots, b_\omega \rangle$ is called a lattice basis for $\mathcal{L}$. The lattice dimension is $\dim(\mathcal{L}) = \omega$. We say that the lattice is full rank if $\omega = n$. If the lattice is full rank, then the determinant of $\mathcal{L}$ is equal to the absolute value of the determinant of the matrix whose rows are the basis vectors $b_1, \cdots, b_\omega$. In 1982, Lenstra, Lenstra and Lovász [86] invented the so-called LLL algorithm to reduce a basis and to find a short lattice vector in time polynomial in

the bit-length of the entries of the basis matrix and in the dimension of the lattice. The following lemma, gives bounds on LLL-reduced basis vectors.

Theorem D.2.1 (Lenstra, Lenstra, Lovász)**.** *Let $\mathcal{L}$ be a lattice of dimension ω. In polynomial time, the LLL- algorithm outputs two reduced basis vectors v_1 and v_2 that satisfy*

$$\|v_1\| \leq 2^{\frac{\omega}{2}} \det(\mathcal{L})^{\frac{1}{\omega}}, \quad \|v_2\| \leq 2^{\frac{\omega}{2}} \det(\mathcal{L})^{\frac{1}{\omega-1}}.$$

Using the LLL algorithm, Coppersmith [34] proposed a method to efficiently compute small roots of bivariate polynomials over the integers or univariate modular polynomials. Howgrave-Graham [65] gave a simple reformulation of Coppersmith's method in terms of the norm of the polynomial $f(x, y) = \sum a_{ij} x^i y^j$ which is defined by

$$\|f(x, y)\| = \sqrt{\sum a_{ij}^2}.$$

Theorem D.2.2 (Howgrave-Graham)**.** *Let $f(x, y) \in \mathbb{Z}[x, y]$ be a polynomial which is a sum of at most ω monomials. Suppose that $f(x_0, y_0) \equiv 0 \pmod{e^m}$ where $|x_0| < X$ and $|y_0| < Y$ and $\|f(xX, yY)\| < \frac{e^m}{\sqrt{\omega}}$. Then $f(x_0, y_0) = 0$ holds over the integers.*

D.2.2 Useful Lemmas

Let $N = pq$ be an RSA modulus. The following lemma is useful to find a value of $ap - bq$ using a known value of $ap + bq$.

Lemma D.2.3. *Let $N = pq$ be an RSA modulus with $q < p < 2q$ and S be a positive integer. Suppose that $ap + bq = S$ where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$. 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. Observe that multiplying $q < p < 2q$ by p gives $N < p^2 < 2N$ and consequently $\sqrt{N} < p < \sqrt{2}\sqrt{N}$. Suppose that $\frac{a}{b}$ is an approximation of $\frac{q}{p}$,

that is $a = \left\lceil \frac{bq}{p} \right\rfloor$. Hence $\left| a - \frac{bq}{p} \right| \leq \frac{1}{2}$, which gives

$$|ap - bq| \leq \frac{p}{2} \leq \frac{\sqrt{2}\sqrt{N}}{2} < 2\sqrt{N}.$$

Next, suppose that $ap+bq = S$. We have $S^2 = (ap+bq)^2 = (ap-bq)^2+4abN$. Since $|ap-bq| < 2\sqrt{N}$, then the quotient and the remainder in the Euclidean division of S^2 by $4N$ are respectively ab and $(ap-bq)^2$. Hence

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

which terminates the proof.

□

The following lemma shows how to factor $N = pq$ using a known value of $ap + bq$.

Lemma D.2.4. *Let $N = pq$ be an RSA modulus with $q < p < 2q$ and S be a positive integer. Suppose that $ap + bq = S$ where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$. Then N can be factored.*

Proof. Suppose that $\frac{a}{b}$ is an approximation of $\frac{q}{p}$ and that $ap + bq = S$. By Lemma D.2.3, we get $ab = \left\lfloor \frac{S^2}{4N} \right\rfloor$ and $|ap - bq| = D$ where

$$D = \sqrt{S^2 - 4abN}.$$

Hence $ap-bq = \pm D$. Combining with $ap+bq = S$, we get $2ap = S \pm D$. Since $a < q$, then $\gcd(N, S \pm D) = \gcd(N, 2ap) = p$. This gives the factorization of N.

□

D.3 The New Attack

Let e, d_1, k_1 be positive integers such that $ed_1 - k_1(N + 1 - ap - bq) = 1$. In this section, we consider the following parameters.

- $2^{m_0} = N^\beta$ where m_0 is a known integer.
- $a < b < N^\alpha$ with $\alpha < \frac{1}{2}$ where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$.
- $ap = 2^{m_0}p_1 + p_0$ where p_0 is a known integer.
- $bq = 2^{m_0}q_1 + q_0$ where q_0 is a known integer.
- $e = N^\gamma$.
- $d_1 = N^\delta$.

The aim in this section is to prove the following result.

Theorem D.3.1. *Suppose that $ap = 2^{m_0}p_1 + p_0$ and $bq = 2^{m_0}q_1 + q_0$ where m_0, p_0 and q_0 are known with $2^{m_0} = N^\beta$ and $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$ satisfying $a, b < N^\alpha$. Let $e = N^\gamma$, $d_1 = N^\delta$ and k_1 be positive integers satisfying an equation $ed_1 - k_1(N + 1 - ap - bq) = 1$. Then one can factor N in polynomial time when*

$$\delta \leq \begin{cases} \delta_1 & \text{if} \quad \gamma \geq \frac{1}{2}(1 + 2\alpha - 2\beta), \\ \delta_2 & \text{if} \quad \gamma \leq \frac{1}{2}(1 + 2\alpha - 2\beta), \end{cases}$$

where

$$\begin{aligned} \delta_1 &= \frac{7}{6} + \frac{1}{3}(\alpha - \beta) - \frac{1}{3}\sqrt{4(\alpha - \beta)^2 + 4(3\gamma + 1)(\alpha - \beta) + 6\gamma + 1}, \\ \delta_2 &= \frac{1}{4}(3 - 2(\alpha - \beta) - 2\gamma). \end{aligned}$$

Proof. Suppose that $ap = 2^{m_0}p_1 + p_0$ and $bq = 2^{m_0}q_1 + q_0$ with known m_0, p_0 and q_0. Then $ap + bq = 2^{m_0}(p_1 + q_1) + p_0 + q_0$. Starting with the variant RSA equation $ed_1 - k_1(N + 1 - ap - bq) = 1$, we get

$$ed_1 - k_1\left(N + 1 - p_0 - q_0 - 2^{m_0}(p_1 + q_1)\right) = 1.$$

Reducing modulo e, we get

$$-2^{m_0}k_1(p_1 + q_1) + (N + 1 - p_0 - q_0)k_1 + 1 \equiv 0 \pmod{e}.$$

Observe that $\gcd(2^{m_0}, e) = 1$. Then multiplying by $-2^{-m_0} \pmod{e}$, we get

$$k_1(p_1 + q_1) + a_1k_1 + a_2 \equiv 0 \pmod{e},$$

where

$$\begin{aligned} a_1 &\equiv -(N+1-p_0-q_0)2^{-m_0} \pmod{e}, \\ a_2 &\equiv -2^{-m_0} \pmod{e}. \end{aligned}$$

Consider the polynomial

$$f(x,y) = xy + a_1 x + a_2.$$

Then $(x,y) = (k_1, p_1+q_1)$ is a modular root of the equation $f(x,y) \equiv 0 \pmod{e}$. Assuming that $\alpha \ll \frac{1}{2}$, we get

$$k_1 = \frac{ed_1 - 1}{N+1-ap-bq} \sim N^{\gamma+\delta-1}.$$

On the other hand, we have

$$p_1 + q_1 < \frac{ap+bq}{2^{m_0}} < N^{\frac{1}{2}+\alpha-\beta}.$$

Define the bounds X and Y as

$$X = N^{\gamma+\delta-1}, \quad Y = N^{\frac{1}{2}+\alpha-\beta}.$$

To find the small modular roots of the equation $f(x,y) \equiv 0 \pmod{e}$, we apply the extended strategy of Jochemsz and May [72]. Let m and t be positive integers to be specified later. For $0 \le k \le m$, define the set

$$M_k = \bigcup_{0 \le j \le t} \{x^{i_1} y^{i_2+j} \;\Big|\; x^{i_1}y^{i_2} \text{ monomial of } f^m(x,y) \text{ and } \frac{x^{i_1}y^{i_2}}{(xy)^k} \text{ monomial of } f^{m-k}\}.$$

Observe that $f^m(x,y)$ satisfies

$$\begin{aligned} f^m(x,y) &= \sum_{i_1=0}^{m} \binom{m}{i_1} x^{i_1} (y+a_1)^{i_1} a_2^{m-i_1} \\ &= \sum_{i_1=0}^{m} \binom{m}{i_1} x^{i_1} \left(\sum_{i_2=0}^{i_1} \binom{i_1}{i_2} y^{i_2} a_1^{i_1-i_2} a_2^{m-i_1} \right) \\ &= \sum_{i_1=0}^{m} \sum_{i_2=0}^{i_1} \binom{m}{i_1} \binom{i_1}{i_2} x^{i_1} y^{i_2} a_1^{i_1-i_2} a_2^{m-i_1}. \end{aligned}$$

Hence, $x^{i_1}y^{i_2}$ is a monomial of $f^m(x,y)$ if

$$i_1 = 0,\ldots,m, \quad i_2 = 0,\ldots,i_1.$$

Consequently, for $0 \le k \le m$, when $x^{i_1}y^{i_2}$ is a monomial of $f^m(x,y)$, then $\frac{x^{i_1}y^{i_2}}{(xy)^k}$ is a monomial of $f^{m-k}(x,y)$ if

$$i_1 = k,\ldots,m, \quad i_2 = k,\ldots,i_1.$$

Hence, for $0 \le k \le m$, we obtain

$$x^{i_1}y^{i_2} \in M_k \quad \text{if} \quad i_1 = k,\ldots,m, \quad i_2 = k,\ldots,i_1+t.$$

Similarly,

$$x^{i_1}y^{i_2} \in M_{k+1} \quad \text{if} \quad i_1 = k+1,\ldots,m, \quad i_2 = k+1,\ldots,i_1+t.$$

For $0 \le k \le m$, define the polynomials

$$g_{k,i_1,i_2}(x,y) = \frac{x^{i_1}y^{i_2}}{(xy)^k} f(x,y)^k e^{m-k} \quad \text{with} \quad x^{i_1}y^{i_2} \in M_k \backslash M_{k+1}.$$

For $0 \le k \le m$, these polynomials reduce to the following sets

$$\begin{cases} k=0,\ldots,m, \\ i_1=k,\ldots,m, \\ i_2=k, \end{cases} \qquad \text{or} \qquad \begin{cases} k=0,\ldots,m, \\ i_1=k, \\ i_2=k+1,\ldots,i_1+t. \end{cases}$$

This gives rise to the polynomials

$$\begin{aligned} G_{k,i_1}(x,y) &= x^{i_1-k} f(x,y)^k e^{m-k}, \quad \text{for} \quad k=0,\ldots m, \quad i_1 = k,\ldots m, \\ H_{k,i_2}(x,y) &= y^{i_2-k} f(x,y)^k e^{m-k}, \quad \text{for} \quad k=0,\ldots m, \quad i_2 = k+1,\ldots,k+t. \end{aligned}$$

Let $\mathcal{L}$ denote the lattice spanned by the coefficient vectors of the polynomials $G_{k,i_1}(xX,yY)$ and $H_{k,i_2}(xX,yY)$. The ordering of two monomials $x^{i_1}y^{i_2}$, $x^{i'_1}y^{i'_2}$ is as in the following rule: if $i_1 < i'_1$, then $x^{i_1}y^{i_2} < x^{i'_1}y^{i_2}$ and if $i_1 = i'_1$ and $i_2 < i'_2$, then $x^{i_1}y^{i_2} < x^{i'_1}y^{i'_2}$. Notice that the matrix is left triangular. For $m = 3$ and $t = 1$, the coefficient matrix for $\mathcal{L}$ is presented in Table D.1. The non-zero elements are marked with an '$\circledast$'.

From the triangular form of the matrix, the $\circledast$ marked values do not contribute in the calculation of the determinant. Hence, the determinant of $\mathcal{L}$ is

$$\det(\mathcal{L}) = e^{n_e} X^{n_X} Y^{n_Y}. \tag{D.1}$$

	1	x	x^2	x^3	y	xy	x^2y	x^3y	xy^2	x^2y^2	x^3y^2	x^2y^3	x^3y^3	x^3y^4
$G_{0,0}$	e^3	0	0	0	0	0	0	0	0	0	0	0	0	0
$G_{0,1}$	0	Xe^3	0	0	0	0	0	0	0	0	0	0	0	0
$G_{0,2}$	0	0	X^2e^3	0	0	0	0	0	0	0	0	0	0	0
$G_{0,3}$	0	0	0	X^3e^3	0	0	0	0	0	0	0	0	0	0
$H_{0,1}$	0	0	0	0	Ye^3	0	0	0	0	0	0	0	0	0
$G_{1,1}$	⊛	⊛	0	0	0	XYe^2	0	0	0	0	0	0	0	0
$G_{1,2}$	0	⊛	⊛	0	0	0	X^2Ye^2	0	0	0	0	0	0	0
$G_{1,3}$	0	0	⊛	⊛	0	0	0	X^3Ye^2	0	0	0	0	0	0
$H_{1,2}$	0	0	0	0	⊛	⊛	0	0	XY^2e^2	0	0	0	0	0
$G_{2,2}$	⊛	⊛	⊛	0	0	⊛	⊛	0	0	X^2Y^2	0	0	0	0
$G_{2,3}$	0	⊛	⊛	⊛	0	0	⊛	⊛	0	0	X^3Y^2e	0	0	0
$H_{2,3}$	0	0	0	0	⊛	⊛	⊛	0	⊛	⊛	0	X^2Y^3e	0	0
$G_{3,3}$	⊛	⊛	⊛	⊛	0	⊛	⊛	⊛	0	⊛	⊛	0	X^3Y^3	0
$H_{3,4}$	0	0	0	0	⊛	⊛	⊛	0	⊛	⊛	⊛	⊛	⊛	X^3Y^4

Table D.1: The coefficient matrix for the case $m = 3$, $t = 1$.

From the construction of the polynomials $G_{k,i_1}(x, y)$ and $H_{k,i_2}(x, y)$, we get

$$n_e = \sum_{k=0}^{m} \sum_{i_1=k}^{m} (m-k) + \sum_{k=0}^{m} \sum_{i_2=k+1}^{k+t} (m-k) = \frac{1}{6} m(m+1)(2m+3t+4).$$

Similarly, we have

$$n_X = \sum_{k=0}^{m} \sum_{i_1=k}^{m} i_1 + \sum_{k=0}^{m} \sum_{i_2=k+1}^{k+t} k = \frac{1}{6} m(m+1)(2m+3t+4),$$

and

$$n_Y = \sum_{k=0}^{m} \sum_{i_1=k}^{m} k + \sum_{k=0}^{m} \sum_{i_2=k+1}^{k+t} i_2 = \frac{1}{6}(m+1)(m^2+3mt+3t^2+2m+3t).$$

Finally, we can calculate the dimension of $\mathcal{L}$ as

$$\omega = \sum_{k=0}^{m} \sum_{i_1=k}^{m} 1 + \sum_{k=0}^{m} \sum_{i_2=k+1}^{k+t} 1 = \frac{1}{2}(m+1)(m+2t+2).$$

For the following asymptotic analysis we let $t = \tau m$. For sufficiently large

m, the exponents n_e, n_X, n_Y and the dimension ω reduce to

$$\begin{aligned} n_e &= \frac{1}{6}(3\tau+2)m^3 + o(m^3), \\ n_X &= \frac{1}{6}(3\tau+2)m^3 + o(m^3), \\ n_Y &= \frac{1}{6}(3\tau^2+3\tau+1)m^3 + o(m^3), \\ \omega &= \frac{1}{2}(2\tau+1)m^2 + o(m^2). \end{aligned}$$

To apply Theorem D.2.2 to the shortest vector in the LLL-reduced basis of $\mathcal{L}$, we have to set

$$2^{\frac{\omega}{2}} \det(\mathcal{L})^{\frac{1}{\omega-1}} < \frac{e^m}{\sqrt{\omega}}.$$

This transforms to

$$\det(\mathcal{L}) < \frac{1}{\left(2^{\frac{\omega}{2}}\sqrt{\omega}\right)^{\omega}} e^{m(\omega-1)} < e^{m\omega}.$$

Using (D.1), we get

$$e^{n_e} X^{n_X} Y^{n_Y} < e^{m\omega}.$$

Plugging n_e, n_X, n_Y, ω as well as the values $e = N^{\gamma}$, $X = N^{\gamma+\delta-1}$, and $Y = N^{\frac{1}{2}+\alpha-\beta}$, we get

$$\begin{aligned} \frac{1}{6}(3\tau+2)m^3\gamma + \frac{1}{6}(3\tau+2)m^3(\gamma+\delta-1) &+ \frac{1}{6}(3\tau^2+3\tau+1)m^3(\frac{1}{2}+\alpha-\beta) \\ &< \frac{1}{2}(2\tau+1)m^3\gamma, \end{aligned}$$

which transforms to

$$3(2\alpha-2\beta+1)\tau^2 + 3(2\alpha+2\delta-2\beta-1)\tau + (2\gamma+2\alpha+4\delta-2\beta-3) < 0 \quad \text{(D.2)}$$

Next, we consider the cases $\tau \neq 0$ and $\tau = 0$ separately. First, we consider the case $\tau > 0$. The optimal value for τ in the left side of (D.2) is

$$\tau = \frac{1+2\beta-2\alpha-2\delta}{2(1+2\alpha-2\beta)}. \quad \text{(D.3)}$$

Observe that for $\alpha < \frac{1}{2}$ and $\beta < \frac{1}{2}$, we have $1+2\alpha-2\beta > 0$. To ensure $\tau > 0$, δ should satisfy $\delta < \delta_0$ where

$$\delta_0 = \frac{1}{2}\left(1 - 2(\alpha - \beta)\right). \tag{D.4}$$

Replacing τ by the optimal value (D.3) in the inequation (D.2), we get

$$-12\delta^2 + 4(7 + 2\alpha - 2\beta)\delta + 4(\alpha - \beta)^2 + 4(4\gamma - 1)(\alpha - \beta) + 8\gamma - 15 < 0,$$

which will be true if $\delta < \delta_1$ where

$$\delta_1 = \frac{1}{3}(\alpha - \beta) + \frac{7}{6} - \frac{1}{3}\sqrt{4(\alpha - \beta)^2 + 4(3\gamma + 1)(\alpha - \beta) + 6\gamma + 1}. \tag{D.5}$$

Since δ has to satisfy both $\delta < \delta_0$ and $\delta < \delta_1$ according to (D.4) and (D.5), let us find the minimum $\min(\delta_0, \delta_1)$. A straightforward calculation shows that

$$\min(\delta_0, \delta_1) = \begin{cases} \delta_0 & \text{if} \quad \gamma \leq \frac{1}{2}(1 + 2\alpha - 2\beta), \\ \delta_1 & \text{if} \quad \gamma \geq \frac{1}{2}(1 + 2\alpha - 2\beta). \end{cases}$$

Now, consider the case $\tau = 0$, that is $t = 0$. Then the inequation (D.2) becomes

$$2\gamma + 2\alpha + 4\delta - 2\beta - 3 < 0,$$

which leads to $\delta < \delta_2$ where

$$\delta_2 = \frac{1}{4}(2\beta + 3 - 2\gamma - 2\alpha). \tag{D.6}$$

To obtain an optimal value for δ, we compare δ_2 as in (D.6) to $\min(\delta_0, \delta_1)$, obtained respectively with $\tau > 0$ and $\tau = 0$. First suppose that $\gamma \leq \frac{1}{2}(1 + 2\alpha - 2\beta)$. Then

$$\min(\delta_0, \delta_1) - \delta_2 = \delta_0 - \delta_2 = \frac{1}{2}\left(g - \frac{1}{2}(1 + 2\alpha - 2\beta)\right) \leq 0.$$

Hence $\min(\delta_0, \delta_1) \leq \delta_2$. Next suppose that $\gamma \geq \frac{1}{2}(1 + 2(\alpha - \beta))$. Then

$$\begin{aligned} \min(\delta_0, \delta_1) - \delta_2 &= \delta_1 - \delta_2 \\ &= \frac{5}{6}(\alpha - \beta) + \frac{1}{2}\gamma + \frac{5}{12} \\ &\quad -\frac{1}{3}\sqrt{4(\alpha - \beta)^2 + 4(3\gamma + 1)(\alpha - \beta) + 6\gamma + 1}. \end{aligned}$$

On the other hand, we have

$$\left(\frac{5}{6}(\alpha-\beta)+\frac{1}{2}\gamma+\frac{5}{12}\right)^2-\left(\frac{1}{3}\sqrt{4(\alpha-\beta)^2+4(3\gamma+1)(\alpha-\beta)+6\gamma+1}\right)^2$$
$$=\frac{1}{16}(1+2(\alpha-\beta)-2\gamma)^2,$$

which implies that $\min(\delta_0,\delta_1)\geq\delta_2$.

Summarizing, the attack will succeed to find k_1, p_1+q_1 and $d_1=N^\delta$ when $\delta<\delta'$ with

$$\delta'=\begin{cases}\delta_1 & \text{if} \quad \gamma\geq\frac{1}{2}(1+2\alpha-2\beta),\\ \delta_2 & \text{if} \quad \gamma\leq\frac{1}{2}(1+2\alpha-2\beta),\end{cases}$$

where δ_1 and δ_2 are given by (D.5) and (D.6).

Next, using the known value of p_1+q_1, we can precisely calculate the value $ap+bq=2^{m_0}(p_1+q_1)+p_0+q_0=S$. Then using Lemma D.2.3 and Lemma D.2.4, we can find p and q. Since every step in the method can be done in polynomial time, then N can be factored in polynomial time. This terminates the proof.

□

For example, consider the standard instance with the following parameters

- $2^{m_0}=N^\beta$ with $\beta=0$.
- $a\leq b\leq N^\alpha$ with $\alpha=0$, that is $ap+bq=p+q$.
- $ap=2^{m_0}p_1+p_0=p_1$, that is $p_0=0$.
- $bq=2^{m_0}q_1+q_0=q_1$, that is $q_0=0$.
- $e=N^\gamma$ with $\gamma=1$.
- $d_1=N^\delta$.

Then $\gamma\geq\frac{1}{2}(1+2\alpha-2\beta)>\frac{1}{2}$ and the instance is insecure if $\delta<\delta_1$, that is if $\delta<\frac{7}{6}-\frac{\sqrt{7}}{3}\approx 0.284$ which is the same boundary as in various cryptanalytic approaches to RSA (see e.g. [17]).

Now suppose that $\gamma = 1$ and that a, b are small. Then $\alpha \approx 0$ and the boundary (D.5) becomes

$$\delta_1 < \frac{7}{6} - \frac{1}{3}\beta - \frac{1}{3}\sqrt{4\beta^2 - 16\beta + 7},$$

where the right side increases from 0.284 to 1 when $\beta \in \left[0, \frac{1}{2}\right[$. This implies that the existence of good approximation $\frac{a}{b}$ of $\frac{q}{p}$ substantially reduces the requirement of LSBs of ap and bq for the new attack. This confirms the recommendation of the X9.31-1997 standard for public key cryptography [1] regarding the generation of primes, namely that $\frac{q}{p}$ shall not be near the ratio of two small integers.

D.4 Experimental Results

We have implemented the new attack for various parameters. The machine was with Windows 7 and Intel(R) Core(TM)2 Duo CPU, 2GHz and the algebra system was Maple 12 [90]. For each set of parameters, we solved the modular equation $f(x, y) \equiv 0 \pmod{e}$ using the method described in Section D.3. We obtained two polynomials $f_1(x, y)$ and $f_2(x, y)$ with the expected root $(k_1, p_1 + q_1)$. We then solved the equation obtained using the resultant of $f_1(x, y)$ and $f_2(x, y)$ in one of the variables. For every instance, we could recover k_1 and $p_1 + q_1$ and hence factor N. The experimental results are shown in Table D.2

In the rest of this section, we present a detailed numerical example. Consider an instance of a 200-bit RSA public key with the following parameters.

- $N = 24632008214381394156795531909533432357612872407468918833633 09.$
- $e = 26662528980140646204174961754108951315840665128320416181615 3.$
  Hence $e = N^\gamma$ with $\gamma = 0.984$.
- $m_0 = 35$. Hence $2^{m_0} = N^\beta$ with $\beta = 0.174$.
- $a < b < N^{0.080}$. Hence $\alpha = 0.080$.

N	γ	β	α	δ	lattice parameters	LLL-time (sec)
2048	0.999	0.219	0.008	0.340	$m = 2$, $t = 1$, dim=9	54
2048	0.999	0.230	0.018	0.340	$m = 3$, $t = 2$, dim=18	2818
2048	0.999	0.172	0.114	0.273	$m = 2$, $t = 1$, dim=9	22
2048	0.999	0.150	0.096	0.272	$m = 2$, $t = 1$, dim=9	20
2048	0.999	0.091	0.019	0.280	$m = 2$, $t = 1$, dim=9	16
1024	0.999	0.326	0.123	0.368	$m = 3$, $t = 2$, dim=18	429
1024	0.999	0.326	0.123	0.339	$m = 2$, $t = 1$, dim=9	7
1024	0.998	0.229	0.050	0.326	$m = 2$, $t = 1$, dim=9	7
1024	0.995	0.102	0.008	0.297	$m = 2$, $t = 1$, dim=9	4
1024	0.999	0.131	0.123	0.239	$m = 2$, $t = 1$, dim=9	4

Table D.2: Experimental results.

- $m = 4$, $t = 2$.

Now suppose we know $p_0 = 28297245379$ and $q_0 = 28341074839$ such that $ap = 2^{m_0}p_1 + p_0$ and $bq = 2^{m_0}q_1 + q_0$. The modular equation to solve is then $f(x, y) = xy + a_1x + a_2 \equiv 0 \pmod{e}$, where

$$\begin{aligned} a_1 &= 396478470953448665961811597015453367067409367\backslash \\ &\quad 997081713297, \\ a_2 &= 230870662106105785001116936880561535466903107\backslash \\ &\quad 3317985538102. \end{aligned}$$

Working with $m = 4$ and $t = 2$, we get a lattice with dimension $\omega = 25$. Using the parameters $\gamma = 0.984$, $\alpha = 0.080$, and $\beta = 0.174$, the method will succeed with the bounds X and Y satisfying

$$\begin{aligned} p_1 + q_1 < X = N^{\gamma+\delta-1} &\approx 2^{52}, \\ k_1 < Y = N^{\frac{1}{2}+\alpha-\beta} &\approx 2^{81}, \end{aligned}$$

if $\delta < 0.356$. Applying the LLL algorithm, we find two polynomials $f_1(x, y)$ and $f_2(x, y)$ sharing the same integer solution. Then solving the resultant equation in y, we get $x = 4535179907267444$ and solving the resultant equation in x, we get $y = 360904506810171729844678 4$. Hence

$$\begin{aligned} p_1 + q_1 &= 4535179907267444, \\ k_1 &= 360904506810171729844678 4. \end{aligned}$$

Next, define

$$S = 2^{m_0}(p_1 + q_1) + p_0 + q_0 = 1240058442982957487861313276493287 30.$$

Then S is a candidate for $ap + bq$, and using Lemma D.2.3, we get

$$\begin{aligned} ab &= \left\lfloor \frac{S^2}{4N} \right\rfloor = 1560718201, \\ |ap - bq| &= D = \sqrt{S^2 - 4abN} = 10892876305854214138340560592 92. \end{aligned}$$

Using S for $ap + bq$ and D for $|ap - bq|$, we get $2ap = S - D$, and finally

$$p = \gcd(N, S - D) = 29735925138042579100455012611 69.$$

Hence $q = \frac{N}{p} = 82835856291783900153334732806 1$. This terminates the factorization of the modulus N. Using the equation $ed_1 = k_1(N + 1 - ap - bq) + 1$, we get $d_1 = 41897971798817657 \approx N^{0.275}$. We notice that, with the standard RSA equation $ed - k\phi(N) = 1$, we have $d \equiv e^{-1} \pmod{\phi(N)} \approx N^{0.994}$ which is out of reach of the attack of Boneh and Durfee as well as the attack of Blömer and May. Also, using $2ap = S - D$, we get $a = \frac{S-D}{2p} = 20851$. Similarly, using $2bq = S + D$, we get $b = \frac{S+D}{2q} = 74851$. We notice that $\gcd(a, b) = 1$ and $\frac{a}{b}$ is not among the convergents of $\frac{q}{p}$. This shows that Nitaj's attack as presented in [107] can not succeed to factor the RSA modulus in this example.

D.5 Conclusion

In this paper, we propose a new polynomial time attack on RSA with a public exponent satisfying an equation $ed_1 - k_1(N + 1 - ap - bq) = 1$ where $\frac{a}{b}$ is an unknown approximation of $\frac{q}{p}$ and where certain amount of the Least Significant Bits of ap and aq are known to the attacker. The attack is based on the method of Coppersmith for solving modular polynomial equations. This attack can be seen as an extension of the well known partial key attack on RSA when $a = b = 1$ and certain amount of the Least Significant Bits of one of the modulus prime factors is known.

Appendix E

Implicit Factorization of Unbalanced RSA Moduli

Journal of Applied Mathematics and Computing 2015
[115] with Muhammad Rezal Kamel Ariffin

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.

E.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 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 [94] 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 [135], 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 [45] 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 [82] 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 $|a_1p_1 - a_2p_2| < \frac{p_1}{2a_2q_1q_2}$. 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 $\beta = 0$, 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 [135] and [45] 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 $N_i = p_iq_i$, $i = 1, \ldots, k$ in the situation that there exist k integers a_i such that the a_ip_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

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

where β is such that $a_i \leq 2^\beta$. Once again, with $\beta = 0$, we improve the bound presented in the attack of [45].

Our third method addresses the implicit factoring problem when two unbalanced RSA moduli $N_1 = p_1q_1$ and $N_2 = p_2q_2$ of arbitrarily sizes are such that there exist two integers a_1 and a_2 such that a_1p_1 and a_2p_2 share t many LSBs. We show that it is possible to factor both N_1 and N_2 if $a_1a_2q_1q_2 < 2^{t-1}$. This method is also based on the continued fraction algorithm, applied to $\frac{T}{2^t}$ where $T \equiv N_2N_1^{-1} \pmod{2^t}$. 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 [94] and [135] and retrieves the bound of [82].

Our fourth method is a generalization of the third method to $k \geq 3$ RSA moduli $N_i = p_iq_i$, $i = 1, \ldots, k$. Assume that there exist k integers a_i such that the a_ip_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 (E.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 [94], [135], [45] and [82]. 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 E.1, we compare the applicability of our methods against the previous methods for the different scenarios.

Table E.1: Applicability of the methods for k RSA moduli.

Method	MSBs		LSBs	
	$k = 2$	$k \geq 3$	$k = 2$	$k \geq 3$
May, Ritzenhofen [94]	No	No	Yes	Yes
Sarkar, Maitra [135]	Yes	No	Yes	No
Faugère et al. [45]	Yes	Yes	No	No
Kurosawa, Ueda [82]	No	No	Yes	No
Our methods	Yes	Yes	Yes	Yes

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 Table E.2 and Table E.3, we compare our results against the former results with k RSA moduli in terms of bounds and dimension of the lattices.

We apply our results to the implicit factorization of $k \geq 2$ RSA for Paranoids [138] $N_i = p_i q_i$, $i = 1, \ldots, k$, 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 Section 2, we introduce some useful background on continued fractions and lattice basis reduction. In section 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 section 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 section 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 Section 6. In Section 7, we present our experiments and we conclude in Section 8.

Table E.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 [94]	Not studied	Not studied
Sarkar, Maitra [135]	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. [45]	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 $q_1 \approx \ldots \approx q_k \approx 2^\alpha$ 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 [82]	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 $q_1 \approx \ldots \approx q_k \approx 2^\alpha$ 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 $t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}(1 + \log_2(\pi e))$ using k-dimensional lattices of $\mathbb{Z}^k$. For $a_1 = \ldots = a_k = 1$, $\beta = 0$ and the the heuristic bound is $t > \frac{k}{k-1}\alpha + \frac{k}{2(k-1)}(1 + \log_2(\pi e))$.

Table E.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 [94]	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 $q_1 \approx \ldots \approx q_k \approx 2^\alpha$ 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 [135]	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. [45]	Not studied.	Not studied.
Kurosawa, Ueda [82]	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 $q_1 \approx \ldots \approx q_k \approx 2^\alpha$ 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 $t > \frac{k}{k-1}\alpha + \frac{k^2}{k-1}\beta + \frac{k}{2(k-1)}(1 + \log_2(\pi e))$ using k-dimensional lattices of $\mathbb{Z}^k$. For $a_1 = \ldots = a_k = 1$, $\beta = 0$ and the the heuristic bound is $t > \frac{k}{k-1}\alpha + \frac{k}{2(k-1)}(1 + \log_2(\pi e))$.

E.2 Preliminaries

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

E.2.1 Continued fractions

First we give the definition of continued fractions and state a related theorem. The details can be referenced in [57]. 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 [57]).

Theorem E.2.1 (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.

E.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 [86].

Theorem E.2.2 (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

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

E.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.

E.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 E.3.1. *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 E.2.1, 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 $q_2 = \gcd(N_2, a_1q_2)$ and $p_2 = \frac{N_2}{q_2}$. □

Remark E.3.2. The result of Theorem E.3.1 is valid even when the RSA moduli are not of the same size. Comparatively, the attacks presented by Sarkar and Maitra in [135] and Faugère et al. in [45] are valid only if $N_1 \approx N_2$ and $q_1 \approx q_2$.

Example E.3.3. Consider the following RSA moduli

$$\begin{aligned} N_1 &= 6343178298641262531091215558254707197227984863447 9, \\ N_2 &= 9946006657067710178027582903059286609914354223. \end{aligned}$$

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

$$\begin{aligned} &[0, 6377, 1, 1, 1, 1, 2, 2, 3, 1, 1, 3, 9, 1, 1, 1, 1, 7, 1, 19, 1, 1, 11, \\ &1, 1, 23, 1, 1, 3, 2, 3, 2, 3, 4, 2, 1, 1, 1, 8, 1, 322, 3, 4, 1, 1, 2, \ldots] \end{aligned}$$

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 E.3.1 are fulfilled. Indeed, the 40th

convergent is $\frac{a}{b} = \frac{1351300027964332}{8618068847003717463}$ and gives

$$\begin{aligned} q_1 &= \gcd(N_1, b) = 2125300178867, \\ p_1 &= \frac{N_1}{q_1} = 298460347470672037864031505763773292 37, \\ q_2 &= \gcd(N_2, a) = 9531501481, \\ p_2 &= \frac{N_2}{q_2} = 104348792022893566794039329416532738 3. \end{aligned}$$

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 [135] and [45] will not give a result in this situation.

E.3.2 Application to unbalanced RSA and RSA for Paranoids

As an application of Theorem E.3.1 to factor two unbalanced RSA moduli of the same bit-size, we get the following result.

Corollary E.3.4. *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 E.3.1, 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 E.3.5. If we consider $\beta = 0$ in Corollary E.3.4, 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 [45].

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 $N_i = p_i q_i$ 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 E.3.4, it is possible to factor N_1 and N_2 if a multiple $a_1 p_1$ and a multiple $a_2 p_2$ share the t MSBs whenever $t \geq 2\alpha + 2\beta + 1$, that is whenever $t \geq 1001 + 2\beta$.

E.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 $N_i = p_i q_i$, $i = 1 \ldots, k$ where the q_i's are α-bit primes and the $a_i p_i$'s share t MSBs. Instead of using the continued fraction algorithm, we use a lattice based method to find simultaneous diophantine approximations.

Theorem E.4.1. *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 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 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 $2 \leq i \leq k$, 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

$$a q_1 N_i - \frac{a a_1 q_i}{a_i} N_1 = \frac{a q_1 q_i x_i}{a_i}.$$

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

$$v = \left(C a q_1, \frac{a q_1 q_2 x_2}{a_2}, \ldots, \frac{a q_1 q_k x_k}{a_k} \right) \in \mathbb{Z}^k. \tag{E.2}$$

Then $v = \left(aq_1, \frac{aa_1q_2}{a_2} \ldots, \frac{aa_1q_k}{a_k}\right) \times M$, 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 $\det(\mathcal{L}) = CN_1^{k-1}$. The Gaussian Heuristic for $\mathcal{L}$ asserts that the length of its shortest non-zero vector is usually $\sigma(\mathcal{L})$ where

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

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 (E.2), we get

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

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

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

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

$$|x_i| = |a_ip_i - a_1p_1| \leq 2^{n-\alpha+\beta-t}.$$

Hence (E.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 (E.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 (E.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 [45]. 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.

E.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 [94], [135], [45] and [82]. In this section, we extend the former attacks to the case where an unknown multiple a_1p_1 of p_1 and an unknown multiple a_2p_2 of p_2 share their t LSBs.

E.5.1 The general attack

Theorem E.5.1. *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

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

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

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

Suppose that $a_1a_2q_1q_2 < 2^{t-1}$. Then dividing by $2^t a_2q_1$, we get

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

Therefore from Theorem E.2.1, 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 (E.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. □

E.5.2 Application to unbalanced RSA and RSA for Paranoids

Here we apply Theorem E.5.1 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 E.5.2. *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 E.5.1, 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 E.5.1, we conclude that $\frac{y}{a_2q_1}$ 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 E.5.3. Here again, if $\beta = 0$, then the condition of Corollary E.5.2 becomes $t \geq 2\alpha + 1$ which improves the bounds found in the former approaches of [94], [135], [45] and retrieves the bound of [82].

As an application of Corollary E.5.2, consider two 1024-bit RSA for Paranoids moduli $N_1 = p_1q_1$, $N_2 = p_2q_2$ where q_1, q_2 are 500-bit primes. Hence $\alpha = 500$ and using Corollary E.5.2, 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_1p_1 and a_2p_2 share t LSBs with $t > 2001 + 2\beta$.

E.6 Factoring k RSA Moduli in the LSB Case

In this section, we assume that we are given $k \geq 3$ different RSA moduli $N_i = p_iq_i$, $i = 1, \ldots, k$ where some unknown multiples a_ip_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 E.6.1. *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 positive 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 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

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

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

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

Observe that the norm of v satisfies

$$\|v\|^2 = a^2 q_1^2 + \sum_{i=2}^{k} \left(\frac{aa_1 q_k}{a_k} \right)^2.$$

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

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

Using (E.8) and (M.9) 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 (E.6), we get $q_1 = \gcd(aq_1, N_1)$ and for $i = 2, \ldots, k$, $q_i = \gcd(\frac{aa_1q_i}{a_i}, N_i)$. This terminates the proof.

□

Once again, if $\beta = 0$, then $a_i = 1$ and the bound of Theorem E.6.1 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 [45].

E.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 E.4.1 and Theorem E.6.1. 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_1p_1 and the a_ip_i's share t MSBs. Then since $a_ip_i \leq 2^{n-\alpha+\beta}$, we see that $|a_ip_i - a_1p_1| \leq 2^{n-\alpha+\beta-t}$. Therefore, $t \leq n - \alpha + \beta$. Similarly, assume that a_1p_1 and the a_ip_i's share t LSBs. Then $|a_ip_i - a_1p_1| = 2^t x_i$ with $t \leq n-\alpha+\beta$. In both cases, combining with the bound of t in Theorem E.4.1 and Theorem E.5.1, 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

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

Consequently, we only consider the situation where the bit-size β of the a_i's satisfies condition (E.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 (E.9). All our experiments were successful and the assumptions on the Gaussian Heuristics were verified. In Table E.4, we notice the experimentally lowest values of t that have 100% success rate.

Table E.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 (E.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	1000
4	400	20	15	617	594	601	1000
10	150	69	50	725	649	674	1000
10	250	48	40	725	667	684	1000
10	350	27	20	614	591	603	1000
10	400	17	12	581	563	570	1000
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

E.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.

Appendix F

Factoring RSA Moduli with Weak Prime Factors

C2SI 2015
[117] with Tajjeeddine Rachidi

Abstract :
In this paper, we study the problem of factoring an RSA modulus $N = pq$ in polynomial time, when p is a weak prime, that is, p can be expressed as $ap = u_0 + M_1u_1 + \ldots + M_ku_k$ for some k integers $M_1, \ldots, M_k$ and $k + 2$ suitably small parameters a, $u_0, \ldots u_k$. We further compute a lower bound for the set of weak moduli, that is, moduli made of at least one weak prime, in the interval $[2^{2n}, 2^{2(n+1)}]$ and show that this number is much larger than the set of RSA prime factors satisfying Coppersmith's conditions, effectively extending the likelihood for factoring RSA moduli. We also prolong our findings to moduli composed of two weak primes.

F.1 Introduction

The RSA cryptosystem, invented in 1978 by Rivest, Shamir and Adleman [131] is undoubtedly one of the most popular public key cryptosystems. In the standard RSA [131], the modulus $N = pq$ is the product of two large primes of the same bit-size. The public exponent e is an integer such that $1 \leq e < \phi(N)$ and $\gcd(e, \phi(N)) = 1$ where $\phi(N) = (p-1)(q-1)$ is the Euler totient function. The corresponding private exponent is the integer d such that $ed \equiv 1 \pmod{\phi(N)}$. In RSA, the encryption, decryption, signature generation, and signature verification require substantial CPU cycles because the time to perform these operations is proportional to the number of bits in public or secret exponents [131]. To reduce CPU time necessary for encryption and signature verification, one may be tempted to use a small public exponent e. This situation has been proven to be insecure against some small public exponent attacks (see [56] and [55]). To reduce the decryption and signature generation time, one may also be tempted to use a small private exponent d. Unfortunately, RSA is also vulnerable to various powerful short secret exponent attacks such as, the attack of Wiener [147], and the attack of Boneh and Durfee [17] (see also [15]). An alternate way for increasing the performance of encryption, decryption, signature generation, and signature verification, without reverting to small exponents, is to use the multi-prime variant of RSA. The multi-prime RSA is a generalization of the standard RSA cryptosystem in which the modulus is in the form $N = p_1 p_2 \cdots p_k$ where $k \geq 3$ and the p_i's are distinct prime numbers. Combined with the Chinese Remainder Theorem, a multi-prime RSA is much more efficient than the standard RSA (see [33]).

