Increment of Insecure RSA Private Exponent Bound Through Perfect Square RSA Diophantine Parameters Cryptanalysis

Wan Nur Aqlili Ruzai Ruzai, Abderrahmane Nitaj, Muhammad Rezal Kamel Ariffin, Zahari Mahad, Muhammad Asyraf Asbullah

Increment of Insecure RSA Private Exponent Bound Through Perfect Square RSA Diophantine Parameters Cryptanalysis

Wan Nur Aqlili Ruzai[a], Abderrahmane Nitaj[b], Muhammad Rezal Kamel Ariffin[a,c,*], Zahari Mahad[a], Muhammad Asyraf Asbullah[a]

[a] *Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[b] *Nicolas Oresme Mathematics Laboratory, University of Caen - Normandy, France*
[c] *Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, Selangor, Malaysia*

Abstract

The public parameters of the RSA cryptosystem are represented by the pair of integers N and e. In this work, first we show that if e satisfies the Diophantine equation of the form $ex^2 - \phi(N)y^2 = z$ for appropriate values of x, y and z under certain specified conditions, then one is able to factor N. That is, the unknown $\frac{y}{x}$ can be found amongst the convergents of $\frac{\sqrt{e}}{\sqrt{N}}$ via continued fractions algorithm. Consequently, Coppersmith's theorem is applied to solve for prime factors p and q in polynomial time. We also report a second weakness that enabled us to factor k instances of RSA moduli simultaneously from the given (N_i, e_i) for $i = 1, 2, \cdots, k$ and a fixed x that fulfills the Diophantine equation $e_i x^2 - y_i^2 \phi(N_i) = z_i$. This weakness was identified by solving the simultaneous Diophantine approximations using the lattice basis reduction technique. We note that this work extends the bound of insecure RSA decryption exponents.

Keywords: RSA cryptosystem, algebraic cryptanalysis, integer factorization problem, Diophantine approximation, lattice basis reduction, kleptography

*Corresponding author
Email address: rezal@upm.edu.my (Muhammad Rezal Kamel Ariffin)

1. Introduction

Since the mid-1990s, the Internet has greatly influenced culture, commerce, and technology, including the advent of the World Wide Web (WWW) with its social networking services, discussion forums, blogs, and online shopping sites [1]. To date, the Internet population has risen from 144 million in 1998 to approximately 4.66 billion as of January 2021. Thus, the execution of information transfer over multiple channels in our daily life has demanded an efficient exchange of secure information [2]. This prime need for information security has led to the emergence of a variety of cryptographic algorithms to implement security in different dimensions and for various purposes [3].

The introduction of asymmetric cryptography in the seminal work of Diffie and Hellman (1976) [4] and consecutively the invention of the first practical asymmetric cryptosystem known as RSA by Turing Award winners; Rivest, Shamir, and Adleman [5] in 1978 are major breakthroughs within the lengthy history of secret communications. Since its development, RSA became one of the widely accepted public-key cryptosystems and is being utilized in most web servers with the goals of providing security, privacy, and authenticity of digital data. RSA is commonly used to secure web traffic, remote login sessions, e-mail, e-commerce, and smart cards. Recently, efforts have been reported for the RSA to be implemented as the underlying security for blockchain technology with the aim to construct a redactable blockchain structure [6].

The RSA cryptosystem comprises of key generation, encryption, and decryption algorithms. The RSA key generation process includes selecting two arbitrary random primes p and q. With these two primes, we calculate Euler's totient function defined by $\phi(N) = (p-1)(q-1)$ and the product $N = pq$ which is referred to as the RSA modulus. Then, an integer $e < \phi(N)$ is chosen such that $\gcd(e, \phi(N)) = 1$. Note that, it satisfies the key equation of RSA defined by $ed - k\phi(N) = 1$ with positive integers d and k. The integers e and d are also known as the public and private exponents of RSA respectively. Finally, the algorithm outputs the public key pair (N, e) and the private key tuple

$(p, q, d, \phi(N))$. To execute the RSA encryption algorithm, one simply computes the ciphertext given by $C \equiv M^e \pmod{N}$ where $M \in \mathbb{Z}_N^*$. On the other hand, one simply computes $M \equiv C^d \pmod{N}$ to execute the RSA decryption process and retrieve back the message M.

Generally, the security of RSA is entrenched in the hardness of solving the integer factorization problem (IFP) from a given large integer N and solving the e^{th} root modulo N problem. In addition, attacks are also proposed upon RSA, which manipulate the key equation structure being used. To date, the general number field sieve (GNFS) is the most efficient algorithm to factor N and is still running in sub-exponential time since the chosen RSA primes p and q are large n-bit primes (usually $n = 1024$) [7]. In practice, the second fastest algorithm is the quadratic sieve (QS) which also runs in sub-exponential time [8]. QS algorithm is considered simpler than GNFS algorithm and still the fastest for integers below 100 decimal digits but not better than GNFS algorithm for integers with 110-120 digits. Thus, no contemporary computers yet can potentially threaten the security of RSA.

Since its invention, studies on improving the efficiency of RSA's decryption execution time and its relation to RSA's overall security features are discussed in-depth by the cryptographic research community. As a result, many variants of RSA were proposed to address all the possible vulnerabilities. There are some variants that are of interest for efficiency reasons [9, 10]. As new weaknesses are revealed, new solutions will be developed [11].

In some applications of RSA, a small decrypting exponent d is desired as it speeds up the modular exponentiation computation during decryption or signing process. However, Wiener [12] discovered that RSA is vulnerable wherever $d < \frac{1}{3}N^{0.25}$ for each exponent e that fulfills $ed - 1 = 0 \mod \phi(N)$, by proving that the unknown $\frac{k}{d}$ is one of the convergents of $\frac{e}{N}$ via the continued fractions. In a different setting, Boneh and Durfee [13] showed that RSA can be heuristically cryptanalysed whenever the decrypting exponent $d < N^{0.292}$ via the Coppersmith's lattice-based technique. Both attacks on RSA manipulated the original structure of the RSA key equation given by $ed - k\phi(N) = 1$.

Another approach is via the modified version of the RSA key equation. In particular, Blömer and May [14] exploited the equation $ex + y = k\phi(N)$ to show that for each public pair (N, e) which fulfills $ex + y = 0 \mod \phi(N)$ with conditions

$$0 < x \leq \frac{1}{3}N^{0.25} \quad \text{and} \quad |y| = \mathcal{O}\left(\frac{ex}{N^{0.75}}\right),$$

one can obtain the prime factors p and q in polynomial time. This strategy combines the continued fractions and Coppersmith's theorem on finding the small solutions of modular polynomial equations.

In 2013, Nitaj [15] reconstructed the attack in [14] and considered the case when the modulus $N = pq$ is a product of unbalanced primes. Nitaj shows that if every public key pair (N, e) fulfills the equation $ex - y\phi(N) = z$ such that $\gcd(x, y) = 1$ with the conditions

$$xy < \frac{N}{4(p+q)} \quad \text{and} \quad |z| < \frac{(p-q)N^{0.25}y}{3(p+q)},$$

then the primes p and q can be determined in polynomial time. As a result, the unknown $\frac{y}{x}$ is found in one of the continued fraction's convergents of $\frac{e}{N}$. Then, it is proven that $\left|p - \frac{U+V}{2}\right| < N^{0.25}$ due to the Coppersmith's theorem where the terms U and V are defined as

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

Once p is known, then it completes the factoring problem of N (i.e. $q = \frac{N}{p}$).

