International Journal of Cryptology Research 8(1): 33 - 53 (2018)

Extending Pollard's $p-1$ Class of Weak Primes and Its Application on RSA

Amir Hamzah Abd Ghafar[*1], **Muhammad Rezal Kamel Ariffin**[1,2], and **Muhammad Asyraf Asbullah**[1,3]

[1]*Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Universiti Putra Malaysia*
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia*
[3]*Centre of Foundation Studies for Agricultural Science, Universiti Putra Malaysia*

E-mail: amirghafar87@gmail.com
[*]*Corresponding author*

ABSTRACT

The Pollard $p-1$ method is able to solve integer factorization problem if the targeted composite number has small prime factors. This is the reason why RSA key generation requires $p-1$ and $q-1$ not to have small factors. In this paper, we identify new weak $p-1$ and $q-1$ structures, even though both $p-1$ and $q-1$ do not have small prime factors. We categorize these primes as the extension of Pollard class of weak primes. We also propose a countermeasure to avoid the weak primes being used in real-world implementation in this paper.

Keywords: Integer factorization problem, RSA cryptosystem, Pollard $p-1$ algorithm

1 INTRODUCTION

The integer factorization problem (IFP) has intrigued scholars for a long time. The problem to find the decomposition of an integer is trivial if the integer is relatively small but becomes more computationally infeasible as the integer increases in value. Trial division is the most arduous approach to solve IFP. Given an integer N to be factored, the method works by dividing N with the integers less than N one by one and outputs the integer as the factor of N if the division produces another integer. In 1974, Pollard introduced a method that can solve IFP if N constitutes specific types of factors (Pollard, 1974). The significance of IFP appeals cryptanalysts when the RSA cryptosystem adopted the IFP to be a source of its security strength (Rivest et al., 1978). Since then, many algorithms were invented either to solve IFP with specific-structured N (similar to Pollard's method) which is called special-purpose algorithms or general-structured of N which is called general-purpose algorithms. The importance of these special-purpose algorithms is that they become the tools to outline criterion in choosing safe RSA parameters which can only be attacked by general-purpose factorization algorithms, where the most efficient known algorithm still runs in sub exponential running time (Lenstra et al., 1993). In this paper, we introduce a new special-purpose factorization algorithm which specifies on the structure of both $p - 1$ and $q - 1$ where $N = pq$ is an RSA modulus but do not fall under the Pollard's class of weak primes. We also propose a countermeasure of the attack achieved by the factorization algorithm by considering beforehand the structures of the primes generated by RSA key generation algorithm.

1.1 RSA Cryptosystem

Before we go further, it is important to identify the parameters of RSA cryptosystem. One of the parameter, known as RSA modulus is $N = pq$ where p and q are two distinct primes. A corresponding Euler's function of N is also calculated, that is, $\phi(N) = (p-1)(q-1)$. To dismiss the possibility of N can be factored using trial division, p and q are chosen such that $p < q < 2p$. Using the value of $\phi(N)$, one select a random $e < \phi(N)$ and compute $d \equiv e^{-1} \pmod{\phi(N)}$. In RSA cryptosystem, N and e are called public keys while

$p, q, d, \phi(N)$ are private keys.

We remark that the secrecy of $\phi(N)$ is crucial to the security of the RSA cryptosystem since knowing such integer leads to factor N in polynomial time. Suppose the integer $\phi(N)$ is known, then we will have

$$\begin{aligned} N - \phi(N) + 1 &= pq - (p-1)(q-1) + 1 \\ &= p + q. \end{aligned}$$

We can see that

$$(X-p)(X-q) = X^2 - (p+q)X + pq.$$

Thus, by applying formula to find roots in quadratic equation, we can find

$$p, q = \frac{(p+q) \pm \sqrt{(p+q)^2 - 4(1)(pq)}}{(2)(1)}.$$

Thus, if there exists an algorithm that is able to find $\phi(N)$, then such algorithm able to efficiently factor the RSA modulus $N = pq$ (Asbullah and Ariffin, 2016).