In Section 4.1.2 of the X9.31-1998 standard for public key cryptography [1], some recommendations are presented regarding the generation of the prime factors of an RSA modulus. For example, it is recommended that the modulus should have $1024 + 256x$ bits for $x \geq 0$. This requirement deters some factorization attacks, such as the Number Field Sieve (NFS) [85] and the Elliptic Curve Method (ECM) [84]. Another recommendation is that the prime difference $|p - q|$ should be large, and $\frac{p}{q}$ should not be near the ratio of two small integers. These requirements guard against Fermat factoring

algorithm [146], as well as Coppersmith's factoring attack on RSA [34] when one knows half of the bits of p. For example, if $N = pq$ and p, q are of the same bit-size with $|p - q| < N^{1/4}$, then $\left|p - \left[\sqrt{N}\right]\right| < N^{1/4}$ (see [104]) where $\left[\sqrt{N}\right]$ is the nearest integer to $\sqrt{N}$, which means that half of the bits of p are those of $[\sqrt{N}]$ which leads to the factorization of N (see [34] and [146]). Observe that the factorization attack of Coppersmith applies provided that one knows half of the bits of p, that is p is in one of the forms

$$p = \begin{cases} M_1 + u_0 & \text{with known } M_1 \text{ and unknown } u_0 \leq N^{\frac{1}{4}}, \\ M_1 u_1 + M_0 & \text{with known } (M_1, M_0) \text{ and unknown } u_1 \leq N^{\frac{1}{4}}. \end{cases}$$

Such primes are called Coppersmith's weak primes. In the case of $p = M_1 u_1 + M_0$ with known M_1 and M_0, the Euclidean division of q by M_1 is in the form $q = M_1 v_1 + v_0$. Hence $N = pq = (M_1 u_1 + M_0)(M_1 v_1 + v_0)$ which gives $M_0 v_0 \equiv N \pmod{M_1}$. Hence, since $gcd(M_0, M_1) = 1$, then $v_0 \equiv N M_0^{-1} \pmod{M_1}$. This means that when p is of the form $p = M_1 u_1 + M_0$ with known M_1 and M_0, then q is necessarily of the form $q = M_1 v_1 + v_0$ with known v_0. Coppersmith's attack is therefore applicable only when small enough parameters M_0 and v_0 can be found such that $p = M_1 u_1 + M_0$ and $q = M_1 v_1 + v_0$. This reduces the applicability of the attack to the set of moduli such that p and q are of the form defined above.

In this paper, we consider the generalization of Coppersmith's attack by considering a more satisfiable decomposition of any of the multipliers of p or q, i.e., ap or aq not just p or q, effectively leading to an increased set of moduli that can be factored. We describe two new attacks on RSA with a modulus $N = pq$. The first attack applies in the situation that, for given positive integers $M_1, \ldots, M_k$, one of the prime factors, p say, satisfies a linear equation $ap = u_0 + M_1 u_1 + \ldots + M_k u_k$ with suitably small integers a and $u_0, \ldots, u_k$. We call such prime factors *weak primes* for the integers $M_1, \ldots, M_k$. The second attack applies when both factors p and q are weak for the integers $M_1, \ldots, M_k$. We note that, for $k = 1$, the weak primes are such that $ap = u_0 + M_1 u_1$. This includes the class of Coppersmith's weak primes. For both attacks, we give an estimation of the RSA moduli $N = pq$ with a prime factor $p \in \left[2^n, 2^{n+1}\right]$ which is weak for the integers $M, M^2, \ldots, M^k$ where $M = \left\lceil 2^{\frac{n}{2k}} \right\rceil$.

The rest of the paper is organized as follows. In Section 2, we give some basic concepts on integer factorization and lattice reduction as well as an overview of Coppersmith's method. In Section 3, we present an attack on an RSA modulus $N = pq$ with one weak prime factor. In Section 4, we present the second attack an RSA modulus $N = pq$ with two weak prime factors. We conclude the paper in Section 5.

F.2 Preliminaries

In this section we give the definitions and results that we need to perform our attacks. These preliminaries include basic concepts on integer factorization and lattice reduction techniques.

F.2.1 Integer factorization: the state of the art

Currently, the most powerful algorithm for factorizing large integers is the Number Field Sieve (NFS) [85]. The heuristic expected time $T_{NFS}(N)$ of the NFS depends on the bitsize of the integer N to be factored:

$$T_{NFS}(N) = \exp\left((1.92 + o(1))(\log N)^{1/3}(\log\log N)^{2/3}\right).$$

If the integer N has small factors, the Elliptic Curve Method (ECM) [84] for factoring is substantially faster than the NFS. It can compute a non-trivial factor p of a composite integer N in an expected runtime T_{ECM}:

$$T_{ECM}(p) = \exp\left(\left(\sqrt{2} + o(1)\right)(\log p)^{1/2}(\log\log p)^{1/2}\right),$$

which is sub-exponential in the bitsize of the factor p. The largest factor found so far with the ECM is a 83 decimal digits (275 bits) prime factor of the special number $7^{337} + 1$ (see [150]).

F.2.2 Lattice reduction

Let m and n be positive integers with $m \leq n$. Let $u_1, \ldots, u_m \in \mathbb{R}^n$ be m linearly independent vectors. The lattice $\mathcal{L}$ spanned by $u_1, \ldots, u_m$ is the set

$$\mathcal{L} = \left\{ \sum_{i=1}^{m} a_i u_i \mid a_i \in \mathbb{Z} \right\}.$$

The set $\{u_1, \ldots, u_m\}$ is called a lattice basis for $\mathcal{L}$. The dimension (or rank) of the lattice $\mathcal{L}$ is $\dim(\mathcal{L}) = m$, and $\mathcal{L}$ is called full rank if $m = n$. It is often useful to represent the lattice $\mathcal{L}$ by the $m \times n$ matrix M whose rows are the coefficients of the vectors $u_1, \ldots, u_m$. The determinant (or volume) of $\mathcal{L}$ is defined as $\det(\mathcal{L}) = \sqrt{M \cdot M^t}$. When $\mathcal{L}$ is full rank, the determinant reduces to $\det(\mathcal{L}) = |\det(M)|$. The Euclidean norm of a vector $v = \sum_{i=1}^{m} a_i u_i \in \mathcal{L}$ is defined as $\|v\| = \sqrt{\sum_{i=1}^{m} a_i^2}$. As a lattice has infinitely many bases, some bases are better than others, and a very important task is to find a basis with small vectors $\{b_1, \ldots, b_m\}$ called the reduced basis. This task is very hard in general, however, the LLL algorithm proposed by Lenstra, Lenstra, and Lovász [86] finds a basis of a lattice with relatively small vectors in polynimial time. The following theorem determines the sizes of the reduced basis vectors obtained with LLL (see [91] for more details).

Theorem F.2.1. *Let $\mathcal{L}$ be a lattice spanned by a basis $\{u_1, \ldots, u_m\}$. The LLL algorithm applied to $\mathcal{L}$ outputs a reduced basis $\{b_1, \ldots, b_m\}$ with*

$$\|b_1\| \leq \|b_2\| \leq \ldots \leq \|b_i\| \leq 2^{\frac{m(m-1)}{4(m-i+1)}} \det(\mathcal{L})^{\frac{1}{m+i-1}}, \text{ for } i = 1, 2, \ldots, m.$$

The existence of a short nonzero vector in a lattice is guaranteed by a result of Minkowski stating that every m-dimensional lattice $\mathcal{L}$ contains a non-zero vector v with $\|v\| \leq \sqrt{m} \det(L)^{\frac{1}{m}}$. On the other hand, the Gaussian Heuristic asserts that the norm γ_1 of the shortest vector of a random lattice satisfies

$$\gamma_1 \approx \sqrt{\frac{\dim(\mathcal{L})}{2\pi e}} \det(\mathcal{L})^{\frac{1}{\dim(\mathcal{L})}}.$$

Hereafter, we will use this result as an estimation for the expected minimum norm of a non-zero vector in a lattice.

F.2.3 Coppersmith's Method

In 1996, Coppersmith [34] presented two techniques based on LLL to find small integer roots of univariate modular polynomials or of bivariate integer polynomials. Coppersmith showed how to apply his technique to factorize an RSA modulus $N = pq$ with $q < p < 2q$ when half of the least or the most significant bits of p is known.

Theorem F.2.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let M_0 and M_1 be two positif integers. If $p = M_1 + u_0$ with $u_0 < N^{\frac{1}{4}}$ or if $p = M_1 u_1 + M_0$ with $u_1 < N^{\frac{1}{4}}$, then N can be factored in time polynomial in $\log N$.*

Coppersmith's technique extends to polynomials in more variables, but the method becomes heuristic. The problem of finding small roots of linear modular polynomials $f(x_1, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + a_{n+1} \pmod{p}$ for some unknown p that divides the known modulus N has been studied using Coppersmith's technique by Herrmann and May [59]. The following result, due to Lu, Zhang and Lin [88] gives a sufficient condition under which modular roots can be found efficiently.

Theorem F.2.3 (Lu, Zhang, Lin)**.** *Let N be a composite integer with a divisor p^u such that $p \geq N^\beta$. Let $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ be a homogenous linear polynomial. Then one can find all the solutions $(y_1, \ldots, y_n)$ of the equation $f(x_1, \ldots, x_n) = 0 \mod p^v$, $v \leq u$ with $\gcd(y_1, \ldots, y_n) = 1$ and $|y_1| < N^{\delta_1}, \ldots, |y_n| < N^{\delta_n}$ if*

$$\sum_{i=1}^{n} \delta_i \leq \frac{u}{v}\left(1 - \left(1 - \frac{u}{v}\beta\right)^{\frac{n}{n-1}} - n\left(1 - \sqrt[n-1]{1 - \frac{u}{v}\beta}\right)\left(1 - \frac{u}{v}\beta\right)\right).$$

The time complexity of the algorithm for finding such sulution $(y_1, \ldots, y_n)$ is polynomial in $\log N$.

F.3 The Attack with One Weak Prime Factor

F.3.1 The Attack

In this section, we present an attack to factor an RSA modulus $N = pq$ when p satisfies a linear equation in the form $ap = u_0 + M_1u_1 + \ldots M_ku_k$ for a suitably small positive integer a and suitably small integers $u_0, u_1, \ldots, u_k$ where $M_1, \ldots, M_k$ are given positive integers. Such prime factor p is called a weak prime for the integers $M_1, \ldots, M_k$.

Theorem F.3.1. *Let $N = pq$ be an RSA modulus such that $p > N^\beta$ and $M_1, \ldots, M_k$ be k positive integers with $M_1 < M_2 < \ldots < M_k$. Suppose that there exists a positive integer a, and $k+1$ integers u_i, $i = 0, \ldots, k$ such that $ap = u_0 + M_1u_1 + \ldots + M_ku_k$ with $\max(u_i) < N^\delta$ and*

$$\delta < \frac{1}{k+1}\left(1 - (1-\beta)^{\frac{k+1}{k}} - (k+1)\left(1 - \sqrt[k]{1-\beta}\right)(1-\beta)\right).$$

Then one can factor N in polynomial time.

Proof. Let $M_1, \ldots, M_k$ be k positive integers such that $M_1 < M_2 < \ldots < M_k$. Suppose that $ap = u_0 + M_1u_1 + \ldots + M_ku_k$, that is $(u_0, \ldots, u_k)$ is a solution of the modular polynomial equation

$$x_0 + M_1x_1 + \ldots + M_kx_k = 0 \pmod{p}. \tag{F.1}$$

Suppose that $|u_i| < N^\delta$ for $i = 0, \ldots, k$. Using $n = k+1$, $u = 1$ and $v = 1$ in Theorem F.2.3, means that the equation (F.1) can be solved in polynomial time, i.e., finding $(u_0, \ldots, u_k)$ if

$$(k+1)\delta < \left(1 - (1-\beta)^{\frac{k+1}{k}} - (k+1)\left(1 - \sqrt[k]{1-\beta}\right)(1-\beta)\right),$$

which gives the bound

$$\delta < \frac{1}{k+1}\left(1 - (1-\beta)^{\frac{k+1}{k}} - (k+1)\left(1 - \sqrt[k]{1-\beta}\right)(1-\beta)\right).$$

This terminates the proof.

□

Remark F.3.2. For a balanced RSA modulus, the prime factors p and q are of the same bit size. Then $p > N^{\beta}$ with $\beta = \frac{1}{2}$. Hence, the condition on δ becomes

$$\delta < \frac{1}{k+1}\left(1 - \left(\frac{1}{2}\right)^{\frac{k+1}{k}}\right) - \frac{1}{2}\left(1 - \left(\frac{1}{2}\right)^{\frac{1}{k}}\right). \tag{F.2}$$

In Table F.1, we give the bound for δ for given β and k.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$
$\beta=0.5$	0.125	0.069	0.047	0.036	0.029	0.024	0.021	0.018	0.016	0.015
$\beta=0.6$	0.180	0.101	0.071	0.054	0.044	0.037	0.032	0.028	0.025	0.022
$\beta=0.7$	0.245	0.142	0.100	0.077	0.063	0.053	0.046	0.046	0.036	0.032

Table F.1: Upper bounds for δ by Theorem F.3.1.

Remark F.3.3. We note that Coppersmith's weak primes correspond to moduli $N = pq$ with $q < p < 2q$ where one of the prime factors is of the form $p = M_1 + u_0$ or $p = M_1 u_1 + M_0$ with $u_0, u_1 < N^{0.25}$ as mentioned in Theorem F.2.2. This a special case of the equation of Theorem F.3.1. Indeed, we can solve the equations $p = M_1 + u_0$ and $p = M_1 u_1 + M_0$ when $|u_0|, |u_1| < N^{\frac{1}{4}}$. Alternatively, Coppersmith's weak primes correspond to the cell $(k, 2\beta) = (1, 0.25)$ in Table F.1.

F.3.2 Numerical Examples

Example F.3.4. Let

$$\begin{aligned} N =& 1000975288631210998802277822755057783708121519200512986478468518574404680187957742118603163855742681296240768835751196370914 1, \end{aligned}$$

be a 412-bit RSA modulus with $N = pq$ where $q < p < 2q$. Then p and q are balanced and $p \approx N^{\frac{1}{2}} \approx 2^{206}$. Hence for $\beta = 0.5$, we have $p > N^{\beta}$. Suppose that p satisfies an equation of the form $ap = u_0 + Mu_1 + M^2 u_2$. Typically, $M^2 \approx N^{\frac{1}{2}}$, that is $M \approx N^{\frac{1}{4}}$. So let $M = 2^{100}$. For $\beta = 0.5$ and $k = 2$, Table (F.1) gives the bound $\delta < 0.069$. Assume therefore that the parameters u_i satisfy $|u_i| < N^{0.069} \approx 2^{28}$ for $i = 0, 1, 2$. By applying

Theorem F.3.1 we should find u_0, u_1 and u_2 as long as $u_0, u_1, u_2 < 2^{28}$. We apply the method of Lu et al. [88] with $m = 4$ and $t = 1$. This gives a 35-dimensional lattice. Applying the LLL algorithm [86], we find a reduced basis with multivariate polynomials $f_i(x_1, x_2, x_3) \in \mathbb{Z}[x_1, x_2, x_3]$, $i = 1, \ldots, 3$. Applying the Gröbner basis technique for solving a system of polynomial equations, we get $u_0 = 9005$, $u_1 = 7123$, $u_2 = 3915$. Using these values, we can compute $ap = u_0 + Mu_1 + M^2u_2$ from which we deduce $p = \gcd(u_0 + Mu_1 + M^2u_2, N)$, that is

$$p = 123356126338704841740132972382836883609800988209539117002682143.$$

Finally, we can compute $q = \frac{N}{p}$, that is

$$q = 811451622502140724659803961925628218026979706614326237650389 87.$$

Note here that there is no linear decomposition of p in the form $p = M_1 + u_0$ nor $p = M_1u_1 + M_0$ with $u_0, u_1 < N^{0.25}$ that makes p vulnerable to the attack of Coppersmith. This shows that the modulus N is vulnerable to our attack, while it is not vulnerable to Coppersmith's attack. Finally, the overall recorded execution time for our attack using an off-the-shelf computer was 17 seconds.

Example F.3.5. In [10], Bernstein et al. discovered many prime factors with special forms. Many of these primes were found by computing the greatest common divisor of a collection of RSA moduli. Others were found by applying Coppersmith's technique. We show below that our attack can find some primes among the list of Bernstein et al. One of these primes is

$$\begin{aligned} p =& 0xc000000000000000000000000000000000000000000000000000000000000000 \\ & 0000000000000000000000000000000000000000000000000000000000000002f9, \\ =& 100558559474569478246805187486543845956095243654442950332926710827913230225551602326014057236251775707675238936398645381403154121089599274598252367545638 33. \end{aligned}$$

Using $M = 2^{510}$, we get $p = 3M + 761 = Mu_1 + u_0$ where $u_1 = 3$ and $u_0 = 761$. We have $u_1, u_0 < N^{\delta}$ with $\delta \approx 0.007$ which is less than the bound 0.125 in Table F.1 for a 1024 bit-size RSA modulus N with $\beta = 0.5$, and $k = 1$. This implies that the conditions for Theorem F.3.1 are satisfied and our method finds p when used in any RSA modulus.

Example F.3.6. Now, consider this other example from the list of Bernstein et al. [10]

$$\begin{aligned} p =& 0xc000b800000000000000000000000000000000000000000000000000000000000000000000\\ & 0000068000000000000000000000000000000000000000000000000000000000000000000251\\ =& 10056002994300661909178585747410296772915190347411207124093761152520749216065545598886037221777994938111659319232428746318812487\\ & 60951383726377211701709393 \end{aligned}$$

Then p has the form $p = 3145774M^7 + 27262976M^3 + 593 = M^7u_7 + M^3u_3 + u_0$ where $M = 2^{70}$. The coefficients u_7, u_3 and u_0 satisfy $u_7, u_3, u_0 < N^\delta$ with $\delta \approx 0.016$ while the bound of Theorem F.3.1 is 0.021 (see Table F.1 for $k = 7$ and $\beta = 0.5$). Again, this shows that our method will find the factorization of any RSA modulus that is a multiple of p.

F.3.3 The Number of Single Weak Primes in an Interval

In this section, we consider two positive integers n and M and present a study of the weak primes with M, that is the primes $p \in \left[2^n, 2^{n+1}\right]$ such that there exists a positive integer a that gives the decomposition

$$ap = \sum_{i=0}^{k} M^i u_i$$

where $|u_i| < N^\delta$ and δ satisfies Theorem F.3.1. We show that the number of the RSA moduli N in the interval $[2^{2n}, 2^{2(n+1)}]$ with a weak prime factor $p \in \left[2^n, 2^{n+1}\right]$ is polynomial in 2^n. That is, this number is lower bounded by 2^η where $\eta > \frac{1}{2}$. We call such a class weak RSA Moduli in the interval $[2^{2n}, 2^{2(n+1)}]$.

Theorem F.3.7. *Let n be a positive integer. For $k \geq 1$, define $M = \left\lceil 2^{\frac{n}{k}} \right\rceil$. Let $\mathcal{N}$ be the set of the weak RSA moduli $N \in \left[2^{2n}, 2^{2(n+1)}\right]$ such that $N = pq$, p and q are of the same bitsize, $p > q$, and $p = \left\lfloor \frac{\sum_{i=0}^{k} M^i u_i}{a} \right\rfloor + b \in \left[2^n, 2^{n+1}\right]$ for some small integers b, $a < N^\delta$ and $|u_i| < N^\delta$ for $i = 0, \ldots, k$ with*

$$\delta = \frac{1}{k+1}\left(1 - \left(\frac{1}{2}\right)^{\frac{k+1}{k}}\right) - \frac{1}{2}\left(1 - \left(\frac{1}{2}\right)^{\frac{1}{k}}\right).$$

Then the cardinality of $\mathcal{N}$ satisfies $\#\mathcal{N} \geq 2^{\eta}$ where

$$\eta = (1 + 2(k+1)\delta)n + \log_2\left(\frac{(n-1)}{n(n+1)\log(2)}\right).$$

Proof. Let N be an RSA moduli. Suppose that $N \in \left[2^{2n}, 2^{2(n+1)}\right]$ with $N = pq$ where p and q are of the same bitsize. Since $p \approx N^{\frac{1}{2}}$, then $p \in \left[2^n, 2^{n+1}\right]$. Suppose further that for some positive integer a, we have $ap = \sum_{i=0}^{k} M^i u_i$. Then

$$M^k = \frac{ap - \sum_{i=0}^{k-1} M^i u_i}{u_k} \approx \frac{a}{u_k} p,$$

which implies $M \approx p^{\frac{1}{k}} \approx N^{\frac{1}{2k}}$. Now, define

$$M = \left\lceil N^{\frac{1}{2k}} \right\rceil = \left\lceil 2^{\frac{n}{k}} \right\rceil,$$

where $\lceil x \rceil$ is the integer greater or equal to x. This yields $2^n \leq M^k \leq 2^{n+1}$. Consider the set $\mathcal{P} \subset \left[2^n, 2^{n+1}\right]$

$$\mathcal{P} = \left\{ p = \left\lfloor \frac{\sum_{i=0}^{k} M^i u_i}{a} \right\rfloor + b,\ p \text{ is prime},\ \mid a < N^{\delta},\ |u_i| < N^{\delta} \right\},$$

where δ satisfies (F.2). Here b is as small as possible so that $\left\lfloor \frac{\sum_{i=0}^{k} M^i u_i}{a} \right\rfloor + b$ is prime. Also, since M^k is the leading term, then observe that

$$\frac{\sum_{i=0}^{k} M^i u_i}{a} - M^k = \frac{u_k - a}{a} M^k + \frac{\sum_{i=1}^{k} M^i u_i}{a}.$$

To ensure $p \in \left[2^n, 2^{n+1}\right]$, we consider only the situation where $u_k \geq a$. Hence, using the bounds $a < N^{\delta}$ and $|u_i| < N^{\delta}$ for $i = 0, \ldots, k-1$, we get a lower bound for the number of possibilities for a and for u_i, which themselves set a lower bound for the cardinality of $\mathcal{P}$ as follows:

$$\#\mathcal{P} \geq \left\lfloor N^{\delta} \right\rfloor \left\lfloor N^{\delta} \right\rfloor^k \approx N^{(k+1)\delta} \approx 2^{2(k+1)n\delta}. \tag{F.3}$$

On the other hand, the prime number theorem asserts that the number $\pi(x)$ of the primes less than x is

$$\pi(x) \approx \frac{x}{\log(x)}.$$

Hence, the number of primes in the interval $\left[2^n, 2^{n+1}\right]$ is approximately

$$\pi\left(2^{n+1}\right) - \pi\left(2^n\right) \approx \frac{2^{n+1}}{\log\left(2^{n+1}\right)} - \frac{2^n}{\log\left(2^n\right)} = \frac{(n-1)2^n}{n(n+1)\log(2)}. \tag{F.4}$$

It follows that the number of RSA moduli $N = pq \in \left[2^{2n}, 2^{2(n+1)}\right]$ with a weak factor $p \in \mathcal{P}$ and $q \in \left[2^n, 2^{n+1}\right]$ is at least $\#(\mathcal{N}) \geq \#\mathcal{P} \times \left(\pi\left(2^{n+1}\right) - \pi\left(2^n\right)\right)$. Using F.3 and F.4, we get

$$\begin{aligned}
\#(\mathcal{N}) &\geq 2^{2(k+1)n\delta} \times \frac{(n-1)2^n}{n(n+1)\log(2)} \\
&= \frac{(n-1)}{n(n+1)\log(2)} \times 2^{(1+2(k+1)\delta)n} \\
&= 2^{\eta},
\end{aligned}$$

where

$$\eta = (1 + 2(k+1)\delta)n + \log_2\left(\frac{(n-1)}{n(n+1)\log(2)}\right).$$

This terminates the proof.

□

Table F.2 presents a list of values of the bound η in terms of k and n. In Table F.2, we see that in the situation $(\beta, k) = (0.5, 1)$, the number $\#(\mathcal{N})$

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
$n = \frac{1}{2}\log_2(N) = 512$	759	715	698	689	684	680	677
$n = \frac{1}{2}\log_2(N) = 1024$	1526	1438	1404	1386	1375	1368	1362
$n = \frac{1}{2}\log_2(N) = 2048$	3061	2885	2818	2782	2759	2744	2733

Table F.2: Lower bounds for η under Theorem F.3.7.

of 1024-bits RSA moduli $N = pq \in \left[2^{1024}, 2^{1026}\right]$ with a weak factor p is at least $\#(\mathcal{N}) \geq 2^{759}$. Observe that the number of RSA moduli with a weak Coppersmith's prime factor in the same interval is approximately $N^{\frac{1}{4}} \cdot N^{\frac{1}{2}} \approx 2^{768}$. Actually, weak Coppersmith's prime are of the form $p = M_1 + u_0$ or $p = M_1 u_1 + M_0$ with one unknown parameter u_0 or u_1, while our weak primes for $k = 1$ are of the form $p = M_1 u_1 + u_0$ with two unknown parameters

u_0 or u_1. This shows that our weak prime factors are different from weak Coppersmith primes.

F.4 The Attack with Two Weak Prime factors

F.4.1 The Attack

In this section, we present an attack on RSA with a modulus $N = pq$ when both the prime factors p and q are weak primes.

Theorem F.4.1. *Let $N = pq$ be an RSA modulus and M be a positive integer. Let $k \geq 1$. Suppose that there exist integers a, b, u_i and v_i, $i = 1, \ldots, k$ such that $ap = \sum_{i=0}^{k} M^i u_i$ and $bq = \sum_{i=0}^{k} M^i v_i$ with $|u_i|, |v_i| < N^\delta$ and*

$$\delta < \frac{1}{2k+1} + \frac{\log\left(2k^3\right)}{2(2k+1)\log(N)} + \frac{\log(2k+1) - \log(2\pi e)}{4\log(N)} - \frac{\log\left(4k^3\right)}{4\log(N)}.$$

Then one can factor N in polynomial time.

Proof. Suppose that $ap = \sum_{i=0}^{k} M^i u_i$ and $bq = \sum_{i=0}^{k} M^i v_i$. Then multiplying ap and bq, we get

$$abN = \sum_{i=0}^{2k} M^i w_i, \quad \text{with} \quad w_i = \sum_{j=0}^{i} u_j v_{i-j}.$$

This can be transformed into the equation

$$M^{2k} x_{2k} + M^{2k-1} x_{2k-1} + \ldots + M x_1 - yN = -x_0, \tag{F.5}$$

with the solution $(x_{2k}, x_{2k-1}, \ldots, x_1, y, x_0) = (w_{2k}, w_{2k}, \ldots, w_1, ab, u_0 v_0)$. For $i = 0, \ldots, k$, suppose that $|u_i|, |v_i| < N^\delta$. Since for $i = 0, \ldots, 2k$, the maximal number of terms in w_i is k, we get

$$|x_i| = |w_i| \leq k \max_j(|u_j|) \cdot \max_j(|v_j|) < kN^{2\delta}. \tag{F.6}$$

Let C be a constant to be fixed later. Consider the lattice $\mathcal{L}$ generated by

the row vectors of the matrix

$$M(\mathcal{L}) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & CM^{2k} \\ 0 & 1 & \dots & 0 & 0 & CM^{2k-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & CM \\ 0 & 0 & 0 & \dots & 0 & -CN \end{bmatrix}. \tag{F.7}$$

The dimension of the lattice $\mathcal{L}$ is $\dim(\mathcal{L}) = 2k+1$ and its determinant is $\det(\mathcal{L}) = CN$. According to the Gaussian Heuristic, the length of the shortest non-zero vector of the lattice $\mathcal{L}$ is approximately $\sigma(\mathcal{L})$ with

$$\sigma(\mathcal{L}) \approx \sqrt{\frac{\dim(\mathcal{L})}{2\pi e}} \det(\mathcal{L})^{\frac{1}{\dim(\mathcal{L})}} = \sqrt{\frac{2k+1}{2\pi e}} (CN)^{\frac{1}{2k+1}}.$$

Consider the vector $v = (x_{2k}, x_{2k-1}, \dots, x_1, -Cx_0)$. Then, using (F.5), we get

$$(x_{2k}, x_{2k-1}, \dots, x_1, -Cx_0) = (x_{2k}, x_{k-1}, \dots, x_1, y) \cdot M(\mathcal{L}).$$

This means that $v \in \mathcal{L}$. Consequently, if C satisfies $\|v\| \leq \sigma(\mathcal{L})$, then, by the Gaussian Heuristic, v is the shortest vector of L. Using the bound (F.6), the length of the vector v satisfies

$$\|v\|^2 = C^2 x_0^2 + \sum_{i=1}^{2k} x_i^2 \leq \left(C^2 + \sum_{i=1}^{2k} k^2 \right) N^{4\delta} = \left(C^2 + 2k^3 \right) N^{4\delta}.$$

Let C be a positive integer satisfying $C \leq \sqrt{2k^3}$. Then the norm of the vector v satisfies $\|v\|^2 < 4k^3 N^{4\delta}$. Hence, using the Gaussian approximation $\sigma(\mathcal{L})$, the inequality $\|v\| \leq \sigma(\mathcal{L})$ is satisfied if

$$2k^{\frac{3}{2}} N^{2\delta} \leq \sqrt{\frac{2k+1}{2\pi e}} \left(2^{\frac{1}{2}} k^{\frac{3}{2}} N \right)^{\frac{1}{2k+1}}.$$

Solving for δ, we get

$$\delta < \frac{1}{2k+1} + \frac{\log\left(2k^3\right)}{2(2k+1)\log(N)} + \frac{\log(2k+1) - \log(2\pi e)}{4\log(N)} - \frac{\log\left(4k^3\right)}{4\log(N)}.$$

If δ satisfies the former bound, then the LLL algorithm, applied to the lattice $\mathcal{L}$ will output the vector $v = (x_{2k}, x_{2k-1}, \ldots, x_1, -Cx_0)$ from which, we deduce

$$w_{2k} = |x_{2k}|,\ w_{2k-1} = |x_{2k-1}|, \ldots, w_1 = |x_1|,\ w_0 = \frac{|-Cx_0|}{C}.$$

Using the coefficients w_i, $i = 1, \ldots, 2k$, we construct the polynomial $P(X) = w_{2k}X^{2k} + w_{2k-1}X^{2k-1} + \ldots + w_1X + w_0$. Factoring $P(X)$, we get

$$P(X) = \left(\sum_{i=0}^{k} M^i u_i\right)\left(\sum_{i=0}^{k} M^i v_i\right),$$

from which we deduce all the values u_i and v_i for $i = 1, \ldots, k$. Using each u_i and v_i for $i = 1, \ldots, k$, we get $ap = \sum_{i=0}^{k} M^i u_i$ and finally obtain $p = \gcd\left(\sum_{i=0}^{k} M^i u_i, N\right)$ which in turn gives $q = \frac{N}{q}$. This terminates the proof.

□

In Table F.3, we give the bound for δ for a given k and a given size of the RSA modulus.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
$\log_2(N) = 1024$	0.332	0.199	0.141	0.109	0.089
$\log_2(N) = 2048$	0.333	0.199	0.142	0.110	0.090

Table F.3: Upper bounds for δ with Theorem F.4.1.

F.4.2 Examples

Example F.4.2. Consider the 234 bits RSA modulus

$$N = 181287275221777294353476345871682929689873183168124359321741177743400 29.$$

Let $M = 2^{50}$. Suppose further that the prime factors p and q are such that $ap = M^2u_2 + Mu_1 + u_0$ and $bq = M^2v_2 + Mv_1 + v_0$, that is $k = 2$ with the notation of Theorem F.4.1. We built the matrix (F.7) with $C = \sqrt{2k^3} = 4$ and applied the LLL algorithm [86]. We got a new basis, where the last row is:

$$(w_4, w_3, w_2, w_1, -Cw_0) =$$
$$(30223231819936, 68646317659290, 109044283791446, 80821741694637, -162291153390444).$$

From this, we form the polynomial $P(X) = w_4X^4+w_3X^3+w_2X^2+w_1X^1+w_0$. which factors as:

$$P(X) = \left(4678994X^2 + 5832048X + 4871673\right) \left(6459344X^2 + 6620037X + 8328307\right).$$

From this, we deduce

$$u_2 = 4678994,\ u_1 = 5832048,\ u_0 = 4871673,$$
$$v_2 = 6459344,\ v_1 = 6620037,\ v_0 = 8328307.$$

Using these values, we compute

$$ap = M^2u_2 + Mu_1 + u_0 = 59313295525642905665289652194515573 69,$$
$$bq = M^2v_2 + Mv_1 + v_0 = 81881912986806196686803624641586187 39.$$

and obtain

$$p = \gcd(ap, N) = 126198501118389160989977983392586327,$$
$$q = \gcd(bq, N) = 143652478924221397696146709897519627.$$

This leads to the factorization of $N = pq$. We note that the first attack described in Section F.3 does not succeed to factor N. Indeed, we have $\frac{\log(\max_i(|v_i|))}{\log N} \approx 0.098$ which is larger than the value $\delta = 0.069$ for $k = 2$ and $\beta = 0.5$ in Table F.1. Finally, the overall recorded execution time for our attack using an off-the-shelf computer was 12 seconds.

F.4.3 The Number of Double Weak Primes in an Interval

In this section, we consider two positive integers n and M and present a study of the double weak primes with M, that is the primes $p, q \in \left[2^n, 2^{n+1}\right]$ such that there exists positive integer a and b that give the decompositions:

$$ap = \sum_{i=0}^{k} M^i u_i, \quad bq = \sum_{i=0}^{k} M^i v_i$$

where $|u_i| < N^\delta$, $|v_i| < N^\delta$ and δ satisfies Theorem F.4.1. We show that the number of the RSA moduli N in the interval $[2^{2n}, 2^{2(n+1)}]$ with a weak prime factors $p, q \in \left[2^n, 2^{n+1}\right]$ is lower bounded by 2^{η_2} where $\eta_2 > \frac{1}{2}$.

Theorem F.4.3. *Let n be a positive integer. For $k \geq 1$, define $M = \left\lceil 2^{\frac{n}{k}} \right\rceil$. Let $\mathcal{N}$ be the set of the weak RSA moduli $N \in \left[2^{2n}, 2^{2(n+1)}\right]$ such that $N = pq$ with $p = \left\lfloor \frac{\sum_{i=0}^{k} M^i u_i}{a} \right\rfloor + u$, $q = \left\lfloor \frac{\sum_{i=0}^{k} M^i v_i}{b} \right\rfloor + v$, $p, q \in \left[2^n, 2^{n+1}\right]$ for some small integers u, v, $a < N^{\delta}$, $b < N^{\delta}$, $|u_i| < N^{\delta}$ and $|v_i| < N^{\delta}$ for $i = 0, \ldots, k$ with*

$$\delta = \frac{1}{k+1}\left(1 - \left(\frac{1}{2}\right)^{\frac{k+1}{k}}\right) - \frac{1}{2}\left(1 - \left(\frac{1}{2}\right)^{\frac{1}{k}}\right).$$

Then the cardinality of $\mathcal{N}$ is at least $\#\mathcal{N} \geq 2^{\eta_2}$ where $\eta_2 = 4(k+1)n\delta$.

Proof. As in the proof of Theorem F.3.7, the number of prime numbers $p \in \left[2^n, 2^{n+1}\right]$ such that $p = \frac{\sum_{i=0}^{k} M^i u_i}{a} + u$ with $|u_i| < 2^{2n\delta}$ is

$$\#\mathcal{P} \geq 2^{2(k+1)n\delta}.$$

Then, the number $\mathcal{N}_2$ of RSA modulus $N \in \left[2^{2n}, 2^{2(n+1)}\right]$ with $N = pq$, where both p and q are weak primes is at least

$$\#\mathcal{N}_2 \geq 2^{4(k+1)n\delta} = 2^{\eta_2},$$

where $\eta_2 = 4(k+1)n\delta$. This terminates the proof.

□

In Table F.3, we present a list of values of the bound η_2 in terms of k and n.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
$n=512$	512	424	390	372	361	353	348
$n=1024$	1024	848	780	744	722	707	696
$n=2048$	2048	1696	1560	1489	1444	1414	1392

Table F.4: Lower bounds for η_2 under Theorem F.4.3.

F.5 Conclusions

In this paper we presented and illustrated two attacks based on factoring RSA moduli with weak primes. We further computed lower bounds for the

sets of weak moduli -that is, moduli made of at least one or two weak prime respectively- in the interval $[2^{2n}, 2^{2(n+1)}]$ and showed that these sets are much larger than the set of RSA prime factors satisfying Coppersmith's conditions, which effectively extending the likelihood for factoring RSA moduli.

Appendix G

New attacks on RSA with Moduli $N = p^r q$

C2SI 2015
[116] with Tajjeeddine Rachidi

Abstract :
We present three attacks on the Prime Power RSA with modulus $N = p^r q$. In the first attack, we consider a public exponent e satisfying an equation $ex - \phi(N)y = z$ where $\phi(N) = p^{r-1}(p-1)(q-1)$. We show that one can factor N if the parameters $|x|$ and $|z|$ satisfy $|xz| < N^{\frac{r(r-1)}{(r+1)^2}}$ thereby extending the recent results of Sakar [132]. In the second attack, we consider two public exponents e_1 and e_2 and their corresponding private exponents d_1 and d_2. We show that one can factor N when d_1 and d_2 share a suitable amount of their most significant bits, that is $|d_1 - d_2| < N^{\frac{r(r-1)}{(r+1)^2}}$. The third attack enables us to factor two Prime Power RSA moduli $N_1 = p_1^r q_1$ and $N_2 = p_2^r q_2$ when p_1 and p_2 share a suitable amount of their most significant bits, namely, $|p_1 - p_2| < \frac{p_1}{2rq_1q_2}$.

G.1 Introduction

The RSA public-key cryptosystem, invented in 1978 by Rivest, Shamir and Adleman [131], is one of the most popular systems in use today. In the RSA cryptosystem, the public key is (N, e) where the modulus $N = pq$ is a product of two primes of the same bitsize, and the public exponent is a positive integer satisfying $ed \equiv 1 \pmod{\phi(N)}$. In RSA, encryption and decryption require executing heavy exponential multiplications modulo the large integer N. To reduce the decryption time, one may be tempted to use a small private exponent d. However, in 1990 Wiener [147] showed that RSA is insecure if $d < \frac{1}{3}N^{0.25}$, and Boneh and Durfee [17] improved the bound to $d < N^{0.292}$. In 2004, Blömer and May [13] combined both Wiener's method and Boneh and Durfee's method to show that RSA is insecure if the public exponent e satisfies an equation $ex + y = k\phi(N)$ with $x < \frac{1}{3}N^{\frac{1}{4}}$ and $|y| \leq N^{-\frac{3}{4}}ex$.

Concurrent to these efforts, many RSA variants have been proposed in order to ensure computational efficiency while maintaining the acceptable levels of security. One such important variant is the Prime Power RSA. In Prime Power RSA the modulus N is in the form $N = p^r q$ for $r \geq 2$. In [145], Takagi showed how to use the Prime Power RSA to speed up the decryption process when the public and private exponents satisfy an equation $ed \equiv 1 \pmod{(p-1)(q-1)}$. As in the standard RSA cryptosystem, the security of the Prime Power RSA depends on the difficulty of factoring integers of the form $N = p^r q$.

Therefore, a Prime Power RSA modulus must be appropriately chosen, since it has to resist factoring algorithms such as the Number Field Sieve [85] and the Elliptic Curve Method [84]. Table G.1, shows the suggested secure Power RSA forms as a function of the size of the modulus back in 2002 (see [33]). Note that, due to the ever increasing development of computing hardware, the form $N = p^2 q$ is no longer recommended for 1024 bit modulus.

Modulus size (bits)	1024	1536	2048	3072	4096	8192
Form of the modulus N	pq, p^2q	pq, p^2q	pq, p^2q	pq, p^2q	pq, p^2q, p^3q	pq, p^2q, p^3q, p^4q

Table G.1: Optimal number of prime factors of a Prime Power RSA modulus [33].

In 1999, Boneh, Durfee, and Howgrave-Graham [16] presented a method for factoring $N = p^r q$ when r is large. Furthermore, Takagi [145] proved that one can factor N if $d < N^{\frac{1}{2(r+1)}}$, and May [92] improved the bound to $d < N^{\frac{r}{(r+1)^2}}$ or $d < N^{\frac{(r-1)^2}{(r+1)^2}}$. Very recently, Lu, Zhang and Lin [88] improved the bound to $d < N^{\frac{r(r-1)}{(r+1)^2}}$, and Sarkar [132] improved the bound for $N = p^2 q$ to $d < N^{0.395}$ and gave explicit bounds for $r = 3, 4, 5$.

In this paper, we focus on the Prime Power RSA with a modulus $N = p^r q$, and present three new attacks: In the first attack we consider a public exponent e satisfying an equation $ex - \phi(N)y = z$ where x and y are positive integers. Using a recent result of Lu, Zhang and Lin [88], we show that one can factor N in polynomial time if $|xz| < N^{\frac{r(r-1)}{(r+1)^2}}$. In the standard situation $z = 1$, the condition becomes $d = x < N^{\frac{r(r-1)}{(r+1)^2}}$ which improves the bound of May [92] for $r \geq 3$ and retrieves the bound of Lu, Zhang and Lin [88]. Note that unlike Sarkar [132] who solves $ex - \phi(N)y = 1$, we solve a more general equation $ex - \phi(N)y = z$. This leads to less constraints on the solution space, which in turn leads to an increase in the number of solutions to the equation. Intuitively speaking, our method has higher likelihood of finding solutions; that is, factoring RSA. In section G.3, we shall present an example supporting this claim.

