Journal of Engineering Science and Technology
Vol. 20, No. 2 (2025) 442 - 457

EFFICIENT FACTORIZATION OF RSA MODULUS: SMALL SOLUTIONS OF WEAK KEY EQUATIONS

WAN NUR AQLILI RUZAI[1], YOU YING[2],
RASYID REDHA MOHD TAHIR[3], MUHAMMAD ASYRAF ASBULLAH[4,*],
MUHAMMAD REZAL KAMEL ARIFFIN[2]

[1]School of Distance Education, Universiti Sains Malaysia, 11800 Penang, Malaysia
[2]Department of Mathematics and Statistics, Faculty of Science,
Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[3]Digital Forensics Department, CyberSecurity Malaysia,
Menara Cyber Axis, Jalan Impact, 63000 Cyberjaya, Selangor, Malaysia
[4]Centre for Foundation Studies in Science of Universiti Putra Malaysia,
43400 UPM Serdang, Selangor, Malaysia
*Corresponding Author: ma_asyraf@upm.edu.my

Abstract

RSA is a widely used asymmetric cryptography method, primarily employed for digital signature validation and message encryption. The security of RSA relies on the computational difficulty of the integer factorization problem, particularly when large security parameters are used. However, vulnerabilities in RSA can be exploited, allowing adversaries to impersonate key holders and decrypt confidential messages. This study presents novel cryptanalysis techniques that target specific weaknesses in the RSA encryption system. We focus on solving key equations of the form $er - (N - p \pm q + u)s = t$, where specific conditions are imposed on the parameters r, s, t, u. By utilizing continued fractions to identify appropriate values for r and s, followed by the application of Coppersmith's method, we successfully factorize the modulus N in polynomial time. The results demonstrate the potential vulnerabilities in RSA when certain key equation structures are exploited.

Keywords: Continued fractions, Coppersmith's method, Diophantine approximations, RSA cryptanalysis, RSA cryptosystem.

1. Introduction

The Information Technology forms the bedrock of modern society, intricately woven into every aspect of our daily lives. Security has consistently taken centre stage in computing, ensuring the safe Internet exchange of information and data. The present-day applications of cryptographic algorithms extend to various domains, including encryption features embedded in web browsers, chat applications, email services, VPNs, and other communication platforms necessitating the secure transmission of data between parties. The basic procedure for message encryption and decryption is illustrated in Fig. 1. The sender transforms the plaintext into ciphertext using the encryption key during encryption. On the other hand, during decryption, the recipient reverses this process by decrypting the received ciphertext back into plaintext using the decryption key.

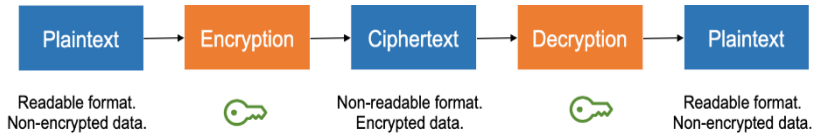

Fig. 1. Encryption and decryption process in cryptography.

Often referred to as the RSA cryptosystem, it derives its name from its creators, Ron Rivest, Adi Shamir, and Leonard Adleman, who presented RSA in their 1977 seminal paper [1]. The RSA public key cryptographic algorithm is prominent in the domain of cryptographic algorithms, being among the earliest and most widely employed concepts of public key cryptography, marking the formal entrance of cryptography into modern cryptography. It is an algorithm that uses a pair of keys: a publicly disclosed public key and a confidential private key. The public key is utilized to encrypt information, whereas the private key is employed for decryption. The ciphertext generated by encrypting plaintext with a public key can only be decrypted with the corresponding private key. The public key cannot be used for decryption. See Fig. 2.

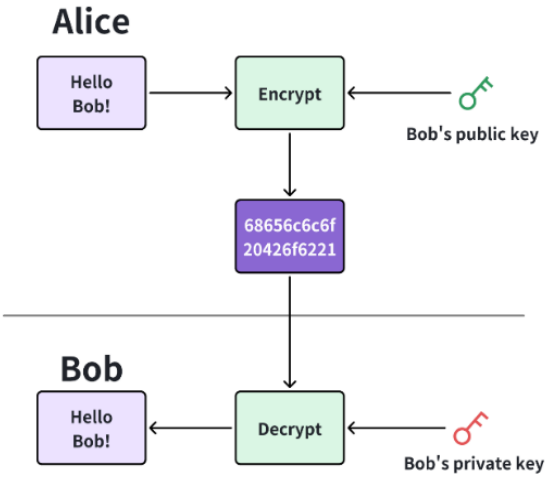

Fig. 2. The concept of public key cryptography.

Before proceeding deeper, it is crucial to define the parameters of the RSA cryptosystem. A fundamental parameter is the RSA modulus, denoted as N, and it is determined by multiplying two large, randomly selected prime numbers, referred to as p and q. Additionally, we compute $\phi(N)$, which is the product of $(p-1)$ and $(q-1)$, representing Euler's totient function for N. We exercise caution in selecting p and q to enhance security and thwart factorisation attempts, ensuring that $q < p < 2q$. Once $\phi(N)$ is established, we select a random integer, e, such that e is less than $\phi(N)$. Subsequently, we calculate the private key component, denoted as d, ensuring that it satisfies the congruence equation $ed \equiv 1 \bmod \phi(N)$. In the RSA cryptosystem, N and e are the public keys, while p, q, d and $\phi(N)$ are designated as the private keys.