1.2 Our Contributions

In this paper, we present a new class of weak primes that fulfill the required standard practice outlined in Section 2.3 but still can be factored in polynomial time. This new class has a specific structure that enables the factorization process and it only requires the public modulus N. Although current existing factoring algorithm may not factor N if N is a very large number but we prove that our proposed method still can solve it in polynomial time.

1.3 Outline of This Paper

In Section 2, we reproduce several special-purpose factoring algorithms that motivates our research. In the same section, we also provide current standard practice guidelines to generate RSA primes. In Section 3, we introduce our attack where we factor RSA modulus in polynomial given the RSA primes satisfy certain structures in general case. Then, in Section 4, we focus our attack on a particular case where we argue that there are infinitely many RSA modulus unknowingly satisfy the structure in the case. The algorithm to implement our attack – both in general and particular case – is presented in Section 5. Finally a countermeasure of the attack is proposed in Section 6 and the paper is concluded in Section 7.

2 SOME PREVIOUS SPECIAL-PURPOSE FACTORING ALGORITHMS AND CURRENT STANDARD OF RSA PRIMES

2.1 Pollard's $p-1$ Algorithm

Pollard $p-1$ algorithm is a special purpose algorithm to factor an integer N where $p-1$ is a B-smooth number. That is, $p-1$ can be broken completely into small prime factors that are less than an integer, B. The algorithm manipulates Fermat's Little Theorem that states for all positive integers k and integers a that coprime to p, we have

$$a^{k(p-1)} \equiv 1 \pmod{p}.$$

To utilize the algorithm, an adversary has to choose a suitable B which is the guessed bound of the smoothness of $p-1$. Then she has to compute

$$L = \prod_{\text{primes } \rho \leq B} \rho^{\lfloor \log_q B \rfloor}. \tag{1}$$

If $p-1$ is a B-smooth number then the adversary can factor N when she computes gcd$(a^L - 1, N) = p$ because

$$a^L - 1 \equiv a^{k(p-1)} - 1 \equiv 0 \pmod{p}$$

for some integer k. The adversary can extend the stages of this algorithm to factor $N = pq$ which $p-1$ has one prime factor larger than B_1 but less than

B_2 while the remaining factors are less than B_1 where $B_2 >> B_1$. To achieve it, she has to compute

$$M = \prod_{\text{primes } \rho' \in (B_1, B_2]} \left(a^L\right)^{\rho'} - 1$$

and compute gcd$(a^M - 1, N)$ to factor N. The cost of computation for Pollard $p-1$ algorithm increases when B (for one-stage algorithm) or B_2 (for two-stages algorithm) increases.

2.2 Elliptic Curve Factoring (ECM) Algorithm

This algorithm was introduced by Lenstra Jr (1987) which fundamentally replaces the multiplicative group used in Pollard $p-1$ algorithm to the group of points in a random elliptic curve. By doing that, the adversary can re-execute the attack by choosing different elliptic curves if the initial elliptic curve chosen is not suitable, until a random elliptic curve with smooth order of $\mathbb{Z} + p$ is found.

2.3 Current Standard of RSA Primes

Up until this point, we can see that if $p-1$ is a B-smooth number then $N = pq$ can be factored in polynomial time by both Pollard $p-1$ and ECM algorithms where B is a suitably small and $B_{pollard} < B_{ECM}$. Hence, several papers have suggested method to find 'strong' primes including by Williams and Schmid (Williams and Schmid, 1979) and Gordon (Gordon, 1984). However in 1999, Rivest and Silverman Rivest and Silverman (1997) noted that as long as the size of the modulus, N used in RSA is large enough (say, 2048-bit) then the necessity of finding such a strong prime is needless.