In the second attack, we consider an instance of the Prime Power RSA with modulus $N = p^r q$. We show that one can factor N if two private keys d_1 and d_2 share an amount of their most significant bits, that is if $|d_1 - d_2|$ is small enough. More precisely, we show that if $|d_1 - d_2| < N^{\frac{r(r-1)}{(r+1)^2}}$, then N can be factored in polynomial time. The method we present is based on a recent result of [88] with Coppersmith's method for solving an univariate linear equation.

In the third attack, we consider two instances of the Prime Power RSA with two moduli $N_1 = p_1^r q_1$ and $N_2 = p_2^r q_2$ such that the prime factors p_1 and p_2 share an amount of their most significant bits, that is $|p_1 - p_2|$ is small. More precisely, we show that one can factor the RSA moduli N_1 and N_2 in polynomial time if $|p_1 - p_2| < \frac{p_1}{2rq_1q_2}$. The method we use for this attack is based on the continued fraction algorithm.

The rest of this paper is organized as follows: In Section 2, we briefly review the preliminaries necessary for the attacks, namely Coppersmith's technique for solving linear equations and the continued fractions theorem. In Section 3, we present the first attack on the Prime Power RSA, which is valid with no conditions on the prime factors. In Section 4, we present the second attack in the situation where two decryption exponents share an amount of their most significant bits. In Section 5, we present the third attack on the Prime Power RSA when the prime factors share an amount of their most significant bits. We then conclude the paper in Section 6.

G.2 Preliminaries

In this section, we present some basics on Coppersmith's method for solving linear modular polynomial equations and an overview of the continued fraction algorithm. Both techniques are used in the crafting of our attacks.

First, observe that if $N = p^r q$ with $q < p$, then $p^{r+1} > p^r q = N$, and $p > N^{\frac{1}{r+1}}$. Hence throughout this paper, we will use the inequality $p > N^\beta$ where $\beta = \frac{1}{r+1}$.

G.2.1 Linear Modular Polynomial Equations

In 1995, Coppersmith [34] developed powerful lattice-based techniques for solving both modular polynomial diophantine equations with one variable and two variables. These techniques have been generalized to more variables, and have served for cryptanalysis of many instances of RSA. More on this can be found in [61, 93]. In [59], Herrmann and May presented a method for finding the small roots of a modular polynomial equation $f(x_1, \ldots, x_n) \equiv 0 \pmod{p}$ where $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ and p is an unknown divisor of a known integer N. Their method is based on the seminal work of Coppersmith [34]. Very recently, Lu, Zhang and Lin [88] presented a generalization for finding the small roots of a modular polynomial equation $f(x_1, \ldots, x_n) \equiv 0 \pmod{p^v}$, where p^v is a divisor of some composite integer N. For the bivariate case, they proved the following result, which we shall use in the crafting of our attacks.

Theorem G.2.1 (Lu, Zhang and Lin). *Let N be a composite integer with a divisor p^u such that $p \geq N^\beta$ for some $0 < \beta \leq 1$. Let $f(x, y) \in \mathbb{Z}[x, y]$ be a homogenous linear polynomial. Then one can find all the solutions (x, y) of the equation $f(x, y) = 0 \mod p^v$ with $\gcd(x, y) = 1$, $|x| < N^{\gamma_1}$, $|y| < N^{\gamma_2}$, in polynomial time if*

$$\gamma_1 + \gamma_2 < uv\beta^2.$$

G.2.2 The Continued Fractions Algorithm

We present here the well known result of Legendre on convergents of a continued fraction expansion of a real number. The details can be found in [57]. Let ξ be a positive real number. 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_i is an integer. This expands ξ as a continued fraction in the following form:

$$\xi = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{\ddots}}}}, \ a_0 \in \mathbb{N}, \text{ and } a_i \in \mathbb{N}^* \text{ for } i \geq 1,$$

which is often rewritten as $\xi = [a_0, a_1, \ldots, a_n, \ldots]$. For $i \geq 0$, the rational numbers $[a_0, a_1, \ldots, a_i]$ are the convergents of ξ. If $\xi = \frac{a}{b}$ is a rational number, then $\xi = [a_0, a_1, \ldots, a_n]$ for some positive integer n, and the continued fraction expansion of ξ is finite with the total number of convergents being polynomial in $\log(b)$. The following result enables one to determine if a rational number $\frac{a}{b}$ is a convergent of the continued fraction expansion of a real number ξ (see Theorem 184 of [57]).

Theorem G.2.2 (Legendre). *Let ξ be a positive real 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.

Note that the continued fractions expansion process is polynomial in time.

G.3 The First Attack on Prime Power RSA with Modulus $N = p^r q$

In this section, we present an attack on the Prime Power RSA when the public key (N, e) satisfies an equation $ex - \phi(N)y = z$ with small parameters x and $|z|$.

Theorem G.3.1. *Let $N = p^r q$ be a Prime Power RSA modulus and e a public exponent satisfying the equation $ex - \phi(N)y = z$ with $y \not\equiv 0 \pmod{pq}$, $1 < e < \phi(N)$ and $\gcd(e, \phi(N)) = 1$. Then one can factor N in polynomial time if*

$$|xz| < N^{\frac{r(r-1)}{(r+1)^2}}.$$

Proof. Suppose that $e < N$ satisfies an equation $ex - \phi(N)y = z$ with $|x| < N^\delta$ and $|z| < N^\gamma$. Then, since $\phi(N) = p^{r-1}(p-1)(q-1)$, we get $ex - z \equiv 0 \pmod{p^{r-1}}$. Applying Theorem G.2.1 with $u = r$, $v = r-1$ and $\beta = \frac{1}{r+1}$, we can solve the equation in polynomial time if

$$\delta + \gamma < uv\beta^2 = \frac{r(r-1)}{(r+1)^2},$$

that is $|xz| < N^{\frac{r(r-1)}{(r+1)^2}}$. Since $\frac{e}{\phi(N)} < 1$, then, using x and z in the equation $ex - \phi(N)y = z$, we get for sufficiently large N comparatively to r,

$$y = \frac{ex - z}{\phi(N)} < \frac{e|x|}{\phi(N)} + \frac{|z|}{\phi(N)} < |x| + |z| \leq 1 + |xz| < 1 + N^{\frac{r(r-1)}{(r+1)^2}} < N.$$

Hence, when $y \not\equiv 0 \pmod{pq}$, we get

$$\gcd(ex - z, N) = \gcd(p^{r-1}(p-1)(q-1)y, p^r q) = g,$$

with $g = p^{r-1}$, $g = p^r$ or $g = p^{r-1}q$. If $g = p^{r-1}$, then $p = g^{\frac{1}{r-1}}$, if $g = p^r$, then $p = g^{\frac{1}{r}}$ and if $g = p^{r-1}q$, then $p = \frac{N}{g}$. This leads to the factorization of N.

□

Example G.3.2. For $r = 2$ and $N = p^r q$, let us take for N and e the 55 digit numbers

$$N = 8138044578297117319482018441148072252199996769522371021,$$
$$e = 1199995230601021126201343651611107957480251354355883029.$$

In order to solve the diophantine equation $ex - \phi(N)y = z$, we transformed it into the equation $ex - z \equiv 0 \pmod{p^{r-1}}$ using Theorem G.3.1. To be able to apply Coppersmith's technique via Theorem G.2.1, we chose the parameters $m = 7$, $t = 6$ so that the dimension of constructed the lattice is 36, and $X = \left\lceil N^{\frac{r(r-1)}{(r+1)^2}} \right\rceil = 1592999974064$. We built the lattice using the polynomial $f(x_1, x_2) = x_1 + ex_2$, then applied the LLL algorithm [86], and used Gröbner basis method to find the smallest solution $x_1 = -11537$ and $x_2 = 7053$ to $f(x_1, x_2) \equiv 0 \pmod{p^{r-1}}$ in 174 seconds using an off-the-shelf computer. From this solution, we deduced $p = gcd(x_1 + ex_2, N) = 2294269585934949239$, and finally recovered $q = \frac{N}{p^2} = 1546077175000723901$. We then computed $\phi(N)$ and $d \equiv e^{-1} \pmod{\phi(N)}$ as follows:

$$\phi(N) = 813804457829711731067122766808956194625789692526157980 0,$$
$$d = 201599474774838877298243639381121331736197186551075626 9.$$

Observe that $d \approx N^{0.98}$ which is out of range of Sarkar's bound [132] which can only retrieve private keys $d < N^{0.395}$ for $r = 2$.

G.4 The Second Attack on Prime Power RSA using Two Decryption Exponents

In this section, we present an attack on the Prime Power RSA when two private exponents d_1 and d_2 share an amount of their most significant bits, that is $|d_1 - d_2|$ is small.

Theorem G.4.1. *Let $N = p^r q$ be an RSA modulus and d_1 and d_2 be two private exponents such that $e_1 e_2 (d_1 - d_2) - (e_2 - e_1) \not\equiv 0 \pmod{N}$. Then, one can factor N in polynomial time, if*

$$|d_1 - d_2| < N^{\frac{r(r-1)}{(r+1)^2}}.$$

Proof. Suppose that $e_1 d_1 - k_1 \phi(N) = 1$ and $e_2 d_2 - k_2 \phi(N) = 1$ with $e_1 > e_2$. Hence $e_1 d_1 \equiv 1 \pmod{\phi(N)}$ and $e_2 d_2 \equiv 1 \pmod{\phi(N)}$. Multiplying the first equation by e_2 and the second by e_1 and subtracting, we get

$$e_1 e_2 (d_1 - d_2) \equiv e_2 - e_1 \pmod{\phi(N)}.$$

Since $\phi(N) = p^{r-1}(p-1)(q-1)$, we get $e_1 e_2(d_1 - d_2) \equiv e_2 - e_1 \pmod{p^{r-1}}$. Now, consider the modular linear equation

$$e_1 e_2 x - (e_2 - e_1) \equiv 0 \pmod{p^{r-1}},$$

$d_1 - d_2$ is a root of such equation. Suppose further that $|d_1 - d_2| < N^{\delta}$, then applying Theorem G.2.1 with $u = r$, $v = r - 1$ and $\beta = \frac{1}{r+1}$ will lead to the solution $x = d_1 - d_2$ obtained in polynomial time if

$$\delta < uv\beta^2 = \frac{r(r-1)}{(r+1)^2}.$$

That is if $|d_1 - d_2| < N^{\frac{r(r-1)}{(r+1)^2}}$. Computing

$$\gcd(e_1 e_2 x - (e_2 - e_1), N) = \gcd\left(p^{r-1}(p-1)(q-1)y, p^r q\right) = g,$$

and assuming that $e_1 e_2(d_1 - d_2) - (e_2 - e_1) \not\equiv 0 \pmod{N}$ will lead to determining p, hence factoring N as follows: $p = g^{\frac{1}{r-1}}$ when $g = p^{r-1}$, or $p = g^{\frac{1}{r}}$ when $g = p^r$, or $p = \frac{N}{g}$ if $g = p^{r-1}q$.

□

Example G.4.2. Let us present an example corresponding to Theorem G.4.1. Consider $N = p^2 q$ with

$$\begin{aligned} N &= 609325385148612087885947195839973772588594652655362621 9, \\ e_1 &= 274960038184748738971596476723561880252967585560637741 1, \\ e_2 &= 357508124495241400931639650151237222654589255889827655 1. \end{aligned}$$

The polynomial equation is $f(x) = e_1 e_2 x - (e_2 - e_1) \equiv 0 \pmod{p^{r-1}}$, which can be transformed into $g(x) = x - a \equiv 0 \pmod{p^{r-1}}$ where $a \equiv (e_2 - e1)(e_1 e_2)^{-1} \pmod{N}$. Using $m = 8$ and $t = 6$, we built a lattice with dimension $\omega = 9$. Applying the LLL algorithm [86] and solving the first reduced polynomials, we get the solution $x_0 = 1826732340$. Hence $\gcd(f(x_0), N) = p = 1789386140116417697$ and finally $q = \frac{N}{p^2} = 1903010275819064491$. The whole process took less than 4 seconds using an off-the-shelf computer. Then, using $\phi(N) = p(p-1)(q-1)$, we retrieved the private exponents $d_1 \equiv e_1^{-1} \pmod{\phi(N)}$ and $d_2 \equiv e_2^{-1} \pmod{\phi(N)}$. Note that again $d_1 \approx d_2 \approx N^{0.99}$ which Sarkar's method with the bound $d < N^{0.395}$ could not possibly retrieve.

G.5 The Third Attack on Prime Power RSA with Two RSA Moduli

In this section, we consider two Prime Power RSA moduli $N_1 = p_1^r q_1$ and $N_2 = p_2^r q_2$, where p_1 and p_2 share an amount of their most significant bits.

Theorem G.5.1. *Let $N_1 = p_1^r q_1$ and $N_2 = p_2^r q_2$ be two RSA moduli with $p_1 > p_2$. If*

$$|p_1 - p_2| < \frac{p_1}{2r q_1 q_2},$$

then, one can factor N in polynomial time.

Proof. Suppose that $N_1 = p_1^r q_1$ and $N_2 = p_2^r q_2$ with $p_1 > p_2$. Then $q_2 N_1 - q_1 N_2 = q_1 q_2 \left(p_1^r - p_2^r\right)$. Hence

$$\left|\frac{N_2}{N_1} - \frac{q_2}{q_1}\right| = \frac{q_1 q_2 \left|p_1^r - p_2^r\right|}{q_1^2 p_1^r}.$$

In order to apply Theorem G.2.2, we need that $\frac{q_1 q_2 |p_1^r - p_2^r|}{q_1^2 p_1^r} < \frac{1}{2q_1^2}$, or equivalently

$$|p_1^r - p_2^r| < \frac{p_1^r}{2 q_1 q_2}. \tag{G.1}$$

Observe that

$$|p_1^r - p_2^r| = |p_1 - p_2| \sum_{i=0}^{r-1} p_1^{r-1-i} p_2^i < r |p_1 - p_2| p_1^{r-1}.$$

Then (G.1) is fulfilled if $r|p_1 - p_2| p_1^{r-1} < \frac{p_1^r}{2 q_1 q_2}$, that is if

$$|p_1 - p_2| < \frac{p_1}{2 r q_1 q_2}.$$

Under this condition, we get $\frac{q_2}{q_1}$ among the convergents of the continued fraction expansion of $\frac{N_2}{N_1}$. Using q_1 and q_2, we get $p_1 = \left(\frac{N_1}{q_1}\right)^{\frac{1}{r}}$ and $p_2 = \left(\frac{N_2}{q_2}\right)^{\frac{1}{r}}$.

□

Example G.5.2. We present here an example corresponding to Theorem G.5.1. Consider $N_1 = p_1^2 q_1$ and $N_2 = p_2^2 q_2$ with

$$N_1 = 17098723391376942050589691743730471981669135383303448246 1,$$
$$N_2 = 12053291181972688288163071400313523776667560282425096592 1.$$

We applied the continued fraction algorithm to compute the first 40 convergents of $\frac{N_2}{N_1}$. Every convergent is a candidate for the ratio $\frac{q_2}{q_1}$ of the prime factors. One of the convergents is $\frac{36443689}{51698789}$ leading to $q_2 = 36443689$ and $q_1 = 51698789$. This gives the prime factors p_1 and p_2

$$p_1 = \sqrt{\frac{N_1}{q_1}} = 181861872438294295146044 3,$$

$$p_2 = \sqrt{\frac{N_2}{q_2}} = 181861872438294303567268 3.$$

G.6 Conclusion

In this paper, we have considered the Prime Power RSA with modulus $N = p^r q$ and public exponent e. We presented three new attacks to factor the modulus in polynomial time. The first attack can be applied if small parameters x, y and z satisfying the equation $ex - \phi(N)y = z$ can be found . The second attack can be applied when two private exponents d_1 and d_2 share an amount of their most significant bits. The third attack can be applied when two Prime Power RSA moduli $N_1 = p_1^r q_1$ and $N_2 = p_2^r q_2$ are such that p_1 and p_2 share an amount of their most significant bits.

Appendix H

A New Attack on the KMOV Cryptosystem

Bulletin of the Korean Mathematical Society
2014
[119]

Abstract :
In this paper, we analyze the security of the KMOV public key cryptosystem. KMOV is based on elliptic curves over the ring $\mathbb{Z}_n$ where $n = pq$ is the product of two large unknown primes of equal bit-size. We consider KMOV with a public key (n, e) where the exponent e satisfies an equation $ex - (p+1)(q+1)y = z$, with unknown parameters x, y, z. Using Diophantine approximations and lattice reduction techniques, we show that KMOV is insecure when x, y, z are suitably small.

00

H.1 Introduction

In 1991, Koyama, Maurer, Okamoto and Vanstone [79] introduced a new public key cryptosystem based on elliptic curves, called KMOV. The KMOV cryptosystem is based on elliptic curves over the ring $\mathbb{Z}_n$ where $n = pq$ is an RSA modulus, that is, the product of two large unknown primes of equal bit-size. Introduced in 1978 by Rivest, Shamir and Adleman, RSA [131] is one of the most popular cryptosystems in research as well as in commercial domain (see [15], [61]). The RSA public key is denoted by (n, e) where $n = pq$ is an RSA modulus and e is an integer satisfying $\gcd(e, (p-1)(q-1)) = 1$. The corresponding private exponent d is an integer satisfying $ed \equiv 1 \pmod{(p-1)(q-1)}$. Then, there exists some integer k such that

$$ed - k(p-1)(q-1) = 1. \tag{H.1}$$

Similarly, the KMOV public key is denoted by (n, e) where $n = pq$ and e is an integer satisfying $\gcd(e, (p+1)(q+1)) = 1$. The corresponding private exponent d is an integer satisfying $ed \equiv 1 \pmod{(p+1)(q+1)}$ which can be reformulated as an equation

$$ed - k(p+1)(q+1) = 1. \tag{H.2}$$

The security of RSA and KMOV is mainly based on the difficulty of factoring the RSA modulus n. To speed up the encryption or decryption one may try to use small public or secret decryption exponent. Many important papers studied RSA and KMOV to explore the weaknesses in using small exponents. In 1990, Wiener [147] showed that using equation (H.1) and the continued fraction algorithm, it is possible to break RSA if the private key d satisfies $d < \frac{1}{3}n^{0.25}$. In 2004, Blömer and May [13] described an attack on RSA starting with the equation

$$ex - k(p-1)(q-1) = y.$$

Using the continued fraction algorithm and lattice reduction techniques, they showed that RSA is insecure if $0 < x < \frac{1}{3}n^{0.25}$ and $|y| = \mathcal{O}\left(n^{-0.75}ex\right)$. In this paper, we consider KMOV with a public exponent e satisfying the more general equation

$$ed - k(p+1)(q+1) = z. \tag{H.3}$$

where x and y are co-prime positive integers. Observe that this equation has infinitely many solutions but we will focus on small solutions. In 1995, Pinch [128] extended the Wiener attack to KMOV using similar techniques applied with equation (H.2), that is when $z = 1$. Similarly, Ibrahimpasic [69], studied the security of KMOV with short secret exponents.

We mainly focus on the equation (H.2) which is a generalization of the equation (H.2). We use Diophantine approximations to find x, y among the convergents of the continued fraction expansion of $\frac{e}{n}$ when x, y and z satisfy

$$|z| < \frac{(p-q)n^{\frac{1}{4}}y}{3(p+q)}, \quad xy < \frac{\sqrt{2}\sqrt{n}}{12}.$$

After finding x and y, one can get an approximation $\tilde{p}$ of p satisfying $|p-\tilde{p}| < n^{\frac{1}{4}}$ where

$$\tilde{p} = \frac{1}{2}\left(\frac{ex}{y} - n - 1\right) + \frac{1}{2}\sqrt{\left|\left(\frac{ex}{y} - n - 1\right)^2 - 4n\right|}.$$

Finally, this approximation leads to the factorization of n by using Coppersmith's Theorem [34].

The rest of this paper is organized as follows. In the next section, we review some necessary definitions and notation on elliptic curves and recall the KMOV cryptosystem. In section 3, we present our new attack on KMOV. In Section 4, we propose a numerical example. We conclude in Section 5.

H.2 Preliminaries

In this section, we give a brief description of the KMOV cryptosystem and elliptic curves (see [140] for more details on elliptic curves).

H.2.1 Elliptic Curves over $\mathbb{F}_p$

An elliptic curve over a field $\mathbb{K}$ is an algebraic curve with no singular points, given by the Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \quad a_i \in \mathbb{K},$$

together with a single element denoted $\mathcal{O}$ and called the point at infinity. The elliptic curve E over $\mathbb{K}$ is denoted $E/\mathbb{K}$ and the set of solutions $(x, y) \in \mathbb{K}^2$ together with $\mathcal{O}$ is denoted $E(\mathbb{K})$. Given two points $P, Q \in E(\mathbb{K})$ we define a third point $P + Q$ so that $E(\mathbb{K})$ forms an abelian group with this addition operation.

- The point $\mathcal{O}$ serves as the identity element.
- The opposite of $P = (x_1, y_1)$, is $-P = (x_1, -y_1 - a_1x_1 - a_3)$.
- If $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ with $Q \neq -P$, then $P + Q = (x_3, y_3)$ where
$$\begin{cases} x_3 &= \lambda^2 - x_1 - x_2 - a_2 + a_1\lambda, \\ y_3 &= -y_1 - (x_3 - x_1)\lambda - a_1x_3 - a_3, \end{cases}$$
where
$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{if} \quad x_1 \neq x_2, \\ \\ \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1^2 + a_1x_1 + a_3} & \text{if} \quad x_1 = x_2, \end{cases}$$

If $\mathbb{K}$ is of characteristic different from 2 or 3, the equation of the elliptic curve E can be transformed into the reduced Weierstrass form

$$y^2 = x^3 + ax + b, \quad a, b \in \mathbb{K},$$

where $4a^3 + 27b^2 \neq 0$. When $\mathbb{K} = \mathbb{F}_p$ for some prime $p > 3$, such a curve will be denoted $E_p(a, b)$.

Theorem H.2.1 (Hasse)**.** *The order of the group $E_p(a, b)(\mathbb{F}_p)$ is given by*

$$\#E_p(a, b) = p + 1 - a_p,$$

where $|a_p| \leq 2\sqrt{p}$.

For the special case $a = 0$, the order $\#E_p(0, b)$ can easily be determined.

Lemma H.2.2. *Let $p > 3$ be a prime satisfying $p \equiv 2 \pmod 3$ and $0 < b < p$. Then*

$$\#E_p(0, b) = p + 1.$$

H.2.2 Elliptic Curves over $\mathbb{Z}_n$

We now consider elliptic curves over the ring $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ where $n = pq$ is the product of two large distinct primes p and q. An elliptic curve $E_n(a, b)$ over $\mathbb{Z}_n$ is the set of points $(x, y) \in \mathbb{Z}_n^2$ satisfying

$$y^2 = x^3 + ax + b \pmod{n}$$

together with the point at infinity $\mathcal{O}$. The addition law can be extended for points in a curve $E_n(a, b)$ over $\mathbb{Z}_n$. Note that the addition law is not always well-defined when using analytical expressions since there are elements in $\mathbb{Z}_n$ which are not invertible. It follows that $E_n(a, b)(\mathbb{Z}_n)$ is not a group. By the Chinese Remainder Theorem, the mapping

$$E_n(a, b) \to E_p(a, b) \times E_q(a, b)$$

defined by the the natural projections is a bijection. Thus, a point (x, y) of the elliptic curve $E_n(a, b)$ is associated to the point

$$((x \pmod{p}, y \pmod{p}), (x \pmod{q}, y \pmod{q})) \in E_p(a, b) \times E_q(a, b).$$

The points $(\mathcal{O}, P)$ and $(P, \mathcal{O})$ can not be represented like this. Finding such a point is, however, very unlikely and would lead to the factorization of n. The Chinese Remainder Theorem leads to the following lemma.

Lemma H.2.3. *Let $n = pq$ be an RSA modulus and $E_n(a, b)$ an elliptic curve over $\mathbb{Z}_n$ with $\gcd\left(4a^3 + 27b^2, n\right) = 1$. Then for any $P \in E_n(a, b)$ and any integer k, we have*

$$(1 + k\#E_p(a, b)\#E_q(a, b))P = P.$$

H.2.3 KMOV Scheme

In 1991, Koyama, Maurer, Okamoto and Vanstone [79] proposed the so called KMOV cryptosystem using elliptic curves defined over the elliptic curve $E_n(a, b)$ where $n = pq$ is an RSA modulus.

- **Key Generation**
  INPUT: The bit-length k of the RSA modulus.
  OUTPUT: The public key (n, e) and the private key (n, d).

1. Find two primes, p and q, of length $k/2$ bits satisfying $p \equiv q \equiv 2 \pmod 3$.
2. Compute the RSA modulus $n = pq$.
3. Choose a public key e co-prime to $(p+1)(q+1)$.
4. Compute the inverse d of e mod $((p+1)(q+1))$.
5. Return the public key (n, e) and the private key (n, d).

- **KMOV Encryption**
  INPUT: The public key (n, e) and the plaintext message m.
  OUTPUT: The cyphertext (c_1, c_2).

1. Represent the message m as a couple $(m_1, m_2) \in \mathbb{Z}_n^2$.
2. Compute $b = m_2^2 - m_1^3 \pmod n$.
3. Compute the point $(c_1, c_2) = e(m_1, m_2)$ on the elliptic curve $y^2 = x^3 + b \pmod n$.
4. Return (c_1, c_2).

- **KMOV Decryption**
  INPUT: The private key (n, d) and the cyphertext (c_1, c_2).
  OUTPUT: The plaintext message (m_1, m_2).

1. Compute $b = c_2^2 - c_1^3 \pmod n$.
2. Compute the point $(m_1, m_2) = d(c_1, c_2)$ on the elliptic curve $y^2 = x^3 + b \pmod n$.
3. Return (m_1, m_2).

The decryption scheme is valid since, using Lemma H.2.2 and Lemma H.2.3, we have

$$\begin{aligned} d(c_1, c_2) &= de(m_1, m_2) \\ &= (1 + k(p+1)(q+1))(m_1, m_2) \\ &= (1 + k\#E_p(0, b)\#E_q(0, b))(m_1, m_2) \\ &= (m_1, m_2), \end{aligned}$$

where k is the integer satisfying $ed = 1 + k(p+1)(q+1)$.

H.3 The New attack on the KMOV Cryptosystem

Let $n = pq$ be an RSA modulus as required by the KMOV Cryptosystem. Suppose that e is an integer satisfying $\gcd(e, (p+1)(q+1)) = 1$. Let x, y be co-prime positive integers. Define z by

$$ex - (p+1)(q+1)y = z.$$

In this section, we show that, under some conditions, it is possible find x, y, p, q which leads to the factorization of the RSA modulus and breaks the system. We shall need the following useful result.

Lemma H.3.1. *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. We have

$$(p+q)^2 = (p-q)^2 + 4n > 4n.$$

Then $p + q > 2\sqrt{n}$. On the other hand, since $q < p < 2q$, then $n < p^2 < 2n$ and $\sqrt{n} < p < \sqrt{2n}$. Hence

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

This terminates the proof. □

We shall also need the following result (see [57], Theorem 184).

Theorem H.3.2. *Let α be a real number. If x and y are positive integers such that $\gcd(x, y) = 1$ and*

$$\left|\alpha - \frac{y}{x}\right| < \frac{1}{2x^2},$$

then $\frac{y}{x}$ is one of the convergents of the continued fraction expansion of α.

Now, we can prove the following theorem which permits to find x and y using the convergents of the continued fraction expansion of $\frac{e}{n}$.

Theorem H.3.3. *Let $n = pq$ be an RSA modulus with $q < p < 2p$. Suppose that the public exponent e satisfies an equation $ex - (p+1)(q+1)y = z$ with* $\gcd(x, y) = 1$ *and*

$$|z| < n^{\frac{1}{4}}y, \quad xy < \frac{\sqrt{2}\sqrt{n}}{12}.$$

Then $\frac{y}{x}$ is one of the convergents of the continued fraction expansion of $\frac{e}{n}$.

Proof. Transforming the equation $ex - (p+1)(q+1)y = z$, we get

$$ex - ny = (p + q + 1)y + z.$$

Dividing by nx, we get

$$\frac{e}{n} - \frac{y}{x} = \frac{(p + q + 1)y + z}{nx}. \tag{H.4}$$

Assume that $|z| < n^{\frac{1}{4}}y$. Then using Lemma H.3.1, we get

$$\begin{aligned} |(p + q + 1)y + z| &\leq (p + q + 1)y + |z| \\ &\leq (p + q + 1)y + n^{\frac{1}{4}}y \\ &= (p + q + 1 + n^{\frac{1}{4}})y \\ &< 2(p + q)y \\ &\leq 3\sqrt{2}\sqrt{n}y. \end{aligned}$$

Now, assume that $xy < \frac{\sqrt{2}\sqrt{n}}{12}$. Then (H.4) implies

$$\left|\frac{e}{n} - \frac{y}{x}\right| = \frac{|(p + q + 1)y + z|}{nx} < \frac{3\sqrt{2}\sqrt{n}y}{nx} < \frac{1}{2x^2}.$$

Then, applying Theorem H.3.2, $\frac{y}{x}$ is a convergent of the continued fraction expansion of $\frac{e}{N}$. This terminates the proof. □

Next assume that x and y are known in the equation $ex-(p+1)(q+1)y = z$. We show how to find p and q. Let us first refer to the following existing result (see [34]).

Theorem H.3.4 (Coppersmith)**.** *Let $n = pq$ be an RSA modulus with $q < p < 2q$. Suppose we know an approximation $\tilde{p}$ of p with $|p - \tilde{p}| < n^{\frac{1}{4}}$. Then n can be factored in time polynomial in* $\log n$.

Next we present the main result.

Theorem H.3.5. *Let $n = pq$ be an RSA modulus with $q < p < 2q$. Suppose that e is an exponent satisfying an equation $ex - (p+1)(q+1)y = z$ with $\gcd(x, y) = 1$ and*

$$|z| < \frac{(p-q)n^{\frac{1}{4}}y}{3(p+q)}, \quad xy < \frac{\sqrt{2}\sqrt{n}}{12}.$$

Then n can be factored in polynomial time.

Proof. Suppose e satisfies an equation $ex-(p+1)(q+1)y = z$. If $|z| < \frac{(p-q)n^{\frac{1}{4}}y}{3(p+q)}$ then $|z| < n^{\frac{1}{4}}y$. In addition if $\gcd(x, y) = 1$ and $xy < \frac{\sqrt{2}\sqrt{n}}{12}$, then, by Theorem H.3.3, we find x and y among the convergents of $\frac{e}{n}$. Next, put

$$U = \frac{ex}{y} - n - 1, \qquad V = \sqrt{|U^2 - 4n|}.$$

Starting with the equation $ex - (p+1)(q+1)y = z$, we get

$$|U - p - q| = \left|\frac{ex}{y} - n - 1 - p - q\right| = \frac{|z|}{y} < \frac{(p-q)n^{\frac{1}{4}}}{3(p+q)}.$$

Hence

$$|U - p - q| < n^{\frac{1}{4}}. \tag{H.5}$$

Now, we have

$$\begin{aligned} \left|(p-q)^2 - V^2\right| &= \left|(p-q)^2 - \left|U^2 - 4n\right|\right| \\ &\leq \left|(p-q)^2 - U^2 + 4n\right| \\ &= \left|(p+q)^2 - U^2\right| \\ &= |p+q-U|\,(p+q+U)\,. \end{aligned}$$

Dividing by $p - q + V$, we get

$$|p - q - V| \leq \frac{|p+q-U|\,(p+q+U)}{p-q+V}. \tag{H.6}$$

Observe that (H.5) implies

$$p + q + U < 2(p+q) + n^{\frac{1}{4}} < 3(p+q).$$

On the other hand, we have $p - q + V > p - q$. Plugging in (H.6), we get

$$|p - q - V| < \frac{3(p+q)(p-q)n^{\frac{1}{4}}}{3(p+q)(p-q)} = n^{\frac{1}{4}}.$$

Combining this with (H.5), we deduce

$$\begin{aligned} \left| p - \frac{U+V}{2} \right| &= \left| \frac{p+q}{2} - \frac{U}{2} + \frac{p-q}{2} - \frac{V}{2} \right| \\ &\leq \left| \frac{p+q}{2} - \frac{U}{2} \right| + \left| \frac{p-q}{2} - \frac{V}{2} \right| \\ &< n^{\frac{1}{4}}. \end{aligned}$$

This implies that $\frac{U+V}{2}$ is an approximation of p up to an error term of at most $n^{\frac{1}{4}}$. Then Coppersmith's Theorem H.3.4 will find p in polynomial time and the factorization of n follows. □

Let us summarize the factorization algorithm.

Algorithm 6 The factorization algorithm

Require: a public key (N, e) satisfying $N = pq$, $q < p < 2q$ and $ex - (p+1)(q+1) = z$ for some parameters x, y, z.

Ensure: the prime factors p and q.

1: Compute the continued fraction expansion of $\frac{e}{n}$.
2: For every convergent $\frac{y}{x}$ of $\frac{e}{n}$ with $x < \sqrt{n}$:
3: Compute $U = \frac{ex}{y} - n - 1$ and $V = \sqrt{|U^2 - 4n|}$.
4: Apply Coppersmith's algorithm with $\frac{U+V}{2}$ as an approximation of p.
5: If Coppersmith's algorithm outputs the factorization of n, then stop.

H.4 A Numerical Example

As an example let us take for n and e the numbers

$$\begin{aligned} n &= 17342828614189479815674825 1, \\ e &= 72375394700973490734223 9. \end{aligned}$$

The first convergents of the continued fraction expansion of $\frac{e}{n}$ are

$$[0, \frac{1}{239}, \frac{1}{240}, \frac{2}{479}, \frac{3}{719}, \frac{5}{1198}, \frac{8}{1917}, \frac{69}{16534}, \frac{146}{34985}, \frac{215}{51519}, \frac{361}{86504}, \frac{5269}{1262575}, \frac{16168}{3874229}, \frac{21437}{5136804}, \frac{80479}{19284641}, \frac{262874}{62990727}, \ldots].$$

Applying the factorization algorithm with the convergent $\frac{x}{y} = \frac{80479}{19284641}$, we get

$$\begin{aligned} U &= \frac{ex}{y} - n - 1 \approx 27457254767091, \\ V &= \sqrt{|U^2 - 4n|} \approx 7758072877807. \end{aligned}$$

Applying Coppersmith's Theorem with $\frac{U+V}{2} = 17607663822449$ as an approximation for p, we get

$$p = 17607663822197, \qquad q = 9849590944783,$$

which leads to the factorization of N. Using p and q, we can compute the secret exponent d satisfying $ed \equiv 1 \pmod{(p+1)(q+1)}$, namely

$$d \equiv e^{-1} \equiv 70154311084917810813949567 \pmod{(p+1)(q+1)},$$

Observe that $d \approx n^{0.985}$. This explains why the attacks on KMOV with small secret exponents do not work in this example.

H.5 Conclusion

We have presented a new attack on the KMOV cryptosystem with a public key (n, e) where $n = pq$ is an RSA modulus and e a public exponent satisfying $\gcd(e, (p+1)(q+1)) = 1$ as required by KMOV. We prove that KMOV is insecure if there exist integers x, y and z with

$$|z| < \frac{(p-q)n^{\frac{1}{4}}y}{3(p+q)}, \quad xy < \frac{\sqrt{2}\sqrt{n}}{12}.$$

and satisfying an equation $ex - (p+1)(q+1)y = z$. The attack combines the continued fraction algorithm and Coppersmith's lattice reduction based method and can be seen as an extension of Pinch's attack on small KMOV secret decryption exponents.

Appendix I

A Generalized Attack on RSA Type Cryptosystems

Theoretical Computer Science 2016
[25] with Martin Bunder, Willy Susilo, Joseph Tonien

Abstract :
Let $N = pq$ be an RSA modulus with unknown factorization. Some variants of the RSA cryptosystem, such as LUC, RSA with Gaussian primes and RSA type schemes based on singular elliptic curves use a public key e and a private key d satisfying an equation of the form $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$. In this paper, we consider the general equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ and present a new attack that finds the prime factors p and q in the case that x, y and z satisfy a specific condition. The attack combines the continued fraction algorithm and Coppersmith's technique and can be seen as a generalization of the attacks of Wiener and Blömer-May on RSA.

I.1 Introduction

In 1978, Rivest, Shamir and Adleman [131] proposed RSA, the first and widely most used public key cryptosystem. The security of RSA is mainly based on the hardness of factoring large composite integers, nevertheless, RSA has been extensively studied for vulnerabilities by various non factorization attacks. The public parameters in RSA are the RSA modulus $N = pq$ which is the product of two large primes of the same bit-size and a public exponent e satisfying $\gcd(e, (p-1)(q-1)) = 1$. The correspondent private exponent is the integer $d < N$ satisfying $ed \equiv 1 \pmod{(p-1)(q-1)}$ which can be rewritten as a key equation $ed - k(p-1)(q-1) = 1$. In RSA, the encryption and decryption time are proportional to the bit-length of the public and the private exponents. To reduce encryption or decryption time, one may be tempted to use small public exponents or private exponents. While a few attacks on RSA with small public exponent e have been launched (see [55]), many attacks on RSA with small or special private exponent d exploit the algebraic properties of the key equation. In 1990, Wiener [147] presented an attack on RSA that solves the key equation and factors N if d is sufficiently small, namely $d < \frac{1}{3}N^{0.25}$. Wiener's attack consists on finding $\frac{k}{d}$ among the convergents of the continued fraction expansion of $\frac{e}{N}$ and then using $\frac{k}{d}$ to factor N. Wiener's attack on RSA has been extended in many ways using lattice reduction and Coppersmith's method [34] (see [15], [61], [91]). In 1997, Boneh and Durfee [17] used lattice reduction and Coppersmith's method to improve the bound to $d < N^{0.292}$. In 2004, Blömer and May studied the variant equation $ex + y \equiv 0 \pmod{(p-1)(q-1)}$ and showed that the RSA modulus can be factored if the unknown parameters satisfy $x < \frac{1}{3}N^{0.25}$ and $|y| \leq cN^{-\frac{3}{4}}ex$ for some constant $c \leq 1$.
In order to improve the implementation of the RSA cryptosystem, many schemes have been presented giving rise to RSA type cryptosystems [20]. One way to extend RSA is to consider a prime-power modulus of the form $N = p^r q$ with $r \geq 2$ (see [145]) or a multi-prime modulus of the form $N = p_1 p_2 \dots p_r$. Another way to extend RSA is to consider the modulus $N = pq$ and the exponent e with specific arithmetical operations such as elliptic curves [81] [79], Gaussian domains [43] and quadratic fields [126].

In 1995, Kuwakado, Koyama and Tsuruoka [81] presented a scheme based on using an RSA modulus $N = pq$ and a singular cubic equation with equation $y^2 = x^3 + bx^2 \mod N$ where a message $M = (m_x, m_y)$ is represented as a point on the singular cubic equation. In this system, the public exponent e and the private exponent d satisfy an equation of the form $ed - k\left(p^2-1\right)\left(q^2-1\right)$.
In 2002, Elkamchouchi, Elshenawy and Shaban [43] adapted RSA to the Gaussian domain by using a modulus of the form $N = PQ$ where P and Q are two Gaussian primes. The public exponent e and the private exponent d satisfy $ed \equiv 1 \pmod{(|P|-1)(|Q|-1)}$. When $P = p$ and $Q = q$ are integer prime numbers, the equation becomes $ed \equiv 1 \pmod{\left(p^2-1\right)\left(q^2-1\right)} = 1$.
In 1993, Smith and Lennon proposed LUC [143], where the public exponent e and the private exponent d are such that $ed \equiv 1 \pmod{\left(p^2-1\right)\left(q^2-1\right)}$.
In 2007, in connection with LUC, Castagnos [29] proposed a scheme that uses an RSA modulus $N = pq$ and a public exponent e. The two public parameters N and e are such that $\gcd\left(e, \left(p^2-1\right)\left(q^2-1\right)\right) = 1$ which implies the existence of two positive integers d and k satisfying the equation $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$.
The former four variants of RSA use a modulus $N = pq$ and a public exponent e satisfying an equation of the form $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$. In [24], an attack is presented that solves the former equation when d satisfies $d < \sqrt{\frac{2N^3-18N^2}{e}}$. The attack, which is related to Wiener's attack on RSA, is based on applying the continued fraction algorithm to find $\frac{k}{d}$ among the covergents of the continued fraction expansion of $\frac{e}{N^2-\frac{9}{4}N+1}$. In this paper, we consider an extension of this attack by studying the more general equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ where the unknown parameters x, y, z satisfy

$$xy < 2N - 4\sqrt{2}N^{\frac{3}{4}} \quad \text{and} \quad |z| < (p-q)N^{\frac{1}{4}}y.$$

The new attack uses the convergents of the continued fraction expansion of $\frac{e}{N^2+1-\frac{9}{4}N}$ to find $\frac{y}{x}$ and then applies Coppersmith's technique [34] to find p and q.

The remainder of the paper is organized as follows. In section 2, we recall some RSA type schemes that are based on a modulus of the form $N = pq$ with a public exponent satisfying $\gcd\left(e, \left(p^2-1\right)\left(q^2-1\right)\right) = 1$. In

Section 3, we briefly review some basic results used in the paper, including continued fractions and Coppersmith's technique. In Section 4, we present some lemmas that will be used in the paper. In Section 5, we present our new method. In Section 6, we give a numerical example. We conclude the paper in Section 7.

I.2 Variant RSA schemes

Let $N = pq$ be an RSA modulus and e a public integer. In this section, we briefly describe three schemes that are variants of the RSA cryptosystem with a modulus $N = pq$ and with a public key e and a private key d satisfying $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$. As this equation does not depend on the underlying variant schemes, we then generalize it to the equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ which is the main focus of this paper.

I.2.1 LUC cryptosystem

In 1993, Smith and Lennon [143] proposed a variant of the RSA cryptosystem, called LUC, based on a Lucas functions. In LUC, the modulus is a RSA modulus $N = pq$ and the public exponent e is a positive integer satisfying $\gcd\left(e, \left(p^2-1\right)\left(q^2-1\right)\right)$ which can be rewritten as an equation $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$. A more general equation is $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ with the unknown parameters x, y and z.