In an era marked by the pervasive use of the Internet and smart devices, incidents of data thefts and breaches have exhibited a persistent upward trajectory. In light of these challenges, researchers and cryptographers are actively working to introduce innovative cryptographic models and improve existing algorithms [2]. These advancements are geared towards practical implementation in real-world applications, aiming to enhance user privacy, fortify data security, strengthen authentication mechanisms, and address many related features.

1.1. Our contribution

The present study introduces a new structure for a weak RSA key equation that facilitates solving the integer factorization problem by using the continued fractions algorithm and Coppersmith's theorem, all accomplished within polynomial time. In the first cryptanalysis, we manipulate the key structure having the form $er-(N-p+q+u)s=t$. Meanwhile, in the second cryptanalysis, we employ the key structure of the form $er-(N-p-q+u)s=t$. We provide a practical demonstration of our new attack method and highlight the uniqueness of our work by comparing it to relevant existing research. It is important to note that this research broadens the scope of vulnerabilities related to insecure RSA decryption exponents.

1.2. Organization of the article

The subsequent sections of this paper are structured as follows. Section 2 provides an overview of essential background information, including previous theorems related to continued fractions, Diophantine approximations, Coppersmith's method, and other theorems relevant to our study. Section 3 presents the core of our work, including the proof of our main attack and illustrative numerical examples to demonstrate our findings. Following that, we also provide a comparative table of our results compared to previously documented attacks. Finally, in Section 4, we draw our conclusions and summarize the key takeaways from this paper.

2. Preliminaries

This section provides an overview of continued fractions, which are used to approximate both rational and irrational numbers. This method forms the basis for creating strategies to attack the RSA cryptosystem and its various versions. Following this, we present an important technique for solving Diophantine equations.

2.1. Diophantine approximation

Definition 1. (Continued Fractions [3]) For any positive $\xi \in \mathbb{R}$, let $\xi_0 = \xi$, and for $i = 1, 2, \cdots, n$, set $\xi_i = \lfloor x_i \rfloor$ and $\xi_{i+1} = \frac{1}{\xi_i - x_i}$ until $\xi_n \in \mathbb{Z}$. Consequently, ξ can be expressed as continued fractions in the form:

$$\xi = x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{\cdots + \cfrac{1}{x_n}}}}} \tag{1}$$

For simplicity, Eq. (1) can be expressed as $\xi = [x_0, x_1, \cdots, x_n, \cdots]$. In the case where ξ is a rational number, the process of computing its continued fractions expansion concludes at a finite index n, resulting in $\xi = [x_0, x_1, \cdots, x_n]$. The convergents $\frac{x}{y}$ of ξ are fractions denoted by $\frac{x}{y} = [x_0, x_1, \cdots, x_i]$ for $i \geq 0$. Notably, if $\xi = \frac{x}{y}$ is a rational number with $\gcd(x, y) = 1$, the continued fractions expansion of ξ can be computed using the Euclidean Algorithm in $O(\log(y))$ time.

The subsequent Theorem 1 assures that the unknown integers y and z belong to the list of convergents in the continued fractions expansion of a rational number X, satisfying the given inequality as in Eq. (2).

Theorem 1. (Legendre's theorem [4]). Consider a rational number X and the positive integers y and z with $\gcd(y, z) = 1$. If the inequality

$$\left| X - \frac{y}{z} \right| < \frac{1}{2z^2} \tag{2}$$

holds, then $\frac{y}{z}$ is convergent in the continued fraction expansion of X.

2.2. Coppersmith's method

Given a sufficiently accurate approximation of any multiple of a divisor of N, the broad application of Coppersmith's result [5] readily provides an effective factorization method, as demonstrated in Theorem 2 and its corresponding variant, Theorem 3. These theorems ensure that revealing half of the leading bits of the prime number p enables the deduction of the remaining bits. Moreover, satisfying these specified conditions results in the factorization of N, achievable within a time frame polynomial in $\log N$.

Theorem 2. [5]. Consider an RSA modulus $N = pq$ where $p > q$. Additionally, let k be an (unknown) integer not divisible by q. Assume we possess an approximation $\tilde{p}$ of kp such that:

$$|kp - \tilde{p}| < N^{\frac{1}{4}} \tag{3}$$

Theorem 3. [6]. Consider an RSA modulus $N = pq$ with $p > q$. Additionally, let k be an (unknown) integer not divisible by q. Assume we possess an approximation $\tilde{p}$ of kp such that:

$$|kp - \tilde{p}| < \sqrt{2} N^{\frac{1}{4}} \tag{4}$$

The significance of Theorem 2 and Theorem 3 in this study lies in its application as a tool to demonstrate that with partial knowledge of bits from p (specifically,

knowledge of the most significant bits (MSBs)), we can successfully solve the factorization of N. As a result, having information about the approximation of p (i.e. $\tilde{p}$) also allows us to estimate the value of p with high accuracy, guaranteeing that the error in this approximation is less than $N^{\frac{1}{4}}$ [7]. It is important to note that, although Theorem 2 and Theorem 3 are based on similar fundamental concepts, these theorems focus on different specific cases and applications within our cryptanalysis framework, as demonstrated in Theorem 4 and Theorem 5.

3. Successful direction of factoring new structure of weak RSA key equations

In this section, we introduce a new structure of weak RSA key equation: $er - (N - p \pm q + u)s$ that is partitioned into two separate equations, presenting the results individually. Subsequently, both results contribute to solving the integer factorization problem efficiently through the continued fractions algorithm and Coppersmith's theorem. In the key generation of the RSA cryptosystem, the chosen primes should be of the same bit size to enhance its security. Allowing uneven factors is a potential security risk because the "small" factor could be easily found [8]. Thus, the notation $q < p < 2q$ is employed to signify that the primes p and q are balanced. In this section, we standardize the sizes of the prime factors p and q of the RSA modulus as stated in Lemma 1, and we also establish the bounds for the term $p + q$ in Lemma 2.