Nitaj [16] also proves that RSA is insecure by providing that the modulus $N = pq$ can be factored in the presence of two or three public exponents e_i which satisfy the Diophantine equations $e_i x_i - \phi(N)y_i = z_i$. That is, one can find the factorization of a common modulus N depending on certain inequalities verified by the parameters x_i, y_i, z_i via continued fractions algorithm and the Coppersmith's lattice based technique. Additionally, Nitaj and Ariffin [17] have designed a strategy to address the problem of implicit factoring of $k \geq 2$ RSA moduli. More precisely, the RSA moduli $N_i = p_i q_i$, $i = 1, \cdots, k$ can be factored simultaneously if some unknown multiples $\alpha_i p_i$ share j amount of Least Significant Bits (LSBs) or share j amount of Most Significant Bits (MSBs).

Their attacks are based on the continued fractions algorithm and lattice basis reduction.

In 2019, Ariffin et al. [18] exploited the original structure of the RSA key equation $ed - k\phi(N) = 1$ to demonstrate that RSA is vulnerable if $d < \frac{\sqrt{3}}{\sqrt{2}} N^{0.75-\delta}$ where $0.25 < \delta \leq 0.5$. This attack utilized the small prime difference method in the form $|b^2 p - a^2 q| < N^{\delta}$ and computed the continued fractions expansion of $\frac{e}{N+1-\left\lceil \frac{a^2+b^2}{ab}\sqrt{N} \right\rceil}$ to obtain the private parameter $\frac{k}{d}$ among its convergents. Additionally, the authors also show that by using $N+1-\left\lceil \frac{a^2+b^2}{ab}\sqrt{N} \right\rceil$ as an approximation of $\phi(N)$, they can factor the k instances of RSA moduli simultaneously in polynomial time from the given public key pairs (N_i, e_i) for $i = 1, 2, \cdots, k$.

Kleptography refers to the study of stealing information securely and subliminally [30]. In particular, our cryptanalytic approaches will identify the conditions for weak keys that should be avoided in the usage of the RSA cryptosystem. Our results show that there are classes of public keys (N, e), which yields the factorization of N. By publishing these weak keys, a rogue certificate authority (CA) can produce a fraudulent RSA digital certificate without being noticed by the users of its peculiarity. The validity of these rogue certificates seems convincing since the weak keys satisfy the conditions dictated during RSA key generation. Suppose an adversary knows about the existence of these particular certificates. Then, the adversary can find the RSA private keys corresponding to the public keys, although no information on the private keys is disclosed to the adversary [31]. This whole situation shows that the adversary (i.e. kleptographic attacker) is able to steal the private information securely, and in an exclusive and subliminal manner.

Our contribution. In this work, we present a new weak RSA key equation structure that wound render the factorization of modulus N using the combination of continued fractions algorithm and Coppersmith's method feasible in polynomial time. Precisely, we show that if e satisfies the Diophantine equation of the form $ex^2 - \phi(N)y^2 = z$ for appropriate values of x, y and z under cer-

tain specified conditions, then one is able to factor the RSA modulus N. We also consider the system of modified generalized RSA key equation that potentially contribute to solve the k instances of RSA moduli simultaneously from the given (N_i, e_i) for $i = 1, 2, \cdots, k$. This attack involves solving the equation $e_i x^2 - y_i^2 \phi(N_i) = z_i$ simultaneously using lattice basis reduction technique.

Organization of the article. Section 1 introduces the RSA cryptosystem and some former attacks launched on RSA. Section 2 provides a brief review of the theory of the continued fractions, Coppersmith, and simultaneous Diophantine approximation and also presents several significant results required in our cryptanalysis method. Section 3 presents the new attacks and also includes numerical examples of the proposed attacks. Section 4 compares our attacks against some existing attacks. Finally, concluding remarks are made in Section 5.

2. Preliminaries

In this section, we provide several results that will be used in the rest of the paper. We begin by reviewing the fundamental knowledge on the continued fractions.

Definition 1. For any $X \in \mathbb{R}$, the continued fractions expansion of X is defined by the expression of the form

$$X = [a_0, a_1, a_2, \cdots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

where $a_0 \in \mathbb{Z}$ and $a_i \in \mathbb{N}$ for $i > 0$.

As noticed from Definition 1, the numbers $a_1, a_2, a_3, \cdots$ represent partial quotients of X. When $i \geq 0$, the fractions $\frac{r_i}{s_i} = [a_0, a_1, a_2, \cdots, a_i]$ represent convergents of an expansion X.

There is a standard way to generate a unique continued fraction from any rational number which is via the Euclidean algorithm [19]. When $X = \frac{a}{b}$ where a and b are co-prime, the Euclidean algorithm will compute the convergents in polynomial time such that its complexity is $\mathcal{O}((\log \max(a, b)))$ [20].

In the notable work of Legendre [21], the next theorem ensures that a rational number $\frac{a}{b}$ can be found amongst the convergents of ξ.

Theorem 1. Suppose ξ is a rational number. Suppose a and b are the positive integers such that $b \neq 0$ and $\gcd(a, b) = 1$. If

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

then $\frac{a}{b}$ is one of the convergents of expansion ξ via continued fractions.

Proof. See [21]. □

The following result by Coppersmith [22] assures that one is able to identify the remaining bits of p given that half of the most significant bits of p is known.

Theorem 2. Let $N = pq$ represents the modulus of RSA with balanced primes (i.e. $q < p < 2q$). Suppose p is approximated to $\tilde{p}$ with the difference between the terms is $|p - \tilde{p}| < N^{0.25}$, then the prime p could be obtained in polynomial time.

Proof. See [23]. □

Consequently, any knowledge on the approximation of $p + q$ will result in the knowledge to approximate p.

Lemma 1. Let $N = pq$ represents the modulus of RSA with balanced primes. Assume $p + q$ is approximated to S where $S > 2\sqrt{N}$ and

$$|p + q - S| < \frac{p - q}{3(p + q)} N^{0.25}.$$

Then p is approximated to $\tilde{P} = \frac{1}{2}\left(S + \sqrt{S^2 - 4N}\right)$ satisfying $|p - \tilde{P}| < N^{0.25}$.

Proof. See [24]. □

Among the highlight of utilizing the LLL algorithm [25] is its ability to solve the simultaneous diophantine approximations problem that can be defined as follows.

Theorem 3. Suppose that for a given rational numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$ and $0 < \varepsilon < 1$, there exist a polynomial time algorithm with respect to $\log(p_i)$ to output a list of integers p_i where $i = 1, 2, \cdots, k$ and $q \in \mathbb{Z}$ when

$$\max_i |q\alpha_i - p_i| < \varepsilon \quad \text{and} \quad q \leq 2^{k(k-3)/4} \cdot 3^k \cdot \varepsilon^{-k}.$$

Proof. See [24]. □

The following theorem provide the ability to factor the primes p_i and q_i from the solution to the system of equations of multiple RSA moduli.