I.2.2 Castagnos cryptosystem

In 2007, Castagnos [29] proposed a cryptosystem related to LUC and RSA where the modulus $N = pq$ and the public exponent e satisfy the condition $\gcd\left(e, \left(p^2-1\right)\left(q^2-1\right)\right)$ or equivalently $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$ for some integers d and k. This equation can be extended to a more general one, namely $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$.

I.2.3 RSA with Gaussian primes

In 2002, Elkamchouchi, Elshenawy and Shaban [43] proposed a generalization of the RSA cryptosystem to the domain of Guassian integers. A Gaussian integer is a complexe number $z = a + bi$ where a and b are both integers. A Gaussian prime is a Gaussian integer that is not the product of two non-unit Gaussian integers, the only units being ± 1 and $\pm i$. The Gaussian primes are of one of the following forms

- $P = \pm 1 \pm i$,
- $P = a$ where $|a|$ is an integer prime with $|a| \equiv 3 \pmod 4$,
- $P = ai$ where $|a|$ is an integer prime with $|a| \equiv 3 \pmod 4$,
- $P = a + ib$ where $|P| = a^2 + b^2 \equiv 1 \pmod 4$ is an integer prime.

In the RSA variant with Gaussian integers, the modulus is $N = PQ$, a product of two Gaussian integer primes P and Q. The Euler totient function is $\phi(N) = (|P| - 1)(|Q| - 1)$ and the public exponent e is a positive integer satisfying $\gcd(e, \phi(N)) = 1$. When $P = p$ and $Q = q$ are integer primes, then $\phi(N) = \left(p^2 - 1\right)\left(q^2 - 1\right)$ and the public exponent satisfies the key equation $ed - k\left(p^2 - 1\right)\left(q^2 - 1\right) = 1$ which can be extended to a more general equation $ex - \left(p^2 - 1\right)\left(q^2 - 1\right)y = z$.

I.2.4 RSA type schemes based on singular cubic curves

Let $N = pq$ be an RSA modulus. For an integer $b \in \mathbb{Z}/n\mathbb{Z}$, consider the cubic curve $E_N(b)$ defined over the ring $\mathbb{Z}/n\mathbb{Z}$ given by the Weierstrass equation

$$E_N(b) : y^2 = x^3 + bx^2 \mod N.$$

In 1995, Kuwakado, Koyama, and Tsuruoka [81] proposed a new cryptosystem based the elliptic curve $E_N(b)$. The encryption key is a positive integer satisfying $\gcd\left(e, \left(p^2 - 1\right)\left(q^2 - 1\right)\right)$ and the decryption key is the integer d satisfying $ed \equiv 1 \pmod{\left(p^2 - 1\right)\left(q^2 - 1\right)}$, or equivalently

$$ed - k\left(p^2 - 1\right)\left(q^2 - 1\right) = 1.$$

The encryption and the decryption procedures use operations on the singular cubic curve $E_N(b)$. Using the continued fraction algorithm, it is possible to

attack the scheme using the key equation $ed-k\left(p^2-1\right)\left(q^2-1\right)=1$. A more general attack on the scheme can be launched by using the equation $ex-\left(p^2-1\right)\left(q^2-1\right)y=z$ and by combining the continued fraction algorithm and Coppersmith's method.

I.3 Preliminaries

In this section, we present the mathematical preliminaries.

I.3.1 Continued fractions

Let x be a real number. Define the sets $(x_0, x_1, \ldots)$ and $[a_0, a_1, \ldots]$ by $x_0 = x$ and by the recurrences

$$a_i = \lfloor x_i \rfloor, \quad x_{i+1} = \frac{1}{x_i - a_i}, \quad i = 0, 1, \ldots .$$

The set $[a_0, a_1, \cdots]$ is the continued fraction expansion of x and satisfies

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}.$$

The convergents of x are the rational numbers $\frac{p_n}{q_n}$, $n = 0, 1, \ldots$ satisfying

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}.$$

Continued fractions have numerous properties and applications in cryptography. The following useful result characterizes the approximations to a real number x (see Theorem 184 of [57]).

Theorem I.3.1 (Legendre)**.** *If a, b be positive integers and*

$$0 < \left| x - \frac{a}{b} \right| < \frac{1}{2b^2}$$

then $\frac{a}{b}$ is a convergent of the continued fraction of x.

Note that when $x = \frac{r}{s}$ is a rational number, then the list of the convergents of the continued fraction expansion of $\frac{r}{s}$ can be done in polynomial time in $\log(\max(a, b))$.

I.3.2 Coppersmith's method

In 1997, Coppersmith [34] introduced an algorithm to find small solutions of univariate modular polynomial equations and another algorithm to find small roots of bivariate polynomial equations. Since then, Coppersmith's method has been applied in various applications in cryptography, mainly to attack the RSA cryptosystem. A typical example is the following result.

Theorem I.3.2. *Let $N = pq$ be the product of two unknown primes such that $q < p < 2q$. Given an approximation $\tilde{p}$ of p with an additive error term at most $N^{\frac{1}{4}}$, one can find p and q in polynomial time in* $\log(N)$.

As a consequence of Coppersmith's Theorem, one can show that if $N = pq$ with $|p - q| < N^{\frac{1}{4}}$, then N can be factored (see [104]). Thus, throughout this paper, we will consider that the prime difference $p - q$ satisfies $|p - q| > N^{\frac{1}{4}}$.

I.4 Useful Lemmas

One of the main RSA standard recommendations for safe parameters is to choose the prime factors factors p, q of the same bit-size. More precisely, p and q should satisfy $1 < \frac{p}{q} < 2$ or equivalently $q < p < 2q$. Under this assumption, one can find intervals for p, q, $p - q$, $p + q$ and $p^2 + q^2$ in terms of N. We begin by the following results (see [104]).

Lemma I.4.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then*

$$\frac{\sqrt{2}}{2}\sqrt{N} < q < \sqrt{N} < p < \sqrt{2}\sqrt{N} \quad \textit{and} \quad 0 < p - q < \frac{\sqrt{2}}{2}\sqrt{N}.$$

We will need the following result.

Lemma I.4.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then*

$$2\sqrt{N} < p + q < \frac{3\sqrt{2}}{2}\sqrt{N} \quad \textit{and} \quad 2N < p^2 + q^2 < \frac{5}{2}N.$$

Proof. Assume that $N = pq$ with $q < p < 2q$. Then $1 < \frac{p}{q} < 2$. The function $f(x) = x + \frac{1}{x}$ is increasing on $[1, +\infty)$. Hence, $f(1) < f(\frac{p}{q}) < f(2)$, that is

$$2 < \frac{p}{q} + \frac{q}{p} < \frac{5}{2}.$$

Multiplying by $N = pq$, we get

$$2N < p^2 + q^2 < \frac{5}{2}N.$$

Similarly, since $1 < \sqrt{\frac{p}{q}} < \sqrt{2}$, then $f(1) < f(\sqrt{\frac{p}{q}}) < f(\sqrt{2})$, or equivalently

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

Hence, multiplying by $\sqrt{N} = \sqrt{pq}$, we get

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

This terminates the proof. ■

I.5 The New Attack

In this section, we present our new attack to solve the equation $ex - \left(p^2 - 1\right)\left(q^2 - 1\right)y = z$ when x, y and z are suitably small. The new method combines two techniques, the continued fraction algorithm and Coppersmith's method.

Theorem I.5.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let e be a public exponent satisfying an equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ with coprime positive integers x and y. If*

$$xy < 2N - 4\sqrt{2}N^{\frac{3}{4}} \quad \textit{and} \quad |z| < (p-q)N^{\frac{1}{4}}y,$$

then one can find p and q in polynomial time in $\log(N)$.

Proof. Suppose that $N = pq$ with $q < p < 2q$ and that a public exponent e satisfies the equation

$$ex - \left(p^2-1\right)\left(q^2-1\right)y = z, \tag{I.1}$$

with $x > 0$, $y > 0$ and $\gcd(x, y) = 1$. Then

$$\begin{aligned} ex - \left(N^2 + 1 - \frac{9}{4}N\right)y &= ex - (p^2-1)(q^2-1)y - \left(p^2+q^2-\frac{9}{4}N\right)y \\ &= z - \left(p^2+q^2-\frac{9}{4}N\right)y. \end{aligned} \tag{I.2}$$

From this we deduce

$$\left|\frac{e}{N^2+1-\frac{9}{4}N} - \frac{y}{x}\right| \leq \frac{|z|}{x\left(N^2+1-\frac{9}{4}N\right)} + \frac{\left|p^2+q^2-\frac{9}{4}N\right|y}{x\left(N^2+1-\frac{9}{4}N\right)}. \tag{I.3}$$

Using Lemma I.4.2, we get that $\left|p^2+q^2-\frac{9}{4}N\right| < \frac{1}{4}N$. Suppose in addition that $|z| < |p-q|N^{\frac{1}{4}}y$. Then, using Lemma I.4.1, we get

$$|z| < |p-q|N^{\frac{1}{4}}y < \frac{\sqrt{2}}{2}\sqrt{N}\cdot N^{\frac{1}{4}}y = \frac{\sqrt{2}}{2}N^{\frac{3}{4}}y. \tag{I.4}$$

Hence (I.3) leads to

$$\begin{aligned} \left|\frac{e}{N^2+1-\frac{9}{4}N} - \frac{y}{x}\right| &< \frac{\frac{\sqrt{2}}{2}N^{\frac{3}{4}}}{N^2+1-\frac{9}{4}N}\cdot\frac{y}{x} + \frac{\frac{1}{4}N}{N^2+1-\frac{9}{4}N}\cdot\frac{y}{x} \\ &= \frac{N+2\sqrt{2}N^{\frac{3}{4}}}{4N^2+4-9N}\cdot\frac{y}{x}. \end{aligned} \tag{I.5}$$

Now, suppose that $xy < 2N - 4\sqrt{2}N^{\frac{3}{4}}$. A straightforward calculation shows that

$$2N - 4\sqrt{2}N^{\frac{3}{4}} < \frac{4N^2+4-9N}{2N+4\sqrt{2}N^{\frac{3}{4}}}.$$

Then $xy < \frac{4N^2+4-9N}{2\left(N+2\sqrt{2}N^{\frac{3}{4}}\right)}$ and $\frac{N+2\sqrt{2}N^{\frac{3}{4}}}{4N^2+4-9N} < \frac{1}{2xy}$. Using this in (I.5), we get

$$\left|\frac{e}{N^2+1-\frac{9}{4}N} - \frac{y}{x}\right| < \frac{N+2\sqrt{2}N^{\frac{3}{4}}}{4N^2+4-9N} \cdot \frac{y}{x} < \frac{1}{2xy} \cdot \frac{y}{x} = \frac{1}{2x^2}.$$

Hence, if this condition is fulfilled, then one can find $\frac{y}{x}$ amongst the convergents of the continued fraction expansion of $\frac{e}{N^2+1-\frac{9}{4}N}$ as stated in Theorem I.3.1. Moreover, since $\gcd(x, y) = 1$, the values of x and y are the denominator and numerator of the convergent. Plugging x and y in (I.1), we get

$$p^2 + q^2 = N^2 + 1 - \frac{ex}{y} + \frac{z}{y}. \tag{I.6}$$

Adding $2N$ to both sides of (I.6), we get

$$(p+q)^2 = (N+1)^2 - \frac{ex}{y} + \frac{z}{y}. \tag{I.7}$$

Similarly, subtracting $2N$ to both sides of (I.6), we get

$$(p-q)^2 = (N-1)^2 - \frac{ex}{y} + \frac{z}{y}. \tag{I.8}$$

Observe that (I.7) can be transformed into

$$\left|p+q-\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|}\right| \times \left|p+q+\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|}\right| = \frac{|z|}{y},$$

from which we deduce

$$\left|p+q-\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|}\right| = \frac{|z|}{\left|p+q+\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|}\right| y} < \frac{|z|}{(p+q)y}.$$

By (I.4) we have $|z| < \frac{\sqrt{2}}{2}N^{\frac{3}{4}}y$ and by Lemma I.4.2 we have $p+q > 2\sqrt{N}$. Then

$$\left|p+q-\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|}\right| < \frac{\frac{\sqrt{2}}{2}N^{\frac{3}{4}}}{2\sqrt{N}} = \frac{\sqrt{2}}{4}N^{\frac{1}{4}} < N^{\frac{1}{4}}.$$

This means that $\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|}$ is an approximation of $p+q$ with error term less than $N^{\frac{1}{4}}$. In a similar way, using (I.8), we get

$$\left|p-q-\sqrt{\left|(N-1)^2-\frac{ex}{y}\right|}\right| \times \left|p-q+\sqrt{\left|(N-1)^2-\frac{ex}{y}\right|}\right| = \frac{|z|}{y},$$

which leads to

$$\left|p-q-\sqrt{\left|(N-1)^2-\frac{ex}{y}\right|}\right| = \frac{|z|}{\left|p-q+\sqrt{\left|(N+1)^2-\frac{ex}{y}\right|}\right| y} < \frac{|z|}{(p-q)y}.$$

Using the assumption $|z| < (p-q)N^{\frac{1}{4}}y$, we get

$$\left|p-q-\sqrt{\left|(N-1)^2-\frac{ex}{y}\right|}\right| < \frac{(p-q)N^{\frac{1}{4}}y}{(p-q)y} = N^{\frac{1}{4}}.$$

Hence, $\sqrt{\left|(N-1)^2 - \frac{ex}{y}\right|}$ is an approximation of $p-q$ with an error term less than $N^{\frac{1}{4}}$. Combing the approximations of $p+q$ and $p-q$, we get

$$\begin{aligned}
&\left|p - \frac{1}{2}\left(\sqrt{\left|(N+1)^2-\frac{ex}{y}\right|} + \sqrt{\left|(N-1)^2-\frac{ex}{y}\right|}\right)\right| \\
&< \frac{1}{2}\left|p+q-\sqrt{\left|(N+1)^2-\frac{ex}{y}\right|}\right| + \frac{1}{2}\left|p-q-\sqrt{\left|(N-1)^2-\frac{ex}{y}\right|}\right| \\
&< \frac{1}{2}N^{\frac{1}{4}} + \frac{1}{2}N^{\frac{1}{4}} \\
&= N^{\frac{1}{4}}.
\end{aligned}$$

This gives an approximation of p with an error term of at most $N^{\frac{1}{4}}$. Hence, using Coppersmith's Theorem I.3.2, one can find p which leads to $q = \frac{N}{p}$. Since every step in the proof can be done in polynomial time in $\log(N)$, then the factorization of N can be obtained in polynomial time in $\log(N)$. ■

We note that, when $\gcd\left(ex, \left(p^2-1\right)\left(q^2-1\right)\right) = 1$, the diophantine equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ is equivalent to the modular equation

$ex \equiv z \pmod{\left(p^2-1\right)\left(q^2-1\right)}$. Moreover, the exponent e satisfies

$$e \equiv \frac{z}{x} \pmod{\left(p^2-1\right)\left(q^2-1\right)}.$$

Hence, Theorem I.5.1 implies that one can factor $N = pq$ for such exponents e in the case where $xy < 2N - 4\sqrt{2}N^{\frac{3}{4}}$ and $|z| < (p-q)N^{\frac{1}{4}}y$.

We now consider an application of Theorem I.5.1 to the private exponent d. We recall that d satisfies $ed \equiv 1 \pmod{\left(p^2-1\right)\left(q^2-1\right)}$. Instead of this modular equation, we consider the key equation

$$ed - k\left(p^2-1\right)\left(q^2-1\right) = 1.$$

Corollary I.5.2. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $e < \left(p^2-1\right)\left(q^2-1\right)$ be a public exponent. If the private exponent d satisfies*

$$d < \sqrt{2N - 4\sqrt{2}N^{\frac{3}{4}}},$$

then one can find p and q in polynomial time in $\log(N)$.

Proof. Suppose that $q < p < 2q$ and $e < \left(p^2-1\right)\left(q^2-1\right)$. Since the private exponent d satisfies $ed - k\left(p^2-1\right)\left(q^2-1\right) = 1$ for a positive integer k, then

$$k = \frac{ed-1}{\left(p^2-1\right)\left(q^2-1\right)} < d \cdot \frac{e}{\left(p^2-1\right)\left(q^2-1\right)} < d.$$

Then $dk < d^2$. Now, assume that $d^2 < 2N - 4\sqrt{2}N^{\frac{3}{4}}$. Then, $dk < 2N - 4\sqrt{2}N^{\frac{3}{4}}$ and d, k fulfill the conditions of Theorem I.5.1 wich leads to the factorization of N in polynomial time in $\log(N)$. ■

I.6 A Numerical Example

In this section we give a detailed numerical example to explain our method as developed in Theorem I.5.1. Let us consider the small public key

$$\begin{aligned} N &= 204645825996541, \\ e &= 26384989321053458213237. \end{aligned}$$

It is obvious that equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ has infinitely many solutions (x, y, z) with positive integers x, y and non zero integer z. Our aim is to find the solution that satisfies the conditions of Theorem I.5.1, if any. Define We want to find $\frac{y}{x}$ among the convergents of the continued fraction expansion of $\frac{e}{N^2+1-\frac{9}{4}N}$. Following the technique of Theorem I.5.1, for each convergent $\frac{y}{x}$ of $\frac{e}{N^2+1-\frac{9}{4}N}$ with $xy < 2N - 4\sqrt{2}N^{\frac{3}{4}} \approx 4.089 \times 10^{14}$, we compute an approximation $\tilde{p}$ of p using

$$\tilde{p} = \frac{1}{2}\left(\sqrt{\left|(N+1)^2 - \frac{ex}{y}\right|} + \sqrt{\left|(N-1)^2 - \frac{ex}{y}\right|}\right),$$

and apply Coppersmith's Theorem I.3.2 with $\tilde{p}$. Using the convergent

$$\frac{y}{x} = \frac{16052}{25478743725},$$

we get $\tilde{p} \approx 19126518$. Coppersmith's Theorem outputs the prime factor $p = 19126831$ from which we deduce the second prime factor $q = \frac{N}{p} = 10699411$. This completes the factorization of N.

I.7 Conclusion

In this paper, we considered some variants of the RSA cryptosystem with a modulus $N = pq$ and an exponent e satisfying $\gcd\left(e, \left(p^2-1\right)\left(q^2-1\right)\right) = 1$. We studied the general equation $ex - \left(p^2-1\right)\left(q^2-1\right)y = z$ and combined the continued fraction algorithm with Coppersmith's technique to find x and y and then to factor the RSA modulus N. Our new method can considered as an extension to some RSA type schemes of two former methods that work for RSA, namely Wiener's attack and Blömer-May attack.

Appendix J

Cryptanalysis of NTRU with two Public Keys

International Journal of Network Security 2014 [118]

Abstract :
NTRU is a fast public key cryptosystem presented in 1996 by Hoffstein, Pipher and Silverman. It operates in the ring of truncated polynomials. In NTRU, a public key is a polynomial defined by the combination of two private polynomials. In this paper, we consider NTRU with two different public keys defined by different private keys. We present a lattice-based attack to recover the private keys assuming that the public keys share polynomials with a suitable number of common coefficients.

J.1 Introduction

The NTRU Public Key Cryptosystem is a ring-based cryptosystem that was first introduced in the rump session at Crypto'96 [63]. It is one of the fastest

public-key cryptosystems, offering both encryption (NTRUencrypt) and digital signatures (NTRUSign). It is a relatively new cryptosystem that appears to be more efficient than the current and more widely used public-key cryptosystems, such as RSA [131] and El Gamal [42]. It is well known that the security of RSA and El Gamal relies on the difficulty of factoring large composite integers or computing discrete logarithms. However, in 1994, Shor [139] showed that quantum computers can be used to factor integers and to compute discrete logarithms in polynomial time. Since NTRU does not rely on the difficulty of factoring or computing discrete logarithms and is still considered secure even against quantum computer attacks, it is a promising alternative to the more established public key cryptosystems. In [63], Hoffstein, Pipher and Silverman have studied different possible attacks on NTRU. The brute force and the meet-in-the-middle attacks may be used against the private key or against a single message but will not succeed in a reasonable time. The multiple transmission attack also will fail for a suitable choice of parameters. However, we notice that NTRU suggests that the public key should be changed very frequently, for each transmission if possible. The most important attack, presented by Coppersmith and Shamir [35] in 1997 makes use of the LLL algorithm of Lenstra, Lenstra and Lovász [86]. Coppersmith and Shamir constructed a lattice generated by the public key and found a factorization of the public key that could be used to break the system if the NTRU parameters are poorly set.

The NTRU cryptosystem depends on three integer parameters (N, p, q) and four sets $\mathcal{L}_f$, $\mathcal{L}_g$, $\mathcal{L}_r$, $\mathcal{L}_m$ of polynomials of degree $N-1$ with small integer coefficients. Let $\mathbb{Z}_q$ denote the ring of integers modulo q. The operations of NTRU took place in the ring of polynomials $\mathbb{Z}_q[X]/\left(X^N-1\right)$. In this ring, the addition of two polynomials is defined as pairwise addition of the coefficients of the same degree and multiplication, noted " $*$ " is defined as convolution multiplication. In NTRU, to create a public key h, one chooses a private key (f, g) composed with two polynomials f and g and computes

$$h = f_q^{-1} * g \in \mathbb{Z}_q[X]/\left(X^N-1\right),$$

where f_q^{-1} is the inverse of f in $\mathbb{Z}_q[X]/\left(X^N-1\right)$.

In this paper, we consider NTRU with two public keys h, h' defined by

the private keys (f, g) and (F', G') with

$$h' = F_q'^{-1} * G' \pmod{q}.$$

Since f is invertible in $\mathbb{Z}_q[X]/\left(X^N - 1\right)$, then we can define $g' = f * h' \pmod{q}$ so that

$$h' = f_q^{-1} * g' \mod q.$$

The main objective of this paper is to show how to find the private key (f, g) when

$$\|g - g'\| < \min(\|g\|, \|g'\|).$$

Using h and h', we construct a lattice $\mathcal{L}(h, h')$ of dimension $2N$, and applying the lattice basis reduction algorithm LLL, we show that short vectors in $\mathcal{L}(h, h')$ can be used to find the private polynomials f, g, g' when $\|g - g'\| < \min(\|g\|, \|g'\|)$. Under this condition, it is important to notice that our method is more efficient than the method of Coppersmith and Shamir to recover the private key (f, g) using the public key h.

We note that when the polynomials g, g' are generated randomly and independently, then with overwhelming probability the condition $\|g - g'\| < \min(\|g\|, \|g'\|)$ is not satisfied. So in practice one can easily avoid this inequality.

Similarly, assume that $h' = F_q'^{-1} * G' \pmod{q}$ is invertible in the ring $\mathbb{Z}_q[X]/\left(X^N - 1\right)$. Then we can define a polynomial f' as

$$f' = h_q'^{-1} * g \pmod{q},$$

where $h_q'^{-1}$ is the inverse of h' in $\mathbb{Z}_q[X]/\left(X^N - 1\right)$. Using lattice reduction techniques, we show that it is possible to recover the private key (f, g) assuming that the condition $\|f - f'\| < \min(\|f\|, \|f'\|)$ is fulfilled.

The paper is organized as follows. In Section 2, we give motivation for our work. Section 3 gives a brief mathematical description of NTRU and introduces the LLL algoritm as well as the attack of Coppersmith and Shamir on NTRU. In Section 4, we present our new attack on NTRU with two private keys (f, g) and (f, g') with $\|g - g'\| < \min(\|g\|, \|g'\|)$ and compare it with the attack of Coppersmith and Shamir. In Section 5, we present our new attack on NTRU when h and h' are invertible and $\|f - f'\| < \min(\|f\|, \|f'\|)$. We conclude the paper in Section 6.

J.2 Motivation

RSA, the most commonly used public-key cryptosystem [131] has stood up remarkably well to years of extensive cryptanalysis and is still considered secure by the cryptographic community (see [15] for more details). Various schemes and digtal signatures are based on the same problem behind RSA (see e.g. [27] and [149]). Indeed, RSA derives its security from the difficulty of factoring large numbers of the shape $N = pq$ where p, q are large unknown primes of the same bit-size. In some cases, the problem can be slightly easier given two RSA modulus $N = pq$, $N' = p'q'$. If $p = p'$, then it is trivial to factor N and N' by computing $\gcd(N, N')$. However, it is possible to factor N and N' when p and p' share a certain amount of bits (see [94], [132]).

The first paper on NTRU was written by Coppersmith and Shamir [35] in 1997. In that paper, they noted that the best way to attack the NTRU cryptosystem was via the techniques of lattice reduction. Nevertheless, the security of NTRU is also based on the following factorization problem: Given a polynomial $h \in \mathbb{Z}[X]/\left(X^N - 1\right)$, find two short polynomials f and g with $f \in \mathbb{Z}[X]/\left(X^N - 1\right)$ and $g \in \mathbb{Z}[X]/\left(X^N - 1\right)$ such that $h = f_q^{-1} * g \pmod q$, where f_q^{-1} is the inverse of f in $\mathbb{Z}_q[X]/\left(X^N - 1\right)$.

Similarly to RSA with two modulus, consider NTRU with two public keys h and h' defined by the same parameters (N, p, q). Assume that $h = f_q^{-1} * g \pmod q$. Then, h' can be expressed as $h' = f_q^{-1} * g' \pmod q$ where $g' = f * h' \pmod q$. The main contribution of this paper is to show how to find the private keys (f, g) when g and g' satisfy $\|g - g'\| < \min(\|g\|, \|g'\|)$.

We notice that lattice-based cryptography is currently seen as one of the most promising alternatives to cryptography based on number theory. Given recent advances in lattice-based cryptography (see [89] and [144]), studying NTRU and related schemes is both useful and timely. In this direction, our work shows that using the same f or the same g in generating public keys h, h' is likely to reduce the security of NTRU.

J.3 Mathematical background

In this section, we give a brief description of the NTRU encryption and the LLL algorithm for lattice reduction and the well known attack of Coppersmith and Shamir on NTRU. Further details can be found in [63] and [35].

J.3.1 Definitions and notations

We start by introducing the ring

$$\mathcal{R} = \mathbb{Z}[X]/(X^N - 1),$$

upon which NTRU operates. We use $*$ to denote a polynomial multiplication in $\mathcal{R}$, which is the cyclic convolution of two polynomials. If

$$\begin{aligned} f &= (f_0, f_1, \cdots, f_{N-1}) = \sum_{i=0}^{N-1} f_i X^i, \\ g &= (g_0, g_1, \cdots, g_{N-1}) = \sum_{i=0}^{N-1} g_i X^i, \end{aligned}$$

are polynomials of $\mathcal{R}$, then $h = f * g$ is given by $h = (h_0, h_1, \cdots, h_{N-1})$, where h_k is defined for $0 \leq k \leq N-1$ by

$$\begin{aligned} h_k &= \sum_{i+j \equiv k \mod N} f_i g_j \\ &= \sum_{i=0}^{k} f_i g_{k-i} + \sum_{i=k+1}^{N-1} f_i g_{N+k-i}. \end{aligned}$$

The Euclidean norm or the length of a polynomial $f = (f_0, f_1, \cdots, f_{N-1})$ is defined as

$$\|f\| = \sqrt{\sum_{i=0}^{N-1} f_i^2}.$$

One more notation is the binary set of polynomials $\mathcal{B}(d)$ defined for a positive integers d by

$$\mathcal{B}(d) = \{f(X) = \sum_{i=0}^{N-1} f_i X^i, \text{ where } f_i \in \{0,1\}, \sum_{i=0}^{N-1} f_i = d\}.$$

In other words, $\mathcal{B}(d)$ is the set of polynomials of $\mathcal{R}$ with d coefficients equal to 1 and all the other coefficients equal to 0.

Different descriptions of NTRUEncrypt and different proposed parameter sets have been in circulation since 1996. The 2005 instantiation of NTRU is set up by six public integers N, p, q, d_f, d_g, d_r and four public spaces $\mathcal{L}_f$, $\mathcal{L}_g$, $\mathcal{L}_m$, $\mathcal{L}_r$ such that

- N is prime and sufficiently large to prevent lattice attacks.
- p and q are relatively prime.
- q is much larger than p.
- $\mathcal{L}_f$ is a set of small polynomials from which the private keys are selected.
- $\mathcal{L}_g$ is a similar set of small polynomials from which other private keys are selected.
- $\mathcal{L}_m$ is the plaintext space. It is a set of polynomials $m \in \mathbb{Z}_p[X]/(X^N - 1)$ that represent encryptable messages.
- $\mathcal{L}_r$ is a set of polynomials from which the blinding value used during encryption is selected.

J.3.2 The NTRU Encryption Scheme

Key pair generation.

To create a NTRU key, one randomly chooses a polynomial $f \in \mathcal{L}_f$ and a polynomial $g \in \mathcal{L}_g$. The polynomial f must satisfy the additional requirement

that it has an inverse f_p^{-1} modulo p and an inverse f_q^{-1} modulo q, that is

$$f * f_p^{-1} = 1 \pmod{p}, \qquad f * f_q^{-1} = 1 \pmod{q}.$$

Then the private key is f and the public key is the polynomial

$$h = f_q^{-1} * g \pmod{q}.$$

We recall that N, p, q are also public.

Encryption.

To encrypt a message $m \in \mathcal{L}_m$, one randomly chooses a polynomial $r \in \mathcal{L}_r$. The ciphertext is the polynomial

$$e = pr * h + m \pmod{q}.$$

Decryption.

To decrypt an encrypted message e using the private key f, one computes

$$a = f * e \mod q,$$

where the coefficients of a lie between $-q/2$ and $q/2$. The message m is then obtained from a by reducing the coefficients of $f_p^{-1} * a$ modulo p.

J.3.3 The LLL algorithm

Since lattice reduction is an essential tool for our attack, let us recall a few facts about lattices and reduced basis. Let $u_1, \ldots, u_n \in \mathbb{R}^m$ be linearly independent vectors with $n \leq m$. The lattice L spanned by $(u_1, \ldots, u_n)$ consists of all integral linear combinations of $u_1, \ldots, u_n$, that is

$$L = \mathbb{Z}u_1 \oplus \cdots \oplus \mathbb{Z}u_n = \left\{ \sum_{i=1}^{n} b_i u_i, \;\middle|\; b_i \in \mathbb{Z} \right\}.$$

The set $(u_1, \ldots, u_n)$ is called a lattice basis. A lattice can be conveniently represented by a matrix B whose rows are the vectors $u_1, \ldots, u_n$. The determinant of the lattice L is defined as

$$\det(L) = \sqrt{\det(BB^T)}.$$

Any two bases of the same lattice L are related by some integral matrix of determinant ± 1.

There are several natural computational problems relating to lattices. An important problem is the shortest vector problem (SVP): given a basis matrix B for L, compute a non-zero vector $v \in L$ such that $\|v\|$ is minimal.

In 1982, Lenstra, Lenstra and Lovász [86] introduced the LLL reduction algorithm which produces an LLL-reduced basis $b_1, \ldots, b_n$ of L with the following property

$$\|b_1\| \leq \|b_2\| \leq \ldots \leq \|b_i\| \leq 2^{\frac{n(n-1)}{4(n+1-i)}} \det(L)^{\frac{1}{n+1-i}},$$

for $i = 1, \ldots, n$. With $i = 1$, this implies that $\|b_1\|$ satisfies $\|b_1\| \leq 2^{\frac{n-1}{4}} \det(L)^{\frac{1}{n}}$. In comparison, a theorem of Minkowski asserts that any lattice L of dimension n contains a non-zero vector v with

$$\|v\| \leq \sqrt{\frac{2n}{e\pi}} \det(L)^{\frac{1}{n}}.$$

On the other hand, the Gaussian heuristic says that the length of the shortest non-zero vector is usually approximately $\sigma(L)$ where

$$\sigma(L) = \sqrt{\frac{n}{2\pi e}} \det(L)^{\frac{1}{n}}.$$

J.3.4 The attack of Coppersmith and Shamir on NTRU

In [35] Coppersmith and Shamir presented a lattice attack on NTRU. They defined a lattice determined by the parameters N, q, h of the system and showed that recovering the secret key (f, g) from the public key h is reduced to finding a shortest vector of the lattice. Let $h = (h_0, h_1, \cdots, h_{N-1})$ be the public key. The NTRU lattice L is the lattice of dimension $2N$ generated by

the row vectors of a matrix of the following form

$$M(L) = \begin{bmatrix} \lambda I_N & H \\ 0 & qI_N \end{bmatrix}$$

$$= \left[\begin{array}{cccc|cccc} \lambda & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{N-1} \\ 0 & \lambda & \cdots & 0 & h_{N-1} & h_0 & \cdots & h_{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & h_1 & h_2 & \cdots & h_0 \\ \hline 0 & 0 & \cdots & 0 & q & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & q & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & q \end{array}\right].$$

Since $h = f_q^{-1} * g \pmod q$, then $f * h - qu = g$ for some $u \in \mathcal{R}$ and

$$(f, -u) * M(L) = (f, -u) * \begin{bmatrix} \lambda I_N & H \\ 0 & qI_N \end{bmatrix} = (\lambda f, g).$$

So the vector $(\lambda f, g)$ is a short vector in the NTRU lattice L, which is with high probability the shortest vector of L. Hence, an attacker uses lattice reduction algorithms to find (f, g) from L, then he can recover the private keys. More precisely, the Gaussian heuristic says that the length of the shortest non-zero vector is usually approximately $\sigma(L)$ where

$$\begin{aligned} \sigma(L) &= \sqrt{\frac{\dim(L)}{2\pi e}} \, (\det L)^{1/\dim(L)} \\ &= \sqrt{\frac{2N}{2\pi e}} (\lambda q)^{\frac{N}{2N}} \\ &= \sqrt{\frac{\lambda q N}{\pi e}}. \end{aligned}$$

Hence, in order to maximize the probability of breaking the NTRU system using lattice reduction, the attacker should choose λ to minimize the ratio

$$c = \frac{\|(\lambda f, g)\|}{\sigma(L)} = \frac{\sqrt{\lambda^2 \|f\|^2 + \|g\|^2}}{\sqrt{\frac{\lambda q N}{\pi e}}}.$$

This occurs for $\lambda = \|g\|/\|f\|$ which leads to

$$c = \sqrt{\frac{2\pi e \|g\| \|f\|}{qN}}. \tag{J.1}$$

The ratio c measures how much smaller the key is compared to the expected smallest vector. If c is very small then we expect a lattice reduction algorithm as LLL to have an easier time finding it.

J.4 The new attack when $\|g - g'\| < \min(\|g\|, \|g'\|)$

J.4.1 The new lattice

Let

$$h(X) = \sum_{i=0}^{N-1} h_i X^i, \quad h'(X) = \sum_{i=0}^{N-1} h'_i X^i,$$

be two NTRU public keys created by the private polynomials (f, g) and (F', G') with the same parameters $(N, p, q, d_f, d_g, d_r, d_m)$, that is

$$\begin{aligned} h &= f_q^{-1} * g \pmod q, \\ h' &= F_q'^{-1} * G' \pmod q. \end{aligned}$$

Let $g' = f * h' \pmod q$. Then

$$h' = f_q^{-1} * g' \pmod q.$$

For a positive constant λ, define the lattice

$$\begin{aligned} &\mathcal{L}(h, h') \\ &= \{(\lambda v, w) \in \mathcal{R}^2 : \\ &\text{where } w = v * (h - h') \pmod q\}. \end{aligned}$$

This is a $2N$-dimension lattice spanned by the matrix

$$M(h, h') = \begin{bmatrix} \lambda I_N & H - H' \\ 0 & qI_N \end{bmatrix},$$

where $H - H'$ is the circulant matrix

$$\begin{bmatrix} h_0 - h'_0 & h_1 - h'_1 & \cdots & h_{N-1} - h'_{N-1} \\ h_{N-1} - h'_{N-1} & h_0 - h'_0 & \cdots & h_{N-2} - h'_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ h_1 - h'_1 & h_2 - h'_2 & \cdots & h_0 - h'_0 \end{bmatrix}.$$

The matrix $M(h, h')$ has the following property.

Proposition J.4.1. *Let h, h' be two NTRU public keys. Assume that*

$$f * h = g + qu, \qquad f * h' = g' + qu'.$$

Then the vector $(\lambda f, g - g')$ is in the lattice $\mathcal{L}(h, h')$ and

$$(f, -u + u') * M(h, h') = (\lambda f, g - g').$$

Proof. Assume that $f * h = g + qu$ and $f * h' = g' + qu'$. Substracting the two equalities, we get

$$f * h - f * h' = f * (h - h') = g - g' \pmod{q}.$$

This implies that the vector $(\lambda f, g - g')$ is in $\mathcal{L}(h, h')$. Next, we have

$$\begin{aligned} &(f, -u + u') * M(h, h') \\ &= (f, -u + u') * \begin{bmatrix} \lambda I_N & H - H' \\ 0 & qI_N \end{bmatrix} \\ &= (\lambda f, g - g'). \end{aligned}$$

This terminates the proof. □

J.4.2 The Gaussian heuristics

For a random lattice L, the Gaussian heuristic says that the length of the shortest non-zero vector is approximately

$$\sigma(L) = \sqrt{\frac{\dim(L)}{2\pi e}} \det L^{1/\dim(L)}.$$

The dimension and determinant of $\mathcal{L}(h, h')$ are given by

$$\dim(\mathcal{L}(h, h')) = 2N, \qquad \det(\mathcal{L}(h, h')) = \lambda^N q^N.$$

Hence for the lattice $\mathcal{L}(h, h')$, we have

$$\sigma(\mathcal{L}(h, h')) = \sqrt{\frac{\lambda N q}{\pi e}}.$$

Let us define the ratio

$$c_1 = \frac{\|(\lambda f, g - g')\|}{\sigma(\mathcal{L}(h, h'))}.$$

So c_1 is the ratio of the length of the target vector to the length of the expected shortest vector. The smaller the value of c_1, the easier it will be to find the target vector. Thus, the idea to increase the chances of LLL to find $(\lambda f, g - g')$ is to choose λ such that $\|(\lambda f, g - g')\|$ is as small as possible compared to $\sigma(\mathcal{L}(h, h'))$. In $\mathcal{L}(h, h')$, we have

$$\|(\lambda f, g - g')\| = \sqrt{\lambda^2 \|f\|^2 + \|g - g'\|^2}.$$

It turns out that we should choose

$$\lambda = \frac{\|g - g'\|}{\|f\|}.$$

This implies that the ratio c_1 satisfies

$$c_1 = \sqrt{\frac{2\pi e \|g - g'\| \|f\|}{qN}}.$$

Let us compare the ratio c_1 and the ratio c as defined by (J.1) in the the attack of Coppersmith and Shamir. Our attack will be more efficient when $c_1 < c$. This leads to the following condition

$$\|g - g'\| < \min(\|g\|, \|g'\|).$$

J.5 The new attack when $\|f - f'\| < \min(\|f\|, \|f'\|)$

J.5.1 The new lattice

Let $h = f_q^{-1} * g \pmod q$ and $h' = F_q'^{-1} * G' \pmod q$ be two NTRU public keys with the same parameters $(N, p, q, d_f, d_g, d_r, d_m)$. In this section, we

assume that h, h' are invertible in $\mathbb{Z}_q[X]/\left(X^N - 1\right)$. Let h_q and h'_q be their inverses. Define $f' = g * h'_q$. We have

$$g * h_q = f \pmod{q}, \qquad g * h'_q = f' \pmod{q}.$$

Let

$$h_q(X) = \sum_{i=0}^{N-1} h_{q,i} X^i, \qquad h'_q(X) = \sum_{i=0}^{N-1} h'_{q,i} X^i,$$

be the representations of $h_q(X)$ and $h'_q(X)$ in $\mathbb{Z}_q[X]/\left(X^N - 1\right)$. For a positive constant λ, define the $2N$ dimension lattice

$$\begin{aligned}&\mathcal{L}_q(h, h') \\ &= \left\{(\lambda v, w) \in \mathcal{R}^2 : w = v * \left(h_q - h'_q\right) \pmod{q}\right\}.\end{aligned}$$

The lattice is generated by the row vectors of the matrix $M_q(h, h')$ given below

$$M_q(h, h') = \begin{bmatrix} \lambda I_N & H_q - H'_q \\ 0 & qI_N \end{bmatrix},$$

where $H_q - H'_q$ is the circulant matrix

$$\begin{bmatrix} h_{q,0} - h'_{q,0} & \cdots & h_{q,N-1} - h'_{q,N-1} \\ h_{q,N-1} - h'_{q,N-1} & \cdots & h_{q,N-2} - h'_{q,N-2} \\ \vdots & \ddots & \vdots \\ h_{q,1} - h'_{q,1} & \cdots & h_{q,0} - h'_{q,0} \end{bmatrix}.$$

The matrix $M_q(h, h')$ has the following property.