Lemma 1. [3]. Let $N = pq$ be an RSA modulus with the same bit size primes written as $q < p < 2q$. Then $\frac{\sqrt{2N}}{2} < q < \sqrt{N} < p < \sqrt{2N}$.

Lemma 2. [3]. Let $N = pq$ be the RSA modulus with $q < p < 2q$. Then $2\sqrt{N} < p + q < \frac{3\sqrt{2N}}{2}$.

In the first cryptanalysis, we manipulate the key structure having the form $er - (N - p + q + u)s = t$ meanwhile, in the second cryptanalysis, we employ the key structure $er - (N - p - q + u)s = t$, as follows.

3.1. Cryptanalysis of key equation $er - (N - p + q + u)s = t$

Theorem 4. Consider an RSA modulus $N = pq$ where $q < p < 2q$. Let e be a public exponent that satisfies the equation $er - (N - p + q + u)s = t$, where $\gcd(r, s) = 1$ and

$$|t| < s|p - q - u| \; and \; rs < \frac{N}{4|p - q - u|} \; and \; \left|\frac{t}{s} + u\right| < N^{\frac{1}{4}}.$$

Under these conditions, the factorization of N can be achieved in polynomial time.

Proof. Consider an RSA modulus $N = pq$ where the RSA primes abide by $q < p < 2q$ and its corresponding public exponent e. We can express the equation $er - (N - p + q + u)s = t$ as follows:

$$er - Ns = t - (p - q - u)s \tag{5}$$

By simplifying Eq. (5) and dividing it by Nr, we obtain:

$$\frac{er}{Nr} - \frac{Ns}{Nr} = \frac{t}{Nr} - \frac{(p-q-u)s}{Nr}$$

which also can be expressed as:

$$\left|\frac{e}{N} - \frac{s}{r}\right| = \frac{|t-(p-q-u)s|}{Nr} \leq \frac{|t|+|p-q-u|s}{Nr} \quad (6)$$

If we suppose $|t| < |p-q-u|s$, we get:

$$\frac{|t|+|p-q-u|s}{Nr} < \frac{2|p-q-u|s}{Nr}$$

Consequently, we can establish that $rs < \frac{N}{4|p-q-u|}$. This allows us to derive the following expression:

$$\frac{2|p-q-u|s}{Nr} < \frac{1}{2r^2}.$$

Hence, we deduce the following

$$\left|\frac{e}{N} - \frac{s}{r}\right| < \frac{1}{2r^2};$$

which satisfies Theorem 1 such that $\frac{s}{r}$ is among the convergent of the continued fraction expansion of $\frac{e}{N}$. Thus, the unknown values of s and r can be determined efficiently in polynomial time.

Subsequently, with the knowledge of s and r, we define

$A = N - \frac{er}{s}$, $B = \sqrt{A^2 + 4N}$.

Dividing both sides of the equation $er - (N-p+q+u)s = t$ by s, we obtain

$$er - (N-p+q+u)s = t$$
$$\frac{er}{s} - N + p - q - u = \frac{t}{s}$$
$$p - q - \left(N - \frac{er}{s}\right) = \frac{t}{s} + u$$
$$p - q - A = \frac{t}{s} + u.$$

Assuming $\left|\frac{t}{s} + u\right| < N^{\frac{1}{4}}$, we can get

$$|p-q-A| = \left|\frac{t}{s} + u\right| < N^{\frac{1}{4}} \quad (7)$$

According to the difference between two squares formula, expressed as:

$$(p-q)^2 - A^2 = (p+q)^2 - 4pq - A^2 = (p+q)^2 - 4N - A^2 = (p+q)^2 - (A^2+4N) = (p+q)^2 - \left(\sqrt{A^2+4N}\right)^2 = (p+q)^2 - B^2 \quad (8)$$

We observe that $(p+q)^2 - B^2 = (p-q)^2 - A^2$. By applying the squared difference formula to expand both sides of Eq. (8), we obtain:

$$(p+q)^2 - B^2 = (p-q)^2 - A^2$$
$$(p+q+B)(p+q-B) = (p-q+A)(p-q-A),$$

which can be simplified as:

$$p + q - B = \frac{(p-q+A)(p-q-A)}{p+q+B} \tag{9}$$

Taking absolute values for both sides of Eq. (9) results in:

$$|p + q - B| = \left|\frac{(p - q + A)(p - q - A)}{p + q + B}\right|$$

Considering the positivity of A and B, and since $p > q$ implies $p - q > 0$, we can deduce that $p - q + A$ and $p + q + B$ are also positive. Therefore:

$$\begin{aligned}|p + q - B| &= \left|\frac{(p - q + A)(p - q - A)}{p + q + B}\right| \\ &= \left|\frac{(p - q + A)|p - q - A|}{p + q + B}\right| \\ &= \frac{(p - q + A)}{p + q + B} \cdot |p - q - A|.\end{aligned}$$

As defined earlier, we have $B = \sqrt{A^2 + 4N}$, hence $A < B$. This implies $p - q + A < p + q + B$, leading to $\frac{p-q+A}{p+q+B} < 1$. Combining this with Eq. (7), we can derive:

$$|p + q - B| = \frac{(p-q+A)}{p+q+B} \cdot |p - q - A| < |p - q - A| < N^{\frac{1}{4}} \tag{10}$$