The latest standard by US National Institute of Standards and Technology (NIST) for generating RSA primes, FIPS PUB 186-4 , stated in Appendix B.3 that p and q of RSA-2048 and above can be generated using random primes as long as the primes are provable or probable primes and satisfy the followings:

1. Size of of p must be $\sqrt{2} \cdot 2^{(nlen/2)-1} < p < 2^{(nlen/2)} - 1$ where $nlen$ is the size of N;

2. Size of of q must be $\sqrt{2} \cdot 2^{(nlen/2)-1} < q < 2^{(nlen/2)} - 1$ where $nlen$ is the size of N; and

3. Size of $|p - q|$ must be greater than $2^{(nlen/2)-100}$ where $nlen$ is the size of N.

More details on the standard can be read here (FIPS, 2013).

3 THE ATTACK IN GENERAL

We discuss our attack in its general form in this section. We begin with a lemma reducing $\sqrt{a^m + 1}$ to its integer and decimal forms.

Lemma 3.1. *Let $a \in \mathbb{Z}^+$ and $m \geq 2$ be a power of 2. Then $\sqrt{a^m + 1} = a^{m/2} + \epsilon$ where $\epsilon < \frac{1}{2}a^{-m/2}$.*

Proof. Let $a^m + 1$ be an integer where $a \in \mathbb{Z}^+$. Then

$$\sqrt{a^m + 1} < \sqrt{a^m + \frac{1}{4}a^m + 1} = \sqrt{(a^{m/2} + \frac{1}{2}a^{-m/2})^2} = a^{m/2} + \frac{1}{2}a^{-m/2}$$

If we assume that $\sqrt{a^m + 1} = a^{m/2} + \epsilon$ then $\epsilon < \frac{1}{2}a^{-m/2}$. This terminates the proof. ■

In the following lemma, we find the upper bound of an intermediate that will be used in our attack.

Lemma 3.2. *Let $a, b \in \mathbb{Z}^+$ and $m \geq 2$ be a power of 2. Suppose $a < b < 2a$. Then $\frac{a^m + b^m}{2(ab)^{m/2}} < \frac{m}{2} + \delta$ where $0 < \delta < 1$.*

Proof. Using binomial expansion, we have

$$\begin{aligned}(b-a)^m &= \binom{m}{0}(-a)^m b^0 + \binom{m}{1}(-a)^{m-1}b^1 + \binom{m}{2}(-a)^{m-2}b^2 \ldots + \\ &\quad \binom{m}{m-2}(-a)^2 b^{m-2} + \binom{m}{m-1}(-a)^1 b^{m-1} + \binom{m}{m}(-a)^0 b^m \\ &= (-a)^m + b^m + C \\ &= a^m + b^m + C \end{aligned} \tag{2}$$

as m will always be a positive even integer and

$$\begin{aligned} C &= \binom{m}{1}(-a)^{m-1}b + \binom{m}{2}(-a)^{m-2}b^2 + \ldots + \binom{m}{m-2}(-a)^2 b^{m-2} \\ &\quad + \binom{m}{m-1}(-a)^1 b^{m-1}. \end{aligned} \tag{3}$$

From (2), we get

$$\frac{a^m + b^m}{2(ab)^{m/2}} = \frac{(b-a)^m - C}{2(ab)^{m/2}}. \tag{4}$$

We can rearrange equation (3) to get

$$\begin{aligned} C =& \binom{m}{2}a^{m-2}b^2 + \binom{m}{4}a^{m-4}b^4 + \ldots + \binom{m}{m-4}a^4 b^{m-4} + \binom{m}{m-2}a^2 b^{m-2} \\ & - \binom{m}{1}a^{m-1}b - \binom{m}{3}a^{m-3}b^3 - \ldots - \binom{m}{m-3}a^3 b^{m-3} - \binom{m}{m-1}ab^{m-1} \\ =& \sum_{\substack{i \text{ is even} \\ 0<i<m}} \binom{m}{i} a^{m-i}b^i - \sum_{\substack{i \text{ is odd} \\ 0<i<m}} \binom{m}{i} a^{m-i}b^i. \end{aligned} \tag{5}$$