Proposition J.5.1. *Let h, h' be two NTRU public keys and h_q, h'_q their inverses in $\mathbb{Z}_q[X]/\left(X^N - 1\right)$. Assume that*

$$g * h_q = f + qv, \qquad g * h'_q = f' + qv'.$$

Then the vector $(\lambda g, f - f')$ is in the lattice $\mathcal{L}_q(h, h')$ and

$$(g, -v + v') * M_q(h, h') = (\lambda g, f - f').$$

Proof. Assume that $g * h_q = f + qv$ and $g * h'_q = f' + qv'$. Then $g * h_q = f \pmod q$ and $g * h'_q = f' \pmod q$. This gives $g * (h_q - h'_q) = f - f' \pmod q$ and it follows that the vector $(\lambda g, f - f')$ is in $\mathcal{L}_q(h, h')$. More precisely,

$$\begin{aligned}
&(g, -v + v') * M_q(h, h') \\
&= (g, -v + v') * \begin{bmatrix} \lambda I_N & H_q - H'_q \\ 0 & qI_N \end{bmatrix} \\
&= (\lambda g, f - f').
\end{aligned}$$

This terminates the proof. □

J.5.2 The Gaussian heuristics

We can apply the the Gaussian heuristic to the lattice $\mathcal{L}_q(h, h')$. The shortest non-zero vector is approximately

$$\begin{aligned}
&\sigma(\mathcal{L}_q(h, h')) \\
&= \sqrt{\frac{\dim(\mathcal{L}_q(h, h'))}{2\pi e}} \det \mathcal{L}_q(h, h')^{1/\dim(\mathcal{L}_q(h,h'))} \\
&= \sqrt{\frac{\lambda N q}{\pi e}}.
\end{aligned}$$

To compare the length of the target vector $(\lambda g, f - f')$ to the length of the expected shortest vector $\sigma(\mathcal{L}_q(h, h'))$, we consider the ratio

$$c_2 = \frac{\|(\lambda g, f - f')\|}{\sigma(\mathcal{L}_q(h, h'))}.$$

In order to increase the chances of LLL to find the vector $(\lambda g, f - f')$, the attacker chooses the balancing constant λ to make c_2 as small as possible. For the lattice $\mathcal{L}_q(h, h')$, we have

$$\|(\lambda g, f - f')\| = \sqrt{\lambda^2 \|g\|^2 + \|f - f'\|^2}.$$

Hence the optimal choice for λ is

$$\lambda = \frac{\|f - f'\|}{\|g\|}.$$

which leads to

$$c_2 = \sqrt{\frac{2\pi e \|f - f'\| \|g\|}{qN}}.$$

To increase the chance of this attack to find $(\lambda g, f - f')$ comparatively to the attack of Coppersmith and Shamir, we should have $c_2 < c$ where c is the constant defined by (J.1). This gives the condition

$$\|f - f'\| < \min\left(\|f\|, \|f'\|\right).$$

J.6 Conclusion

We have shown that choosing two NTRU public keys $h = f_q^{-1} * g \pmod q$ and $h' = F_q'^{-1} * G' \pmod q$ could be insecure in some cases. Rewriting h' as $h' = f_q^{-1} * g' \pmod q$, where $g' = f * h' \pmod q$, we have shown, that using lattice reduction techniques, it is possible to find the private key (f, g) when $\|g - g'\| < \min\left(\|g\|, \|g'\|\right)$. We have shown that the same techniques apply when h' is invertible modulo q and $\|f - f'\| < \min\left(\|f\|, \|f'\|\right)$. Here f' is defined by the equality $f' * h' = g \pmod q$. For implementations of NTRU key pair generation we recommend to build in a check for $\|g - g'\| > \min\left(\|g\|, \|g'\|\right)$ and $\|f - f'\| > \min\left(\|f\|, \|f'\|\right)$. This is very easy to implement, and will only in extremely rare cases imply that the key pair is to be rejected. The main reason is that when f, g, F' and G' are generated randomly, the probability that g and $g' = f * h' \pmod q$ share an important amount of monomials is negligible. Similarly, the probability that f and $f' = g * h'^{-1} \pmod q$ share an important amount of monomials is also negligible.

Appendix K

Dirichlet Product for Boolean Functions

Journal of Applied Mathematics and Computing 2016
[120] with Willy Susilo and Joseph Tonien

Abstract :
Boolean functions play an important role in many symmetric cryptosystems and are crucial for their security. It is important to design boolean functions with reliable cryptographic properties such as balancedness and nonlinearity. Most of these properties are based on specific structures such as Möbius transform and Algebraic Normal Form. In this paper, we introduce the notion of Dirichlet product and use it to study the arithmetical properties of boolean functions. We show that, with the Dirichlet product, the set of boolean functions is an Abelian monoid with interesting algebraic structure. In addition, we apply the Dirichlet product to the sub-family of coincident functions and exhibit many properties satisfied by such functions.

K.1 Introduction

Boolean functions are used in logic and in many cryptographic applications such as blocks of symmetric key cryptosystems, stream cipher systems, coding theory and hash functions. Boolean functions are important for the security of such systems. So, for security reason, one seeks boolean functions having good properties such as nonlinearity, balancedness and algebraic immunity [38, 136] (see [37] for more properties). A boolean function is a mapping $\{0,1\}^n \to \{0,1\}$, often characterized by its truth table. The number of boolean functions with n variables is 2^{2^n} and it is impracticable to exhaustively exhibit a boolean function with optimal properties. One way to tackle this problem is to study the arithmetical structure of boolean functions and test their cryptographic reliability by the mean of algebraic tools such as Möbius transform and Algebraic Normal Form. For this reason, a lot of effort has been given to find ways to construct boolean functions with strong cryptographic properties.

For $n \geq 1$, we set $GF(2) = \{0,1\}$ and $GF(2)^n = \{0,1\}^n$. Any vector $x \in GF(2)^n$ is represented by its coordinates as $x = (x_1, \ldots, x_n)$ or simply $x = x_1 \ldots x_n$. The Hamming weight $w_H(x)$ of $x \in GF(2)^n$ is the number of non zero coordinates of x. An n-boolean function f is a mapping from $GF(2)^n$ into $GF(2)$. A boolean function is completely determined by its truth table

$$f(0,0,0\ldots,0), f(0,1,0,\ldots,0), f(0,1,0,\ldots,0), \ldots, f(1,1,1,\ldots,1),$$

and can be represented uniquely by the algebraic normal form (ANF)

$$f(x_1,\ldots,x_n) = \sum_{(\epsilon_1,\ldots,\epsilon_n)\in GF(2)^n} \hat{f}(\epsilon_1,\ldots,\epsilon_n) x_1^{\epsilon_1} \ldots x_n^{\epsilon_n},$$

where $\hat{f}$ is also a boolean function, called the Möbius transform of f. The transformation of f to its ANF can be performed using the truth table of f (see [28] and [127]).

Boolean functions have been intensively studied and various arithmetical properties are known such as Möbius transforms [127], Fourier transforms [28] and some cryptographic applications [136]. In this paper, we improve much

further such arithmetic properties by introducing the concept of *Dirichlet product*. Usually, Dirichlet product is well defined for arithmetical functions. An arithmetical function is a real-valued function defined on the positive integers [5]. The classical Dirichlet product $F*G$ for two arithmetical functions $F, G : \mathbb{N} \to \mathbb{R}$ is defined by

$$(F * G)(n) = \sum_{d|n} F(d)G(\frac{n}{d}) = \sum_{xy=n} F(x)G(y).$$

Dirichlet product is commutative $F * G = G * F$, associative $F * (G * H) = (F * G) * H$, and it has an identity

$$I(n) = \begin{cases} 1 \text{ if } n = 1 \\ 0 \text{ if } n > 1 \end{cases} \tag{K.1}$$

where $F * I = I * F = F$. So the set of all arithmetical functions $\mathbb{N} \to \mathbb{R}$ together with the Dirichlet product form an Abelian monoid. What more is that if $F(1) \neq 0$ then F has an inverse. So the subset of all arithmetical functions such that $F(1) \neq 0$ is an Abelian group with respect to the Dirichlet multiplication. The classical Dirichlet product provides great inside into some of the classical theorems in number theory. Many identities involving the Möbius function μ and the Euler totient function ϕ can be seen more intuitively in the language of Dirichlet product. For example, we have this identity

$$\sum_{d|n} \mu(d) = \begin{cases} 1 \text{ if } n = 1 \\ 0 \text{ if } n > 1 \end{cases} \tag{K.2}$$

where μ is the the Möbius function

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n = p_1 \cdot p_2 \cdots p_k \\ 0 & \text{otherwise.} \end{cases}$$

In the language of Dirichlet product, the identity (K.2) is $\mu * 1 = I$, it means that the Möbius function μ is the Dirichlet inverse of the constant function 1 where $1(n) = 1$. Similarly, Euler's totient function satisfies the following result.

$$\phi(n) = \sum_{d|n} \mu(d)\frac{n}{d}, \tag{K.3}$$

In the language of Dirichlet product, the identity (K.3) is $\mu * N = \phi$ where N is the function $N(n) = n$. In the language of group theory, it implies that $N = \phi * \mu^{-1} = \phi * 1$, that is

$$\sum_{d|n} \phi(d) = n. \tag{K.4}$$

So under the notion of Dirichlet product, two isolated results, (K.3) and (K.4) are ultimately related: (K.3) means $\phi = \mu * N$, whereas (K.4) means $N = \phi * 1 = \phi * \mu^{-1}$.

For two boolean functions f and g, we define the concept of Dirichlet product by setting for all $x \in GF(2)^n$

$$(f * g)(x) = \sum_{u \preceq x} f(u)g(x - u)$$

where, for $u = (u_1, \ldots, u_n) \in GF(2)^n$ and $x = (x_1, \ldots, x_n) \in GF(2)^n$, $u \preceq x$ if and only if for each $i \in \{1, \ldots, n\}$, $u_i \leq x_i$. We show that the Dirichlet product for boolean functions is commutative, associative and that the set of all boolean functions is an Abelian monoid and has the identity function I satisfying

$$I(x) = \begin{cases} 1 \text{ if } x = 0 \\ 0 \text{ if } x \neq 0 \end{cases}$$

Moreover, we link a boolean function f to its Möbius transform $\hat{f}$ using the Dirichlet products $f = \hat{f} * 1$ and $\hat{f} = f * 1$ where 1 is the constant function $1(x) = 1$. We show that the set of all boolean functions f such that $f(0, 0, \ldots, 0) = 1$ under the Dirichlet product form an Abelian group and the inverse of any such function f is f itself.

Finally, we will study the set of coincident functions and its algebraic structure. A coincident function is a boolean function f such that $\hat{f} = f$. Under the Dirichlet product, we show that the set of all coincident functions is a 2^{n-1} subspace with cardinality $2^{2^{n-1}}$.

The rest of this paper is organized as follows. In Section 2, we review the basic properties of boolean functions. In Section 3, we introduce the new notion of Dirichlet product for boolean functions and study its arithmetic

properties. In Section 4, we study the arithmetical and algebraic structure of the set of all coincident boolean functions. We conclude the paper in Section 5.

K.2 Boolean functions

Let $n \geq 1$. A boolean function f on n variables is a mapping from $\{0,1\}^n$ into $\{0,1\}$. It can be defined by its truth table, that is by $f(x_1,\ldots,x_n)$ for each $(x_1,\ldots,x_n) \in \{0,1\}^n$. For $x_i, \epsilon_i \in GF(2)$, we define $x_i^{\epsilon_i}$

$$x_i^{\epsilon_i} = \begin{cases} x_i & \text{if } \epsilon_i = 1, \\ 1 & \text{if } \epsilon_i = 0 \end{cases}$$

with the convention that $0^0 = 1$.

The set of all boolean functions on n variables is denoted $\mathcal{B}_n$ and any boolean function $f \in \mathcal{B}_n$ can be uniquely represented by an n-multivariate polynomial over $GF(2)$, called *algebraic normal form* (ANF),

$$f(x) = \sum_{\epsilon \in GF(2)^n} f_\epsilon \; x^\epsilon,$$

where $f_\epsilon \in GF(2)$ is the coefficient of the term $x^\epsilon = x_1^{\epsilon_1} x_2^{\epsilon_2} \ldots x_n^{\epsilon_n}$. In $GF(2)$, the addition operation is simply the XOR.

The summand $x^\epsilon = x_1^{\epsilon_1} \ldots x_n^{\epsilon_n}$ is called a monomial (term) in the ANF of f. The summand x^ϵ is said to appear in f if $f_\epsilon \neq 0$. The degree of this summand x^ϵ is the Hamming weight $w_H(\epsilon)$ of ϵ, that is the number of non-zero elements in it. The (*algebraic*) degree of f, denoted by $\deg(f)$, is the maximum degree of all summands that appear in f, that is maximum of all Hamming weights. For a constant zero function, we assume its degree is 0. The coefficient f_ϵ of the summand x^ϵ is related the Möbius transformation.

Definition K.2.1. Let $f \in \mathcal{B}_n$ with a polynomial

$$f(x) = \sum_{\epsilon \in GF(2)^n} f_\epsilon \; x^\epsilon.$$

The *Möbius transformation* of f is the boolean function $\hat{f} : GF(2)^n \to GF(2)$ defined as

$$\hat{f}(\epsilon) = f_\epsilon.$$

Using this definition, the polynomial $f(x)$ becomes

$$f(x) = \sum_{\epsilon \in GF(2)^n} \hat{f}(\epsilon)\ x^\epsilon.$$

We now define a partial ordering $\preceq$ in $GF(2)^n$ in the following definition.

Definition K.2.2. Let $u = (u_1, u_2, \ldots, u_n) \in GF(2)^n$ and $x = (x_1, x_2, \ldots, x_n) \in GF(2)^n$. We define the ordering

$$u \preceq x \Leftrightarrow u_i \leq x_i \quad \text{for all} \quad i \quad \text{with} \quad 1 \leq i \leq n.$$

The following simple result gives an expression of a boolean function f in terms of its Möbius transform $\hat{f}$.

Theorem K.2.3. *For $x = (x_1, \ldots, x_n) \in GF(2)^n$ and $u = (u_1, \ldots, u_n) \in GF(2)^n$,*

$$f(x) = \sum_{u \preceq x} \hat{f}(u), \tag{K.5}$$

Take an example, let $n = 3$,

$$\begin{aligned} f(x_1, x_2, x_3) = {} & \hat{f}(0,0,0) + \hat{f}(1,0,0)x_1 + \hat{f}(0,1,0)x_2 + \hat{f}(0,0,1)x_3 + \\ & \hat{f}(1,1,0)x_1x_2 + \hat{f}(0,1,1)x_2x_3 + \hat{f}(1,0,1)x_1x_3 + \hat{f}(1,1,1)x_1x_2x_3. \end{aligned}$$

So

$$\begin{aligned} f(0,0,0) &= \hat{f}(0,0,0) \\ f(1,0,0) &= \hat{f}(0,0,0) + \hat{f}(1,0,0) \\ f(0,1,0) &= \hat{f}(0,0,0) + \hat{f}(0,1,0) \\ f(0,0,1) &= \hat{f}(0,0,0) + \hat{f}(0,0,1) \\ f(1,1,0) &= \hat{f}(0,0,0) + \hat{f}(1,0,0) + \hat{f}(0,1,0) + \hat{f}(1,1,0) \\ &\ldots \end{aligned}$$

Solving these equations, we have the dual equations

$$\begin{aligned}
\hat{f}(0,0,0) &= f(0,0,0) \\
\hat{f}(1,0,0) &= f(0,0,0) + f(1,0,0) \\
\hat{f}(0,1,0) &= f(0,0,0) + f(0,1,0) \\
\hat{f}(0,0,1) &= f(0,0,0) + f(0,0,1) \\
\hat{f}(1,1,0) &= f(0,0,0) + f(1,0,0) + f(0,1,0) + f(1,1,0) \\
&\cdots
\end{aligned}$$

In matrix form, these equations become

$$\begin{pmatrix} f(0,0,0) \\ f(1,0,0) \\ f(0,1,0) \\ f(0,0,1) \\ f(1,1,0) \\ f(1,0,1) \\ f(0,1,1) \\ f(1,1,1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \hat{f}(0,0,0) \\ \hat{f}(1,0,0) \\ \hat{f}(0,1,0) \\ \hat{f}(0,0,1) \\ \hat{f}(1,1,0) \\ \hat{f}(1,0,1) \\ \hat{f}(0,1,1) \\ \hat{f}(1,1,1) \end{pmatrix}, \tag{K.6}$$

and

$$\begin{pmatrix} \hat{f}(0,0,0) \\ \hat{f}(1,0,0) \\ \hat{f}(0,1,0) \\ \hat{f}(0,0,1) \\ \hat{f}(1,1,0) \\ \hat{f}(1,0,1) \\ \hat{f}(0,1,1) \\ \hat{f}(1,1,1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} f(0,0,0) \\ f(1,0,0) \\ f(0,1,0) \\ f(0,0,1) \\ f(1,1,0) \\ f(1,0,1) \\ f(0,1,1) \\ f(1,1,1) \end{pmatrix}. \tag{K.7}$$

In the above example, we can see the duality between f and $\hat{f}$

$$\hat{f}(x) = \sum_{u \preceq x} f(u). \tag{K.8}$$

This is not accidental. The duality between (K.5) and (K.8) is explained by the fact that $\hat{f} = f * 1$ and $f = \hat{f} * 1$ as in Theorem K.3.9.

K.3 Dirichlet product for boolean functions

In this section, we define the Dirichlet product $f * g$ for two boolean functions f and g and study several properties of the monoid $(\mathcal{B}_n, *)$. In the rest of this paper, the term $(0, 0, \ldots, 0) \in GF(2)^n$ is often denoted as 0.

Lemma K.3.1. *Let $x = (x_1, x_2, \ldots, x_n) \in GF(2)^n$. Then there are $2^{w_H(x)}$ terms $u = (u_1, u_2, \ldots, u_n) \in GF(2)^n$ such that $u \preceq x$ where $w_H(x)$ is the Hamming weight of x.*

Proof. Let $x = (x_1, x_2, \ldots, x_n)$. For each i with $1 \leq i \leq n$, we have

$$u_i \leq x_i \quad \text{for} \quad \begin{cases} u_i = 0 & \text{if } x_i = 0 \\ u_i \in \{0, 1\} & \text{if } x_i = 1 \end{cases}$$

It follows that the number of terms $u \in GF(2)^n$ satisfying $u \preceq x$ is

$$\prod_{i=1}^{n} 2^{x_i} = 2^{w_H(x)},$$

$w_H(x)$ is the Hamming weight of x. □

Example K.3.2. Let $n = 3$ and $x = (1, 0, 1) \in GF(2)^3$. Then the set of all $u \in GF(2)^3$ such that $u \preceq x$ is

$$\{(0, 0, 0), (0, 0, 1), (1, 0, 0), (1, 0, 1)\}.$$

Now, we define the notion of Dirichlet product of two boolean functions.

Definition K.3.3. The Dirichlet product of two boolean functions $f, g \in \mathcal{B}_n$ is defined as

$$(f * g)(x) = \sum_{u \preceq x} f(u) g(x - u)$$

Example K.3.4. Let $n = 3$ and $x = (0, 1, 1) \in GF(2)^3$. Let $f, g \in \mathcal{B}_3$. Then the Dirichlet product of f and g is

$$\begin{aligned}(f * g)(0,1,1) =& f(0,0,0)g(0,1,1) + f(0,1,0)g(0,0,1) \\ &+ f(0,0,1)g(0,1,0) + f(0,1,1)g(0,0,0).\end{aligned}$$

The following result shows that the set $\mathcal{B}_n$ is an abelian monoid with respect to the Dirichlet product.

Theorem K.3.5. $(\mathcal{B}_n, *)$ *is an Abelian monoid with the identity*

$$I(x) = \begin{cases} 1 \text{ if } x = 0 \\ 0 \text{ if } x \neq 0 \end{cases} \tag{K.9}$$

Proof. We have

$$\begin{aligned}(f * g)(x) &= \sum_{u \preceq x} f(u)g(x-u) \\ &= \sum_{u,v \preceq x, u+v=x} f(u)g(v) \\ &= \sum_{v \preceq x} g(v)f(x-v) = (g * f)(x),\end{aligned}$$

so the Dirichlet product is commutative: $f * g = g * f$.
We also have

$$((f * g) * h)(x) = \sum_{u,v,w \preceq x, u+v+w=x} f(u)g(v)h(w) = (f * (g * h))(x)$$

so the Dirichlet product is associative.
Finally,

$$(f * I)(x) = \sum_{u,v \preceq x, u+v=x} f(u)I(v) = f(x)I(0) = f(x),$$

and I is the identity. □

The following result shows that the Dirichlet product is distributive over the addition operation in $\mathcal{B}_n$.

Lemma K.3.6. *For $f, g \in \mathcal{B}_n$, define addition operation $f + g \in \mathcal{B}_n$ as*

$$(f + g)(x) = f(x) + g(x).$$

Then the Dirichlet product is distributive over this addition operation.

Proof. We have

$$(f * (g + h))(x) = \sum_{u \preceq x} f(u)(g + h)(x - u) = \sum_{u \preceq x} f(u)(g(x - u) + h(x - u))$$

$$= \sum_{u \preceq x} f(u)g(x - u) + \sum_{u \preceq x} f(u)h(x - u) = (f * g)(x) + (f * h)(x)$$

so $f * (g + h) = f * g + f * h$. □

The next result gives one of the basic properties of the Dirichlet product.

Lemma K.3.7. *For any functions $f, g \in \mathcal{B}_n$,*

$$(f * g)(0) = f(0)g(0)$$

Proof. Since $u \preceq 0$ happens only for $u = 0$, we have

$$(f * g)(0) = \sum_{u \preceq 0} f(u)g(0 - u) = f(0)g(0).$$

□

The next result defines the constant boolean function 1 and links it to the identity function I.

Lemma K.3.8. *Let $1 \in \mathcal{B}_n$ denote the constant function*

$$1(x) = 1, \quad \forall x \in GF(2)^n \tag{K.10}$$

then

$$1 * 1 = I.$$

It means that 1 is its own inverse under Dirichlet multiplication.

Proof. By Theorem K.3.7, we have $(1 * 1)(0) = 1(0)1(0) = 1$. For $x \neq 0$, we have

$$(1 * 1)(x) = \sum_{u \preceq x} 1(u)1(x - u) = \sum_{u \preceq x} 1.$$

Since, by Lemma K.3.1, there are $2^{w_H(x)}$ terms u with $u \preceq x$, we have $(1 * 1)(x) = 0$ for $x \neq 0$. In conclusion, $1 * 1 = I$. □

The following result shows that the ANF of a boolean function is related to the Dirichlet product.

Theorem K.3.9. *For any function $f \in \mathcal{B}_n$, we have*

$$f = \hat{f} * 1, \quad \hat{f} = f * 1, \quad \hat{\hat{f}} = f.$$

Proof. First, we have

$$(\hat{f} * 1)(x) = \sum_{u \preceq x} \hat{f}(u)1(x - u) = \sum_{u \preceq x} \hat{f}(u).$$

Therefore, by Theorem K.2.3, $f = \hat{f} * 1$.
Combining this with Lemma K.3.8, we get

$$f * 1 = (\hat{f} * 1) * 1 = \hat{f} * (1 * 1) = \hat{f} * I = \hat{f}.$$

Applying the former results, we get

$$\hat{\hat{f}} = \hat{f} * 1 = (f * 1) * 1 = f * (1 * 1) = f * I = f.$$

This terminates the proof. □

The mysterious duality between a boolean function and its Möbius transformation is actually a manifestation of a simple fact in Dirichlet product, that is $1 * 1 = I$. The relationship between the results of Theorem K.3.9 is liken to that of (K.3) and (K.4).

Theorem K.3.10. *For any function $f \in \mathcal{B}_n$,*

$$\hat{f}(0) = f(0).$$

Proof. The proof follows from Lemma K.3.7 and Theorem K.3.9. □

The following result shows that $f * f$ is either the identity I or the constant function 0.

Theorem K.3.11. *For any function $f \in \mathcal{B}_n$,*

$$f * f = f(0)I = \begin{cases} I \text{ if } f(0) = 1 \\ 0 \text{ if } f(0) = 0 \end{cases} \tag{K.11}$$

Proof. Applying Lemma K.3.7, we get $(f * f)(0) = f(0)f(0) = f(0)$. When $x \neq 0$,

$$(f * f)(x) = \sum_{u \preceq x} f(u)f(x - u).$$

Since $u \preceq x$ and $x - u \preceq x$, everything in the sum appear twice. Hence, $(f * f)(x) = 0$. So $f * f = f(0)I$. □

Theorem K.3.12. *For any function $f \in \mathcal{B}_n$,*

$$f * \hat{f} = \hat{f} * f = f(0), \tag{K.12}$$

where $f(0)$ is the constant function defined by $f(0)(x) = f(0)$.

Proof. By Theorem K.3.9 and Theorem K.3.11, we have

$$f * \hat{f} = f * (f * 1) = (f * f) * 1 = f(0)I * 1 = f(0)1 = f(0),$$

□

In the following result, we give a characterization of a reversible boolean function with respect to the Dirichlet product.

Theorem K.3.13. *For any function $f \in \mathcal{B}_n$, f has a Dirichlet inverse if and only if $f(0) = 1$, and in this case, f is the Dirichlet inverse of itself.*

Proof. Suppose that f is invertible with an inverse g. Then $f * g = I$ and $(f*g)(0) = f(0)g(0) = 1$. Then $f(0) = 1$. Conversely, suppose that $f(0) = 1$, then $f * f = f(0)I = I$. Hence f is invertible and f is the Dirichlet inverse of itself. □

Next, we show that the set of Dirichlet invertible boolean functions is an Abelian group.

Theorem K.3.14. *Let $\mathcal{B}_n^+$ denote the set*

$$\mathcal{B}_n^+ = \{f \in \mathcal{B}_n \ : \ f(0) = 1\}.$$

Then $(\mathcal{B}_n^+, *)$ is an Abelian group.

Proof. Let $f \in \mathcal{B}_n^+$ and $g \in \mathcal{B}_n^+$ be two invertible boolean functions. By Theorem K.3.13, we know that $f(0) = g(0) = 1$. Then $(f*g)(0) = f(0)g(0) = 1$, which implies that $f * g \in \mathcal{B}_n^+$. Moreover, the inverse of $f \in \mathcal{B}_n^+$ is itself and $\mathcal{B}_n^+$ contains the identity function I. These properties show that $(\mathcal{B}_n^+, *)$ is an Abelian subgroup of $(\mathcal{B}_n, *)$. □

The following result is related to the degree of boolean functions. Recall the degree of a boolean function f is defined as the maximum number of variables of the terms $x^\epsilon = x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}$ in the ANF of f.

Theorem K.3.15. *For any $f, g \in \mathcal{B}_n$, we have*

$$\deg(f) + \deg(g) \geq \deg(f * g * 1) \quad \textit{and} \quad \deg(f) + \deg(\hat{f}) \geq n.$$

Proof. To prove the first assertion, first, if $\deg(f) + \deg(g) \geq n$ then this assertion is obviously true. We only need to prove it for the case $\deg(f) + \deg(g) < n$. If $w_H(x) > \deg(f) + \deg(g)$, then for any $u \preceq x$, $w_H(u) + w_H(x - u) = w(x) > \deg(f) + \deg(g)$, so $w_H(u) > \deg(f)$ or $w_H(x - u) > \deg(g)$. If $w_H(u) > \deg(f)$ then $\hat{f}(u) = 0$, and if $w_H(x-u) > \deg(g)$ then $\hat{g}(x-u) = 0$, so in either case, we have $\hat{f}(u)\hat{g}(x - u) = 0$. It follows that

$$(\hat{f} * \hat{g})(x) = \sum_{u \preceq x} \hat{f}(u)\hat{g}(x - u) = 0$$

holds for any $x \in GF(2)^n$ such that $w_H(x) > \deg(f) + \deg(g)$. Therefore,

$$\deg((\hat{f} * \hat{g}) * 1) \leq \deg(f) + \deg(g).$$

Finally, $(\hat{f} * \hat{g}) * 1 = f * 1 * g * 1 * 1 = f * g * 1$. This gives $\deg(f) + \deg(g) \geq \deg(f * g * 1)$.
Next, we have

$$\deg(f) + \deg(\hat{f}) \geq \deg(f * \hat{f} * 1).$$

But $f * \hat{f} * 1 = f * f * 1 * 1 = f(0)I * I = f(0)I = I$, so $\deg(f * \hat{f} * 1) = \deg(I) = n$ and we obtain the inequality $\deg(f) + \deg(\hat{f}) \geq n$. □

K.3.1 Basis for $(\mathcal{B}_n, +)$

For $f, g \in \mathcal{B}_n$, the function $f+g \in \mathcal{B}_n$ is defined as $(f+g)(x) = f(x)+g(x)$. With this addition operation, $\mathcal{B}_n$ is a free Abelian group. There are two natural ways to choose a basis for $\mathcal{B}_n$. We will describe them in Theorem K.3.16 and Theorem K.3.17.

Theorem K.3.16. *For each $a \in GF(2)^n$, define the function $\delta_a \in \mathcal{B}_n$ as follows*

$$\delta_a(x) = (x_1 + a_1 + 1)(x_2 + a_2 + 1)\dots(x_n + a_n + 1) = \begin{cases} 1 \text{ if } x = a \\ 0 \text{ if } x \neq a \end{cases} \quad \text{(K.13)}$$

Then $\{\delta_a\}_{a \in GF(2)^n}$ forms a basis for the vector space $(\mathcal{B}_n, +)$. Each function $f \in \mathcal{B}_n$ can be written as a linear combination of basis functions δ_a as

$$f = \sum_{a \in GF(2)^n} f(a)\ \delta_a. \quad \text{(K.14)}$$

Proof. If $x = a$, then for each $i = 1, 2, \dots, n$, $x_i + a_i + 1 = 1$ and $\delta_a(x) = 1$. If $x \neq a$, then $x_i \neq a_i$ for some i. Hence $x_i + a_i + 1 = 0$ and $\delta_a(x) = 0$.

We have

$$\sum_{a \in GF(2)^n} f(a)\ \delta_a(x) = f(x)\ \delta_x(x) + \sum_{a \neq x} f(a)\ \delta_a(x) = f(x),$$

so $f = \sum_{a \in GF(2)^n} f(a)\ \delta_a$. □

Note that, δ_0 is the Dirichlet identity function I:

$$I(x) = \delta_0(x) = (x_1 + 1)(x_2 + 1)\dots(x_n + 1) = \begin{cases} 1 \text{ if } x = 0 \\ 0 \text{ if } x \neq 0 \end{cases} \quad \text{(K.15)}$$

Theorem K.3.17. *For each $a \in GF(2)^n$, define the function $\rho_a \in \mathcal{B}_n$ as follows*

$$\rho_a(x) = x^a = x_1^{a_1} x_2^{a_2} \dots x_n^{a_n} = \begin{cases} 1 \text{ if } a \preceq x \\ 0 \text{ if } a \not\preceq x \end{cases} \quad \text{(K.16)}$$

Then $\{\rho_a\}_{a\in GF(2)^n}$ *forms a basis for the vector space* $(\mathcal{B}_n, +)$*. Each function* $f \in \mathcal{B}_n$ *can be written as a linear combination of basis functions* ρ_a *as*

$$f = \sum_{a\in GF(2)^n} \hat{f}(a)\ \rho_a. \tag{K.17}$$

Proof. If $a \preceq x$ then $a_i \leq x_i$ for each $i = 1, 2, \ldots, n$. If $x_i = 0$, then $a_i = 0$ and $x_i^{a_i} = 0^0 = 1$. If $x_i = 1$, then $x_i^{a_i} = 1$. In all cases, $x_i^{a_i} = 1$ and $\rho_a(x) = 1$. Next, suppose that $a \not\preceq x$. Then there exists i with $1 \leq i \leq n$ such that $a_i > x_i$. This implies that $x_i = 0$ and $a_i = 1$. Hence $x_i^{a_i} = x_i = 0$ and $\rho_a(x) = 0$.
Now, we have for $x \in GF(2)^n$,

$$\sum_{a\in GF(2)^n} \hat{f}(a)\ \rho_a(x) = \sum_{a\preceq x} \hat{f}(a)\ \rho_a(x) + \sum_{a\not\preceq x} \hat{f}(a)\ \rho_a(x) = \sum_{a\preceq x} \hat{f}(a) = f(x),$$

by Theorem K.2.3. □

Theorem K.3.18. *For any* $a \in GF(2)^n$*, the basis functions* δ_a *and* ρ_a *satisfy the following relations:*

- $\delta_a * 1 = \rho_a$ *and* $\rho_a * 1 = \delta_a$,
- $\delta_a * \delta_b = \rho_a * \rho_b = \rho_a\, \rho_b\, \delta_{a+b}$.

Proof. First, observe that since $\rho_a(x) = x^a$, the function ρ_a in ANF has only one monomial term x^a, so its ANF coefficient function is δ_a. That is $\rho_a * 1 = \delta_a$, and so $\delta_a * 1 = \rho_a * 1 * 1 = \delta_a * I = \delta_a$.
Next, for any a and b, we have

$$\begin{aligned}(\delta_a * \delta_b)(x) &= \sum_{u,v\preceq x, u+v=x} \delta_a(u)\delta_b(v) \\ &= \begin{cases} 1 & \text{if } a \preceq x,\ b \preceq x,\ a+b = x. \\ 0 & \text{otherwise} \end{cases} \\ &= \rho_a(x)\rho_b(x)\delta_{a+b}(x)\end{aligned}$$

Therefore,

$$\delta_a * \delta_b = \rho_a\, \rho_b\, \delta_{a+b}.$$

Finally,

$$\rho_a * \rho_b = \delta_a * 1 * \delta_b * 1 = \delta_a * \delta_b.$$

□

K.4 Coincident functions

In this section, we study a special family of boolean functions, called coincident functions which was first introduced in [127].

Definition K.4.1. A coincident function is a function $f : GF(2)^n \to GF(2)$ such that $\hat{f} = f$.

Example K.4.2. For $n = 3$, let f be the function

$$\begin{aligned} & f(x_1, x_2, x_3) \\ = \; & \hat{f}(0,0,0) + \hat{f}(1,0,0)x_1 + \hat{f}(0,1,0)x_2 + \hat{f}(0,0,1)x_3 + \\ & \hat{f}(1,1,0)x_1x_2 + \hat{f}(0,1,1)x_2x_3 + \hat{f}(1,0,1)x_1x_3 + \hat{f}(1,1,1)x_1x_2x_3 \\ = \; & 0 + x_1 + x_2 + x_3 + x_1x_2 + x_2x_3 + x_1x_3 + x_1x_2x_3. \end{aligned}$$

Then

$$\begin{aligned} f(0,0,0) &= \hat{f}(0,0,0) = 0, \\ f(1,0,0) &= \hat{f}(1,0,0) = 1, \\ &\dots, \\ f(1,1,1) &= \hat{f}(1,1,1) = 1, \end{aligned}$$

that is $f = \hat{f}$ and f is coincident.

Theorem K.4.3. *For any coincident function f,*

$$f(0) = 0.$$

Proof. Suppose that f is a coincident function, that is $f = \hat{f}$. Then, using Theorem K.2.3, we get

$$f(0,0,\dots,0,1) = \hat{f}(0) + \hat{f}(0,0,\dots,0,1).$$

Since $f(0,0,\dots,0,1) = \hat{f}(0,0,\dots,0,1)$, then $\hat{f}(0) = f(0) = 0$. □

Let $\mathcal{C}_n$ denote the set of all such coincident functions.

Theorem K.4.4. *A function $f \in \mathcal{B}_n$ is a coincident function if and only if*

$$(1 + I) * f = 0.$$

Thus, $\mathcal{C}_n$ is the annihilator of $1 + I$ in $\mathcal{B}_n$.

Proof. Suppose that f is a coincident function, that is $f = \hat{f}$. Then

$$0 = \hat{f} + f = f * 1 + f * I = f * (1 + I).$$

Conversely, suppose that $f * (1 + I) = 0$. Then, using Theorem K.3.9, we get $f*1+f*I = \hat{f}+f = 0$. This implies that $\hat{f} = f$ and then f is coincident. □

Observe that for any $x \in GF(2)^n$, we have

$$\begin{aligned}(1 + I)(x) &= (x_1 + 1)(x_2 + 1) \dots (x_n + 1) + 1,\\ \delta_{1\dots1}(x) &= x_1 x_2 \dots x_n,\\ \rho_{1\dots1}(x) &= x_1 x_2 \dots x_n.\end{aligned}$$

Theorem K.4.5. *The boolean functions $1 + I$, $\delta_{1\dots1}$ and $\rho_{1\dots1}$ are coincident functions.*

Proof. Combining Theorem K.4.4 and Theorem K.3.8, we get

$$(1 + I) * (1 + I) = 1 * 1 + I * I = I + I = 0.$$

Hence $1+I$ is coincident. Next, combining Theorem K.4.4 and Lemma K.3.18, we get for any $x \in GF(2)^n$,

$$(1 + I) * \delta_{1\dots1}(x) = (1 * \delta_{1\dots1})(x) + (I * \delta_{1\dots1})(x) = \rho_{1\dots1}(x) + \delta_{1\dots1}(x).$$

Then, using Theorem K.3.16 and Theorem K.3.17, we get

$$\rho_{1\dots1}(x) + \delta_{1\dots1}(x) = \begin{cases} 1 + 1 = 0 & \text{if } x = 1 \dots 1,\\ 0 + 0 = 0 & \text{if } x \neq 1 \dots 1.\end{cases}$$

It follows that $(1 + I) * \delta_{1\dots1} = 0$ and $\delta_{1\dots1}$ is coincident. □

Theorem K.4.6. *For any $u \in GF(2)^n$, $\delta_u + \rho_u$ is a coincident function.*

Proof. Combining Theorem K.4.4 and Theorem K.3.18, we get

$$(1 + I) * (\delta_u + \rho_u) = 1 * \delta_u + 1 * \rho_u + \delta_u + \rho_u = 2\delta_u + 2\rho_u = 0$$

Hence $\delta_u + \rho_u$ is a coincident function. □

Theorem K.4.7. *A function $f \in \mathcal{B}_n$ is a coincident function if and only if for any $(x_2, \ldots, x_n) \in GF(2)^{n-1}$,*

$$f(0, x_2, \ldots, x_n) = \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(1, u_2, \ldots, u_n), \tag{K.18}$$

where $u \prec x$ means $u \preceq x$ and $u \neq x$.

Proof. Since

$$(1 + I)(x) = \begin{cases} 0 \text{ if } x = 0 \\ 1 \text{ if } x \neq 0 \end{cases} \tag{K.19}$$

we have

$$((1 + I) * f)(x) = \sum_{u \preceq x} f(u)(1 + I)(x - u) = \sum_{u \prec x} f(u).$$

Therefore, $(1 + I) * f = 0$ if and only if for any $x \in GF(2)^n$,

$$\sum_{u \prec x} f(u) = 0.$$

Consider two cases, $x_1 = 0$ and $x_1 = 1$.
When $x_1 = 0$, the condition becomes

$$\sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(0, u_2, \ldots, u_n) = 0.$$

When $x_1 = 1$, the condition becomes

$$f(0, x_2, \ldots, x_n) + \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(0, u_2, \ldots, u_n) + \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(1, u_2, \ldots, u_n) = 0.$$

Therefore, if f is a coincident function then for any $(x_2, \ldots, x_n) \in GF(2)^{n-1}$, we must have

$$f(0, x_2, \ldots, x_n) = \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(1, u_2, \ldots, u_n).$$

Conversely, suppose that for any $(x_2, \ldots, x_n) \in GF(2)^{n-1}$,

$$f(0, x_2, \ldots, x_n) = \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(1, u_2, \ldots, u_n).$$

Then

$$\sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(0, u_2, \ldots, u_n) = \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} \sum_{(v_2,\ldots,v_n) \prec (u_2,\ldots,u_n)} f(1, v_2, \ldots, v_n).$$

The above sum is equal to 0 because for any term $f(1, v_2, \ldots, v_n)$, the number of its occurrences in the sum is equal to the number of $(u_2, \ldots, u_n)$ such that $(v_2, \ldots, v_n) \prec (u_2, \ldots, u_n) \prec (x_2, \ldots, x_n)$, and this is always an even number for any $(v_2, \ldots, v_n) \prec (x_2, \ldots, x_n)$. Hence for any $(x_2, \ldots, x_n) \in GF(2)^{n-1}$, we have

$$\sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(0, u_2, \ldots, u_n) = 0. \tag{K.20}$$

Therefore,

$$f(0, x_2, \ldots, x_n) + \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(0, u_2, \ldots, u_n) + \sum_{(u_2,\ldots,u_n) \prec (x_2,\ldots,x_n)} f(1, u_2, \ldots, u_n) = 0. \tag{K.21}$$

Combining (K.20) and (K.21), we see that

$$\sum_{u \prec x} f(u) = 0,$$

that is $(1 + I) * f = 0$ and f is a coincident function. □

The following theorem reveals a relationship between the set of coincident functions $\mathcal{C}_n$ and the set of all boolean functions $\mathcal{B}_n$.

Theorem K.4.8. *It holds that*

1. *A coincident function $f \in \mathcal{C}_n$ is specified freely and uniquely by its values on 2^{n-1} points $(1, u_2, \ldots, u_n) \in GF(2)^n$.*
2. *There are exactly $2^{2^{n-1}}$ coincident functions in total.*
3. *$(\mathcal{C}_n, +)$ is a 2^{n-1}-dimensional linear subspace of $(\mathcal{B}_n, +)$.*

Proof. To prove the first assertion, observe that by Theorem K.4.7, a coincident function $f \in \mathcal{C}_n$ is specified freely by its values on 2^{n-1} points $(1, u_2, \ldots, u_n) \in GF(2)^n$, and its values on 2^{n-1} other points $(0, u_2, \ldots, u_n) \in GF(2)^n$ are uniquely determined by (K.18). The second assertion follows since there are exactly 2 choices for choosing $f(1, u_2, \ldots, u_n) \in \{0, 1\}$, then there are exactly $2^{2^{n-1}}$ coincident functions in total.
To prove the third assertion, observe that if $f \in \mathcal{C}_n$ and $g \in \mathcal{C}_n$, then $f + g \in \mathcal{C}_n$. On the other hand, the relation (K.18) defines any coincident function $f \in \mathcal{C}_n$. It follows that $(\mathcal{C}_n, +)$ is a 2^{n-1}-dimensional linear subspace of $(\mathcal{B}_n, +)$. □

K.4.1 Basis for $(\mathcal{C}_n, +)$

By Theorem K.4.8, we know that $(\mathcal{C}_n, +)$ is a 2^{n-1}-dimensional linear subspace of $(\mathcal{B}_n, +)$. The following result gives an explicit basis for $(\mathcal{C}_n, +)$.