Summarizing the conclusion from Eqs. (7) and (10),

$$|p - q - A| < N^{\frac{1}{4}}, \qquad |p + q - B| < N^{\frac{1}{4}} \tag{11}$$

Therefore, using the conclusions given by Eq. (11), we have the following:

$$\begin{aligned}\left|p - \frac{A + B}{2}\right| &= \left|\frac{p}{2} - \frac{A}{2} + \frac{p}{2} - \frac{B}{2}\right| \\ &= \left|\tfrac{p}{2} - \tfrac{q}{2} - \tfrac{A}{2} + \tfrac{p}{2} + \tfrac{q}{2} - \tfrac{B}{2}\right| \\ &= \left|\tfrac{p-q-A}{2} + \tfrac{p+q-B}{2}\right| \\ &\leq \left|\tfrac{p-q-A}{2}\right| + \left|\tfrac{p+q-B}{2}\right| \\ &= \tfrac{1}{2} \cdot |p - q - A| + \tfrac{1}{2} \cdot |p + q - B| \\ &< \tfrac{1}{2} \cdot N^{\frac{1}{4}} + \tfrac{1}{2} \cdot N^{\frac{1}{4}} \\ &= N^{\frac{1}{4}}.\end{aligned}$$

As a result, $\left|p - \frac{A+B}{2}\right| < N^{\frac{1}{4}}$. Consequently, $\frac{A+B}{2}$ serves as an approximation of p with an additive term of at most $N^{\frac{1}{4}}$. By employing Coppersmith's technique from Theorem 2, this approximation facilitates the factorization of N.

Following this, we present Algorithm 1 to illustrate the procedure for factorizing $N = pq$ using the method outlined in Theorem 4.

Next, Example 1 illustrates the attack proposed in Theorem 4 using Algorithm 1. The factorization of N detailed in Example 1 was performed on a Windows 10

environment, using a computer equipped with an Intel (R) Core (TM) i5-8265U CPU running at 1.60 GHz and having 12.0 GB of RAM.

Algorithm 1: Factorization of Weak RSA Modulus Using Theorem 4

Input: RSA moduli N and public exponent e.

Output: The corresponding prime numbers $p, q, or \perp$.

1. Find the continued fraction expansion of $\frac{e}{N}$.
2. For every convergent $\frac{s}{r}$ of $\frac{e}{N}$, calculate $A = N - \frac{er}{s}, B = \sqrt{A^2 + 4N}$, and $\tilde{p} = \frac{A+B}{2}$,
3. Define the function $g(x)$ as $g(x) = (x + \tilde{p})$ and identify polynomials with same root modulo p.
4. Using the polynomials identified in Step 3, construct a matrix $\mathcal{M}$.
5. Apply the LLL algorithm to $\mathcal{M}$ yields the result denoted as $\mathcal{M}_{LLL}$.
6. Based on the outcomes of matrix $\mathcal{M}_{LLL}$'s first row, construct polynomial $\mathcal{M}'(x)$.
7. Determine the small solution x_0 by computing the roots of $\mathcal{M}'(x)$.
8. Determine the values of p and q by $p = \tilde{p} + x_0$ and $q = \frac{N}{P}$ respectively.
9. **if** q is an integer, **then** output values of p and q.
10. **else,** output $\perp$.

Next, Example 1 illustrates the attack proposed in Theorem 4 using Algorithm 1. The factorization of N detailed in Example 1 was performed on a Windows 10 environment, using a computer equipped with an Intel (R) Core (TM) i5-8265U CPU running at 1.60 GHz and having 12.0 GB of RAM.

Example 1. Consider a scenario where an adversary gains access to an RSA-1024 modulus, along with its respective public exponent e, as outlined below:

$N = 140238670550774504191147474274507082449562292053320102834166870719509260999106031045236643724073285564699329962589310658784025949148036097122968882592669088781358786718568318594808775805832509925131221853693485649813437652738111066836982426077576404323285810707084274299366266631462469082622609121926110391069.$

$e = 297734257440572237796827536300150598507025796686940519613049817333205448491727830363463132856060287587381503929963782796242710782652836287640492790412679650247193657457051701377390283311356668590397766683046160281736845538534181880531942344712074245795899417714946441042228487641041315963841107083747.$

The convergents of the continued fraction expansion $\frac{e}{N}$ are outlined below.

$[0, \frac{1}{4710195990}, \frac{11}{51812155891}, \frac{12}{56522351881}, \frac{23}{108334507772}, \frac{35}{164856859653}, \frac{583}{2746044262220}, \cdots, \frac{1565313}{7372931015829113}, \frac{3012892}{14191311816961469}, \frac{4578205}{21564242832790582}, \frac{94576992}{445476168472773109}, \frac{99155197}{467040411305563691}, \frac{193732189}{912516579778336800}, \frac{3586334599}{16892338847315626091}, \cdots]$

By analysing this list, we can determine the values of r and s. From the above list of convergents, the values of $\frac{s}{r} = \frac{3586334599}{16892338847315626091}$. Subsequently, we calculate $A = N - \frac{er}{s}$ and $B = \sqrt{A^2 + 4N}$ such that:

A = 191440888909789329716524931959085175597651535079563644102471662691756564110236572356605636654980319877354293074195781350715639984772374530759732584082384 4,

B = 237617264439635161515314407329499469618544722848626884256322770395137774152472096390231792970784097422052844607947209343436398762210266932368509996068521 83.