Theorem 4. Let $N_i = p_i q_i$ be k RSA moduli where $N = \min_i N_i$ and e_i be k public exponents such that $i = 1, 2, \cdots, k$ for $k \geq 2$. Define $\delta = \frac{k}{2(k+1)}$. If there exist an integer $x < N^\delta$ and k integers $y_i < N^\delta$ and $|z_i| < \frac{p_i - q_i}{3(p_i + q_i)} y_i N^{\frac{1}{4}}$ such that $e_i x - y_i \phi(N_i) = z_i$ for $i = 1, \cdots, k$, then one can factor the k RSA moduli $N_1, \cdots, N_k$ within polynomial time.

Proof. See [24]. □

Next, we provide significant lemmas that enabled us to identify the new weaknesses.

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

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

Proof. See [26]. □

Based on Lemma 2, we easily obtain

$$\phi(N) = N + 1 - (p + q) > N + 1 - \frac{3\sqrt{2}\sqrt{N}}{2} > \frac{1}{2}N.$$

Lemma 3. Let $N = pq$ represents the modulus of RSA with $q < p < 2q$. Then

$$\sqrt{N} - \sqrt{\phi(N)} < \sqrt{2}.$$

Proof. See Appendix A. □

3. The New Weaknesses

In this section, we provide results that outline new weaknesses of the RSA cryptosystem. We discuss the weaknesses in the following two sections.

3.1. The First Weakness

We now present the newly proposed cryptanalytic strategies. The main idea is to find the unknown $\frac{y}{x}$ that can be found amongst the convergents of $\frac{\sqrt{e}}{\sqrt{N}}$ via continued fractions algorithm. Then, the unknown term $p+q$ can be approximated to S by computing $S = \left[N + 1 - \frac{ex^2}{y^2}\right]$ with the obtained values of x and y. By knowing S, it has led to an approximation of p given by $\tilde{p} = \left[\frac{1}{2}(S+\sqrt{S^2-4N})\right]$. Hence, we obtain the prime factor p via Coppersmith's theorem. Then, $q = \frac{N}{p}$. Our cryptanalytic strategy is formally described as follows.

Theorem 5. Let $N = pq$ and e where $e > \sqrt{N}$ be RSA public parameters. Let $ex^2 - \phi(N)y^2 = z$ where

$$x < \frac{\sqrt{2}}{4}N^{\frac{1}{2}}e^{-\frac{1}{4}}, \quad y < 2^{\frac{1}{4}}N^{\frac{-3}{8}}\sqrt{e}x, \quad \text{and} \quad |z| < N^{\frac{1}{4}}y^2,$$

then N can be factored within polynomial time.

Proof. See Appendix B. □

Next, we want to estimate the number of exponents e satisfying the conditions stated in Theorem 5.

Proposition 1. The number of exponents e satisfying an equation of the form

$$ex^2 - \phi(N)y^2 = z$$

is at least $\frac{1}{4}N^{\frac{1}{2}}$.

Proof. See Appendix C. □

We now provide Algorithm 1 to describe the process of factoring $N = pq$ via the strategy specified in Theorem 5.

Algorithm 1: Factoring the modulus $N = pq$ via Theorem 5

Input: The public parameters N and e.

Output: The primes p and q or $\perp$.

1: Calculate the continued fractions expansion of $\frac{\sqrt{e}}{\sqrt{N}}$.

2: For each convergent $\frac{y}{x}$ of $\frac{\sqrt{e}}{\sqrt{N}}$,
calculate $S = \left\lceil N + 1 - \frac{ex^2}{y^2} \right\rceil$ and $\tilde{p} = \left[\frac{1}{2}(S + \sqrt{S^2 - 4N})\right]$.

3: Define $F(b) = (b + \tilde{p})$ and consider the polynomials with the same small root modulo p.

4: Construct a matrix $\mathcal{M}$ corresponding to the polynomials in Step 3.

5: Apply the lattice basis reduction algorithm onto $\mathcal{M}$.

6: Form the polynomial $\mathcal{M}^{'}(b)$ via the first row of output in Step 5.

7: Calculate the roots of $\mathcal{M}^{'}(b)$ to obtain small solution b_0.

8: Calculate $p = b_0 + \tilde{p}$ and $q = \frac{N}{p}$.

9: **if** q is an integer, **then** return the primes p, q.

10: **else** output $\perp$.

3.1.1. A Working Example

This section shows a numerical illustration of our first attack by running Algorithm 1.

Example 1. On input of the RSA public key pair (N, e) with respect to the conditions specified in Theorem 5,

$N =$380796130864151153360147597642268363473847778072739943117355700808322662820686653060328997 3,

$e =$169242724828510773524799469193129126896720838736474743950760396714587897009294134056628747 9,

we proceed to calculate the continued fractions expansion of $\frac{\sqrt{e}}{\sqrt{N}}$ to obtain a list

of its convergents

$$\Bigg[0, 1, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{132709599234851}{199064398852277}, \tfrac{398128797704555}{597193196556834}, \tfrac{14863475114303386}{22295212671455135}, \tfrac{15261603912007941}{22892405868011969},$$

$$\ldots, \tfrac{325677584517478536492}{488516376776219031769}, \tfrac{935463993074503788901}{1403195989611759207829}, \tfrac{2196605570666486114294}{3294908355999737447427},$$

$$\tfrac{3132069563740989903195}{4698104345611496655256}, \tfrac{1159281426188945582387 9}{1738922139283422741319 5}, \tfrac{1472488382563044572707 4}{2208732573844572406845 1},$$

$$\tfrac{4104258191315034727802 7}{6156387286972567555009 7}, \tfrac{5576746573878079300510 1}{8365119860817139961854 8}, \tfrac{5429497735621774843239 36}{8144246603432682721170 29},$$

$$\tfrac{2770516333549668214624781}{4155774500324512760203693}, \tfrac{6083982440661513913573498}{9125973660992293792524415}, \tfrac{3373895505699329334611671 71}{5060843258549006713490465 18}, \ldots\Bigg].$$

Then, we calculate the approximation S of $p + q$ by taking the convergent $\frac{y}{x} = \frac{1472488382563044572707 4}{2208732573844572406845 1}$ which result in

$$\begin{aligned} S &= \left\lfloor N + 1 - \frac{ex^2}{y^2} \right\rfloor \\ &= 401255480710077217056953576233736692947773391 2, \end{aligned}$$

from which we obtain an integer $\tilde{P}$ such that

$$\begin{aligned} \tilde{P} &= \left\lfloor \frac{1}{2}\left(S + \sqrt{S^2 - 4N}\right) \right\rfloor \\ &= 247231142736771125909332226957764115316428870 3. \end{aligned}$$

Now we apply the Coppersmith's theorem to factor $N = pq$ with the partial knowledge of p from $\tilde{P}$. Let $F(b) = (b + \tilde{P})$. Suppose the upper bound of the unknown $|p - \tilde{P}|$ is $B = 87005818903900$.