Then (4) will become

$$\begin{aligned}
\frac{a^m+b^m}{2(ab)^{m/2}} &= \frac{1}{2(ab)^{m/2}}\left((b-a)^m-\left(\sum_{\substack{i \text{ is even}\\ 0<i<m}}\binom{m}{i}a^{m-i}b^i-\sum_{\substack{i \text{ is odd}\\ 0<i<m}}\binom{m}{i}a^{m-i}b^i\right)\right)\\
&= \frac{1}{2(ab)^{m/2}}\left((b-a)^m+\sum_{\substack{i \text{ is odd}\\ 0<i<m}}\binom{m}{i}a^{m-i}b^i-\sum_{\substack{i \text{ is even}\\ 0<i<m}}\binom{m}{i}a^{m-i}b^i\right)\\
&= \frac{(b-a)^m}{2(ab)^{m/2}}+\frac{\binom{m}{1}a^{m-1}b^1}{2(ab)^{m/2}}+\sum_{\substack{i \text{ is odd}\\ 3\le i<m}}\frac{\binom{m}{i}a^{m-i}b^i}{2(ab)^{m/2}}-\sum_{\substack{i \text{ is even}\\ 0<i<m}}\frac{\binom{m}{i}a^{m-i}b^i}{2(ab)^{m/2}}\\
&= \frac{(b-a)^m}{2(ab)^{m/2}}+\frac{m}{2}\left(\frac{a}{b}\right)^{\frac{m}{2}-1}+\sum_{\substack{i \text{ is odd}\\ 3\le i<m}}\frac{\binom{m}{i}a^{m-i}b^i}{2(ab)^{m/2}}-\sum_{\substack{i \text{ is even}\\ 0<i<m}}\frac{\binom{m}{i}a^{m-i}b^i}{2(ab)^{m/2}}
\end{aligned} \tag{6}$$

If $a<b<2a$ then $\frac{a}{b}<1$. Thus (6) will become

$$\begin{aligned}
\frac{a^m+b^m}{2(ab)^{m/2}} &< \frac{(b-a)^m}{2(ab)^{m/2}}+\frac{m}{2}\left(1\right)^{\frac{m}{2}-1}+\sum_{\substack{i \text{ is even}\\ 0<i<m}}\frac{\binom{m}{i}a^{m-i}b^i}{2(ab)^{m/2}}-\sum_{\substack{i \text{ is even}\\ 0<i<m}}\frac{\binom{m}{i}a^{m-i}b^i}{2(ab)^{m/2}}\\
&= \frac{m}{2}+\frac{(b-a)^m}{2(ab)^{m/2}}.
\end{aligned} \tag{7}$$

To complete the proof, we need to show that $\frac{(b-a)^m}{2(ab)^{m/2}}=\delta$ where $0<\delta<1$. Observe that if $a<b<2a$ then

$$a-a<b-a<2a-a \Rightarrow 0<(b-a)^m<a^m. \tag{8}$$

Observe also that

$$a^2<ab \Rightarrow a<(ab)^{1/2} \Rightarrow a^m<(ab)^{m/2}. \tag{9}$$

By comparing (8) and (9), we get $(b-a)^m<(ab)^{m/2}$ hence $0<\frac{(b-a)^m}{2(ab)^{m/2}}<1$. This completes the proof. ■

With the results from Lemma 3.1 and 3.2, we propose a theorem that helps our attack later.