Upon obtaining the values of A and B, we are able to compute the approximation of p; which in this case is $\tilde{p} = \frac{A+B}{2}$, such that

$\tilde{p}$ = 128380676665307047243483450262703993589154938178291624333284968332156715281747876812946178318141064704894136957683393739253981380343752192722241627238380 13.

Define $g(x) = (x + \tilde{p})$. Given X as an upper bound for the unknown $|p - \tilde{p}|$, where

X = 108822060920318407088886168471413236072491824349520042767740845925196781544801.

Let x_0 be the common root of polynomials $N, g(x), x \cdot g(x)$, and $x^2 \cdot g(x)$ modulo p. Here, we construct a matrix $\mathcal{M}$ to represent these polynomials.

$$\mathcal{M} = \begin{bmatrix} N & 0 & 0 & 0 \\ \tilde{p} & X & 0 & 0 \\ 0 & \tilde{p} \cdot X & X^2 & 0 \\ 0 & 0 & \tilde{p} \cdot X^2 & X^3 \end{bmatrix}$$

Next, applying the LLL algorithm to $\mathcal{M}$ yields the result denoted as $\mathcal{M}_{LLL}$. Based on matrix $\mathcal{M}_{LLL}$'s first row outcomes, construct polynomial $\mathcal{M}'(x)$ as follows.

$\mathcal{M}'(x)$ = x^3 −189395102662987171583253662071594618256x^2 −101485358174393913981911589373652189489051291155841896516817079611964859540951396426825027723983649155021179829218 20x +109281913368915060516834985702480456183871166361314409836079257460991123038703820020867730795409668161106209218533353562230024960912.

Then, we determine the integer root of $\mathcal{M}'(x)$, resulting in:

x_0 = 1076824433935814.

Thus, we obtain the values of p denoted as $p = \tilde{p} + x_0$ which returns:

p = 128380676665307047243483450262703993589154938178291624333284968332156715281747876812946178318141064704894136957683393739253981380343752192733009871577738 27.

Finally, we complete the factorization of N by calculating $q = \frac{N}{p}$ which outputs:

q = 109236587774328114271830957066795476029389784670335259923037802062981058870724219577285614652643032717158707650263815604182417381866514739637105881931756 47.

3.2. Cryptanalysis of key equation $er - (N - p - q + u)s = t$

Theorem 5. Consider an RSA modulus $N = pq$ where $q < p < 2q$. Let e be a public exponent satisfying the equation $er - (N - p - q + u)s = t$, with $\gcd(r, s) = 1$ and

$|t| < s|p + q - u|$ and $rs < \frac{N}{4|p+q-u|}$ and $\left|\frac{t}{s} + u\right| < \frac{p-q}{p+q} \cdot N^{\frac{1}{4}}$.

Under these conditions, the factorization of N can be accomplished in polynomial time.

Proof. Consider an RSA modulus $N = pq$ where $q < p < 2q$ and a public exponent e. The equation $er - (N - p - q + u)s = t$ can be expressed as follows:

$$er - Ns = t - (p + q - u)s \tag{12}$$

Simplifying Eq. (12) and dividing it by Nr, we obtain:

$$\frac{er}{Nr} - \frac{Ns}{Nr} = \frac{t}{Nr} - \frac{(p + q - u)s}{Nr}$$

which can be expressed as:

$$\left|\frac{e}{N} - \frac{s}{r}\right| = \frac{|t-(p+q-u)s|}{Nr} \leq \frac{|t|+|p+q-u|s}{Nr} \tag{13}$$

Assuming $|t| < s|p + q - u|$, then

$$\frac{|t| + |p + q - u|s}{Nr} < \frac{2|p + q - u|s}{Nr}$$

Consequently, we can establish that $rs < \frac{N}{4|p+q-u|}$. This leads to the following expression:

$$\frac{2|p + q - u|s}{Nr} < \frac{1}{2r^2}.$$

Hence, we deduce the following

$$\left|\frac{e}{N} - \frac{s}{r}\right| < \frac{1}{2r^2};$$

which satisfies Theorem 1, indicating that $\frac{s}{r}$ is a convergent of the continued fraction expansion of $\frac{e}{N}$. Thus, the unknown values of s and r can be found in polynomial time.

As a result, with the knowledge of s and r, we define

$$A' = N - \frac{er}{s}, \qquad B' = \sqrt{|A'^2 - 4N|}\,.$$

By dividing both sides of the equation $er - (N - p - q + u)s = t$ by s, we derive

$$er - (N - p - q + u)s = t$$
$$\frac{er}{s} - N + p + q - u = \frac{t}{s}$$
$$p + q - \left(N - \frac{er}{s}\right) = \frac{t}{s} + u$$
$$p + q - \mathrm{A}' = \frac{t}{s} + u \tag{14}$$

If $\left|\frac{t}{s} + u\right| < \frac{p-q}{p+q} \cdot N^{\frac{1}{4}}$, we can establish

$$|p + q - A'| = \left|\frac{t}{s} + u\right| < \frac{p-q}{p+q} \cdot N^{\frac{1}{4}} \tag{15}$$

According to the difference between two squares formula, expressed as:

$$(p+q)^2 - A'^2 = (p-q)^2 + 4pq - A'^2 = (p-q)^2 + 4N - A'^2 = (p-q)^2 - \left(A'^2 - 4N\right) \geq (p-q)^2 - \left(\sqrt{\left|A'^2 + 4N\right|}\right)^2 = (p-q)^2 - B'^2 \quad (16)$$