Let us consider the polynomials $N, F(b), b \cdot F(b)$ and $b^2 \cdot F(b)$, which all share similar root b_0 modulo p [32]. Then, we construct a matrix $\mathcal{M}$ corresponding to these polynomials. In particular,

$$\mathcal{M} = \begin{bmatrix} N & 0 & 0 & 0 \\ \tilde{P} & B & 0 & 0 \\ 0 & \tilde{P} \cdot B & B^2 & 0 \\ 0 & 0 & \tilde{P} \cdot B^2 & B^3 \end{bmatrix}.$$

Let $\mathcal{M}_{LLL}$ be the LLL-reduced basis matrix. We take each element from the first row of matrix $\mathcal{M}_{LLL}$ which denotes the coefficients of polynomial $\mathcal{M}'(b)$ where

$$\mathcal{M}'(b) = 2b^3 - 2656983079459352b^2 - 28336351428068984089405787092b + 3159496701891952077970076005170894373378560.$$

From here, we find the integer root of $\mathcal{M}'(b)$ which yields

$$b = 87005818903680.$$

Observe that,

$$p = b + \tilde{P} = 247231142736771125909332226957772815898319238 3.$$

Next, we solve for q such that

$$q = \frac{N}{p} = 154024337973306091147621349275967157192192673 1.$$

Thus, we complete the factorization of N.

Remark 1. In our case, observe that $x^2 > N^{0.25}$ where $x^2 \approx N^{0.493}$. This shows that the attack of Blömer and May in [14] will not succeed to find the primes p and q.

3.2. The Second Weakness

In 2003, the Taiwan government introduced the Taiwanese digital ID for all citizens as an initiative to support the national public key infrastructure [27]. The RSA keys are generated and inserted into the cards, digitally signed by a government authority, and also stored into an online database called "Citizen Digital Certificates (CDCs)". The physical cryptographic roll out for the Taiwanese ID was advertised as having passed the FIPS 140-2 Level 2 and the Common Criteria standards. However, in 2013, Bernstein et al. [27] showed that an adversary can efficiently solve the factorization of 184 distinct 1024-bit RSA keys downloaded from Taiwan's CDSs database. This attack was a success due to Coppersmith's partial-key-recovery attack on a bulk of weak RSA

keys used. One has to take note that the underlying idea is to identify weak characteristics within a collection of public RSA keys.

An exact scenario occured on Estonia's digital identity card. In 2017, Nemec et al. [28] crippled millions of high-security crypto keys, specifically the RSA keys, by providing a practical factorization methodology on RSA moduli utilized. Precisely, [28] proposed a feasible method based on a refined version of Coppersmith's method (i.e. Howgrave-Graham) [29] to factor a collection of RSA moduli N with various key lengths including 1024 and 2048 bits. The authors reported that all the vulnerable keys can be quickly identified, even in very large datasets.

As such, continuous research on scenarios which involves collection of RSA public keys is of utmost important in order to avoid vulnerable situations in real life applications.

By the above motivation, in this paper we highlight a second weakness on the RSA. This newly identified weakness is the case when the adversary obtains a set of weak RSA public key pairs that would render the factorisation of each N simultaneously in polynomial time feasible. That is, the collection of RSA keys will have its parameters satisfying the system of equations given by $e_i x^2 - y_i^2 \phi(N_i) = z_i$ where $x^2 \in \mathbb{Z}$. Note that, this weakness was identified through solving the simultaneous Diophantine approximations using the lattice basis reduction technique. Our cryptanalytic strategy is formally described as follows.

Theorem 6. Let k RSA moduli be denoted by $N_i = p_i q_i$ for $i = 1, 2, \cdots, k$. Let $N = \min\{N_i\}$ and e_i be k public exponents. If there exist an integer $x^2 < N^{\delta}$, k integers $y_i^2 < N^{\delta}$ and $|z_i| < \frac{p_i - q_i}{3(p_i+q_i)} y_i^2 N^{\frac{1}{4}}$ where $\delta = \frac{k}{2(k+1)}$ that satisfies the equation $e_i x^2 - y_i^2 \phi(N_i) = z_i$, then one can factor the k RSA moduli simultaneously within polynomial time.

Proof. See Appendix D. □

Next, we provide Algorithm 2 to demonstrate the process of factoring $N = pq$ via the strategy specified in Theorem 6.

Algorithm 2: Factoring k RSA moduli simultaneously via Theorem 6

Input: The public RSA key pairs (N_i, e_i) for $i = 1, 2, 3, \cdots, k$.

Output: The prime factors p_i and q_i or $\perp$.

1: Set $N = \min(N_1, N_2, \cdots, N_k)$.
2: Compute $\delta = \frac{k}{2(k+1)}$.
3: Compute $\varepsilon = \sqrt{5}N^{\delta - \frac{1}{2}}$.
4: Compute $C = \left[3^{k+1} \cdot 2^{\frac{(k+1)(k-4)}{4}} \cdot \varepsilon^{-k-1}\right]$.
5: Compute lattice $\mathcal{L}$ spanned by the the rows of the matrix $\mathcal{M}$ as shown in the proof of Theorem 3.
6: Compute matrix $\mathcal{K}$ by applying LLL algorithm onto $\mathcal{M}$.
7: Compute matrix $\mathcal{H} = \mathcal{K}\mathcal{M}^{-1}$.
8: Assign each element in the first row of $\mathcal{H}$ (starting from most left) as $X, Y_1, \cdots, Y_k$ respectively.
9: **for** $i = 2, , 3, \cdots, k$ **do**
10: Compute $S_i = \left[N + i + 1 - \frac{e_i X}{Y_i}\right]$.
11: Compute $\tilde{P}_i = \frac{1}{2}\left(S_i + \sqrt{S_i^2 - 4N_i}\right)$.
12: Apply the Coppersmith's method in Theorem 2 onto P_i to output p_i.
13: Compute $q_i = \frac{N_i}{p_i}$.
14: **if** $q_i \in \mathbb{Z}$, **then** output p_i, q_i.
15: **else** Algorithm fails or $\perp$.
16: **end for**

3.2.1. A Working Example

Next, we provide a numerical illustration of the proposed attack according to Theorem 6 and Algorithm 2.

Example 2. Let us consider the following three pairs of RSA moduli and its corresponding public exponents:

$$N_1 = 595320594314653717, \quad e_1 = 87693103457412687,$$
$$N_2 = 8324526487133150381, \quad e_2 = 4884469043442535176,$$
$$N_3 = 1939786546012035661, \quad e_3 = 192778371459086335.$$

Then, we set $N = \min(N_1, N_2, N_3) = 595320594314653717$. Since in this case of k equal to 3, we obtain $\delta = 0.3750$ and $\varepsilon = \sqrt{5}N^{\delta-0.5} \approx 0.01342$.