Theorem K.4.9. *For each $(u_2, \ldots, u_n) \in GF(2)^{n-1}$, define the function $\gamma_{(u_2,\ldots,u_n)} \in \mathcal{B}_n$ as follows*

$$\gamma_{(u_2,\ldots,u_n)} = \delta_{(0,u_2,\ldots,u_n)} + \delta_{(1,u_2,\ldots,u_n)} + \rho_{(0,u_2,\ldots,u_n)} + \rho_{(1,u_2,\ldots,u_n)}$$

Then $\{\gamma_{(u_2,\ldots,u_n)}\}_{(u_2,\ldots,u_n)\in GF(2)^{n-1}}$ forms a basis for the subspace $(\mathcal{C}_n, +)$, and each coincident function $f \in \mathcal{C}_n$ can be written as a linear combination of basis functions as

$$f = \sum_{(u_2,\ldots,u_n)\in GF(2)^{n-1}} f(1, u_2, \ldots, u_n)\ \gamma_{(u_2,\ldots,u_n)}.$$

Proof. A coincident function $f \in \mathcal{B}_n$ is specified freely and uniquely by its values on 2^{n-1} points $(1, u_2, \ldots, u_n) \in GF(2)^n$. For each $(u_2, \ldots, u_n) \in GF(2)^{n-1}$, define the coincident function $c_{(u_2,\ldots,u_n)} : GF(2)^n \to GF(2)$ as follows

$$c_{(u_2,\ldots,u_n)}(x) = \begin{cases} 1 \text{ if } (x_2, \ldots, x_n) = (u_2, \ldots, u_n) \\ 0 \text{ otherwise} \end{cases}$$

then the collection of these functions $c_{(u_2,\ldots,u_n)}$ will form a basis for the vector space $\mathcal{C}_n$ and

$$f = \sum_{(u_2,\ldots,u_n) \in GF(2)^{n-1}} f(1, u_2, \ldots, u_n)\ c_{(u_2,\ldots,u_n)}.$$

We need to show that

$$c_{(u_2,\ldots,u_n)} = \gamma_{(u_2,\ldots,u_n)}.$$

Indeed, by Theorem K.4.6, $\gamma_{(u_2,\ldots,u_n)}$ is a coincident function, so it suffices to show that $\gamma_{(u_2,\ldots,u_n)}$ and $c_{(u_2,\ldots,u_n)}$ agree on 2^{n-1} points $(1, x_2, \ldots, x_n)$. We have

$$\delta_{(0,u_2,\ldots,u_n)}(1, x_2, \ldots, x_n) = 0$$

$$\delta_{(1,u_2,\ldots,u_n)}(1, x_2, \ldots, x_n) = \begin{cases} 1 \text{ if } (x_2, \ldots, x_n) = (u_2, \ldots, u_n) \\ 0 \text{ otherwise} \end{cases}$$

$$\rho_{(0,u_2,\ldots,u_n)}(1, x_2, \ldots, x_n) = \rho_{(1,u_2,\ldots,u_n)}(1, x_2, \ldots, x_n)$$

Therefore,

$$\gamma_{(u_2,\ldots,u_n)}(1, x_2, \ldots, x_n) = \begin{cases} 1 \text{ if } (x_2, \ldots, x_n) = (u_2, \ldots, u_n) \\ 0 \text{ otherwise} \end{cases}$$

and thus, $\gamma_{(u_2,\ldots,u_n)} = c_{(u_2,\ldots,u_n)}$. □

Example K.4.10. When $n = 3$, the following 4 coincident functions form a basis for the subspace of all coincident functions:

$$\begin{aligned} \gamma_{(0,0)} &= \delta_{(0,0,0)} + \delta_{(1,0,0)} + \rho_{(0,0,0)} + \rho_{(1,0,0)} \\ &= (x_1 + 1)(x_2 + 1)(x_3 + 1) + x_1(x_2 + 1)(x_3 + 1) + 1 + x_1 \\ &= x_1 + x_2 + x_3 + x_2x_3 \end{aligned}$$

$$\begin{aligned}\gamma_{(1,0)} &= \delta_{(0,1,0)} + \delta_{(1,1,0)} + \rho_{(0,1,0)} + \rho_{(1,1,0)} \\ &= (x_1+1)x_2(x_3+1) + x_1x_2(x_3+1) + x_2 + x_1x_2 \\ &= x_1x_2 + x_2x_3\end{aligned}$$

$$\begin{aligned}\gamma_{(0,1)} &= \delta_{(0,0,1)} + \delta_{(1,0,1)} + \rho_{(0,0,1)} + \rho_{(1,0,1)} \\ &= (x_1+1)(x_2+1)x_3 + x_1(x_2+1)x_3 + x_3 + x_1x_3 \\ &= x_1x_3 + x_2x_3\end{aligned}$$

$$\begin{aligned}\gamma_{(1,1)} &= \delta_{(0,1,1)} + \delta_{(1,1,1)} + \rho_{(0,1,1)} + \rho_{(1,1,1)} \\ &= (x_1+1)x_2x_3 + x_1x_2x_3 + x_2x_3 + x_1x_2x_3 \\ &= x_1x_2x_3.\end{aligned}$$

These 4 functions can be seen to be coincident in the following table

	$\gamma_{(0,0)}$	$\gamma_{(1,0)}$	$\gamma_{(0,1)}$	$\gamma_{(1,1)}$
$(0,0,0)$	0	0	0	0
$(0,1,0)$	1	0	0	0
$(0,0,1)$	1	0	0	0
$(0,1,1)$	1	1	1	0
$(1,0,0)$	1	0	0	0
$(1,1,0)$	0	1	0	0
$(1,0,1)$	0	0	1	0
$(1,1,1)$	0	0	0	1

Theorem K.4.11. *For each $f \in \mathcal{C}_n$ define*

$$f_\delta = \sum_{(u_2,\ldots,u_n)\in GF(2)^{n-1}} f(1,u_2,\ldots,u_n)\ (\delta_{(0,u_2,\ldots,u_n)} + \delta_{(1,u_2,\ldots,u_n)}).$$

and

$$f_\rho = \sum_{(u_2,\ldots,u_n)\in GF(2)^{n-1}} f(1,u_2,\ldots,u_n)\ (\rho_{(0,u_2,\ldots,u_n)} + \rho_{(1,u_2,\ldots,u_n)}).$$

then

$$f = f_\delta + f_\rho = (1+I) * f_\delta = (1+I) * f_\rho.$$

Proof. By Theorem K.4.9,

$$f = f_\delta + f_\rho$$

and by Theorem K.3.18,

$$f_\delta * 1 = f_\rho, \quad f_\rho * 1 = f_\delta,$$

therefore,

$$f = (1 + I) * f_\delta = (1 + I) * f_\rho.$$

□

Theorem K.4.12. *A function $f \in \mathcal{B}_n$ is a coincident function if and only if $f = (1 + I) * g$ for some function $g \in \mathcal{B}_n$.*

Proof. Suppose that $f = (1 + I) * g$. Then, using Theorem K.4.5, we get

$$(1 + I) * f = (1 + I) * (1 + I) * g = 0 * g = 0,$$

so f is a coincident function.
Conversely, suppose that f is a a coincident function. Then by Theorem K.4.11, we have $f = (1 + I) * g$ with $g = f_\delta$. □

K.5 Conclusion and Future Work

In this paper, we have introduced a new notion, called *Dirichlet product* for boolean functions. We have intensively studied the arithmetical and the algebraic structures of the set of all boolean functions under this Dirichlet product. We have presented the affects of the Dirichlet product on a boolean function and its Möbius transform. We have applied the Dirichlet product to coincident boolean functions and exhibited new properties and characterizations of such functions.
The results presented in this paper on the new notion of Dirichlet product for boolean functions are not exhaustive. They are only the first steps toward further applications of the Dirichlet product, especially in cryptography. We leave it as future work to investigate possible applications of the Dirichlet product to find useful results to compute the algebraic degree of a boolean function and to characterize cryptographic properties such as nonlinearity, balancedness, correlation immunity and algebraic immunity.

Appendix L

New Attack on RSA and Demytko's Elliptic Curve Cryptosystem

Mathematics in Computer Science, 2016
[121] with Emmanuel Fouotsa

Abstract :
Let $N = pq$ be an RSA modulus and e a public exponent. We show how to factor the RSA modulus if e satisfies an equation of the form $eu - (p - s)(q - r)v = w$ with suitably small unknown integers u, v, w, r and s under the condition that $p - s$ is factorable using the Elliptic Curve Method for factorization ECM. As an application, we propose an attack on Demytko's elliptic curve cryptosystem. Our method is based on Coppersmith's technique for solving multivariate polynomial modular equations.

L.1 Introduction

In 1976, Diffie and Hellman [40] invented the concept of the public-key cryptosystem. Since then, various schemes have been proposed as public-key

cryptosystems.

In 1978, Rivest, Shamir, and Adleman [131] proposed RSA, the most widely used public-key cryptosystem. The public parameters in RSA are the modulus $N = pq$ and the public exponent e satisfying $\gcd(e, (p-1)(q-1)) = 1$ where p, q are large prime numbers of the same bit-size. The decryption exponent is the integer d such that $ed \equiv 1 \pmod{(p-1)(q-1)}$.

In 1985, Koblitz [75] and Miller [95] independently suggested the use of elliptic curves in cryptography, mainly for the Diffie-Hellman [40] key exchange protocol and the El Gamal cryptosystem [42]. Let $p > 3$ be a prime number and a, b be two integers such that $\gcd(4a^3 + 27b^2, p) = 1$. The elliptic curve $E_p(a, b)$ over the field $\mathbb{F}_p$ is the set of points $P = (x, y)$ such that $y^2 \equiv x^3 + ax + b \pmod{p}$ together with the point at infinity. The number of points in $E_p(a, b)$ is $\#E_p(a, b) = p + 1 - t_p$ where t_p is an integer satisfying the Hasse bound $|t_p| \leq 2\sqrt{p}$. Elliptic curves can be extended over the ring $\mathbb{Z}/n\mathbb{Z}$ where n is a composite integer. Such elliptic curves can serve to find small prime factors of n as in the Elliptic Curve Method (ECM) for factorization [84].

In 1994, Demytko [39] developed a cryptosystem using an elliptic curve $E_N(a, b)$ over the ring $\mathbb{Z}/N\mathbb{Z}$ where $N = pq$ is an RSA modulus. In the Demytko system, the public parameters are N, a, b together with a public exponent e satisfying $\gcd\left(e, \left(p^2 - t_p^2\right)\left(q^2 - t_q^2\right)\right) = 1$. The decryption exponent is an integer d satisfying $ed \equiv 1 \pmod{\operatorname{lcm}(p + 1 \pm t_p, q + 1 \pm t_q)}$ where $t_p = p + 1 - \#E_p(a, b)$ and $t_q = q + 1 - \#E_q(a, b)$.

The RSA cryptosystem is deployed in many commercial systems for providing privacy and authenticity. If RSA is deployed in a device with small computing power, it is desirable to use a small public exponent e or a small private exponent d. Unfortunately, in 1990, Wiener [147] showed that RSA is insecure if $d < \frac{1}{3}N^{\frac{1}{4}}$. In 1999, Boneh and Durfee [17] improved this bound up to $d < N^{0.292}$. Their method is based on Coppersmith's method [34] for solving modular polynomial equations and uses the RSA key equation $ed - k(p-1)(q-1) = 1$. Afterwards, many attacks on RSA or variants of RSA have been presented using Coppersmith's method or other techniques (see [61], [93], [13]).

In this paper, using a variant RSA equation, we present a new attack on RSA by combining Coppersmith's method and the Elliptic Curve Method for factorization ECM. Let B be a positive integer. An integer n is said to be B-smooth if all prime factors are less than B. We say that B is an efficiency bound for ECM if every prime factor less than B of an integer n can be found by ECM.

Suppose that the public exponent $e = N^\beta$ satisfies a variant equation of the form $eu - (p-s)(q-r)v = w$ with suitably small unknown integers $u < N^\delta$, $|w| < N^\gamma$, $|r| < N^\alpha$ and $|s| < N^\alpha$ with $\alpha < \frac{1}{4}$. We show that the RSA modulus $N = pq$ can be factored under two conditions. The first condition is that $p - s$ is B-smooth for some efficiency bound B of ECM and the second condition is that δ satisfies the following inequality

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)} - \varepsilon,$$

where ε is a small positive constant. Our method is based on combining Coppersmith's method and ECM. We use Coppersmith's method to find the small solutions $(u, v, w, (p-s)(q-r))$ of the equation $eu-(p-s)(q-r)v = w$ and ECM to factor $(p-s)(q-r)$ and to extract the value of $p-s$ from the B-smooth part of $(p-s)(q-r)$. Finally reusing Coppersmith's method, we can find p from the value of $p - s$.

We apply the new method to present a new attack on Demytko's scheme. In this scheme, the public exponent e and the private exponent d satisfy one of the four modular equations $ed \equiv 1 \pmod{\operatorname{lcm}(p+1\pm t_p, q+1\pm t_q)}$. This gives rise to an equation of the form $eu - (p+1\pm t_p)(q+1\pm t_q)v = w$. Let $e = N^\beta$. Suppose that $|u| < N^\delta$, $|w| < N^\gamma$, $|t_p| < N^\alpha$ and $|t_q| < N^\alpha$ with $\alpha < \frac{1}{4}$ and that $p+1\pm t_p$ or $q+1\pm t_q$ is B-smooth. Then applying the new method as for RSA, one can factor the RSA modulus $N = pq$.

The rest of this paper is organized as follows. In Section 2, we review Coppersmith's method, the theory of elliptic curves, Demytko's elliptic curve cryptosystem and the Elliptic Curve Method ECM for factorization. In Section 3, we present the new attack on RSA, and in Section 4, we present the new attack on Demytko's scheme. We conclude in Section 5.

L.2 Preliminaries

The following classical result is useful for the proof of our new attack (see [104]).

Lemma L.2.1. *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Then*

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

L.2.1 Coppersmith's method

In 1996, Coppersmith [34] describes a technique to find small modular roots of univariate polynomials and small integer roots of bivariate polynomials. This method has been extended to more variables and has many surprising results in cryptanalysis. A typical example is the following result [91].

Theorem L.2.2 (Coppersmith)**.** *Let $N = pq$ be an RSA modulus with $q < p < 2q$. Let $\widetilde{S}$ be an approximation of an unknown multiple pr of p with $r \neq q$ and $|pr - \widetilde{S}| < N^{\frac{1}{4}}$. Then one can factor N in polynomial time.*

Let $h(x, y, z) \in \mathbb{Z}[x, y, z]$ be a polynomial with ω monomials of the form

$$h(x, y, z) = \sum_{i,j,k} a_{i,j,k} x^i y^j z^k.$$

The Euclidean norm of $h(x, y, z)$ is defined as

$$\|h(x, y, z)\| = \sqrt{\sum_{i,j,k} a_{i,j,k}^2}.$$

Under some conditions, a modular polynomial equation can be solved over the integers as presented in the following result [65].

Theorem L.2.3 (Howgrave-Graham)**.** *Let e be a positive integer and $h(x, y, z) \in \mathbb{Z}[x, y, z]$ be a polynomial with at most ω monomials. Suppose that*

$$h(x_0, y_0, z_0) \equiv 0 \pmod{e^m} \quad \text{and} \quad \|h(xX, yY, zZ)\| < \frac{e^m}{\sqrt{\omega}},$$

where $|x_0| < X$, $|y_0| < Y$, $|z_0| < Z$. Then $h(x_0, y_0, z_0) = 0$ holds over the integers.

To find polynomials with small coefficients that can be used in Howgrave-Graham's Theorem L.2.3, Coppersmith's method uses a lattice and a lattice reduction algorithm such as the LLL algorithm [86]. This reduction algorithm can be applied to find a basis of lattice vectors with relatively small norms (see [91]).

Theorem L.2.4 (LLL)**.** *Let $\mathcal{L}$ be a lattice spanned by a basis $(u_1, \ldots, u_\omega)$, then the LLL algorithm produces a new basis $(b_1, \ldots, b_\omega)$ satisfying*

$$\|b_1\| \leq \ldots \leq \|b_i\| \leq 2^{\frac{\omega(\omega-1)}{4(\omega+1-i)}} \det(\mathcal{L})^{\frac{1}{\omega+1-i}}, \quad i = 1, \ldots, \omega - 1.$$

Under the condition of Howgrave-Graham's Theorem, some modular polynomial equations derived from the reduced basis can be transformed to polynomial equations over the integers. For multivariate modular equations, solving the system of these polynomials is heuristic and depends on some extra assumptions such as the following one.

Assumption L.2.5. *Let $h_1, h_2, h_3 \in \mathbb{Z}[x, y, z]$ be the polynomials that are found by Coppersmith's method. Then the ideal generated by the polynomial equations $h_1(x, y, z) = 0$, $h_2(x, y, z) = 0$, $h_3(x, y, z) = 0$ has dimension zero.*

Under this assumption, a system of polynomials sharing the root can be solved by using Gröbner basis computation or resultant techniques (see [8] for more details).

L.2.2 Elliptic curves

Let $N = pq$ be an RSA modulus and let a and b be two integers such that $\gcd(4a^3 + 27b^2, N) = 1$. An elliptic curve $E_N(a, b)$ is the set of points (x, y) such that

$$y^2 \equiv x^3 + ax + b \pmod{N},$$

together with the point at infinity $\mathcal{O}$. It is well known that chord-and-tangent method in the case of elliptic curves $E_p(a, b)$ defined over the finite filed $\mathbb{F}_p$ still hold for $E_n(a, b)$ unless the inversion of a non-zero number Q does not exist modulo N. This case would lead to find a factor of N by computing

$\gcd(Q, N)$. When the prime factors p, q in $N = pq$ are large, then with overwhelming probability the inversion of a non-zero number will exist modulo N.

Let p be a prime number. Under modulo p, the cardinality of $E_p(a, b)$ is denoted $\#E_p(a, b)$ and satisfies the following result (see [140], p. 131).

Theorem L.2.6 (Hasse)**.** *The order of an elliptic curve $E_p(a, b)$ over $\mathbb{F}_p$ is given by*

$$\#E_p(a, b) = p + 1 - t_p, \quad \textit{where} \quad |t_p| \leq 2\sqrt{p}.$$

When the prime number p and the elliptic curve $E_p(a, b)$ are given, one can find the value of t_p using computational methods such the Schoof-Elkies-Atkin algorithm (SEA) (see [137]). Conversely, let p be a prime number and t an integer with $|t| < 2\sqrt{p}$. Let $H(d)$ denote the Kronecker class number (see Section 1.6 of [84]). Deuring's theory of CM-elliptic curves implies that there are $H(t^2-4p)$ elliptic curves on $\mathbb{Z}/p\mathbb{Z}$ having $p+1-t$ points. Note that when $|t| < \sqrt{p}$, $H(t^2-4p)$ satisfies the following inequalities (see Proposition 1.9 of [84])

$$c_1 \frac{\sqrt{p}}{\log p} < H(t^2 - 4p) < c_2 \sqrt{p} (\log p)(\log \log p)^2,$$

where c_1 and c_2 are effectively computable positive constants. This shows that the number of elliptic curves with known cardinality is non negligible.

Let p be a prime number and $E_p(a, b)$ be an elliptic curve with equation $y^2 \equiv x^3 + ax + b \pmod{p}$ and cardinality $\#E_p(a, b) = p + 1 - t_p$. The twist of $E_p(a, b)$ is the elliptic curve $E'_p(a, b)$ defined by the equation $cy^2 \equiv x^3 + ax + b \pmod{p}$ where c is a fixed quadratic non-residue modulo p. Then the cardinality of $E'_p(a, b)$ is $\#E'_p(a, b) = p + 1 + t_p$.

L.2.3 Demytko's elliptic curve cryptosystem

In 1994, Demytko [39] proposed a new cryptosystem defined over the field $\mathbb{Z}/N\mathbb{Z}$ where $N = pq$ is an RSA modulus such that $p \equiv q \equiv 2 \pmod{3}$. Demytko's scheme uses fixed integers a and b and a fixed modulus N. Demytko's scheme uses only the x-coordinate of a point $P = (x, y) \in E_N(a, b)$ to

compute a multiple $eP \in E_N(a,b)$ (see Lemma 2 in [82]). Demytko's scheme can be summarized as follows.

1. **Key Generation:**
   - Choose two distinct prime numbers p and q of similar bit-length.
   - Compute $N = pq$.
   - Select two integers $a, b < p$ such that $\gcd\left(n, 4a^3 + 27b^2\right) = 1$.
   - Choose e such that $\gcd\left(e, \left(p^2 - t_p^2\right)\left(q^2 - t_q^2\right)\right) = 1$.
   - Keep p, q secret and publish N, e, a, b.
2. **Encryption:**
   - Transform the message m as the x-coordinate of a point $P = (m_x, m_y)$ on the elliptic curve $E_N(a,b)$.
   - Compute the ciphertext point $C = eP = (c_x, c_y) = e(m_x, m_y)$ on the elliptic curve $y^2 = x^3 + ax + b \pmod N$.
3. **Decryption:**
   - Compute $u = c_x^3 + ac_x + b \pmod N$.
   - Compute the Legendre symbols $u_p = \left(\frac{u}{p}\right)$ and $u_q = \left(\frac{u}{q}\right)$.
   - If $(u_p, u_q) = (1,1)$, then compute $d \equiv e^{-1} \pmod{\operatorname{lcm}(p+1-t_p, q+1-t_q)}$.
   - If $(u_p, u_q) = (1,-1)$, then compute $d \equiv e^{-1} \pmod{\operatorname{lcm}(p+1-t_p, q+1+t_q)}$.
   - If $(u_p, u_q) = (-1,1)$, then compute $d \equiv e^{-1} \pmod{\operatorname{lcm}(p+1+t_p, q+1-t_q)}$.
   - If $(u_p, u_q) = (-1,-1)$, then compute $d \equiv e^{-1} \pmod{\operatorname{lcm}(p+1+t_p, q+1+t_q)}$.
   - Compute m as the x-coordinate of $dC = deP = P = (m_x, m_y)$ on the elliptic curve $y^2 = x^3 + ax + b \pmod N$.

A variant of Demytko's scheme is to consider $d \equiv e^{-1} \pmod{(p+1 \pm t_p, q+1 \pm t_q)}$ instead of modulo $\operatorname{lcm}(p+1 \pm t_p, q+1 \pm t_q)$. Then e and d satisfy an

equation of the form

$$ed - k\,(p - s)\,(q - r) = 1, \quad s = \mp t_p - 1, \quad r = \mp t_q - 1.$$

This equation matches the RSA variant key equation that will be studied in this paper.

L.2.4 The Elliptic Curve Method

An integer m is said to be B-*smooth* if all the prime factors of m are less than or equal to B. Smooth numbers are used in cryptography by many factoring and discrete logarithm algorithms (see [84] and [85]). The counting function of B-smooth numbers in an interval $[1, x]$ is defined as

$$\psi(x, B) = \#\left\{m : 1 \leq m \leq x, m \text{ is } B\text{-smooth}\right\}.$$

In the particular case $x = B^u$, Hildebrand [60] gave the asymptotic formula $\psi(x, B) = x\rho(u)$ where $\rho(u)$ is the Dikman rho-function defined as the solution of the differential equation $u\rho'(u) = -\rho(u - 1)$ for $u \geq 1$ with the initial condition $\rho(u) = 1$ for $0 \leq u \leq 1$. For $1 \leq u \leq 2$, the Dikman function satisfies $\rho(u) = 1 - \log u$ so that $\psi(x, B) = x(1 - \log u)$. The Elliptic Curve method (ECM) is a probabilistic method for integer factorization and was discovered by H.W. Lenstra [84] in 1987. It is a fast partially factoring algorithm, especially for finding small prime factors p, in a heuristic running time $\mathcal{O}\left(\exp\left(c(\log p)^{1/2}\right)\left(\log\log p\right)^{1/2}\right)\right)$, for some constant $c > 0$. The ECM algorithm is based on the property of the Chinese Remainder Theorem, that is, for any elliptic curve $E(a, b)$, if $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$, then

$$E\left(\mathbb{Z}/n\mathbb{Z}\right) = E\left(\mathbb{Z}/p_1^{e_1}\mathbb{Z}\right) \times E\left(\mathbb{Z}/p_2^{e_2}\mathbb{Z}\right) \times \cdots \times E\left(\mathbb{Z}/p_k^{e_k}\mathbb{Z}\right).$$

Suppose that the order of $E\left(\mathbb{Z}/p_1^{e_1}\mathbb{Z}\right)$ is B-smooth and let m be a multiple of $\left|E\left(\mathbb{Z}/p_1^{e_1}\mathbb{Z}\right)\right|$, typically $m = \operatorname{lcm}(2, \ldots, B)$. Then, for every $P \in E\left(\mathbb{Z}/n\mathbb{Z}\right)$, we have $mP = (0 : 1 : 0) \pmod{p_1}$. Consequently, computing mP where $P \in E\left(\mathbb{Z}/n\mathbb{Z}\right)$, using the addition formulas on $E\left(\mathbb{Z}/n\mathbb{Z}\right)$, we must get $mP = (x : y : z) = (0 : 1 : 0) \pmod{p_1}$. This implies that $z \equiv 0 \pmod{p_1}$ and that $\gcd(z, n) = p_1^r$ for some positive integer r which will reveal p_1.

L.3 The Attack on RSA

In this section, we present an attack on RSA when the public key (N, e) satisfies an equation $eu-(p-s)(q-r)v = w$ with suitably small parameters u, v, w, r, s under the condition that one of the factors $(p-s)$ or $(q-r)$ is B-smooth for some ECM-efficiency bound B.

L.3.1 The attack

Theorem L.3.1. *Let $N = pq$ be an RSA modulus and $e = N^\beta$ be a public exponent. Suppose that e satisfies the equation $eu-(p-s)(q-r)v = w$ with $|r|, |s| < N^\alpha$, $u < N^\delta$ and $|w| < N^\gamma$. If*

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)} - \varepsilon,$$

then one can find the product $(p-s)(q-r)$ in polynomial time.

Proof. Suppose that $N = pq$ is an RSA modulus and e is a public exponent satisfying $eu-(p-s)(q-r)v = w$. Since $(p-s)(q-r) = N - pr - qs + rs$, then $-v(N - pr - qs + rs) - w \equiv 0 \pmod{e}$, which can be rewritten as $v(pr+qs-rs) - Nv - w \equiv 0 \pmod{e}$. Consider the polynomial $f(x, y, z) = xy - Nx + z$, Then $(x, y, z) = (v, pr + qs - rs, -w)$ is a solution of the modular polynomial equation $f(x, y, z) \equiv 0 \pmod{e}$. The small solutions of this modular equation can be found by applying Coppersmith's method [34]. Let m and t be two positive integers. Consider the polynomials

$$\begin{aligned} G_{k,i_1,i_2,i_3}(x, y, z) &= x^{i_1-k} z^{i_3} f(x, y, z)^k e^{m-k}, \\ &\qquad \text{for} \quad k = 0, \ldots m,\ i_1 = k, \ldots, m,\ i_2 = k,\ i_3 = m - i_1, \\ H_{k,i_1,i_2,i_3}(x, y, z) &= y^{i_2-k} z^{i_3} f(x, y, z)^k e^{m-k}, \\ &\qquad \text{for} \quad k = 0, \ldots m,\ i_1 = k,\ i_2 = k+1, \ldots, i_1 + t,\ i_3 = m - i_1. \end{aligned}$$

Let $\mathcal{L}$ denote the lattice spanned by the coefficient vectors of the polynomials $G_{k,i_1,i_2,i_3}(Xx, Yy, Zz)$ and $H_{k,i_1,i_2,i_3}(Xx, Yy, Zz)$. We can get a left triangular matrix if the ordering of the rows follows the ordering of the k's and the ordering of the the monomials of a polynomial follows the natural ordering following the i_1's, then the i_2's, then the i_3's. Hence, using the triangular form

of the matrix, the determinant of $\mathcal{L}$ is in the form $\det(\mathcal{L}) = e^{n_e} X^{n_X} Y^{n_Y} Z^{n_Z}$. For $m = 2$ and $t = 1$, the coefficient matrix for $\mathcal{L}$ is presented in Table L.1. The non-zero elements are marked with an '$\circledast$'.

	z^3	xz^2	x^2z	x^3	xyz^2	x^2yz	x^3y	x^2y^2z	x^3y^2	x^3y^3	xy^2z^2	x^2y^3z	x^2yz	x^3y^4
G_{k,i_1,i_2,i_3}														
$G_{0,0,0,3}$	Z^3e^3	0	0	0	0	0	0	0	0	0	0	0	0	0
$G_{0,1,0,2}$	0	XZ^2e^3	0	0	0	0	0	0	0	0	0	0	0	0
$G_{0,2,0,1}$	0	0	X^2Ze^3	0	0	0	0	0	0	0	0	0	0	0
$G_{0,3,0,0}$	0	0	0	X^3	0	0	0	0	0	0	0	0	0	0
$G_{1,1,1,2}$	⊛	0	0	0	XYZ^2e^2	0	0	0	0	0	0	0	0	0
$G_{1,2,1,1}$	0	⊛	⊛	0	0	X^2YZe^2	0	0	0	0	0	0	0	0
$G_{1,3,1,0}$	0	0	⊛	⊛	0	0	X^3Ye^2	0	0	0	0	0	0	0
$G_{2,2,2,1}$	⊛	⊛	⊛	0	⊛	⊛	0	X^2Y^2Ze	0	0	0	0	0	0
$G_{2,3,2,0}$	0	⊛	⊛	⊛	0	⊛	⊛	0	X^3Y^2e	0	0	0	0	0
$G_{3,3,3,0}$	⊛	0	⊛	⊛	⊛	⊛	⊛	⊛	⊛	X^3Y^3	0	0	0	0
H_{k,i_1,i_2,i_3}														
$H_{0,0,1,3}$	0	0	0	0	⊛	0	0	0	0	⊛	$XY^2Z^2e^2$	0	0	0
$H_{1,1,2,2}$	0	0	0	0	⊛	⊛	0	⊛	0	0	⊛	X^2Y^3Ze	0	0
$H_{2,2,3,1}$	0	0	0	⊛	0	0	⊛	⊛	0	0	0	0	X^2YZe	0
$H_{3,3,4,0}$	0	0	0	0	⊛	0	0	⊛	⊛	0	0	0	0	X^3Y^4

Table L.1: The coefficient matrix for the case $m = 2$, $t = 1$.

To find the values of the exponents, define $S(x)$ to be

$$S(x) = \sum_{k=0}^{m} \sum_{i_1=k}^{m} \sum_{i_2=k}^{k} \sum_{i_3=m-i_1}^{m-i_1} x + \sum_{k=0}^{m} \sum_{i_1=k}^{k} \sum_{i_2=k+1}^{i_1+t} \sum_{i_3=m-i_1}^{m-i_1} x.$$

Using the construction of the polynomials G and H, we get

$$\begin{aligned} n_e &= S(m-k) = \frac{1}{6}m(m+1)(2m+3t+4), \\ n_X &= S(i_1) = \frac{1}{6}m(m+1)(2m+3t+4), \\ n_Y &= S(i_2) = \frac{1}{6}(m+1)\left(m^2+3mt+3t^2+2m+3t\right), \\ n_Z &= S(i_3) = \frac{1}{6}m(m+1)(m+3t+2), \\ \omega &= S(1) = \frac{1}{2}(m+1)(m+2t+2). \end{aligned} \tag{L.1}$$

Let $t = \tau m$ for some positive τ to be optimized later. The dominant terms of the exponents in (L.1) are

$$\begin{aligned} n_e &\approx \frac{1}{6}(3\tau+2)m^3 + o(m^3), \\ n_X &\approx \frac{1}{6}(3\tau+2)m^3 + o(m^3), \\ n_Y &\approx \frac{1}{6}\left(3\tau^2+3\tau+1\right)m^3 + o(m^3), \\ n_Z &\approx \frac{1}{6}(3\tau+1)m^3 + o(m^3), \\ w &\approx \frac{1}{6}(6\tau+3)m^2 + o(m^2). \end{aligned} \tag{L.2}$$

Applying the LLL algorithm L.2.4 to the lattice $\mathcal{L}$, we get a reduced basis where the three first vectors h_i, $i = 1, 2, 3$ satisfy

$$\|h_1\| \leq \|h_2\| \leq \|h_3\| \leq 2^{\frac{\omega(\omega-1)}{4(\omega-2)}} \det(\mathcal{L})^{\frac{1}{\omega-2}}.$$

To apply Howgrave-Graham's Theorem L.2.3 to h_1, h_2 and h_3, we set

$$2^{\frac{\omega(\omega-1)}{4(\omega-2)}} \det(\mathcal{L})^{\frac{1}{\omega-2}} < \frac{e^m}{\sqrt{\omega}}.$$

This can be transformed to

$$\det(\mathcal{L}) < 2^{-\frac{\omega(\omega-1)}{4}} \frac{1}{(\sqrt{\omega})^{\omega-2}} e^{m(\omega-2)} < e^{m\omega},$$

or equivalently $e^{n_e} X^{n_X} Y^{n_Y} Z^{n_Z} < e^{m\omega}$. Suppose that $e = N^\beta$, $u < N^\delta$, $|w| < N^\gamma$ and $\max(|r|, |s|) < N^\alpha < N^{\frac{1}{4}}$. Since $q < p < \sqrt{2}\sqrt{N}$ by Lemma L.2.1, then

$$|pr + qs - rs| < 2\max(p|r|, q|s|) < 2\sqrt{2}\sqrt{N} \cdot N^\alpha = 2\sqrt{2}N^{\frac{1}{2}+\alpha}.$$

On the other hand, since $(p-s)(q-r) \approx N$ and $|w| < eu$, we get

$$v = \frac{eu - w}{(p-s)(q-r)} < \frac{eu + |w|}{(p-s)(q-r)} < \frac{2eu}{N} < 2N^{\beta+\delta-1}, \tag{L.3}$$

Let $X = 2N^{\beta+\delta-1}$, $Y = 2\sqrt{2}N^{\frac{1}{2}+\alpha}$ and $Z = N^\gamma$. Then the target solution (x, y, z) satisfies $|x| < X$, $|y| < Y$ and $|z| < Z$. Using the approximations of n_e, n_X, n_Y, n_Z and ω given in (L.2), the inequality $e^{n_e} X^{n_X} Y^{n_Y} Z^{n_Z} < e^{m\omega}$ can be transformed to

$$(3\tau+2)\beta + (3\tau+2)(\beta+\delta-1) + \left(3\tau^2+3\tau+1\right)\left(\frac{1}{2}+\alpha\right) + (3\tau+1)\gamma < (6\tau+3)\beta.$$

The optimal value for τ is

$$\tau_0 = \frac{1 - 2\delta - 2\alpha - 2\gamma}{2(1+2\alpha)},$$

and, plugging this value in the former inequality, we obtain

$$4\alpha^2 + 16\alpha\beta + 8\alpha\delta - 8\alpha\gamma - 12\delta^2 - 24\delta\gamma - 12\gamma^2 - 4\alpha + 8\beta + 28\delta + 20\gamma < 15,$$

and consequently

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)}.$$

Under this condition, the LLL algorithm applied to the lattice $\mathcal{L}$ outputs three vectors v_i, $i = 1, 2, 3$. These vectors represent the coefficients of three polynomials $h_i(Xx, Yy, Zz)$, $i = 1, 2, 3$ sharing the root $(x, y, z) = (v, pr + qs + rs, -w)$. Then, applying Gröbner basis computations, we get the expected solution, from which we deduce $(p-s)(q-r) = N - (pr+qs+rs)$. This terminates the proof. □

Remark L.3.2. If $r = s = w = 1$, then the equation $eu-(p-s)(q-r)v = w$ is the classical RSA key equation $ed-(p-1)(q-1)k = 1$ with $d < N^{\delta}$. Using $\alpha = 0$, $\beta = 1$ and $\gamma = 0$, the bound of Theorem L.3.1 gives $\delta < \frac{7}{6} - \frac{\sqrt{7}}{3}$. This retrieves the classical bound on the private exponent d (see [17]).

Theorem L.3.3. *Let $N = pq$ be an RSA modulus and $e = N^{\beta}$ be a public exponent. Suppose that e satisfies the equation $eu-(p-s)(q-r)v = w$ with $|r|, |s| < N^{\alpha} < N^{\frac{1}{4}}$, $u < N^{\delta}$ and $|w| < N^{\gamma}$. Let B be an ECM-efficiency bound for the Elliptic Curve Method. If $(p-s)$ or $(q-r)$ is B-smooth and*

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)},$$

then one can find p and q in polynomial time.

Proof. Suppose that, in the equation $eu-(p-s)(q-r)v = w$, the parameters satisfy $|r|, |s| < N^{\alpha} < N^{\frac{1}{4}}$, $e = N^{\beta}$, $u < N^{\delta}$, $|w| < N^{\gamma}$ and that the exponent parameters satisfy $\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha+1)(2\alpha+6\beta-6\gamma+1)}$. Then, by applying Theorem L.3.1, we can find the exact value of $(p-s)(q-r)$. Next, suppose that $(p-s)$ is B-smooth where B is a bound for the efficiency of the Elliptic Curve Method (ECM). Hence, ECM will reveal a partial factorization of $(p-s)(q-r)$ as

$$(p-s)(q-r) = M \cdot \prod_{i=1}^{\omega((p-s)(q-r))} p_i^{e_i},$$

were $\omega(p-s)$ is the number of distinct prime factors of $p-s$ and M is such that $M = 1$ or all prime factors of M are greater than B. The average order of the number of prime factors of an integer n is $\omega(n) \approx \frac{\log n}{\log\log n}$ (see [57], pp. 355). Since $|r|, |s| < N^{\alpha}$ and $\sqrt{N} < p < \sqrt{2N}$, then

$$\sqrt{N} - N^{\alpha} < p - s < \sqrt{2N} + N^{\alpha}. \tag{L.4}$$

Hence, the average number of the prime factors of $p-s$ satisfies

$$\omega(p-s) \approx \frac{\log(p-s)}{\log\log(p-s)} \approx \frac{\log N}{2\log\log N}.$$

On the other hand, according to the factorization

$$(p-s) = \prod_{i=1}^{\omega((p-s))} p_i^{e_i},$$

the number of distinct divisors of $p-s$ is exactly $\prod_{i=1}^{\omega(p-s)}(e_i+1)$. However, the average number of divisors of an integer n is $\log n$ (see Theorem 319 of [57]). Hence, the average number of divisors of $p-s$ is approximately $\log(p-s) \approx \frac{1}{2}\log N$. Let d be a divisor of $(p-s)(q-r)$ such that $d = p-s$. Then

$$d = \prod_{i=1}^{\omega(p-s)} p_i^{x_i}, \quad 0 \leq x_i \leq e_i.$$

Using (L.4), we get

$$\log\left(\sqrt{N} - N^{\alpha}\right) < \sum_{i=1}^{\omega(p-s)} x_i \log p_i < \log\left(\sqrt{2N} + N^{\alpha}\right).$$

The former inequalities can be solved by applying linear programming algorithms such as PSLQ [47] and LLL [86], and using a solution $(x_1, \ldots, x_{\omega(p-s)})$, we compute $d = \prod_{i=1}^{\omega(p-s)} p_i^{x_i}$ which is then a candidate for $p-s$. Since $|s| < N^{\alpha} < N^{\frac{1}{4}}$, then d is an approximation of the prime factor p of N with an error term less than $N^{\frac{1}{4}}$. Hence, using Theorem L.2.2, this leads to the exact value of p if d is the good candidate. Repeating this process sequentially for the factors d of $(p-s)(q-r)$ in the range $\sqrt{N} - N^{\alpha} < d < \sqrt{2N} + N^{\alpha}$, we will find p and then get $q = \frac{N}{p}$. This achieves the factorization of the RSA modulus. □

L.3.2 A numerical example

Consider the following RSA 265 bit-size modulus N with the public exponent e,

N =43115265506687226436196728756959707266402158394261294759458139340520129183826747,

e =44291096833783216353731643543595440193954966593379368311328977066819711783511 39.

Suppose that $N = pq$ with unknown factorization and e satisfies an equation $eu-(p-s)(q-r)v = w$ with the suitably small unknown parameters u, v, w, r and s. Then applying the method of Theorem L.3.1 to solve the equation $eu - (p - s)(q - r)v = w$, with the bounds

$$u < N^{\delta} = N^{0.15},\ |w| < N^{\gamma} = N^{0.15},\ |r|, |s| < N^{\alpha} = N^{0.15},\ e = N^{\beta} = N^{0.987},$$

we get

$$\begin{aligned} v =& 8330878683394 \\ w =& 2516643, \\ ps + qr - rs =& 456241034994533467152256390448296889414536571 47, \end{aligned}$$

Since $(p - s)(q - r) = N - (pr + qs - rs)$, we get

$$\begin{aligned} (p-s)(q-r) =& 43115265506687226436196728756959661642298658940914579 53 \\ & 38190945108311877301696 00. \end{aligned}$$

Then, using the Elliptic Curve Method with the bound $B = N^{\frac{1}{10}} \approx 91931238$, we get the factorization

$$\begin{aligned} (p-s)(q-r) = & 2^8 \cdot 3 \cdot 5^2 \cdot 13 \cdot 23 \cdot 53 \cdot 89 \cdot 181 \cdot 1663 \cdot 2833 \cdot 2969 \cdot 5197 \cdot 5233 \cdot \\ & 6481 \cdot 12007 \cdot 18439 \cdot 36973 \cdot 435876180528100336114933071348569. \end{aligned}$$

Using the factorization of $(p - s)(q - r)$, we can find the set of the factors d such that $\sqrt{N} - N^{\alpha} < d < \sqrt{2N} + N^{\alpha}$. Such divisors are candidate for $p - s$, that is we $p - s = d$ for one of these factors. Then by applying Coppersmith's Theorem L.2.2, we can find p using the correct candidate. For the divisor $d = 66722240146623401785797214743267281856 00$, we apply Coppersmith's Theorem L.2.2 and find

$$p = 66722240146623401785797214743267341527 49.$$

Then $q = \frac{N}{p} = 64619031693091544838337970117858865065 03$.

L.4 Application to Demytko's Scheme

In this section, we show how to apply the technique of Theorem L.3.1 to break the Demytko scheme in some situations and provide a numerical example.