Consequently, we derive $(p-q)^2 - B'^2 \leq (p+q)^2 - A'^2$. Next, apply the squared difference formula to expand both sides of Eq. (16):

$$(p-q)^2 - B'^2 \leq (p+q)^2 - A'^2$$

$$(p-q+B')(p-q-B') \leq (p+q+A')(p+q-A')$$

$$p-q-B' \leq \frac{(p+q+A')(p+q-A')}{p-q+B'} \quad (17)$$

Taking absolute values for both sides of Eq. (17) results in:

$$|p-q-B'| \leq \left|\frac{(p+q+A')(p+q-A')}{p-q+B'}\right| \quad (18)$$

From Eq. (15), we can deduce:

$$p+q-\frac{p-q}{p+q}\cdot N^{\frac{1}{4}} \leq A' \leq p+q+\frac{p-q}{p+q}\cdot N^{\frac{1}{4}}$$

$$p+q-(p-q)\cdot\frac{N^{\frac{1}{4}}}{p+q} \leq A' \leq p+q+(p-q)\cdot\frac{N^{\frac{1}{4}}}{p+q} \quad (19)$$

Since $p+q > N^{\frac{1}{4}}$, we get $\frac{N^{\frac{1}{4}}}{p+q} < 1$, so $(p-q)\cdot\frac{N^{\frac{1}{4}}}{p+q} < p-q$. Considering that both A' and B' are positive, and since $p > q$ implies $p-q > 0$, we can deduce that $p-q+A'$ and $p+q+B'$ are also positive. Therefore:

$$p+q-(p-q) \leq A' \leq p+q+(p-q)$$

$$2q \leq A' \leq 2p \quad (20)$$

Given $A' > 0, B' > 0$ and $q < p$, we can deduce that $p+q+A' > 0$ and $p-q+B' > 0$, allowing further discussion of Eq. (18):

$$|p-q-B'| < \frac{p+q+A'}{p-q+B'}\cdot|p+q-A'|$$

Here, $p+q+A' < p+q+2p = 3p+q$, $p-q+B' > p-q$, $|p+q-A'| < \frac{p-q}{p+q}\cdot N^{\frac{1}{4}}$. Consequently, we can derive:

$$|p-q-B'| < \frac{3p+q}{p-q}\cdot\frac{p-q}{p+q}\cdot N^{\frac{1}{4}} = \frac{3p+q}{p+q}N^{\frac{1}{4}} \quad (21)$$

Summarizing the conclusions from Eqs. (15) and (21), we obtain:

$$|p+q-A'| < \frac{p-q}{p+q}\cdot N^{\frac{1}{4}}, \quad |p-q-B'| < \frac{3p+q}{p+q}\cdot N^{\frac{1}{4}}$$

Using these conclusions, we can derive:

$$\left|p-\frac{A'+B'}{2}\right| = \left|\frac{p}{2}-\frac{A'}{2}+\frac{p}{2}-\frac{B'}{2}\right| = \left|\frac{p}{2}+\frac{q}{2}-\frac{A'}{2}+\frac{p}{2}-\frac{q}{2}-\frac{B'}{2}\right| = \left|\frac{p+q-A'}{2}+\frac{p-q-B'}{2}\right|$$

$$
\begin{aligned}
&\leq \left|\frac{p+q-A'}{2}\right| + \left|\frac{p-q-B'}{2}\right| \\
&= \frac{1}{2}\cdot|p+q-A'| + \frac{1}{2}\cdot|p-q-B'| \\
&< \frac{1}{2}\cdot\frac{p-q}{p+q}\cdot N^{\frac{1}{4}} + \frac{1}{2}\cdot\frac{3p+q}{p+q}N^{\frac{1}{4}} \\
&= \frac{1}{2}\cdot\left(\frac{p-q+3p+q}{p+q}\right)\cdot N^{\frac{1}{4}} \\
&= \frac{1}{2}\cdot\left(\frac{4p}{p+q}\right)\cdot N^{\frac{1}{4}} \\
&= \frac{2p}{p+q}\cdot N^{\frac{1}{4}}
\end{aligned}
$$

Now, we have

$$\left|p - \frac{A\prime+B\prime}{2}\right| < \frac{2p}{p+q}\cdot N^{\frac{1}{4}} \tag{22}$$

Considering Lemma 1 and Lemma 2, this implies $p+q > 2N^{\frac{1}{2}}$ and $p < \sqrt{2}N^{\frac{1}{2}}$, respectively. Substituting these into Eq. (22), we get

$$\left|p - \frac{A'+B'}{2}\right| < \frac{2p}{p+q}\cdot N^{\frac{1}{4}} < \frac{2\cdot\sqrt{2}N^{\frac{1}{2}}}{2N^{\frac{1}{2}}}\cdot N^{\frac{1}{4}} = \sqrt{2}N^{\frac{1}{4}}.$$

Therefore, $\left|p - \frac{A\prime+B\prime}{2}\right| < \sqrt{2}N^{\frac{1}{4}}$. Thus, $\frac{A\prime+B\prime}{2}$ serves as an approximation of p up with an additive term at most $\sqrt{2}N^{\frac{1}{4}}$. By utilizing the Coppersmith's technique of Theorem 3 enables the factorization of N.

Following this, we present Algorithm 2 to illustrate the procedure for factorizing $N = pq$ using the method outlined in Theorem 5.

Algorithm 2: Factorization of Weak RSA Modulus Using Theorem 5

Input: RSA moduli N and public exponent e.
Output: The corresponding prime numbers $p, q, or\ \perp$.