Next, we use the formula stated in proof of Theorem 4 in [24] to compute

$$C = [3^{k+1} \cdot 2^{\frac{(k+1)(k-4)}{4}} \cdot \varepsilon^{-k-1}] = 4674184425.$$

We also compute $c_i = \left[-\frac{C \cdot e_i}{(N_i+1)} \right]$ for $i = 1, 2, 3$ and get

$$c_1 = -688526052, \quad c_2 = -2742607542, \quad c_3 = -464526194.$$

Then, we compute lattice $\mathcal{L}$ spanned by the following matrix

$$M = \begin{bmatrix} 1 & c_1 & c_2 & c_3 \\ 0 & C & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & C \end{bmatrix} = \begin{bmatrix} 1 & -688526052 & -2742607542 & -464526194 \\ 0 & 4674184425 & 0 & 0 \\ 0 & 0 & 4674184425 & 0 \\ 0 & 0 & 0 & 4674184425 \end{bmatrix}.$$

Afterwards, we compute a reduced basis matrix by applying the LLL algorithm into $\mathcal{L}$ to obtain

$$K = \begin{bmatrix} -2042599 & -7058127 & -4156242 & -2199094 \\ 15186560 & 25305 & -2716845 & -5661565 \\ 5093487 & 11394351 & -18084654 & 3083247 \\ -7588606 & 5196912 & 4492677 & -33934561 \end{bmatrix}.$$

Next, we compute matrix $N = K \cdot M^{-1}$

$$N = \begin{bmatrix} -2042599 & -300883 & -1198508 & -202996 \\ 15186560 & 2237041 & 8910811 & 1509259 \\ 5093487 & 750291 & 2988636 & 506197 \\ -7588606 & -1117832 & -4452663 & -754165 \end{bmatrix},$$

which later we observe from its first row the following

$$\begin{bmatrix} -2042599 & -300883 & -1198508 & -202996 \end{bmatrix}.$$

Observe the first row of matrix N, hence we deduce

$$X = 2042599, \quad Y_1 = 300883, \quad Y_2 = 1198508 \quad \text{and} \quad Y_3 = 202996.$$

We note here that, we assigned the term $X \in \mathbb{Z}$ to denote the term x^2 whilst $Y_i \in \mathbb{Z}$ to denote y_i^2 as formally described in Theorem 6.
Next, we compute $S_i = \left\lfloor N + i + 1 - \frac{e_i X}{Y_i} \right\rfloor$ for each $i = 1, 2, 3$ which yields

$$S_1 = 3157932677, \quad S_2 = 6152899180 \quad S_3 = 4582853369.$$

We also compute $\tilde{P}_i = \frac{1}{2}(S_i + \sqrt{S_i^2 - 4N_i})$ for each $i = 1, 2, 3$ which returns

$$\tilde{P}_1 = 2956578078, \quad \tilde{P}_2 = 4144164717 \quad \tilde{P}_3 = 4111000730.$$

For $i = 1, 2, 3$, we apply the Coppersmith's approximation of p (Theorem 2) onto each $\tilde{P}_i$ which result in

$$p_1 = 2956578083, \quad p_2 = 4144164721 \quad p_3 = 4111000753.$$

We completed the factorization of N_1, N_2 and N_3, respectively, by solving $q_i = \frac{N_i}{p_i}$ for each $i = 1, 2, 3$ such that

$$q_1 = 201354599, \quad q_2 = 2008734461 \quad q_3 = 471852637.$$

Remark 2. We remark that in our case, $X \approx N^{0.355}$ is larger than Blömer-May's bound $x < \frac{1}{3}N^{0.25}$ as reported in [14], Nitaj et al.'s bound $x \approx N^{0.344}$ as proposed in [24] and Ariffin et al.'s bound $d \approx N^{0.345}$ as observed in [18].

4. Comparison with Former Attacks

In this section, we provide a comparative analysis in terms of the structure of key equation being manipulated and its specified conditions in Table 1.

Attacks	**The Structure of Key Equation**	**Conditions**
Wiener (1990, [12])	$ed - k\phi(N) = 1$	$d < \frac{1}{3}N^{0.25}$
Blömer and May (2004,[14])	$ex - k\phi(N) = -y$	$x < \frac{1}{3}N^{0.25}$ and $\lvert y\rvert = \mathcal{O}(N^{-0.75}ex)$
Nitaj (2013,[15])	$ex - y\phi(N) = z$	$xy < \frac{N}{4(p+q)}$ and $\lvert z\rvert < \frac{(p-q)N^{0.25}y}{3(p+q)}$
Nitaj et al. (2014,[24])	$e_i x - y_i\phi(N_i) = z_i$	$N = \min_i N_i$, $x < N^{\delta}, \quad y_i < N^{\delta}$, $\lvert z_i\rvert < \frac{p_i-q_i}{3(p_i+q_i)}y_i N^{0.25}$, where $\delta = \frac{k}{2(k+1)}$, $x, y_i \in \mathbb{Z}$
Ariffin et al. (2019,[18])	$ed - k\phi(N) = 1$	$\lvert b^2p - a^2q\rvert < N^{\delta}$, $d < \frac{\sqrt{3}}{\sqrt{2}}N^{0.75-\delta}$, where $0.25 < \delta \leq 0.5$
Ariffin et al. (2019,[18])	$e_i d - k_i\phi(N_i) = z_i$	$N = \max_i N_i$, $d < N^{\delta}, \quad k_i < N^{\delta}, \quad z_i < N^{\delta}$, where $\delta = \frac{3k}{2(4k+1)}$
Our Attack (Theorem 5)	$ex^2 - \phi(N)y^2 = z$	$e > \sqrt{N}$, $x < \frac{\sqrt{2}}{4}N^{0.5}e^{-0.25}$, $y < 2^{\frac{1}{4}}N^{-0.375}\sqrt{e}x$, and $\lvert z\rvert < N^{0.25}y^2$
Our Attack (Theorem 6)	$e_i x^2 - y_i^2\phi(N_i) = z_i$	$N = \min_i N_i$, $x^2 < N^{\delta}, \quad y_i^2 < N^{\delta}$, $\lvert z_i\rvert < \frac{p_i-q_i}{3(p_i+q_i)}y_i^2 N^{0.25}$, where $\delta = \frac{k}{2(k+1)}$, $x^2, y_i^2 \in \mathbb{Z}$

Table 1: Comparison of Our Attacks Against Existing Attacks

In Table 1, it shows an evolution of results with the sole objective to increase the security bound of the unknown parameters of the RSA Diophantine key equation. Initial results revolved around the original RSA Diophantine key equation.

It then came to the attention of researchers on the importance to observe generalizations of the RSA Diophantine key equation in order to properly understand the consequences upon the security of the RSA modulus. The first attempt by Blömer and May [14], which utilized continued fractions and Coppersmith's method, reaffirmed result by Wiener [12]. Then in 2013, Nitaj [15] observed the generalized RSA key equation in the form of $ex - \phi(N)y = z$ where the prime factors can be unbalanced. Suppose p and q are balanced primes. Based upon the assumptions, $z = 1$ is a valid parameter. Following through the assumptions, the value of y can be taken as small as possible, hence increasing the value of x to approximately $N^{0.5}$. At the same time, if y is increased up till $N^{0.25}$, the value of x is approximately $N^{0.25}$. This reaffirms result by Wiener [12]. As a note, if y is further increased, the value of x gets smaller. In 2019, Ariffin et al. [18] performed an attack on RSA by manipulating the standard RSA key equation and managed to improve the upper bound of d to $d < \frac{\sqrt{3}}{\sqrt{2}} N^{0.28125}$ when $\delta = \frac{15}{32}$.