L.4.1 The attack on Demytko's Scheme

In Demytko's scheme, the RSA modulus is $N = pq$ and the elliptic curve $E_N(a, b)$ is such that $\#E_p(a, b) = p + 1 - t_p$ and $\#E_q(a, b) = q + 1 - t_q$ where, according to Hasse Theorem, $|t_p| < 2\sqrt{p}$ and $|t_q| < 2\sqrt{q}$. Also, the public exponent e and the private exponent d satisfy one of the four equations

$$eu - (p + 1 \pm t_p)(q + 1 \pm t_q)v = w.$$

These equations can be transformed into one of the form $eu-(p-s)(q-r)v = w$ where $s = \mp t_p - 1$ and $t = \mp t_q - 1$, which can be studied using the technique of Theorem L.3.1.

Corollary L.4.1. *Let (N, e, a, b) the public parameters of a Demytko's instance where $N = pq$. Suppose that $e = N^\beta$ satisfies an equation of the form $eu - (p + 1 \pm t_p)(q + 1 \pm t_q)v = w$ with $| \pm t_p - 1|, | \pm t_q - 1| < N^\alpha < N^{\frac{1}{4}}$, $u < N^\delta$ and $|w| < N^\gamma$. Let B be an ECM-efficiency bound for the Elliptic Curve Method. If $p + 1 \pm t_p$ or $q + 1 \pm t_q$ is B-smooth and*

$$\delta < \frac{7}{6} + \frac{1}{3}\alpha - \gamma - \frac{1}{3}\sqrt{(2\alpha + 1)(2\alpha + 6\beta - 6\gamma + 1)},$$

then one can find p and q in polynomial time.

Proof. Since the equation $eu-(p+1\pm t_p)(q+1\pm t_q)v = w$ can be transformed into $eu - (p - s)(q - r)v = w$ with $s = \mp t_p - 1$ and $t = \mp t_q - 1$, then this equation can be solved under the conditions of Theorem L.3.1 when $|t_p - 1| < N^\alpha$ and $|t_q - 1| < N^\alpha$. □

L.4.2 A numerical example

Example L.4.2. Let us consider the Demytko public parameters (N, e, a, b) where N is an 510-bit RSA modulus

$$\begin{aligned}
N =&244564152049718837289391032953867582433145492152016390042656239363441852689757568224991629341622126967445954070062427486023623868460973836075181541009161 7,\\
e =&207753540686843587408555602893982678168821441852165899252123932416370824148707563812033872059010473801740084336709522813588017197501164099322578137710783,\\
a =&0,\\
b =&9,
\end{aligned}$$

with the elliptic curve $E_N(a, b)$ with equation $y^2 \equiv x^3 + 9 \pmod N$. We suppose that e satisfies the equation $eu - (p + 1 \pm t_p)(q + 1 \pm t_q)v = w$ with $t_p, t_q < N^\alpha = N^{0.1}$. Then applying the method of Theorem L.3.1 to solve the equation $eu - (p - s)(q - r)v = w$ where $s = \mp t_p - 1$ and $r = \mp t_p - 1$, we get for $e = N^\beta \approx N$, $u < N^\delta = N^{0.1}$, $|w| < N^\gamma = N^{0.1}$

$$\begin{aligned}
v =&6889077569105,\\
w =&2916646,\\
pr + qs - rs =&784357999339618220094311636350013903165826707133763 3,\\
&244222164466922717093026565590439040792,
\end{aligned}$$

Then

$$\begin{aligned}
&N - (pr + qs - rs) = (p - s)(q - r)\\
&= 244564152049718837289391032953867582433145492152016\\
&\quad 390042656238579083853350139348130679992991608223801 6\\
&\quad 192469362991030638071771761892645334186224971050825.
\end{aligned}$$

Applying the Elliptic Curve Method for factorization with the bound

$B = 2^{80} \approx N^{0.16}$, we get the factorization

$$\begin{aligned}(p-s)(q-r) =& 3^6 \cdot 5^2 \cdot 7^2 \cdot 13^3 \cdot 43^2 \cdot 103^2 \cdot 277 \cdot 674^2 \cdot 1021 \cdot 4177 \cdot 15061 \\ & \cdot 21737^2 \cdot 27109^2 \cdot 52291^2 \cdot 84991 \cdot 90841 \cdot 132661 \cdot 347329^2 \\ & \cdot 3834631 \cdot 29327821 \cdot 69689551 \cdot 30404961633073956301 \\ & \cdot 305196537135675591605491.\end{aligned}$$

Any divisor d of $(p-s)(q-r)$ is a candidate for $p-s$ or $q-r$. Using the divisor

$$\begin{aligned}d =& 3^3 \cdot 13^2 \cdot 277 \cdot 1021 \cdot 15061 \cdot 21737^2 \cdot 27109^2 \cdot 52291^2 \cdot 90841 \\ & \cdot 305196537135675591605491,\end{aligned}$$

as a candidate for $p-s$ in Coppersmith's Theorem L.2.2, we get p and then $q = \frac{N}{p}$ as follows

$$\begin{aligned}p =& 685920425598306143251778583414905266471238279458502857598279318189925533951 71, \\ q =& 356548869114654893865594787391255957316985729824840925802871758605570764820 27,\end{aligned}$$

which completes the factorization of N.

L.5 Conclusion

In this paper, we consider an instance of RSA where the public exponent satisfies a generalized key equation with many unknown parameters. Under suitable conditions, we combine Coppersmith's method and the Elliptic Curve Method for factorization ECM, we solve the equation and find the prime factors of the RSA modulus. We apply the same technique to launch an attack on Demytko's Elliptic Curve Cryptosystem when the secret parameters are suitably small.

Appendix M

Lattice Attacks on the DGHV Homomorphic Encryption Scheme

Mathematics in Computer Science, 2016
[122] with Tajjeeddine Rachidi

Abstract :
In 2010, van Dijk, Gentry, Halevi, and Vaikuntanathan described the first fully homomorphic encryption over the integers, called DGHV. The scheme is based on a set of m public integers $c_i = pq_i + r_i$, $i = 1, \cdots, m$, where the integers p, q_i and r_i are secret. In this paper, we describe two lattice-based attacks on DGHV. The first attack is applicable when $r_1 = 0$ and the public integers c_i satisfy a linear equation $a_2c_2 + \ldots + a_mc_m = a_1q_1$ for suitably small integers a_i, $i = 2, \ldots, m$. The second attack works when the positive integers q_i satisfy a linear equation $a_1q_1 + \ldots + a_mq_m = 0$ for suitably small integers a_i, $i = 1, \ldots, m$. We further apply our methods for the DGHV recommended parameters as specified in the original work of van Dijk, Gentry, Halevi, and Vaikuntanathan.

M.1 Introduction

In the last ten years, cloud computing has gained major importance and widespread. Yet, a very important concern of cloud computing remains the security and privacy of data. A useful solution to this concern is the use of fully homomorphic encryption (FHE) to encrypt data stored remotely. Indeed, a fully homomorphic encryption scheme supports the computation of arbitrary functions on encrypted data, possibly distributed across the cloud, without the need to resort to decryption. Unfortunately, not all encryption schemes are fully homomorphic. For example, RSA [131] is only multiplicatively homomorphic: given two ciphertexts $c_1 \equiv m_1^e \pmod N$ and $c_2 \equiv m_2^e \pmod N$, one can compute the encrypted form of $m_1 m_2$, that is $(m_1 m_2)^e \pmod N$, without having to recover the plaintexts m_1 and m_2, simply by applying $c_1 c_2 \equiv m_1^e m_2^e \equiv (m_1 m_2)^e \pmod N$. Similarly, ElGamal [42] is multiplicatively homomorphic. By contrast, Paillier [124] is additively homomorphic: given two ciphertexts $c_1 = g^{m_1} r_1^N \pmod{N^2}$ and $c_2 = g^{m_2} r_2^N \pmod{N^2}$, one can perform $c_1 c_2 = g^{m_1+m_2}(r_1 r_2)^N \pmod{N^2}$, which gives $m_1 + m_2$ without having to resort to the decryption of the ciphertexts c_1 and c_2. Another example of an additively homomorphic scheme is the Goldwasser-Micali scheme [52].

In 2009, Gentry [49], presented the first construction of a FHE scheme. Gentry's scheme supports both addition and multiplication on ciphertexts and consists of three main steps. The first step constructs a *somewhat* homomorphic scheme, which is limited to evaluating low-degree polynomials over encrypted data. The second step slightly modifies the *somewhat* homomorphic scheme to make it bootstrappable, i.e., capable of evaluating its own decryption circuit (operations). The third step transforms the bootstrappable *somewhat* homomorphic encryption scheme into a fully homomorphic encryption through a recursive self-embedding. The security of Gentry's scheme has been determined to be based on the worst-case hardness of solving specific problems in an ideal lattice, namely the shortest independent vector problem (SIVP) over ideal lattices in the worst-case (see [50]).

A key disadvantage of Gentry's scheme, however, is its computational inefficiency. Therefore, much effort has been made by the research community

to find alternative efficient FHE schemes. In 2010, van Dijk, Gentry, Halevi and Vaikuntanathan [41] presented DGHV, a computationally efficient FHE scheme over the integers. This scheme is based on a set of public integers, $c_i = pq_i + r_i$, $i = 1, \ldots, m$, where the parameters p, q_i and r_i are secrets with the following size constraints:

- p is a prime number.
- η is the bit-length of the secret key p.
- ρ is the bit-length of the secret noises r_i.
- γ is the bit-length of the public integers c_i.

In [36, 41, 87], the security of DGHV has been studied against several attacks, which served the purpose of improving its security by defining optimal bounds for its parameter bit size (η, ρ, and γ). As reported in [87], these attacks can be categorized according to their underling techniques:

- **Brute force search** [30, 41]: When $c_1 = pq_1$, this technique consists in removing the noise, say r_2 from c_2 by trying all possibilities for $r_2 \in (-2^\rho, 2^\rho)$ and computing $\gcd(c_1, c_2 - r_2)$ which gives p with overwhelming probability.
- **Continued fractions** [41, 87]: This consists on recovering q_i/q_j from c_i/c_j using continued fractions, which yields immediate calculation of $p = \lfloor c_i/q_i \rceil$.
- **Attacks on the Approximate-GCD assumption** [41, 87]: The recovery of p through the recovery of r_i or q_i, $i = 1, \ldots, m$, using a combination of lattice reduction and other techniques. These attacks include Coppersmith's technique [34], the method for solving simultaneous diophantine equations [86] and the orthogonal lattice attacks [41, 87] (See Section M.3 for more on these attacks).

Yet, a more direct way to break the DGHV scheme when $r_1 = 0$ consists in finding p and q_1 by factoring $c_1 = pq_1$. To date, the most efficient known methods to factor c_1 are the Number Field Sieve (NFS) [85] and the Elliptic

Curve Method (ECM) [84]. As shown in [87] (p. 82, Table 7.1), the DGHV factorization problem of c_1 is considered as untractable if $p > 2^{261}$ and $c_1 > 2^{2911}$.

M.1.1 Our Contribution

In this paper, we propose two new attacks on the DGHV scheme. The starting point of both attacks are the existance of two linear equations involving the public integers c_i for $i = 2, \ldots, m$ and the secret parameter q_1 on the one hand, and the secret parameters q_i, $i = 1, \ldots, m$ on the other.

In the first attack, we suppose that $c_1 = pq_1$ and that $c_i = pq_i + r_i$ with $r_i \neq 0$ for $i = 2, \ldots, m$. To avoid factoring attacks on c_1, we also suppose that q_1 is prime. If q_1 is not coprime with one of the integers c_i for $2 \leq i \leq m$, then $\gcd(c_1, c_i) = q_1$ which will reveal $p = \frac{c_1}{q_1}$. Hence, q_1 is coprime c_i for $i = 2, \ldots, m$. Therefore for any integers $a_2, \ldots, a_{m-1}$, the integer $a_m \equiv -(a_2c_2 + \ldots + a_{m-1}c_{m-1})(c_m)^{-1} \pmod{q_1}$ exists and satisfies the linear integer relation $a_2c_2 + \ldots + a_mc_m = a_1q_1$ for an integer a_1. We will leverage this relationship and show that one can find the DGHV parameters p, q_i and r_i in polynomial time if the coefficients a_i, $i = 1, \cdots, m$ are suitably small. The attack uses Coppersmith's method for solving multivariate linear modular equations, as presented by Herrmann and May in [59].

In the second attack, we suppose that $c_i = pq_i + r_i$ for $i = 1, \ldots, m$. Let $G = \gcd(q_1, \ldots, q_m)$. Then $\frac{q_{m-1}}{G}$ is coprime with one $\frac{q_i}{G}$, $i \neq m-1$. Assume that $\frac{q_{m-1}}{G}$ is coprime with $\frac{q_m}{G}$. Let $a_1, \ldots, a_{m-2}$ be arbitrary integers. Define

$$a_{m-1} \equiv -\left(a_1\frac{q_1}{G} + \ldots + a_{m-2}\frac{q_{m-2}}{G}\right)\left(\frac{q_{m-1}}{G}\right)^{-1} \pmod{\frac{q_m}{G}}.$$

Then there exists an integer a_m such that

$$a_1\frac{q_1}{G} + \ldots + a_{m-2}\frac{q_{m-2}}{G} + a_{m-1}\frac{q_{m-1}}{G} + a_m\frac{q_m}{G} = 0,$$

or equivalently $a_1q_1 + \ldots + a_mq_m = 0$. This shows that the integers $q_1, \ldots, q_m$ are linked by infinitely many linear integer relations. We exploit this relation, and show that if the coefficients a_i, $i = 1, \ldots, m$ are sufficiently small, then one can efficiently find all the DGHV parameters. Unlike the first attack,

this attack is based solely on lattice reduction techniques, namely the LLL algorithm [86].

For both attacks, we carry out experiments to verify the validity and the effectiveness of our methods. We also define the new bounds for DGVH secret parameters that resist our attacks, effectively improving on previously proposed optimal bounds [41], [87].

M.1.2 Organization

The rest of this paper is organized as follows: In Section 2, we briefly review the preliminaries necessary for both our attacks. Section 3 is dedicated to leading attacks on the DGHV scheme. In Section 4, we present our first lattice-based attack on DGHV, that is when $r_1 = 0$ and the numbers c_i satisfy a linear equation $a_2c_2+\ldots+a_mc_m = a_1q_1$ for suitably small integers a_i, $i = 2, \ldots, m$. In section 5, we present our second lattice-based attack, which is applicable when the integers q_i satisfy a linear equation $a_1q_1+\ldots+a_mq_m = 0$ for suitably small integers a_i. We then conclude the paper in Section 6.

M.2 Preliminaries

In this section, we review the DGHV scheme parameters and the Approximate-GCD assumption upon which its security is based. We also recall Coppersmith's method for solving linear diophantine equations, and review the lattice reduction technique used in our new attacks on the DGHV scheme.

M.2.1 The DGHV Scheme over the Integers

In 2010, van Dijk, Gentry, Halevi and Vaikuntanathan [41] proposed a fully homomorphic encryption scheme based on m public integers $c_i = pq_i + r_i$ where the secret parameters p, q_i, r_i are such that:

- For $i = 1, \ldots, m$, c_i is a public integer of bit-length γ.
- p is a private prime number of bit-length η.

- For $i = 1, \ldots, m$, q_i is a private integer of bit-length $\gamma - \eta$.
- For $i = 1, \ldots, m$, r_i is a private random integer with $|r_i| < 2^\rho$.

In [41], it is shown that the scheme is semantically secure under the Approximate-GCD assumption which states the following:

Definition M.2.1 (Approximate-GCD assumption)**.** Let γ, η, ρ be positive integers. For any η-bit prime number p, given m many positive integers $c_i = pq_i + r_i$ with m many $(\gamma - \eta)$-bit integers q_i and m many integers r_i satisfying $|r_i| < 2^\rho$, it is hard to find p.

The hardness of the Approximate-GCD assumption has been studied by Howgrave-Graham [65], and used in the study of the security of the DHGV scheme in [41] and [87], leading to the establishment of typical integer sizes that guarantee high security levels of DHGV. Therein, the values $\rho \approx \sqrt{\eta}$, $\gamma = \eta^3 + \eta$ are considered secure (see [41]).

M.2.2 Lattice reduction

Here we 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 lattice $\mathcal{L}$ spanned by $b_1 \ldots, b_d$ is the set of all integer linear combination $x_1 b_1 + \ldots + x_d b_d$ of $b_1 \ldots, b_d$ with $x_1, \ldots, x_d \in \mathbb{Z}$. The set of vectors $(b_1 \ldots, b_d)$ is called a basis of $\mathcal{L}$ and d is its dimension. If B is the matrix of $b_1 \ldots, b_d$ in the canonical basis of $\mathbb{R}^n$, then the determinant of $\mathcal{L}$ is $\det(\mathcal{L}) = \sqrt{B^t B}$, and the Euclidean norm of a vector $v \in \mathcal{L}$ is defined using the scalar product $\|v\| = \sqrt{v \cdot v}$.

Of interest to many applications and algorithms is the shortest non-zero vector in a lattice. Finding the shortest non-zero vector is a computationally hard problem known as the *Shortest Vector Problem (SVP)* that guarantees the security of many cryptographic schemes. However, Minkowski's theorem, which dates back to 1889, guarantees the existence of short vectors, i.e., non-zero vectors whose length is not too large as in the following theorem.

Theorem M.2.2 (Minkowski)**.** *Let $\mathcal{L}$ be a lattice. Then there exists a non-zero vector $v \in \mathcal{L}$ such that*

$$\|v\| \leq \sqrt{\dim(\mathcal{L})}\, \det(\mathcal{L})^{\frac{1}{\dim(\mathcal{L})}}.$$

Given a lattice $\mathcal{L}$ and its original basis $b_1 \ldots, b_d$, lattice reduction consists in finding another basis, where a short non-zero vector is easily determined. This can be achieved through different algorithms, whose running time is usually at least exponential in the dimension of the lattice d. However, the LLL algorithm of Lenstra, Lenstra, and Lovász [86] can find, in polynomial time, short non-zero vectors in a lattice with reasonable dimension.

Theorem M.2.3 (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 L satisfying*

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

in polynomial time.

Thus, finding a reduced basis using LLL leads to finding reasonably short vectors in polynomial time.

M.2.3 Coppersmith's method for solving linear diophantine equations

The LLL algorithm has many applications in cryptography, including solving diophantine equations. Using the LLL algorithm, Coppersmith [34] derived a method for finding small roots of univariate modular equations and bivariate equations. This strategy is know as Coppersmith's technique and has been heuristically generalized for finding small roots of multivariate linear equations. The following result by Herrmann and May [59] gives a sufficient condition under which small roots of a modular linear equation can be found in polynomial time.

Theorem M.2.4 (Herrmann-May)**.** *Let N be a composite integer of unknown factorization with a divisor $p \geq N^\beta$. Let $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ be a linear polynomial in n variables. One can find in polynomial time all solutions $\left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ of the equation $f(x_1, \ldots, x_n) \equiv 0 \pmod{p}$ with $\left|x_1^{(0)}\right| < N^{\lambda_1}, \ldots, \left|x_n^{(0)}\right| < N^{\lambda_n}$ if*

$$\sum_{i=1}^{n} \lambda_i < 1 - (1-\beta)^{\frac{n+1}{n}} - (n+1)\left(1 - \sqrt[n]{1-\beta}\right)(1-\beta).$$

M.3 Former attacks on the DGHV Scheme

We recall here the main existing attacks on the DGHV scheme. For more details, we refer to [41] and [87]. Assume that we have an instance of DGHV with $c_1 = pq_1$ and $c_i = pq_i + r_i$, $i = 2, \ldots, m$ where the parameters p, q_i and r_i are secret. The task is to recover p. We recall that p is a η-bit prime and $0 < r_i < 2^\rho$.

M.3.1 Brute force on the remainder

A simple way to recover p is to remove the noise, say from c_2, by finding r_2, and then compute $p = \gcd(c_1, c_2 - r_2)$. This can be achieved by trying all integers r_2 with $0 < |r_2| < 2^\rho$. The complexity of this attack is obviously $\mathcal{O}(2^\rho)$. However, applying the method of Chen and Nguyen [30], one can find p with complexity $\mathcal{O}\left(2^{\rho/2}\right)$. As a consequence, removing the noise to recover p does not work in practice when ρ is sufficiently large.

M.3.2 Continued fractions

Using $c_1 = pq_1$ and $c_2 = pq_2 + r_2$, and given that q_1 and q_2 are prime numbers, one gets

$$\left|\frac{c_2}{c_1} - \frac{q_2}{q_1}\right| = \frac{|r_2|}{c_1}.$$

To recover $\frac{q_2}{q_1}$ as a convergent of the continued fraction expansion of $\frac{c_2}{c_1}$, we need $\frac{|r_2|}{c_1} < \frac{1}{2q_1^2}$, that is $2q_1|r_2| < p$. This is not possible if q_1 is much larger than p as for the recommended values for the DGHV parameters where q_1 is η^3-bit size while p is η-bit size.

M.3.3 Simultaneous Diophantine approximation

In [86], it is shown that the LLL algorithm can find a solution for the simultaneous diophantine approximations. That is, given n rational numbers $\alpha_1, \cdots, \alpha_n$ and ε with $0 < \varepsilon < 1$, one can efficiently find integers $p_1, \ldots, p_n$,

and q such that, for $i = 1, \ldots, n$,

$$|q\alpha_i - p_i| \leq \varepsilon, \quad \text{and} \quad 1 \leq q \leq 2^{\frac{n(n+1)}{4}} \varepsilon^{-n}.$$

This can be applied to the DHGV scheme. Using $c_1 = pq_1$ and $c_i = pq_i + r_i$, we get for $i = 2, \ldots, m$

$$\left| q_1 \frac{c_i}{c_1} - q_i \right| = \frac{|r_i|}{p} < 2^{\rho - \eta}.$$

This gives $m - 1$ simultaneous diophantine approximations which can be solved by applying the LLL algorithm [86] to reduce a basis of a lattice of dimension m. The LLL algorithm will succeed under the condition:

$$q_1 \leq 2^{\frac{(m-1)m}{4}} \cdot 2^{-(\rho-\eta)(m-1)} = 2^{\frac{(m-1)m}{4} + (\eta-\rho)(m-1)}.$$

Since $q_1 \approx 2^{\gamma-\eta}$, then $\gamma - \eta \leq \frac{m(m-1)}{4} + (\eta - \rho)(m-1)$, which can be achieved if

$$m > -2\eta + 2\rho + \frac{1}{2} + \frac{1}{2}\sqrt{16\eta^2 - 32\eta\rho + 16\rho^2 + 16\gamma - 8\eta - 8\rho + 1}.$$

For secure DHGV parameters, such as $\gamma = \eta^3 + \eta$, $\rho \approx \sqrt{\eta}$ with a sufficiently large η, this gives a large lower bound for m, and in this case lattice reduction will not recover the shortest vector. For example, for $\eta = 200$ we get $m > 5297$, which makes the lattice reduction totally inefficient according to the optimal complexity bound $\mathcal{O}\left(m^4 \log B \mathcal{M}(m \log(B))\right)$ where B is an upper bound of the Euclidean norms of the basis vectors and $\mathcal{M}(k)$ denotes the time required to multiply k-bit integers (see [102]).

M.3.4 Orthogonal lattice attack

Another attack on DGHV is the orthogonal lattice attack [41, 87]. Let $c_1 = pq_1$ and $c_i = pq_i + r_i$, for $i = 2, \ldots, m$. Then there exist $m - 1$ integers a_i, $i = 2, \ldots, m$ such that $a_2 c_2 + \ldots + a_m c_m \equiv 0 \pmod{c_1}$. This can be rewritten as

$$p(a_2 q_2 + \ldots + a_m q_m) + a_2 r_2 + \ldots + a_m r_m \equiv 0 \pmod{pq_1}.$$

Hence $a_2r_2+\ldots+a_mr_m \equiv 0 \pmod{p}$, and when the integers a_i, $i=2,\ldots,m$, satisfy $|a_i| \leq \frac{2^{\eta-1-\rho}}{m-1}$, then

$$\begin{aligned}|a_2r_2+\ldots+a_mr_m| &\leq |a_2||r_2|+\ldots+|a_m||r_m| \\ &\leq (m-1)\cdot \max_i |a_i| \cdot \max_i |r_i| \\ &\leq (m-1)\cdot \frac{2^{\eta-1-\rho}}{m-1}\cdot 2^{\rho} \\ &\leq 2^{\eta-1}.\end{aligned}$$

Since $p > 2^{\eta-1}$, then $|a_2r_2+\ldots+a_mr_m| < p$, that is $a_2r_2+\ldots+a_mr_m = 0$. Finding many such a_i's, leads to recovering p using $\gcd(c_1, a_2c_2+\ldots+a_mc_m) = p$.

M.4 Our First Lattice-based Attack on DGHV

In this section, we present our first attack on the DGHV scheme. We exploit the existence of a linear relation between the $c_2,\ldots,c_m$ and the factor q_1 of c_1 in the form

$$a_2c_2+\ldots+a_mc_m = a_1q_1,$$

where $a_1,\ldots,a_m$ are integers (see section M.1.1). We derive a condition on the size of each $|a_i|$ under which the above equation can be solved leading to the cryptanalysis of the scheme. After presenting the attack, we will present a comparison with the orthogonal lattice attack [41], and show that our attack significantly increases the bound of the parameters a_i leading to more successful attacks.

M.4.1 The attack

Theorem M.4.1. *Let $c_1 = pq_1$ and $c_i = pq_i + r_i$, $i=2,\ldots,m$, be m positive integers with $2^{\eta-1} < p < 2^{\eta}$, $2^{\gamma-1} < c_i < 2^{\gamma}$ and $|r_i| < p$ for $i=2,\ldots,m$. Let $a_1,\ldots,a_m$ be m integers satisfying $|a_i| < 2^{\alpha_i}$ for $i=2,\ldots,m$ and $a_2c_2+\ldots+a_mc_m = a_1q_1$. Define $\beta = \frac{\gamma-\eta-1}{\gamma}$. If*

$$\sum_{i=2}^{m}\alpha_i < \left(1-(1-\beta)^{\frac{m}{m-1}} - m\left(1-\sqrt[m-1]{1-\beta}\right)(1-\beta)\right)(\gamma-1),$$

then, one can find p, $q_1, \ldots, q_m$, $r_2, \ldots, r_m$ *in polynomial time.*

Proof. Suppose that $c_1 = pq_1$ and $c_i = pq_i + r_i$ for $i = 2, \ldots, m$. Let $a_1, \ldots, a_m$ be m integers satisfying $a_2c_2 + \ldots + a_mc_m = a_1q_1$. Then

$$a_2c_2 + \ldots + a_mc_m \equiv 0 \pmod{q_1}, \tag{M.1}$$

where q_1 is an unknown divisor of c_1. Suppose that $2^{\eta-1} < p < 2^{\eta}$ and $2^{\gamma-1} < c_1 < 2^{\gamma}$. Then, since $q_1 = \frac{c_1}{p}$, we get

$$2^{\gamma-\eta-1} < q_1 < 2^{\gamma-\eta+1}.$$

Define $\beta = \frac{\gamma-\eta-1}{\gamma}$. Then

$$q_1 > 2^{\gamma-\eta-1} = 2^{\gamma\beta} > c_1^{\beta}.$$

Using Herrman-May's Theorem M.2.4, we can solve the equation (M.1) if the unknown parameters a_i satisfy $|a_i| < c_1^{\lambda_i}$ for $i = 2, \ldots, m$ where

$$\sum_{i=2}^{m} \lambda_i < 1 - (1-\beta)^{\frac{m}{m-1}} - m\left(1 - \sqrt[m-1]{1-\beta}\right)(1-\beta). \tag{M.2}$$

For $i = 2, \ldots, m$, define $\alpha_i = (\gamma - 1)\lambda_i$. Then

$$c_1^{\lambda_i} > 2^{(\gamma-1)\lambda_i} = 2^{\alpha_i}.$$

Now, suppose that $|a_i| < 2^{\alpha_i}$ for $i = 2, \ldots, m$. Then $|a_i| < c_1^{\lambda_i}$ and plugging $\alpha_i = (\gamma-1)\lambda_i$ in equation (M.2), one can find the parameters a_i, $i = 2, \ldots, m$ if

$$\sum_{i=2}^{m} \alpha_i < \left(1 - (1-\beta)^{\frac{m}{m-1}} - m\left(1 - \sqrt[m-1]{1-\beta}\right)(1-\beta)\right)(\gamma - 1).$$

Using the recovered values of the parameters a_i for $i = 2, \ldots, m$, we compute

$$q_1 = \gcd(c_1, a_2c_2 + \ldots + a_mc_m), \quad p = \frac{c_1}{q_1}.$$

Next, for $i = 2, \ldots, m$, we find $r_i \equiv c_i \pmod{p}$ and $q_i = \frac{c_i - r_i}{p}$.

□

Let us summarize the whole method in Algorithm 7.

Algorithm 7 : The first attack

Require: A set of **public values** $c_1 = pq_1$, $c_i = pq_i + r_i$, $i = 2, \ldots, m$.

Ensure: The set of private parameters p, q_i, $i = 1, \ldots, m$ if the conditions of Theorem M.4.1 are fulfilled.

1: Set $f(x_2, \ldots, x_m) = c_2x_2 + \ldots + c_mx_m$.
2: Apply Coppersmith's technique and Herrman-May's Theorem M.2.4 to solve the polynomial equation $f(x_2, \ldots, x_m) \equiv 0 \pmod{q_1}$.
3: **for** each solution $(x_2, \ldots, x_m)$ **do**
4: Compute $g = \gcd(c_1, x_2c_2 + \ldots + x_mc_m)$.
5: **if** $g > 1$ **then**
6: Set $q_1 = g$ and $p = \frac{c_1}{q_1}$.
7: **for** $i = 2, \ldots, m$ **do**
8: Compute $r_i \equiv c_i \pmod{p}$.
9: Compute $q_i = \frac{c_i - r_i}{p}$.
10: **end for**
11: Output p, q_i, $i = 1, \ldots, m$, r_i, $i = 2, \ldots, m$.
12: Halt
13: **end if**
14: **end for**

M.4.2 Comparison with the orthogonal lattice attack

Let us now compare our method with the orthogonal lattice attack of [41]. Suppose that $c_1 = pq_1$ and $c_i = pq_i + r_i$ for $i = 2, \ldots, m$. Let $a_2, \ldots, a_m$ be $m - 1$ integers satisfying $a_2c_2 + \ldots + a_mc_m \equiv 0 \pmod{c_1}$ and $|a_i| < 2^\alpha$ as required in the orthogonal attack. Then, since $c_1 = pq_1$, we get $a_2c_2 + \ldots + a_mc_m \equiv 0 \pmod{q_1}$ which means that the equation can be exploited in our attack. Using Theorem M.4.1, our attack can recover all the parameters p, q_1, q_i, r_i for $i = 2, \ldots, m$ if

$$\alpha < \frac{1}{m-1}\left(1 - (1-\beta)^{\frac{m}{m-1}} - m\left(1 - \sqrt[m-1]{1-\beta}\right)(1-\beta)\right)(\gamma - 1), \quad \text{(M.3)}$$

where $\beta = \frac{\gamma - \eta - 1}{\gamma}$. Recall that the orthogonal attack of [41], as explained in Section M.3.4, will find p if

$$|a_i| \leq \frac{2^{\eta-1-\rho}}{m-1} = 2^{\eta-1-\rho-\log_2(m-1)},$$

for $i = 2, \ldots, m$. So, define the bound for the orthogonal lattice attack of [41]

$$\alpha_0 = \eta - 1 - \rho - \log_2(m-1),$$

and the bound for our attack

$$\alpha_{\text{new}} = \frac{1}{m-1}\left(1-(1-\beta)^{\frac{m}{m-1}} - m\left(1-\sqrt[m-1]{1-\beta}\right)(1-\beta)\right)(\gamma-1).$$

Let us compare α_0 and α_{new} in the optimal situation where $\eta \geq 200$, $\gamma = \eta^3+\eta$ and $\rho \approx \sqrt{\eta}$ as recommended by [41]. These parameter sizes are believed to resist currently known attacks including factorization, diophantine and lattice-based attacks. In Table M.1, we show the maximal values of α_0 for which the orthogonal attack of [41] works, and the maximal values of α_{new} for which our attack works. Clearly, our method *significantly* increases the bounds of the size of the unknown integers a_i, $i = 2, \ldots, a_m$ for which DGHV is vulnerable.

	$m=2$		$m=3$		$m=5$		$m=10$		$m=15$	
η	α_0	α_{new}	α_0	α_{new}	α_0	α_{new}	α_0	α_{new}	α_0	α_{new}
200	184.8	7.9×10^6	183.8	3.9×10^6	182.8	1.9×10^6	181.6	8.8×10^5	181	5.7×10^5
300	284.8	2.6×10^7	283.8	1.3×10^7	282.8	6.7×10^6	281.6	2.9×10^6	281	1.9×10^6
400	384.8	6.3×10^7	383.8	3.1×10^7	382.8	1.5×10^7	381.6	7.1×10^6	381	4.5×10^6
500	484.8	1.2×10^8	483.8	6.2×10^7	482.8	3.1×10^7	481.6	1.3×10^7	481	8.9×10^6

Table M.1: Comparison of α_0 and α_{new} for certain values of η and m.

M.4.3 Deriving new parameter sizes

To avoid the new attack, it is sufficient to make the inequality (M.3) impossible or hard to occur. Since γ is large, this could be possible if

$$1-(1-\beta)^{\frac{m}{m-1}} - m\left(1-\sqrt[m-1]{1-\beta}\right)(1-\beta) \approx 0,$$

where $\beta = \frac{\gamma-\eta-1}{\gamma}$. Therefore, for $m > 1$, our attack will fail if $\beta \approx 0$, or equivalently $\gamma \approx \eta$. However, our attack is likely to be successful when $\beta \approx 1$ and the number m of public integers c_i, $i = 1, \ldots, m$ is not very large. In this situation, the inequality (M.3) reduces to $\alpha < \frac{\gamma-1}{m-1}$. Note that $\beta \approx 1$ implies that γ is much larger than η which is the case for the currently recommended parameters. Therefore, for the recommended parameters $\gamma = \eta^3 + \eta$ with large η, our attack will be successful as long as $\alpha < \frac{\gamma-1}{m-1}$.

M.4.4 Experimental Results

We implemented our attack and experimented it with 100 instances of DGHV. All the 100 attacks were successful. For efficiency reasons, we considered only instances of DGHV where the sizes of the parameters are small, typically $\eta \leq 60$ and $\gamma \leq 200$. The recommended DHGV parameters $\eta \geq 200$ and $\gamma = \eta^3 + \eta$ i.e., $\gamma \geq 8000200$ are not suitable for experimentation using an off-the-shelf computer.
The following example is presented as a concrete illustration of our attack.

Consider the following situation with $m = 4$ public integers:

$$c_1 = pq_1 = 115681713396549343702207914242260837695350516124613657,$$
$$c_2 = pq_2 + r_2 = 108225557677193859451749518166560930564055519997881978,$$
$$c_3 = pq_3 + r_3 = 870086279935814181906531631207348759267570817322424 10,$$
$$c_4 = pq_4 + r_4 = 639007350722203683834523048430478564768424234694 73333,$$

where, for $i = 2, 3, 4$, $c_i < 2^\gamma$ with $\gamma = 177$. According to Theorem M.4.1, we can solve the linear equation $a_2c_2 + a_3c_3 + a_4c_4 = a_1q_1$ if the unknown parameters a_2, a_3 and a_3 are suitably small. Combining the method of Herrmann and May [59] for solving the equation $a_2c_2 + a_3c_3 + a_4c_4 \equiv 0 \pmod{q_1}$, and the LLL algorithm [86], we get at least two polynomials sharing the solutions a_2, a_3, a_4. Then applying Gröbner Basis computation for solving systems of polynomial equations, we get the solution

$$a_2 = 130722418993,\ a_3 = 16613347,\ a_4 = 27131339.$$

Using these values, we get

$$\begin{aligned} q_1 &= \gcd(c_1, a_2c_2 + a_3c_3 + a_4c_4) \\ &= 293929964541029095109322043966679684364726508 1, \\ p &= \frac{c_1}{q_1} = 39356897. \end{aligned}$$

Using the value of p, we get

$$\begin{aligned} r_2 &\equiv c_2 \equiv 13835383 \pmod{p}, \\ r_3 &\equiv c_3 \equiv 37261850 \pmod{p}, \\ r_4 &\equiv c_4 \equiv 1283090 \pmod{p}. \end{aligned}$$

Finally, we get

$$q_2 = \frac{c_2 - r_2}{p} = 27498498592811790891885993493481188459561616 35,$$
$$q_3 = \frac{c_3 - r_3}{p} = 22107593490813419104319419064143922709851104 81,$$
$$q_4 = \frac{c_4 - r_4}{p} = 16236222858783904733000750756099459893101436 19.$$

The whole process, including Gröbner Basis computation, took less than one minute. Note that since $a_2c_2+a_3c_3+a_4c_4 \not\equiv 0 \pmod{p}$, the orthogonal attack of [41] and [87] is not applicable to this DGHV instance.

M.5 Our Second Lattice Attack on DGHV

In this section, we consider the situation where the DGHV public values are of the general form $c_i = pq_i + r_i$, $i = 1, \ldots, m$, and there exists a linear relation between the q_i's of the form $a_1q_1 + \ldots + a_mq_m = 0$. We show that it is possible to solve the equation and recover all the private parameters, when specific conditions on the size of the unknown coefficients a_i, $i = 1, \ldots, m$ are fulfilled.

M.5.1 The attack

Theorem M.5.1. *Let $c_i = pq_i + r_i$, $i = 1, \ldots, m$, be m positive integers with $c_1 < \ldots < c_m$ and $|r_i| < 2^\rho$ for $i = 1, \ldots, m$. Let $a_1, \ldots, a_m$ be m integers satisfying $|a_i| < 2^\alpha$ for $i = 1, \ldots, m$ and $a_1q_1 + \ldots + a_mq_m = 0$. If*

$$\alpha < \frac{1}{m}\log_2(c_m) + \log_2\left(\frac{\sqrt{m}}{m+1}\right) - \rho,$$

then, one can find p, $q_1, \ldots, q_m$, $r_1, \ldots, r_m$ in polynomial time.

Proof. Let $c_i = pq_i + r_i$ for $i = 1, \ldots, m$ with $r_i \neq 0$. Then, there exist m integers a_i, $i = 1, \ldots m$ such that $a_1q_1 + \ldots + a_mq_m = 0$. Combining the values of c_i for $i = 1, \ldots, m$, we get:

$$a_1c_1 + \ldots + a_mc_m = a_1r_1 + \ldots + a_mr_m. \tag{M.4}$$

Consider the $m \times m$ lattice $\mathcal{L} \subset \mathbb{Z}^m$ defined by the rows of the matrix

$$M = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & c_1 \\ 0 & 1 & 0 & \dots & 0 & c_2 \\ 0 & 0 & 1 & \dots & 0 & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & c_{m-1} \\ 0 & 0 & 0 & \dots & 0 & c_m \end{bmatrix}.$$

The dimension of $\mathcal{L}$ is $\dim(\mathcal{L}) = m$ and the determinant is $\det(\mathcal{L}) = c_m$. Let $v \in \mathcal{L}$ be a target vector generated from the vector $u = (a_1, \dots, a_m) \in \mathbb{Z}^m$, that is,

$$v = uM = (a_1, \dots, a_{m-1}, c_1a_1 + \dots + c_ma_m). \tag{M.5}$$

Minkowski's Theorem M.2.2 for $\mathcal{L}$ asserts that there exists short non-zero vectors of size at most $\sigma(\mathcal{L})$ where

$$\sigma(\mathcal{L}) = \sqrt{\dim(\mathcal{L})}\det(\mathcal{L})^{\frac{1}{\dim(\mathcal{L})}} = \sqrt{m}c_m^{\frac{1}{m}}. \tag{M.6}$$

For our target vector v to be among the shortest non-zero vectors of the lattice $\mathcal{L}$, the inequality $\sigma(\mathcal{L}) > \|v\|$ must hold. Assume further that for $i = 1, \dots, m$, we have $|a_i| \leq 2^\alpha$ and $|r_i| \leq 2^\rho$. Using (M.5) with $a_1q_1 + \dots + a_mq_m = 0$, we get

$$\begin{aligned} \|v\| &= \left(\sum_{i=1}^{m-1} a_i^2 + (c_1a_1 + \dots + c_ma_m)^2\right)^{1/2} \\ &= \left(\sum_{i=1}^{m-1} a_i^2 + \left(\sum_{i=1}^{m} a_ir_i\right)^2\right)^{1/2} \\ &< \left(2^{2\alpha}(m-1) + \left(2^{\alpha+\rho}m\right)^2\right)^{1/2} \\ &< \left(2^{2(\alpha+\rho)}\,(m+1)^2\right)^{1/2} \\ &= (m+1)2^{\alpha+\rho}. \end{aligned}$$

Therefore, the inequality $\sigma(L) > \|v\|$ is fullfiled if $\sqrt{m}c_m^{\frac{1}{m}} > (m+1)2^{\alpha+\rho}$, from which we deduce the following condition on α.

$$\alpha < \frac{1}{m}\log_2(c_m) + \log_2\left(\frac{\sqrt{m}}{m+1}\right) - \rho. \tag{M.7}$$

If the condition (M.7) holds, then applying lattice reduction to $\mathcal{L}$ yields the vector $v = (a_1, \ldots, a_{m-1}, c_1a_1 + \ldots + c_ma_m)$ with $c_1a_1 + \ldots + c_ma_m = a_1r_1 + \ldots + a_mr_m$, as in (M.4). Combining the obtained values from the reduction, that is $a_1, \ldots, a_{m-1}$ and $c_1a_1 + \ldots + c_ma_m$, and the known public values c_i, $i = 1, \cdots, m$, one can calculate a_m as follows:

$$a_m = \frac{(c_1a_1 + \ldots + c_ma_m) - (c_1a_1 + \ldots + c_{m-1}a_{m-1})}{c_m}.$$

The next step in the attack is to solve the equation

$$a_1r_1 + \ldots + a_mr_m = c_1a_1 + \ldots + c_ma_m, \tag{M.8}$$

with the unknown parameters $r_1, \ldots, r_m$ with $|r_i| < 2^\rho$ for $i = 1, \ldots, m$. To do so, we consider the $(m+1) \times (m+1)$ lattice $\mathcal{L}' \subset \mathbb{Z}^{m+1}$ defined by the rows of the matrix

$$M' = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & Ca_1 \\ 0 & 1 & 0 & \ldots & 0 & Ca_2 \\ 0 & 0 & 1 & \ldots & 0 & Ca_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & Ca_m \\ 0 & 0 & 0 & \ldots & 0 & C(c_1a_1 + \ldots + c_ma_m) \end{bmatrix},$$

where C is a given parameter to be optimized later. The determinant of $\mathcal{L}'$ is $\det(\mathcal{L}') = C|c_1a_1 + \ldots + c_ma_m|$ and its dimension is $\dim(\mathcal{L}') = m+1$. Let $v' \in \mathcal{L}'$ be a vector. Then there exists a vector $u' = (y_1, \ldots, y_{m+1}) \in \mathbb{Z}^{m+1}$ such that

$$v' = u'M' = (y_1, \ldots, y_m, C(a_1y_1 + \ldots + a_my_m) + C(c_1a_1 + \ldots + c_ma_m)y_{m+1}).$$

We set our target vector to be $v' = (r_1, \ldots, r_m, 0)$, therefore

$$\begin{aligned} &y_1 = r_1, \ldots, y_m = r_m, \\ &(a_1y_1 + \ldots + a_my_m) + (c_1a_1 + \ldots + c_ma_m)y_{m+1} = 0. \end{aligned}$$

In addition, we need $y_{m+1} = -1$ so that $a_1r_1 + \ldots + a_mr_m = c_1a_1 + \ldots + c_ma_m$ which provides a solution to equation (M.8). Now recall that Minkowski's Theorem M.2.2 asserts that there exist short non-zero vectors in the lattice $\mathcal{L}'$ of size at most $\sigma(\mathcal{L}')$ where

$$\sigma(\mathcal{L}') = \sqrt{\dim(\mathcal{L}')}\det(\mathcal{L}')^{\frac{1}{\dim(\mathcal{L}')}} = \sqrt{m+1}\cdot C^{\frac{1}{m+1}}\cdot|c_1a_1 + \ldots + c_ma_m|^{\frac{1}{m+1}}.$$

Since $c_1a_1 + \ldots + c_ma_m = a_1r_1 + \ldots + a_mr_m$ with $|a_i| < 2^\alpha$ and $|r_i| < 2^\rho$, then,

$$\sigma(\mathcal{L}') < \sqrt{m+1}\cdot C^{\frac{1}{m+1}}\cdot\left(2^{\alpha+\rho}m\right)^{\frac{1}{m+1}}. \tag{M.9}$$

The norm of our target vector $v' = (r_1, \ldots, r_m, 0)$ with $|r_i| < 2^\rho$ for $i = 1, \ldots, m$, satisfies

$$\|v'\| = \left(\sum_{i=1}^{m} r_i^2\right)^{1/2} < 2^\rho\sqrt{m}.$$

Therefore, for our target vector v' to be among the short vectors, the inequality $\sigma(\mathcal{L}') > \|v'\|$ must be satisfied. For this, it is sufficient that ρ satisfies

$$\sqrt{m+1}\cdot C^{\frac{1}{m+1}}\cdot\left(2^{\alpha+\rho}m\right)^{\frac{1}{m+1}} > 2^\rho\sqrt{m}$$

which leads to the following condition on C

$$C > m^{\frac{m-1}{2}}\cdot(m+1)^{-\frac{m+1}{2}}\cdot 2^{m\rho-\alpha}. \tag{M.10}$$

So, under condition M.10, applying lattice reduction to $\mathcal{L}'$ recovers a short non zero vector $v' = (r_1, \ldots, r_m, 0)$ which yields the r_i's. Next, using r_1 and r_2, we get $p = \gcd(c_1 - r_1, c_2 - r_2)$ and for $i = 1, \ldots, m$, we get $q_i = \frac{c_i - r_i}{p}$. This terminates the proof. □

We can summarize the whole method in Algorithm 8.