1. Find the continued fraction expansion of $\frac{e}{N}$.
2. Calculate $A' = N - \frac{er}{s}$, $B' = \sqrt{A'^2 + 4N}$, and $\tilde{p} = \frac{A\prime+B\prime}{2}$, for every convergent $\frac{s}{r}$ of $\frac{e}{N}$.
3. Define the function $g(x)$ as $g(x) = (x + \tilde{p})$ and identify polynomials with same roo modulo p.
4. Using the polynomials identified in Step 3, construct a matrix $\mathcal{M}$.
5. Apply the LLL algorithm to $\mathcal{M}$ yields the result denoted as $\mathcal{M}_{LLL}$.
6. Based on the outcomes of matrix $\mathcal{M}_{LLL}$'s first row, construct polynomial $\mathcal{M}'(x)$.
7. Determine the small solution x_0 by computing the roots of $\mathcal{M}'(x)$.
8. Determine the values of p and q by $p = \tilde{p} + x_0$ and $q = \frac{N}{p}$ respectively.
9. **if** q is an integer, **then** output values of p and q.
10. **else**, output $\perp$.

Next, Example 2 illustrates the attack proposed in Theorem 5 using Algorithm 2 as follows.

Example 2. Suppose the adversary has been granted access to an RSA-1024 modulus, along with its respective public exponent e, as outlined below:

$N = 140238670550774504191147474274507082449562292053320102834166870719509260999106031045236643724073285564699329962589310658784$0

259491480360971229688825926690887813587867185683185948087758058325099251312218536934856498134376527381110668369824260775764043232858107070842742993662666314624690826226091219261103910691.

e =320861722579068925318686905744593994178297164789293265309311049397652889299419578956004098031255178546757376537365709496548957301500764341534835614843154330664455987544597841774649728463718100716760619181769027328382484352493718313152894905469682110828974277077381672029269385449453692417836231579 27.

The list of convergent of the continued fraction expansion $\frac{e}{N}$ are as follows.

$$[0, \frac{1}{4370688701}, \frac{6}{26224132207}, \frac{7}{30594820908}, \frac{20}{87413774023}, \frac{127}{555077465046}, \frac{274}{1197568704115},$$
$$\cdots, \frac{33750269}{147511919379060766}, \frac{34573563}{151110281162585053}, \frac{172044521}{751953044029400978}, \frac{206618084}{903063325191986031},$$
$$\cdots, \frac{585280689}{2558079694413373040}, \frac{3305066050}{14445414841288252209}, \frac{3890346739}{17003494535701625249}, \cdots]$$

From the above list of convergent, we can determine the values of s and r such that $\frac{s}{r} = \frac{3890346739}{17003494535701625249}$. Next, we proceed to compute the values A' and B' using the formulas $A' = N - \frac{er}{s}$ and $B' = \sqrt{|A'^2 - 4N|}$. After acquiring the values of A' and B', we can estimate p by calculating the expression $\frac{A'+B'}{2}$.

In this case, the approximation of p is given by $\tilde{p} = \frac{A'+B'}{2}$ where: $\tilde{p}$ = 128380676665307047243483450262703993589154938178291624333284968332156715281747876812946178318141064704894136957683393739259398138034375219273300721807265449.

Next, we apply Step 3 – Step 7 of Algorithm 2, which returns $b_0 = 265350508378$. Hence the value of prime $p = \tilde{p} + b_0$ is:

p =128380676665307047243483450262703993589154938178291624333284968332156715281747876812946178318141064704894136957683393739253981380343752192733009871577738 27.

Thus, we complete the factorization of N by computing the prime $q = \frac{N}{p}$ given by:

q = 109236587774328114271830957066795476029389784670335259923037802062981058870724219577285614652643032717158707650263815604182417381866514739637105881931756 47.

3.3. Comparison with Existing Results

Researchers have increasingly recognized the importance of exploring generalizations of the RSA Diophantine key equation to deepen our understanding of their impact on RSA modulus security. Building on this foundation, we conducted a comparative analysis of the structure and specific conditions of the key equations outlined in Theorems 4 and 5, which are summarized in Table 1.

Table 1 provides a detailed comparison of our findings with those from existing attacks in the literature. Since Wiener's pioneering cryptanalysis [4], numerous

researchers have dedicated efforts to developing new attacks that expand the boundaries of the private key d. Notable contributions in this area include many works [11-15]. While these studies have significantly focus on the standard modulus $N = pq$, the work of [16] specifically addressing the weaknesses on the RSA-type modulus $N = p^2q$. Additionally, [17] proposed small private key attacks on common prime RSA, further illustrating the diverse approaches to compromising RSA security.

Table 1. Comparison of our results with related existing attacks.

Reference	Key Equation's Structure	Conditions
[4]	$ed - \phi(N)k = 1$	$d < \frac{1}{3}N^{0.25}$
[9]	$ed - \phi(N)k = 1$	$d < N^{0.292}$
[10]	$ex - \phi(N)y = z$	$x < \frac{1}{3}N^{0.25}$ and $\lvert y\rvert = \mathcal{O}(N^{-0.75}ex)$
[3]	$ex - \phi(N)y = z$	$xy < \frac{N}{4(p+q)}, \lvert z\rvert < \frac{(p-q)N^{0.25}y}{3(p+q)}$
Theorem 4	$er - (N - p + q + u)s = t$	$\lvert t\rvert \le s\lvert p - q - u\rvert, rs \le \frac{N}{4\lvert p-q-u\rvert}$, $\left\lvert \frac{t}{s} + u \right\rvert < N^{\frac{1}{4}}$
Theorem 5	$er - (N - p - q + u)s = t$	$\lvert t\rvert \le s\lvert p + q - u\rvert, rs \le \frac{N}{4\lvert p+q-u\rvert}$, $\left\lvert \frac{t}{s} + u \right\rvert < \frac{p-q}{p+q} \cdot N^{\frac{1}{4}}$