In retrospect, as observed from our attack upon our defined generalized RSA Diophantine key equation, when considering the public exponent e of size approximately $N^{0.5}$, we are able to conclude that the chosen size of parameter x is approximate to $N^{0.375}$, y is approximate to $N^{0.25}$ and z is approximate to $N^{0.75}$. Our result differs from Nitaj [15] because initially the value of x does not represent a generalization of the RSA decryption exponent. Rather, it represents a variable, where if it exists together with the variables (y, z) according to the hypothesis, the RSA modulus can be factored.

On the other hand, since $z = 1 < N^{0.75}$ is a valid parameter, our analysis can correctly be viewed as an RSA Diophantine key equation where x^2 is approximate to $N^{0.75}$. This is a significant improvement in identifying the lower bound for the value of the RSA decryption exponent d to be secure.

5. Conclusion

Taken as a whole, this paper focuses on a cryptanalytic approach to factor the RSA modulus of $N = pq$ with generalized RSA Diophantine key equation of the form $ex^2 - \phi(N)y^2 = z$ for suitably small integers x, y and z under certain defined conditions. We proved that the unknown $\frac{y}{x}$ can be determined via the continued fractions expansion of $\frac{\sqrt{e}}{\sqrt{N}}$. Then, we utilized the obtained unknown values to complete the prime factoring of N via Coppersmith's method. In addition, we formulated our work on the scenario when k samples of weak RSA public key pairs are given that satisfy the equation $e_i x^2 - y_i^2 \phi(N_i) = z_i$, we can simultaneously retrieve the unknown primes of each modulus N within polynomial time. This attack combines the simultaneous Diophantine approximations and lattice basis reduction methods.

Acknowledgements

The present research was partially supported by the Putra Grant with Project Number GP-IPS/2018/9657300.

Conflict of Interest

None of the authors of this paper has a financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of the paper.

References

References

[1] Hamlen, K. W., Thuraisingham, B.: Data security services, solutions and standards for outsourcing. Comput. Stand. Interfaces. 35(1), 1–5 (2013) https://doi.org/10.1016/j.csi.2012.02.001

[2] Lancaster, S., Yen, D. C., Huang, S.-M.: Public key infrastructure: a micro and macro analysis. Comput. Stand. Interfaces. 25(5), 437–446 (2003) https://doi.org/10.1016/S0920-5489(03)00043-6

[3] Leszczyna, R.: Cybersecurity and privacy in standards for smart grids– A comprehensive survey. Comput. Stand. Interfaces. 56, 62-73 (2018) http://dx.doi.org/10.1016/j.csi.2017.09.005

[4] Diffie, W., Hellman, M.: New directions in cryptography. IEEE Trans. Inform. Theory. 22(6), 644-654 (1976)

[5] Rivest, R., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM. 21(2), 17–28 (1978)

[6] Grigoriev, D., Shpilrain, V.: RSA and redactable blockchains. Int. J. Comput. Math. Comput. Syst. Theory. 6(1), 1–6 (2021)

[7] Ghafar, A. H. A., Ariffin, M. R. K., Yasin, S. M., Sapar, S. H.: Partial key attack given MSBs of CRT-RSA private keys. Mathematics. 8(12), 2188 (2020) https://doi.org/10.3390/math8122188

[8] Herman, T. R., Walter, L., Winter, D.: Factoring with the quadratic sieve on large vector computers. J. Comput. Appl. Math. 27(1-2), 267–278 (1989) https://doi.org/10.1016/0377-0427(89)90370-1

[9] Asaar, M. R., Salmasizadeh, M. Susilo, W.: A short identity-based proxy ring signature scheme from RSA. Comput. Stand. Interfaces. 38, 144–151 (2015) https://doi.org/10.1016/j.csi.2014.10.002

[10] Susilo, W., Tonien, J., Yang, G.: The Wiener attack on RSA revisited: a quest for the exact bound. In: Proceedings of the 24th Australasian Conference on Infor- mation Security and Privacy, Lecture Notes in Computer Science 11547, pp. 381–398. Springer, Cham (2019) https://doi.org/10.1007/978-3-030-21548-4_21

[11] Evans, D. M., Yen, D. C.: Private key infrastructure: balancing computer transmission privacy with changing technology and security demands. Comput. Stand. Interfaces. 27(4), 423–437 (2005) https://doi.org/10.1016/j.csi.2004.10.010

[12] Wiener, M. J.: Cryptanalysis of short RSA secret exponents. IEEE Trans. Inform. Theory. 36(3), 553-558 (1990)

[13] Boneh, D., Durfee, G.: Cryptanalysis of RSA with private key d less than $N^{0.292}$. IEEE Trans. Inform. Theory. 46(4), 1339-1349 (2000)

[14] Blömer, J., May, A.: A generalized Wiener attack on RSA. In: International Workshop on Public Key Cryptography, pp. 1–13. Springer, Berlin, Heidelberg (2004)

[15] Nitaj A.: Diophantine and lattice cryptanalysis of the RSA cryptosystem. In: Artificial Intelligence, Evolutionary Computing and Metaheuristics, pp. 139–168. Springer, Berlin, Heidelberg (2013)

[16] Nitaj A.: A new attack on RSA with two or three decryption exponents. J. Appl. Math. Comput. 42(1), 309–319 (2013)

[17] Nitaj, A., Ariffin, M. R. K.: Implicit factorization of unbalanced RSA moduli. J. Appl. Math. Comput. 48(1), 349–363 (2015)

[18] Ariffin, M. R. K., Abubakar, S. I., Yunos, F., Asbullah, M. A.: New cryptanalytic attack on RSA modulus $N = pq$ using small prime difference method. Cryptography. 3(1), 1–25 (2019)

[19] Susilo, W., Tonien, J., Yang, G.: Divide and capture: An improved cryptanalysis of the encryption standard algorithm RSA. Comput. Stand. Interfaces. 74, 103470 (2021) https://doi.org/10.1016/j.csi.2020.103470

[20] Susilo, W., Tonien, J., Yang, G.: A generalised bound for the Wiener attack on RSA. Journal of Information Security and Applications. 53, 102531 (2020) https://doi.org/10.1016/j.jisa.2020.102531

[21] Hardy, G. H., Wright, E. M.: An Introduction to the Theory of Numbers. Oxford University Press, Oxford (1979)

[22] Coppersmith, D.: Small solutions to polynomial equations, and low exponent RSA vulnerabilities. J. Cryptology. 10(4), 233–260 (1997)

[23] May, A.: New RSA vulnerabilities using lattice reduction methods. Ph.D. diss., University of Paderborn (2003)