Theorem 3.1. *Let $m \geq 2$ is a power of 2 and $a < b < 2a$ where a, b are n-bit numbers. If $\epsilon_1 < \frac{1}{2}a^{-m/2}$ and $\epsilon_2 < \frac{1}{2}b^{-m/2}$ then $1 < \sqrt{(a^m+1)(b^m+1)} - (ab)^{m/2} < \frac{m}{2} + \delta$ where $\delta \in (0,1)$.*

Proof. To prove the lower bound, first we need to show that $a^m + b^m > 2(ab)^{m/2}$. Observe that

$$(a^{m/2} - b^{m/2})^2 = a^m + b^m - 2(ab)^{m/2}. \tag{10}$$

As $(a^{m/2} - b^{m/2})^2$ will always be positive value, it implies that $a^m + b^m > 2(ab)^{m/2}$. Then

$$\begin{aligned} \sqrt{(a^m+1)(b^m+1)} &= \sqrt{(ab)^m + a^m + b^m + 1} \\ &> \sqrt{(ab)^m + 2(ab)^{m/2} + 1} \\ &= \sqrt{\left((ab)^{m/2} + 1\right)^2} \\ &= (ab)^{m/2} + 1 \end{aligned} \tag{11}$$

Thus, $\sqrt{(a^m+1)(b^m+1)} - (ab)^{m/2} > 1$. To prove the upper bound, by Lemma 3.1,

$$\begin{aligned} \sqrt{(a^m+1)(b^m+1)} &= \sqrt{a^m+1}\sqrt{b^m+1} \\ &= (a^{m/2} + \epsilon_1)(b^{m/2} + \epsilon_2) \\ &= (ab)^{m/2} + a^{m/2}\epsilon_2 + b^{m/2}\epsilon_1 + \epsilon_1\epsilon_2 \\ &< (ab)^{m/2} + a^{m/2}\left(\frac{1}{2}b^{-m/2}\right) + b^{m/2}\left(\frac{1}{2}a^{-m/2}\right) \\ &\quad + \left(\frac{1}{2}a^{-m/2}\right)\left(\frac{1}{2}b^{-m/2}\right). \end{aligned} \tag{12}$$

(12) can be rewritten as

$$\begin{aligned} \sqrt{(a^m+1)(b^m+1)} - (ab)^{m/2} &< a^{m/2}\left(\frac{1}{2}b^{-m/2}\right) + b^{m/2}\left(\frac{1}{2}a^{-m/2}\right) \\ &\quad + \left(\frac{1}{2}a^{-m/2}\right)\left(\frac{1}{2}b^{-m/2}\right). \end{aligned}$$

We can see that

$$\begin{aligned}\left(\frac{1}{2}a^{-m/2}\right)\left(\frac{1}{2}b^{-m/2}\right) &= \frac{1}{4}\left((ab)^{-m/2}\right)\\ &= \delta_1 < 1 \qquad (13)\end{aligned}$$

for some $\delta_1 \in (0,1)$. Now we show that

$$a^{m/2}\left(\frac{1}{2}b^{-m/2}\right) + b^{m/2}\left(\frac{1}{2}a^{-m/2}\right) < m + \delta.$$

for some $\delta \in (0,1)$. If $a < b < 2a$, then

$$\begin{aligned}a^{m/2}\left(\frac{1}{2}b^{-m/2}\right) + b^{m/2}\left(\frac{1}{2}a^{-m/2}\right) &= \frac{1}{2}\left(\frac{a^{m/2}}{b^{m/2}} + \frac{b^{m/2}}{a^{m/2}}\right)\\ &= \frac{a^m + b^m}{2(ab)^{m/2}}.\end{aligned}$$

Based on Lemma 3.2,

$$\frac{a^m + b^m}{2(ab)^{m/2}} < \frac{m}{2} + \delta_2$$

for some

$$\delta_2 = \frac{(b-a)^m}{2(ab)^{m/2}} \in (0,1). \qquad (14)$$

To ensure that $\delta = \delta_1 + \delta_2 \in (0,1)$, we need to find the summation of (13) and (14). Given $a < b < 2a$,

$$\begin{aligned}\frac{1}{4(ab)^{m/2}} + \frac{(b-a)^m}{2(ab)^{m/2}} &= \frac{2(b-a)^m + 1}{4(ab)^{m/2}}\\ &< \frac{2a^m + 1}{4(ab)^{m/2}}\\ &< \frac{2a^m + 1}{4(a^2)^{m/2}}\\ &= \frac{1}{2} + \frac{1}{4a^m} < 1.\end{aligned}$$

Hence $1 < \sqrt{(a^m + 1)(b^m + 1)} - (ab)^{m/2} < \frac{m}{2} + \delta$ for some $\delta \in (0,1)$. This terminates the proof. ■

Using information form Theorem 3.1, we can find the distance between $\left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor$ and $ab^{m/2}$ in the next proposition.