In our research, we identified that when the parameter u is set to 1, the equation aligns with the weak key equation equation $ex - \phi(N)y = z$ as proposed by [9]. Moreover, when both $u = 1$ and $t = 1$, the equation simplifies to the original RSA key equation $ed - \phi(N)k = 1$. Our work builds on and extends existing RSA cryptanalysis, particularly in the context of non-standard RSA moduli and key equations. By exploring these generalizations and comparing them with established attacks, we contribute to a more comprehensive understanding of the vulnerabilities inherent in RSA cryptosystems.

4. Conclusions

In conclusion, this work introduces a novel approach to exploit a weak RSA key equation structure, leading to a polynomial-time solution for the integer factorization problem using the continued fractions algorithm and Coppersmith's theorem. Two distinct attacks are proposed, targeting key structures of the form $er - (N - p + q + u)s = t$ and $er - (N - p - q + u)s = t$.

This study includes a practical demonstration illustrating the proposed attack and conducts a comparative evaluation against relevant existing research. The findings are based on the RSA cryptographic system with modulus $N = pq$. Significantly, our research broadens the scope of insecure RSA decryption exponents, providing valuable insights into cryptographic security. The employed techniques exhibit generality and applicability, suggesting their potential adaptation to variant RSA moduli, such as $N = p^vq$ or $N = p^vq^w$ where $v \neq w$. Examining different moduli contributes to a more comprehensive understanding of RSA system security.

Acknowledgements

This work was supported by the Universiti Sains Malaysia Short-Term Grant, Project No. 304/PJJAUH/6315667.

References

1. Rivest, R.L.; Shamir, A.; and Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2), 120-126.
2. Nitaj, A.; Kamel Ariffin, M.R.B.; Adenan, N.N.H.; Lau, T.S.C.; and Chen, J. (2022). Security issues of novel RSA variant. *IEEE Access*, 10, 53788-53796.
3. Nitaj, A. (2013). *Diophantine and lattice cryptanalysis of the RSA* cryptosystem. In Yang, X.-S. (Ed.), *Artificial intelligence, evolutionary computing and metaheuristics*: *In the footsteps of alan turing*. Springer, Berlin, Heidelberg, 139-168.
4. Wiener, M.J. (1990). Cryptanalysis of short RSA secret exponents. *IEEE Transactions on Information Theory*, 36(3), 553-558.
5. Coppersmith, D. (1997). Small solutions to polynomial equations, and low exponent RSA vulnerabilities. *Journal of Cryptology*, 10, 233-260.
6. Hinek, M.J. (2009). *Cryptanalysis of RSA and its variants*. Chapman and Hall/CRC.
7. May, A. (2003). *New RSA vulnerabilities using lattice reduction methods*. PhD Thesis, University of Paderborn.
8. Nitaj, A. (2008). *Another generalization of Wiener's attack on RSA*. In Vaudenay, S. (Ed.), *Progress in Cryptology-AFRICACRYPT* 2008. Springer Berlin, Heidelberg, 174-190.
9. Boneh, D.; and Durfee, G. (1999). *Cryptanalysis of RSA with private key d less than* $N^{0.292}$. In Stern, J. (Ed.), *Advances in Cryptology-EUROCRYPT'*99. Springer Berlin Heidelberg.
10. Blőmer, J.; and May, A. (2004). *A generalized Wiener attack on RSA*. In Bao, F.; Deng, R.; and Zhou, J. (Eds.), *Public key cryptography-PKC* 2004. Springer Berlin Heidelberg, 1-13.
11. Wan Mohd Ruzai, W.N.A.; Nitaj, A.; Kamel Ariffin, M.R.; Mahad, Z.; and Asbullah, M.A. (2022). Increment of insecure RSA private exponent bound through perfect square RSA Diophantine parameters cryptanalysis. *Computer Standards & Interfaces*, 80, 103584.
12. Ruzai, W.N.A.; Ariffin, M.R.K.; Asbullah, M.A.; and Ghafar, A.H.A. (2024). New simultaneous Diophantine attacks on generalized RSA key equations. *Journal of King Saud University-Computer and Information Sciences*, 36(5), 102074.
13. Miller, S.D.; Narayanan, B.; and Venkatesan, R. (2021). Coppersmith's lattices and "focus groups": An attack on small-exponent RSA. *Journal of Number Theory*, 222, 376-392.
14. Kamel Ariffin, M.R.; Abubakar, S.I.; Yunos, F.; and Asbullah, M.A. (2018). New cryptanalytic attack on RSA modulus $N = pq$ using small prime difference method. *Cryptography*, 3(1), 2.

15. Nitaj, A.; Adenan, N.N.H.; and Ariffin, M.R.K. (2024). *Cryptanalysis of a new variant of the RSA cryptosystem*. In Vaudenay, S.; and Petit, C. (Eds.), *Progress in Cryptology-AFRICACRYPT* 2024. Springer Nature Switzerland, 327-345.
16. Nek Abd Rahman, N. (2024). Successful cryptanalysis on RSA type modulus $N = p^2q$. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, 8, 100466.
17. Zheng, M. (2024). Revisiting small private key attacks on common prime RSA. *IEEE Access*, 12, 5203-5211.