[24] Nitaj A., Ariffin M. R. K., Nassr D. I., Bahig H. M.: New attacks on the RSA cryptosystem. In: International Conference on Cryptology in Africa, pp. 178–198. Springer, Cham (2014)

[25] Lenstra, A. K., Lenstra, H.W., Lovász, L.: Factoring polynomials with rational coefficients. Mathematische Annalen, 261, 515–534 (1982)

[26] Nitaj A.: Another generalization of Wieners attack on RSA. In: International Conference on Cryptology in Africa, pp. 174–190. Springer, Berlin, Heidelberg (2008)

[27] Bernstein, D. J., Chang, Y.-A., Cheng, C.-M., Chou, L.-P., Heninger, N., Lange, T., Van Someren, N.: Factoring RSA keys from certified smart cards: Coppersmith in the wild. In: International Conference on the Theory and Application of Cryptology and Information Security, pp. 341–360, Springer (2013) https://ia.cr/2013/599

[28] Nemec, M., Sys, M., Svenda, P., Klinec, D., Matyas, V.: The return of Coppersmith's attack: Practical factorization of widely used RSA moduli. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 1631–1648 (2017) https://doi.org/10.1145/3133956.3133969

[29] Howgrave-Graham, N.: Finding small roots of univariate modular equations revisited. In: Proceedings of the 6th IMA International Conference on Cryptography and Coding, pp. 131–142, Springer-Verlag (1997)

[30] Young, A., Yung, M.: Kleptography: using cryptography against cryptography. In: Fumy W. (eds) Advances in Cryptology EUROCRYPT 97. EUROCRYPT 1997. Lecture Notes in Computer Science, vol 1233, pp. 62–74, Springer Berlin Heidelberg (1997) https://doi.org/10.1007/3-540-69053-0_6

[31] Ghafar A. H. A.: New compendium of RSA vulnerabilities. Doctoral thesis, Universiti Putra Malaysia. (2020)

[32] Galbraith, S. D.: Mathematics of Public Key Cryptography. Cambridge University Press, Cambridge (2012)

Appendix A. The proof of Lemma 3

Proof. We have $0 < N - \phi(N) = p + q - 1 < p + q$. Then

$$0 < \sqrt{N} - \sqrt{\phi(N)} < \frac{p+q}{\sqrt{N} + \sqrt{\phi(N)}}.$$

Now we have $\sqrt{N} + \sqrt{\phi(N)} > \sqrt{N} + \frac{1}{2}\sqrt{N} = \frac{3}{2}\sqrt{N}$. Hence, by combining with Lemma 2, we obtain

$$0 < \sqrt{N} - \sqrt{\phi(N)} < \frac{3\sqrt{2}\sqrt{N}}{3\sqrt{N}} = \sqrt{2}.$$

□

Appendix B. The proof of Theorem 5

Proof. Consider the equation $ex^2 - \phi(N)y^2 = z$, then divide with $x^2\phi(N)$ to get

$$\begin{aligned} \frac{e}{\phi(N)} - \frac{y^2}{x^2} &= \frac{z}{x^2\phi(N)} \\ \left|\left(\frac{\sqrt{e}}{\sqrt{\phi(N)}} - \frac{y}{x}\right)\left(\frac{\sqrt{e}}{\sqrt{\phi(N)}} + \frac{y}{x}\right)\right| &= \frac{|z|}{x^2\phi(N)} \\ \left|\frac{\sqrt{e}}{\sqrt{\phi(N)}} - \frac{y}{x}\right| &= \frac{|z|}{\left(x\sqrt{e} + y\sqrt{\phi(N)}\right)x\sqrt{\phi(N)}}. \end{aligned} \tag{B.1}$$

Since $x > 0$ and $y > 0$, then $x\sqrt{e} + y\sqrt{\phi(N)} > x\sqrt{e}$. Combining this with $\phi(N) > \frac{1}{2}N$, (B.1) becomes

$$\left|\frac{\sqrt{e}}{\sqrt{\phi(N)}} - \frac{y}{x}\right| < \frac{|z|}{\sqrt{e\phi(N)}x^2} < \frac{\sqrt{2}|z|}{\sqrt{eN}x^2}. \tag{B.2}$$

Next, we have

$$\begin{aligned}\left|\frac{\sqrt{e}}{\sqrt{N}} - \frac{y}{x}\right| &\leq \left|\frac{\sqrt{e}}{\sqrt{N}} - \frac{\sqrt{e}}{\sqrt{\phi(N)}}\right| + \left|\frac{\sqrt{e}}{\phi(N)} - \frac{y}{x}\right| \\ &< \sqrt{e}\frac{\sqrt{N} - \sqrt{\phi N}}{\sqrt{N\phi(N)}} + \frac{\sqrt{2}|z|}{\sqrt{eN}x^2}. \end{aligned}\tag{B.3}$$

By Lemma 3 and using $\phi(N) > \frac{1}{2}N$, (B.3) yields

$$\left|\frac{\sqrt{e}}{\sqrt{N}} - \frac{y}{x}\right| < \frac{2\sqrt{e}}{N} + \frac{\sqrt{2}|z|}{\sqrt{eN}x^2}. \tag{B.4}$$

Now, suppose that $|z| < N^{\frac{1}{4}}y^2$. Then, (B.4) becomes

$$\left|\frac{\sqrt{e}}{\sqrt{N}} - \frac{y}{x}\right| < \frac{2\sqrt{e}}{N} + \frac{\sqrt{2}N^{\frac{1}{4}}y^2}{\sqrt{eN}x^2}. \tag{B.5}$$

Next, suppose that $y < 2^{\frac{1}{4}}N^{\frac{-3}{8}}\sqrt{e}x$. Then, (B.5) yields

$$\left|\frac{\sqrt{e}}{\sqrt{N}} - \frac{y}{x}\right| < \frac{4\sqrt{e}}{N}. \tag{B.6}$$

Finally, suppose that $x < \frac{\sqrt{2}}{4}N^{\frac{1}{2}}e^{-\frac{1}{4}}$. Then

$$\left|\frac{\sqrt{e}}{\sqrt{N}} - \frac{y}{x}\right| < \frac{1}{2x^2}. \tag{B.7}$$

As a result, by Theorem 1, $\frac{y}{x}$ is a convergent of the continued fractions expansion of $\frac{\sqrt{e}}{\sqrt{N}}$. Afterwards, we get an approximation term for $p + q$ with the obtained values of x and y,

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

Since $|z| < N^{\frac{1}{4}}y^2$, then $\left|p + q - \left(N + 1 - \frac{ex^2}{y^2}\right)\right| < N^{\frac{1}{4}}$, and by Lemma 1, we can determine the primes p and q. □

Appendix C. The proof of Proposition 1

Proof. We want to estimate the number of exponents e satisfying an equation of the form

$$ex^2 - \phi(N)y^2 = z,$$

with the conditions

$$x < \frac{\sqrt{2}}{4} N^{\frac{1}{2}} e^{-\frac{1}{4}}, \quad y < 2^{\frac{1}{4}} N^{\frac{-3}{8}} \sqrt{e}x, \quad |z| < N^{\frac{1}{4}} y^2.$$