M.5.2 Application with the DGHV recommended parameters

Let us consider the recommended optimal parameters for a secure DGHV, as stated in [41], that is $\gamma = \eta^3 + \eta$, $\rho \approx \sqrt{\eta}$ and $\eta \geq 200$. Then, the condition of Theorem M.5.1 becomes

$$\alpha < \frac{\eta^3 + \eta}{m} + \log_2\left(\frac{\sqrt{m}}{m+1}\right) - \sqrt{\eta}.$$

Algorithm 8 : The second attack

Require: A set of ciphertexts $c_i = pq_i + r_i$, $i = 1, \ldots, m$.

Ensure: The set of private parameters p, q_i, $i = 1, \ldots, m$ if the conditions of Theorem M.5.1 are fulfilled.

Define the lattice $\mathcal{L}$ with the basis matrix

$$M = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & c_1 \\ 0 & 1 & 0 & \ldots & 0 & c_2 \\ 0 & 0 & 1 & \ldots & 0 & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & c_{m-1} \\ 0 & 0 & 0 & \ldots & 0 & c_m \end{bmatrix}.$$

Apply the LLL algorithm to reduce the basis matrix.
for each row $(a_1, \ldots, a_{m-1}, R)$ of the reduced matrix **do**
  Compute $a_m = \frac{R-(c_1a_1+\ldots+c_{m-1}a_{m-1})}{c_m}$.
  Compute $\alpha = \max_i \left(\log_2(|a_i|)\right)$.
  Compute $\rho = \frac{1}{m}\log_2(c_m) + \log_2\left(\frac{\sqrt{m}}{m+1}\right) - \alpha$.
  Let C be the integral part of $m^{\frac{m-1}{2}} \cdot (m+1)^{-\frac{m+1}{2}} \cdot 2^{m\rho-\alpha} + 1$.
  Define the lattice $\mathcal{L}'$ with the basis matrix

$$M' = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & Ca_1 \\ 0 & 1 & 0 & \ldots & 0 & Ca_2 \\ 0 & 0 & 1 & \ldots & 0 & Ca_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & Ca_m \\ 0 & 0 & 0 & \ldots & 0 & C(c_1a_1 + \ldots + c_ma_m) \end{bmatrix}$$

  Apply the LLL algorithm to reduce the basis matrix.
  for each row $(r_1, \ldots, r_{m+1})$ of the reduced matrix **do**
    if $r_{m+1} = 0$ **then**
      Compute $p = \gcd(c_1 - r_1, c_2 - r_2)$.
      if $p > 1$ **then**
        for $i = 1, \ldots, m$ **do**
          Compute $q_i = \frac{c_i - r_i}{p}$.
        end for
      end if
    end if
    Output p, q_i, $i = 1, \ldots, m$, r_i, $i = 1, \ldots, m$.
    Halt
  end for
end for

On the other hand, the condition on the constant C in (M.10) becomes

$$C > m^{\frac{m-1}{2}} \cdot (m+1)^{-\frac{m+1}{2}} \cdot 2^{m\sqrt{\eta}-\alpha}.$$

In Table M.2, we present the upper bounds for α in terms of η and m under which our second method will solve the equation $a_1q_1 + \ldots + a_mq_m = 0$ and then find all the DGHV parameters. For all cases, we use $C = 1$, wich fullfils condition (M.10).

η	$m = 2$	$m = 3$	$m = 5$	$m = 10$	$m = 15$
200	4×10^6	2.6×10^6	1.6×10^6	8×10^5	5.3×10^5
300	1.3×10^7	9×10^6	5.4×10^6	2.7×10^6	1.8×10^6
400	3.2×10^7	2.1×10^7	1.2×10^7	6.4×10^6	4.2×10^6
500	6.2×10^7	4.1×10^7	2.5×10^7	1.5×10^7	8.3×10^6

Table M.2: Optimal values for α for different values of η and m.

M.5.3 Experimental Results

For our second attack, we also experimented with 100 DHGV instances with various practical sizes of the parameters η, ρ, γ and m. When the conditions of Theorem M.5.1 are satisfied, we always succeeded in finding the solutions of our equations and recovered the secret parameters. We illustrate the steps of our attack through the following detailed example.
Consider the following DHGV instance:

$$c_1 = pq_1 + r_1 = 564058455074945300209410084805729402861816892372588 54,$$
$$c_2 = pq_2 + r_2 = 399048214644609484947002841923365255233574075450676 68,$$
$$c_3 = pq_3 + r_3 = 562949913454332849006128056132490607872372793280225 19,$$

with the bounds $c_i < 2^\gamma$, $i = 1, 2, 3$, with $\gamma = 176$. According to Theorem M.5.1, one can solve the equation $a_1q_1 + a_2q_2 + a_3q_3 = 0$ if the unknown coefficients a_i, $i = 1, 2, 3$ satisfy $|a_i| < 2^\alpha$ with

$$\alpha + \rho < \frac{1}{3}\log_2(c_3) + \log_2\left(\frac{\sqrt{3}}{4}\right) \approx 57.459,$$

where ρ is the bit size of the noise r_i, $i = 1, 2, 3$. Let $\mathcal{L}$ be the lattice spanned by the rows of the matrix

$$\begin{pmatrix} 1 & 0 & c_1 \\ 0 & 1 & c_2 \\ 0 & 0 & c_3 \end{pmatrix}.$$

Applying the LLL algorithm [86] for reduction, yields a reduced basis, where the first vector is $(3991298341123, 3713241313153, 18196712614595893)$. From this, we deduce

$$\begin{aligned} a_1 &= 3991298341123, \\ a_2 &= 3713241313153, \\ a_3 &= \frac{18196712614595893 - (a_1c_1 + a_2c_2)}{c_3} = -6631296680887. \end{aligned}$$

In this example, we have $|a_i| < 2^\alpha$ for $i = 1, 2, 3$ with $\alpha = 43$. Next, the aim is to solve the equation

$$a_1r_1 + a_2r_2 + a_3r_3 = a_1c_1 + a_2c_2 + a_3c_3 = 18196712614595893,$$

with the unknown coefficients r_1, r_2, and r_3. Let C be a constant, and consider the lattice $\mathcal{L}'$ spanned by the rows of the matrix

$$\begin{pmatrix} 1 & 0 & 0 & Ca_1 \\ 0 & 1 & 0 & Ca_2 \\ 0 & 0 & 1 & Ca_3 \\ 0 & 0 & 0 & C(a_1c_1 + a_2c_2 + a_3c_3) \end{pmatrix}.$$

Then, using $C = 1$ and applying the LLL algorithm, we get the following short vector $(-23593, -18617, -21881, 0)$. This leads to the values of

$$r_1 = 23593, \qquad r_2 = 18617, \qquad r_3 = 21881.$$

Hence, in this example, we have $|r_i| < 2^\rho$ for $i = 1, 2, 3$ with $\rho = 15$. We then deduce

$$\begin{aligned}
p &= \gcd(c_1 - r_1, c_2 - r_2) = 706549229, \\
q_1 &= \frac{c_1 - r_1}{p} = 79832859753208406686890615063671579331921809, \\
q_2 &= \frac{c_2 - r_2}{p} = 56478472874338029310481752988136833029305319, \\
q_3 &= \frac{c_3 - r_3}{p} = 79675964582268739409540570899482307392394422.
\end{aligned}$$

In this example, we have $p < 2^\eta$ with $\eta = 30$. We notice that the dimensions of the underlying lattices are small and that the computation took less than 30 seconds using an off-the-shelf computer. Also, we notice that the condition of Theorem M.5.1 is satisfied since

$$\alpha + \rho \approx \frac{1}{m}\log_2(c_m) + \log_2\left(\frac{\sqrt{m}}{m+1}\right) \approx 58.$$

More importantly, this example shows that while our second attack was successful, the existing attacks of [41], as described in Section M.3, fail to recover the parameter p: the continued fraction attack fails because we need $\frac{|r_2|}{c_1} < \frac{1}{2q_1^2}$, which is not the case in this example, the simultaneous diophantine approximation attack fails too, because the condition on m should be

$$m > -2\eta + 2\rho + \frac{1}{2} + \frac{1}{2}\sqrt{16\eta^2 - 32\eta\rho + 16\rho^2 + 16\gamma - 8\eta - 8\rho + 1} > 9,$$

while $m = 3$ in this example. Finally, the orthogonal attack can not work since none of the $r_i = 0$.

M.6 Conclusion

In this paper, we presented two new lattice-based attacks on the DHGV encryption scheme using Coppersmith's technique and the LLL algorithm for the first attack, and only the LLL algorithm for the second attack. The first attack is applicable when $c_1 = pq_1$ and the $m - 1$ public integers c_i, $i = 2, \ldots, m$ satisfy a linear equation $a_2c_2 + \ldots + a_mc_m = a_1q_1$ for suitably small

integers a_i, $i = 2, \ldots, m$. The second attack works even with $c_1 = pq_1 + r_1$ when the integers q_i satisfy a linear equation $a_1q_1 + \ldots + a_mq_m = 0$ for suitably small integers a_i, $i = 1, \ldots, m$. We illustrated our attacks by providing experimental results and examples, and further computed the bounds for DGHV recommended parameters for which our attacks are applicable, thus effectively extending on previously proposed optimal parameter bounds for p, c_i and r_i, $i = 1, \ldots, m$.

Bibliography

[1] ANSI Standard X9.31-1998, *Digital Signatures Using Reversible Public Key Cryptography for the Financial Services Industry (rDSA).*

[2] Accredited Standards Committee, *Lattice-Based Polynomial Public Key Establishment Algorithm for the Financial Services Industry*, ANSI X9.98-2010, American National Standards Institute, 2010.

[3] M. Ajtai. *The shortest vector problem in L2 is NP-hard for randomized reductions*, In STOC'98, pp. 10-19 (1998)

[4] M. Albrecht, S. Bai, and L. Ducas. *A subfield lattice attack on overstretched NTRU assumptions: Cryptanalysis of some FHE and Graded Encoding Schemes*. Cryptology ePrint Archive, Report 2016/127, 2016. `http://eprint.iacr.org/`.

[5] T.M. Apostol. *Introduction to Analytic Number Theory*, Springer, 1976.

[6] Frederik Armknecht, Colin Boyd, Christopher Carr, Kristian Gjøsteen, Angela Jäschke, Christian A. Reuter, Martin Strand. *A Guide to Fully Homomorphic Encryption*. IACR Cryptology ePrint Archive 2015: 1192 (2015)`https://eprint.iacr.org/2015/1192.pdf`

[7] A. Baker. *Linear forms in the logarithms of algebraic numbers IV*. Mathematika 15, 204–216, 1966.

[8] A. Bauer, A. and Joux. *Toward a rigorous variation of Coppersmith's algorithm on three variables*. In Proceedings of Eurocrypt'07, volume 4515 of Lecture Notes in Computer Science, Springer-Verlag, pp. 361–378 (2007)

[9] Josh Benaloh, Melissa Chase, Eric Horvitz, Kristin Lauter. Patient-controlled encryption: patient privacy in electronic medical records. In Proceedings of the 2009 ACM workshop on Cloud computing security

[10] D.J. Bernstein, Y.A., Chang, C.M. Cheng, L.P. Chou, N. Heninger, T. Lange, N. van Someren. *Factoring RSA keys from certified smart cards: Coppersmith in the wild.* In Advances in Cryptology-ASIACRYPT 2013. Springer, 2013, pp. 341–360 (2013)

[11] D.J. Bernstein, C. Chuengsatiansup, T. Lange and C. van Vredendaal. *NTRU Prime*, Cryptology ePrint Archive, Report 2016/461, 2016. `https://eprint.iacr.org/2016/461`

[12] D. Bleichenbacher, A. May. *New attacks on RSA with small secret CRT-exponents*. In: Yung, M., Dodis, Y., Kiayias, A., Malkin, T.G. (eds.) PKC 2006. LNCS, vol. 3958, pp. 1–13. Springer, Heidelberg (2006)

[13] J. Blömer, A. May. *A generalized Wiener attack on RSA*, In Public Key Cryptography - PKC 2004, volume 2947 of LNCS, pp. 1-13. Springer-Verlag (2004)

[14] J. Blömer, A. May. *New partial key exposure attacks on RSA*, Proceedings of CRYPTO 2003, LNCS 2729 [2003], pp. 27-43. Springer Verlag (2003)

[15] D. Boneh. *Twenty years of attacks on the RSA cyptosystem*, Notices of the AMS 46 (2) (February 1999) pp. 203-213 (1999)

[16] D. Boneh, G. Durfee, N. Howgrave-Graham. *Factoring $N = p^r q$ for Large r*. In M. Weiner, ed., Proceedings of Crypto'99, vol. 1666 of LNCS, pp. 326–337. Springer-Verlag, Aug. (1999)

[17] D. Boneh, G. Durfee. *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, 1–11 (1999)

[18] D. Boneh, G. Durfee, Y. Frankel. *An attack on RSA given a small fraction of the private key bits*. In: Ohta, K., Pei, D. (eds.) Advances in Cryptology Asiacrypt'98. Lecture Notes in Computer Science, vol. 1514, pp. 25–34. Springer-Verlag (1998)

[19] D. Boneh, E.J. Goh, and K. Nissim. *Evaluating 2-DNF formulas on ciphertexts*. In Theory of Cryptography - TCC'05, volume 3378 of Lecture Notes in Computer Science, pp. 325–341. Springer, 2005.

[20] D. Boneh, H. Shacham. *Fast Variants of RSA*, CryptoBytes, Vol. 5, No. 1, pp. 1–9, (2002)

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

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

[23] M. Bunder, J. Tonien. *A new improved attack on RSA based on Wiener's technique of continued fractions*, submitted manuscript.

[24] M. Bunder, A. Nitaj, W. Susilo, J. Tonien. *A new attack on three variants of the RSA cryptosystem*, Proceedings of ACISP, the 21st Australasian Conference on Information Security and Privacy, Lecture Notes in Computer Science 9723, 2016, 258-268.

[25] M. Bunder, A. Nitaj, W. Susilo, J. Tonien. *A generalized attack on RSA type cryptosystems*, To appear in Theoretical Computer Science.

[26] C. Cachin, S. Micali, M. Stadler. *Computationally private information retrieval with polylogarithmic communication*, J. Stern (Ed.): EUROCRYPT'99, LNCS 1592, pp. 402-414, 1999, Springer-Verlag (1999)

[27] Z. Cao. *Universal Encrypted Deniable Authentication Protoco*, International Journal of Network Security (IJNS), Vol.8, No.2, pp. 151–158, 2009.

[28] C. Carlet. *Boolean Functions for Cryptography and Error Correcting Codes*, In Yves Crama and Peter L. Hammer (eds.) Boolean Models and Methods in Mathematics, Computer Science, and Engineering, pages 257–397, Cambridge University Press, 2010.

[29] G. Castagnos. *An efficient probabilistic public-key cryptosystem over quadratic field quotients*, 2007, Finite Fields and Their Applications, 07/2007, 13(3-13), p. 563-576. `http://www.math.u-bordeaux1.fr/~gcastagn/publi/crypto_quad.pdf`

[30] Y. Chen, P.Q. Nguyen. *Faster algorithms for approximate common divisors: Breaking fully-homomorphic-encryption challenges over the integers*. In Pointcheval and Johansson (Eds.), EUROCRYPT 2012 Proceedings, volume 7237 of Lecture Notes in Computer Science. Springer, 2012, pp. 502–519 (2012)

[31] J.H. Cheon, J. Jeong, C. Lee. *An Algorithm for NTRU Problems and Cryptanalysis of the GGH Multilinear Map without an encoding of zero*. Cryptology ePrint Archive, Report 2016/139, 2016. `http://eprint.iacr.org/`.

[32] H. Cohen. *A Course in Computational Algebraic Number Theory*, Graduate Texts in Mathematics, Vol. 138, Springer-Verlag (1993)

[33] Compaq Computer Corperation. *Cryptography using Compaq multiprime technology in a parallel processing environment*, 2002. Availabe online at
`ftp://ftp.compaq.com/pub/solutions/CompaqMultiPrimeWP.pdf`

[34] D. Coppersmith. *Small solutions to polynomial equations, and low exponent RSA vulnerabilities*, Journal of Cryptology, 10(4), pp. 233–260, 1997.

[35] D. Coppersmith and A. Shamir. *Lattice attacks on NTRU*. In Advances in cryptology-EUROCRYPT'97, volume 1233 of Lecture Notes in Comput. Sci., pages 52-61. Springer, Berlin, 1997.

[36] J.S. Coron, A. Mandal, D. Naccache, M. Tibouchi. *Fully homomorphic encryption over the integers with shorter public keys*. CRYPTO 2011, In Rogaway (Eds.), Proceedings, volume 6841 of Lecture Notes in Computer Science. Springer, 2011, pp. 487–504 (2011)

[37] Y. Crama, P.L. Hammer. *Boolean Functions, Theory, Algorithms, and Applications*, Cambridge University Press, 2010.

[38] D.K. Dalai, K.C. Gupta and S. Maitra. *Cryptographically significant boolean functions: construction and analysis in terms of algebraic immunity*, In INDOCRYPT 2004, pages 92–106, number 3348, Lecture Notes in Computer Science, Springer-Verlag.

[39] N. Demytko. *A new elliptic curve based analogue of RSA*, in T. Helleseth (ed.), EUROCRYPT 1993, Lecture Notes in Computer Science 765, Springer-Verlag, pp. 40–49 (1994)

[40] W. Diffie, E. Hellman. *New directions in cryptography*, IEEE Transactions on Information Theory, 22, 5 (1976), pp. 644-654, 1976.

[41] M. van Dijk, C. Gentry, S. Halevi, V. Vaikuntanathan. *Fully homomorphic encryption over the integers*. In H. Gilbert (Ed.), EUROCRYPT 2010, LNCS, vol. 6110, Springer, 2010, pp. 24–43 (2010)

[42] T. El Gamal. *A public key cryptosystem and signature scheme based on discrete logarithms*, IEEE Transactions on Information Theory IT-31, pp. 496–473, 1985.

[43] H. Elkamchouchi, K. Elshenawy, H. Shaban. *Extended RSA cryptosystem and digital signature schemes in the domain of Gaussian integers*, in Proceedings of the 8th International Conference on Communication Systems, (2002) pp. 91–95.

[44] M. Ernst, E. Jochemsz, A. May, B. de Weger. *Partial key exposure attacks on RSA up to full size exponents*. In: Cramer, R. (ed.) Advances in Cryptology Eurocrypt 2005. Lecture Notes in Computer Science, vol. 3494, pp. 371–386. Springer-Verlag (2005)

[45] J-C. Faugère, R. Marinier, G. Renault. *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) pp. 70–87.

[46] J. Feigenbaum, M. Merritt. *Open Questions, Talk Abstracts, and Summary of Discussions*. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol 2, pp. 1–45, (1991).

[47] H.R.P. Ferguson, D.H. Bailey. *A polynomial time, numerically stable integer relation algorithm*. RNR Technical Report RNR-91-032, NASA Ames Research Center, Moffett Field, CA. (December 1991)

[48] J. Friedlander,A. Granville. *Smoothing "Smooth" Numbers*, Philos. Trans. Roy. Soc. London Ser. A 345, 339–347, 1993.

[49] C. Gentry. *Fully homomorphic encryption using ideal lattices*. In Symposium on Theory of Computing-STOC 2009, ACM, pp. 169–178, 2009.

[50] C. Gentry. *Toward basing fully homomorphic encryption on worst-case hardness*, Crypto 2010, LNCS 6223, pp. 116–137 (2010)

[51] M. Girault, P. Toffin, and B. Vallée. *Computation of approximate L-th roots modulo n and application to cryptography*. In CRYPTO'88, pp. 100–117 (1988)

[52] S. Goldwasser, S. Micali. *Probabilistic Encryption*. Journal of Computer and System Sciences, Vol 28, Issue 2, pp. 270–299 (1984)

[53] A. Granville. *Smooth numbers: computational number theory and beyond*, Proc. MSRI Conf. Algorithmic Number Theory: Lattices, Number Fields, Curves, and Cryptography, Berkeley, 2000, J. Buhler and P. Stevenhagen, eds., Cambridge University Press.

[54] R.R. Hall, G. Tenenbaum. *Divisors*. Cambridge Tracts in Mathematics, 90, Cambridge University Press, 1988.

[55] J. Hastad. *Solving simultaneous modular equations of low degree*, SIAM J. of Computing, Vol. 17, p.336-341, 1988.

[56] J. Hastad. *On Using RSA with Low Exponent in a Public Key Network*, in Proceedings of CRYPTO'85, Springer-Verlag, pp. 403–408 (1986)

[57] G.H. Hardy, E.M. Wright. *An Introduction to the Theory of Numbers*, Oxford University Press, London.

[58] M. Herrmann. *Improved cryptanalysis of the multi-prime Φ - hiding assumption*, A. Nitaj and D. Pointcheval (Eds.): AFRICACRYPT 2011, LNCS 6737, pp. 92-99. Springer Verlag (2011)

[59] M. Herrmann, A. May. *Solving linear equations modulo divisors: On factoring given any bits*. In Advances in Cryptology-ASIACRYPT 2008. Springer, 2008, pp. 406–424 (2008)

[60] A. Hildebrand. *On the number of positive integers $\leq x$ and free of prime factors $> y$*, J. Number Theory 22, 289–307, 1986.

[61] M.J. Hinek. *Cryptanalysis of RSA and its Variants*. Chapman & Hall/CRC Cryptography and Network Security. CRC Press, Boca Raton, FL, (2010)

[62] M.J. Hinek, C.C.Y. Lam. *Common modulus attacks on small private exponent RSA and some fast variants (in Practice)*, J. Math. Cryptology, Volume 4, Issue 1, July 2010, pp. 58–93 (2010)

[63] J. Hoffstein, J. Pipher, and J. H. Silverman. *NTRU: A Ring Based Public Key Cryptosystem*, in Algorithmic Number Theory. Lecture Notes in Computer Science 1423, Springer-Verlag, pages 267–288, 1998.

[64] J. Hoffstein, J. Pipher, J.M. Schanck, J. H. Silverman, W. Whyte, and Z. Zhang. *Choosing parameters for NTRUEncrypt*. Cryptology ePrint Archive, Report 2015/708, 2015. `http://eprint.iacr.org/2015/708`.

[65] N. Howgrave-Graham. *Finding small roots of univariate modular equations revisited*, In Cryptography and Coding, LNCS 1355, pp. 131-142, Springer-Verlag.

[66] N. Howgrave-Graham. *Approximate integer common divisors*, CaLC 2001, LNCS vol. 2146, pp. 51-66, Springer-Verlag.

[67] N. Howgrave-Graham, *A Hybrid Lattice-Reduction and Meet-in-the-Middle Attack Against NTRU*, Advances in Cryptology - CRYPTO 2007 Volume 4622 of the series Lecture Notes in Computer Science pp. 150–169.

[68] N. Howgrave-Graham, J-P. Seifert. *Extending Wiener's attack in the presence of many decrypting exponents*. In Secure Networking - CQRE (Secure)'99, volume 1740 of Lecture Notes in Computer Science, pp. 153–166. Springer-Verlag, (1999)

[69] B. Ibrahimpasic. *Cryptanalysis of KMOV cryptosystem with short secret exponent*, Central European Conference on Information and Intelligent Systems, CECIIS - 2008.

[70] IEEE Computer Society, *IEEE Standard Specification for Public Key Cryptographic Tech- niques Based on Hard Problems over Lattices*,

IEEE Std 1363.1-2008, The Institute of Electrical and Electronics Engineers, 2009.

[71] E. Jochemsz. *Cryptanalysis of RSA variants using small roots of polynomials*, *Ph.D. Dissertation*. TU Eindhoven. 2007

[72] E. Jochemsz,A. May. *A strategy for finding roots of multivariate polynomials with new applications in attacking RSA variants*, in: ASIACRYPT 2006, LNCS, vol. 4284, 2006, pp. 267–282, Springer-Verlag (2006)

[73] P. Kirchner and P-A. Fouque. *An improved BKW algorithm for LWE with applications to cryptography and lattices*. In Rosario Gennaro and Matthew J. B. Robshaw, editors, Advances in Cryptology - CRYPTO 2015, Part I, volume 9215 of Lecture Notes in Computer Science, pp. 43–62, Santa Barbara, CA, USA, August 16-20, 2015. Springer, Heidelberg, Germany.

[74] P. Kirchner and P-A. Fouque. *Comparison between subfield and straightforward attacks on NTRU*, Cryptology ePrint Archive, Report 2016/717, 2016. `http://eprint.iacr.org/2016/717`.

[75] N. Koblitz. *Elliptic Curve Cryptosystems*. Mathematics of Computation, no. 177, 1987, pp. 203–209, 1987.

[76] P. Kocher. *Timing attacks on implementations of Diffie-Hellman, RSA, DSS and other systems*. Crypto 1996, LNCS 1109, pp. 104–113 (1996)

[77] P. Kocher, J. Jaffe, and B. Jun. *Differential power analysis*. Crypto 1999, LNCS 1666, pp. 388–397 (1999)

[78] K. Konrad. *The Gaussian integers*, preprint, available at `http://www.math.uconn.edu/~kconrad/blurbs/ugradnumthy/Zinotes.pdf`.

[79] K. Koyama, U. M. Maurer, T. Okamoto and S. A. Vanstone, *New public-key schemes based on elliptic curves over the ring* $\mathbb{Z}_n$, Advances in Cryptology - Crypto'91, Lecture Notes in Computer Science, Springer-Verlag, 252-266 (1991).

[80] K. Kurosawa, T. Ueda. *How to factor N_1 and N_2 when $p_1 = p_2$ (mod 2^t)*. In K. Sakiyama and M. Terada (Eds.): IWSEC 2013, Lecture Notes in Computer Science, Springer **8231** (2013) 217–225

[81] H. Kuwakado, K. Koyama, and Y. Tsuruoka. *A new RSA-type scheme based on singular cubic curves $y^2 = x^3 + bx^2$ (mod n)*, IEICE Transactions on Fundamentals, vol. E78-A (1995) pp. 27–33.

[82] Kurosawa, K., Okada, K., Tsujii, S.: Low exponent attack against elliptic curve RSA, Low exponent attack against elliptic curve RSA. Inform. Process. Lett. 53, no. 2, pp. 77-83 (1995)

[83] R.S. Lehman. *Factoring large integers.* Mathematics of Computation, Vol. 28, 637–646, (1974)

[84] H. Lenstra. *Factoring integers with elliptic curves*, Annals of Mathematics, Vol. 126, pp. 649–673 (1987)

[85] A.K., Lenstra, H.W. Lenstra, M.S. Manasse, J.M. Pollard. *The number field sieve*, Proc. 22nd Annual ACM Conference on Theory of Computing, pp. 564–572, Baltimore, Maryland (May 1990)

[86] A.K., Lenstra, H.W. Lenstra, L. Lovász. *Factoring polynomials with rational coefficients*, Mathematische Annalen, Vol. 261, pp. 513-534 (1982)

[87] T. Lepoint. *Design and Implementation of Lattice-Based Cryptography*, Ph.D. thesis (2014)
`https://www.cryptoexperts.com/tlepoint/thesis/lepoint-phd-thesis.pdf`.

[88] Y. Lu, R. Zhang,D. Lin. *New Results on Solving Linear Equations Modulo Unknown Divisors and its Applications*, Cryptology ePrint Archive, Report 2014/343, 2014 `https://eprint.iacr.org/2014/343`.

[89] V. Lyubashevsky, C. Peikert, O. Regev. *On Ideal Lattices and Learning with Errors over Rings.* Advances in cryptology-Eurocrypt 2010, pp. 1–23, Lecture Notes in Comput. Sci., 6110, Springer, Berlin, 2010.

[90] Maple, http://www.maplesoft.com/products/maple/

[91] A. May. *New RSA Vulnerabilities Using Lattice Reduction Methods*, Ph.D. thesis, Paderborn, 2003, `http://www.cits.rub.de/imperia/md/content/may/paper/bp.ps`

[92] A. May. *Secret Exponent Attacks on RSA-type Schemes with Moduli $N = p^r q$*. Public Key Cryptography–PKC 2004, LNCS 2947, pp. 218–230, (2004)

[93] A. May. *Using LLL-reduction for solving RSA and factorization problems: a survey*. In: LLL+25 Conference in Honour of the 25th Birthday of the LLL Algorithm. Springer, Berlin, Heidelberg (2007)

[94] A. May., R. Ritzenhofen. *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** pp. 1–14, (2009)

[95] V. Miller. *Uses of elliptic curves in cryptography*. Advances in Cryptology-CRYPTO'85, LNCS 218, 1986, pp. 417–426, 1986.

[96] P.L. Montgomery. S*peeding the Pollard and elliptic curve methods of factorization*, Mathematics of Computation, vol. 48, 243–264 (1987)

[97] M.A. Morrison, J. Brillhart. *A method of factoring and the factorization of F7*, Math. of Comput., t. 29, pp. 183–205 (1975)

[98] M. Naehrig, K. Lauter, V. Vaikuntanathan. *Can homomorphic encryption be practical?*, Proceeding CCSW'11 Proceedings of the 3rd ACM workshop on Cloud computing security workshop, pp. 113–124.

[99] D.I. Nassr, H.M. Bahig, A. Bhery, S.S. Daoud. *A new RSA vulnerability using continued fractions*. In Proceedings of AICCSA. 2008, pp. 694-701.

[100] D.I. Nassr, H.M. Bahig, A. Bhery, A. Nitaj. Another Look at Private Exponent Attack on RSA using Lattices, To appear in International Journal of Applied and Computational Mathematics.

[101] P.Q. Nguyen and B. Vallée (Eds.), *The LLL Algorithm: Survey and Applications*, Springer (Series: Information Security and Cryptography), 2009.

[102] P.Q. Nguyen and D. Stehlé. *An LLL algorithm with quadratic complexity*. SIAM J. of Computing, 39(3), pp. 874–903 (2009).

[103] NIST Special Publication 800-56B Revision 1 `http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-56Br1.pdf`

[104] A. Nitaj. *Another generalization of Wiener's attack on RSA*, in Vaudenay, S. (ed.) Africacrypt 2008. Lecture Notes in Computer Science, Springer-Verlag Vol. 5023, pp. 174–190 (2008)

[105] A. Nitaj. *Cryptanalysis of RSA with constrained keys*. Int. J. Number Theory 5 (2009), no. 2, pp. 311–325 (2009)

[106] A. Nitaj. *Application of ECM to a class of RSA keys*, Journal of Discrete Mathematical Sciences & Cryptography , 12, No. 2, pp. 121–137 (2009)

[107] A. Nitaj. *Cryptanalysis of RSA using the ratio of the primes*. Progress in cryptology-AFRICACRYPT 2009, pp. 98–115, Lecture Notes in Comput. Sci., 5580, Springer, Berlin, (2009)

[108] A. Nitaj. *A new vulnerable class of exponents in RSA*. JP J. Algebra Number Theory Appl. 21 (2011), no. 2, pp. 203–220 (2011).

[109] A. Nitaj. *New weak RSA keys*, JP Journal of Algebra, Number Theory and Applications. Volume 23, Number 2, 2011, pp. 131–148 (2011)

[110] A. Nitaj. *A new attack on RSA and CRT-RSA*. Progress in cryptology-AFRICACRYPT 2012, pp. 221–233, Lecture Notes in Comput. Sci., 7374, Springer, Heidelberg, (2012)

[111] A. Nitaj. *A new attack on RSA with two or three decryption exponents*. J. Appl. Math. Comput. 42 (2013), no. 1-2, pp. 309–319 (2013)

[112] A. Nitaj. *An attack on RSA using LSBs of multiples of the prime factors*. Progress in Cryptology-AFRICACRYPT 2013, pp. 297–310, Lecture Notes in Comput. Sci., 7918, Springer, Heidelberg, 2013.

[113] A. Nitaj, M. Ould Douh. *A new attack on RSA with a composed decryption exponent*, International Journal on Cryptography and Information Security (IJCIS), Vol.3, No. 4, December 2013

[114] A. Nitaj, M.R.K. Ariffin, D.I. Nassr, H.M. Bahig. *New attacks on the RSA cryptosystem. Progress in cryptology-AFRICACRYPT 2014*, pp. 178–198, Lecture Notes in Comput. Sci., 8469, Springer, Cham, (2014)

[115] A. Nitaj, M.R.K. Ariffin. *Implicit factorization of unbalanced RSA moduli*. J. Appl. Math. Comput. 48 (2015), no. 1-2, pp. 349–363 (2015)

[116] A. Nitaj, T. Rachidi. *New attacks on RSA with moduli $N = p^r q$*. Codes, Cryptology, and Information Security, pp. 352–360, Lecture Notes in Comput. Sci., 9084, Springer, Cham, (2015)

[117] A. Nitaj, T. Rachidi. *Factoring RSA moduli with weak prime factors*. Codes, Cryptology, and Information Security, 361-374, Lecture Notes in Comput. Sci., 9084, Springer, Cham, 2015.

[118] A. Nitaj. *Cryptanalysis of NTRU with two Public Keys*, International Journal of Network Security, 16(2), pp. 112-117 (2014)

[119] A. Nitaj. *A new attack on the KMOV cryptosystem*. Bull. Korean Math. Soc. 51 (2014), no. 5, pp. 1347–1356.

[120] A. Nitaj, W. Susilo, J. Tonien. *Dirichlet Product for Boolean Functions*, To appear in Journal of Applied Mathematics and Computing, 2016.

[121] A. Nitaj, E. Fouotsa. *A New Attack on RSA and Demytko's Elliptic Curve Cryptosystem*, Submitted to Mathematics in Computer Science.

[122] A. Nitaj, T. Rachidi. *Lattice Attacks on the Homomorphic DGHV Scheme*, Submitted to Discrete Applied Mathematics.

[123] A. Nitaj, M.R.K. Ariffin: Generalizations of Former Attacks on RSA, Submitted to Journal of Mathematical Cryptology.

[124] P. Paillier. *Public-key cryptosystems based on composite degree residuosity classes*. In J. Stern (Ed.), EUROCRYPT'99, LNCS, vol. 1592, Spring, 1999, pp. 223–238 (1999)

[125] PARI/GP, version `2.1.7`, Bordeaux, 2007, `http://pari.math.u-bordeaux.fr/`

[126] S. Paulus, T. Takagi, T.: *A new public key cryptosystem over quadratic orders with quadratic decryption time*. J. Cryptology 13, 263–272 (2000)

[127] J. Pieprzyk and X.M. Zhang, *Computing Möbius Transforms of Boolean Functions and Characterizing Coincident Boolean functions*, *Proceedings of the International Conference on Boolean Functions: Cryptography and Applications 2007*

[128] R.G.E. Pinch, *Extending the Wiener attack to RSA-type cryptosystems*, Electronics Letters 31, 1736-1738 (1995).

[129] J.M. Pollard. *A Monte Carlo method for factorization.* BIT 15, pp. 331–334 (1975)

[130] R. L. Rivest, L. Adleman, and M. L. Dertouzos. *On data banks and privacy homomorphisms.* Foundations of Secure Computation, Academia Press, pp. 169–179, 1978.

[131] R. Rivest, A. Shamir, L. Adleman. *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM, Vol. 21 (2), 120–126 (1978)

[132] S. Sarkar. *Small secret exponent attack on RSA variant with modulus* $N = p^r q$, Designs, Codes and Cryptography, Volume 73, Issue 2 , pp 383–392 (2015)

[133] S. Sarkar, S. Maitra. *Cryptanalysis of RSA with two decryption exponents*, Inform. Process. Lett. 110 (5) pp. 178–181 (2010)

[134] S. Sarkar, S. Maitra. *Cryptanalysis of RSA with more than one decryption exponent*, Inform. Process. Lett. 110 (8-9) pp. 336–340 (2010)

[135] S. Sarkar, S. Maitra. *Further results on implicit factoring in polynomial time.* Advances in Mathematics of Communications **3** (2009) 205–217

[136] P. Sarkar, S. Maitra: *Construction of Nonlinear Boolean Functions with Important Cryptographic Properties*, In EUROCRYPT 2000, number 1807 in Lecture Notes in Computer Science, pages 485–506. Springer Verlag, May 2000.

[137] R. Schoof. *Counting points on elliptic curves over finite fields*, Journal de Théorie des Nombres de Bordeaux, 7(1):219-254, (1995)

[138] A. Shamir. *RSA for Paranoids*. RSA Laboratories CryptoBytes **1** (1995) pp. 3–4.

[139] P.W. Shor. *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, *SIAM J. Computing* 26, pp. 1484–1509, 1997.

[140] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag, GTM 106, 1986. Expanded 2nd Edition, (2009).

[141] J. Hoffstein, J. Pipher, J.H. Silverman. *An Introduction to Mathematical Cryptography*, Springer-Verlag - Undergraduate Texts in Mathematics, 2014.

[142] R.D. Silverman. *The multiple polynomial quadratic sieve*, Mathematics of Computation, Vol. 48, pp. 329–339 (1987)

[143] P.J. Smith, G.J.J. Lennon. *LUC: a new public-key cryptosystem*, Ninth IFIP Symposium on Computer Science Security, Elseviver Science Publishers, 1993, 103-117.

[144] D. Stehlé and R. Steinfeld. *Making NTRU as Secure as Worst-Case Problems over Ideal Lattices*, in Kenneth G. Paterson (Ed.): *Advances in Cryptology - Eurocrypt 2011*, Lecture Notes in Comput. Sci., 6632, pp. 27–47, Springer, 2011.

[145] T. Takagi. *Fast RSA-Type Cryptosystem Modulo $p^k q$*, in Advances in Cryptography - Proceedings of CRYPTO 1998, pp. 318–326, volume 1462, Lecture Notes in Computer Science series, (1998).

[146] B. de Weger. *Cryptanalysis of RSA with small prime difference*, Applicable Algebra in Engineering, Communication and Computing, Vol. 13(1), pp. 17-28.

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

[148] S.Y. YAN. *Cryptanalytic Attacks on RSA*, Springer 2008.

[149] B. Yang, H. Ma and S. Zhu, *A Traitor Tracing Scheme Based on the RSA System*, International Journal of Network Security (IJNS), Vol.5, No.2, pp.182–186, 2007.

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