First, let y be an integer with $1 \leq y \leq N^{\frac{1}{4}}$. Let x be an integer satisfying $\gcd(x, y) = 1$, and $x < y$. Define e by the nearest integer function

$$e = \left[\phi(N) \frac{y^2}{x^2} \right].$$

Set $ex^2 - \phi(N)y^2 = z$. Then

$$|z| < \frac{1}{2}x^2 < \frac{1}{2}y^2 < N^{\frac{1}{4}} y^2.$$

Next, suppose that there are two couples (x, y) and (x', y') such that

$$e = \left[\phi(N) \frac{y^2}{x^2} \right] = \left[\phi(N) \frac{y'^2}{x'^2} \right].$$

Then

$$\left| \phi(N) \frac{y^2}{x^2} - \phi(N) \frac{y'^2}{x'^2} \right| \leq 1,$$

or equivalently

$$\phi(N) \left| y^2 x'^2 - y'^2 x^2 \right| \leq x^2 x'^2.$$

Since $x < y < N^{\frac{1}{4}}$, and $x' < y' < N^{\frac{1}{4}}$, and $\frac{1}{2}N < \phi(N) < N$, then

$$\phi(N) \left| y^2 x'^2 - y'^2 x^2 \right| \leq x^2 x'^2 < N^{\frac{2}{4}} \times N^{\frac{2}{4}} = N,$$

which is possible only if $y^2 x'^2 - y'^2 x^2 = 0$, that is $yx' = y'x$. Since $\gcd(x, y) = \gcd(x', y') = 1$, this implies that $(x, y) = (x', y')$. This allows us to count the number of exponents e with the former conditions. We have

$$\#e = \sum_{\substack{y=1 \\ x<y, \gcd(y,x)=1}}^{N^{\frac{1}{4}}} 1 = \sum_{y=1}^{N^{\frac{1}{4}}} \phi(y),$$

where $\phi(y)$ is the Euler totient function of y. Theorem 330 of [21] states that $\sum_{y=1}^{X} \phi(y) > \frac{3}{\pi^2} X^2$. For $X = N^{\frac{1}{4}}$, we get

$$\#e > \frac{3}{\pi^2} N^{\frac{1}{2}} > \frac{1}{4} N^{\frac{1}{2}}.$$

It follows that the number of public exponents e satisfying the equation is at least $\frac{1}{4} N^{\frac{1}{2}}$. □

Appendix D. The proof of Theorem 6

Proof. For $k \geq 2$, let $N_i = p_i q_i$ be defined as k RSA moduli for $i = 1, 2, \cdots, k$. Then, we rewrite $e_i x^2 - y_i^2 \phi(N_i) = z_i$ as

$$\begin{aligned} e_i x^2 - y_i^2 (N_i + 1 - (p_i + q_i)) &= z_i \\ e_i x^2 - y_i^2 (N_i + 1) &= z_i - y_i^2 (p_i + q_i). \end{aligned}$$

Hence

$$\left| \frac{e_i x^2}{N_i + 1} - y_i^2 \right| = \frac{|z_i - y_i^2 (p_i + q_i)|}{N_i + 1}. \tag{D.1}$$

Now, we choose $N = \min\{N_i\}$ and suppose that $y_i^2 < N^{\delta}$ and $|z_i| < \frac{p_i - q_i}{3(p_i + q_i)} y_i^2 N^{\frac{1}{4}}$. Then $|z_i| < y_i^2 N^{\frac{1}{4}} < N^{\delta} N^{\frac{1}{4}} < N^{\delta + \frac{1}{4}}$. Since we have $p_i + q_i < \frac{3\sqrt{2}\sqrt{N}}{2}$ from Lemma 2, we will obtain

$$\begin{aligned} \frac{|z_i - y_i^2 (p_i + q_i)|}{N_i + 1} &\leq \frac{|z_i| + y_i^2 (p_i + q_i)}{N} \\ &< \frac{N^{\delta + \frac{1}{4}} + N^{\delta} \cdot \frac{3\sqrt{2}}{2} \sqrt{N}}{N} \\ &= \frac{N^{\delta + \frac{1}{4}} + \frac{3\sqrt{2}}{2} N^{\delta + \frac{1}{2}}}{N} \\ &< \frac{\sqrt{5} N^{\delta + \frac{1}{2}}}{N} = \sqrt{5} N^{\delta - \frac{1}{2}}. \end{aligned} \tag{D.2}$$

Plugging (D.2) into (D.1), we get

$$\left| \frac{e_i x^2}{N_i + 1} - y_i^2 \right| < \sqrt{5} N^{\delta - \frac{1}{2}}.$$

Here, we continue to show the existence of an integer x^2. Define $\delta = \frac{k}{2(k+1)}$ and $\varepsilon = \sqrt{5} N^{\delta - \frac{1}{2}}$. We have

$$N^{\delta} \cdot \varepsilon^k = N^{\delta} \cdot N^{k\delta - \frac{k}{2}} \cdot (\sqrt{5})^k = N^{\delta(1+k) - \frac{k}{2}} \cdot (\sqrt{5})^k. \tag{D.3}$$

Since $\delta = \frac{k}{2(k+1)}$, (D.3) becomes

$$N^{\frac{k}{2(k+1)}(1+k)-\frac{k}{2}} \cdot (\sqrt{5})^k = N^0 \cdot (\sqrt{5})^k = (\sqrt{5})^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k. \tag{D.4}$$

Combining (D.3) and (D.4), we obtain

$$N^{\delta} < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}.$$

It follows that if $x^2 < N^{\delta}$, then $x^2 < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$. To sum up, for $i = 1, 2, \cdots, k$, we have

$$\left| \frac{e_i x^2}{N_i + 1} - y_i^2 \right| < \varepsilon, \quad x^2 < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$$

which fulfills the conditions mentioned in Theorem 3 and leads to successfully find $x^2 \in \mathbb{Z}$ and $y_i^2 \in \mathbb{Z}$ for $i = 1, 2, ..., k$. Hence, rearrange $e_i x^2 - y_i^2 \phi(N_i) = z_i$ as

$$p_i + q_i = N_i + 1 - \frac{e_i x^2}{y_i^2} + \frac{z_i}{y_i^2}.$$

Since $|z_i| < \frac{p_i - q_i}{3(p_i+q_i)} y_i^2 N^{\frac{1}{4}}$, then $\frac{z_i}{y_i^2} < \frac{p_i - q_i}{3(p_i+q_i)} N^{\frac{1}{4}}$ and $S_i = N_i + 1 - \frac{e_i x^2}{y_i^2}$ is an integer close to $p_i + q_i$ with absolute different less than $\frac{p_i - q_i}{3(p_i+q_i)} N^{\frac{1}{4}}$. Thus, we can find an approximation $\tilde{P}_i = \frac{1}{2}(S_i + \sqrt{S_i^2 - 4N_i})$ of p_i satisfying $|p_i - \tilde{P}_i| < N_i^{\frac{1}{4}}$. Based on Theorem 2, the prime factors of the RSA moduli $N_1, N_2, \cdots, N_k$ can simultaneously be obtained in polynomial time, respectively. □