Proposition 3.1. *If a and b are two distinct integers such that $a<b<2a$ then* $\left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor-(ab)^{m/2}=\frac{m}{2}$.

Proof. Let

$$\begin{aligned}\sqrt{(a^m+1)(b^m+1)}-(ab)^{m/2} &= \left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor+\epsilon_1 \\ &\quad -(ab)^{m/2} \qquad (15)\end{aligned}$$

where $\epsilon_1\in(0,1)$. From Theorem 3.1,

$$1<\sqrt{(a^m+1)(b^m+1)}-(ab)^{m/2}<\frac{m}{2}+\delta. \qquad (16)$$

where $\delta\in(0,1)$. Combining (15) and (16), we get

$$\left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor+\epsilon_1-(ab)^{m/2}<\frac{m}{2}+1 \qquad (17)$$

Observe

$$\begin{aligned}\left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor+\epsilon_1-(ab)^{m/2}-1 &< \left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor+1 \\ &\quad -(ab)^{m/2}-1 \\ &= \left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor \\ &\quad -(ab)^{m/2}. \qquad (18)\end{aligned}$$

Inequalities in (17) and (18) shows that there are two cases of $\left\lfloor\sqrt{(a^2+1)(b^2+1)}\right\rfloor-ab$:

Case 1: We consider that

$$\left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor-(ab)^{m/2}\geq\frac{m}{2} \qquad (19)$$

and

$$\begin{aligned}\sqrt{(a^m+1)(b^m+1)}-(ab)^{m/2}-\epsilon_1 &< \frac{m}{2}+\delta-\epsilon_1 \\ \Rightarrow \left\lfloor\sqrt{(a^m+1)(b^m+1)}\right\rfloor-(ab)^{m/2} &< \frac{m}{2}+\epsilon_2. \qquad (20)\end{aligned}$$

where $\epsilon_2 = \delta - \epsilon_1$. Combining both (19) and (20), we have

$$\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} \in \left[\frac{m}{2}, \frac{m}{2} + \epsilon_2\right).$$

Since $\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} \in \mathbb{Z}$, thus

$$\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} = \frac{m}{2}. \tag{21}$$

Case 2: We also consider that

$$\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} \leq \frac{m}{2}$$

Since $\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} \in \mathbb{Z}$, thus

$$\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} = \frac{m}{2} \tag{22}$$

Combining (21) and (22) completes the proof. ■

Proposition 3.1 shows that the difference between $\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2}$ is indeed $\leq$ than m. The next theorem shows how this fact can be manipulated into attacking RSA modulus, $N = pq = (a^m+1)(b^m+1)$ where $m \geq 2$ is a power of 2.

Theorem 3.2. *Let $N = pq$ be a valid RSA modulus where $p = a^m + 1$ and $q = b^m + 1$ for n-bit positive integer a, b such that $a < b < 2a$ and $m \geq 2$ is a power of 2. Then N can be factored in polynomial time.*

Proof. Observe

$$N = pq = (a^m+1)(b^m+1) \tag{23}$$

and

$$\begin{aligned} \phi(N) &= (p-1)(q-1) = a^m b^m = (ab)^m \\ \phi(N)^{1/2} &= ((p-1)(q-1))^{1/2} = a^{m/2} b^{m/2} = (ab)^{m/2}. \end{aligned} \tag{24}$$

From Proposition 3.1 we know that

$$\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - (ab)^{m/2} = \frac{m}{2} \tag{25}$$

where a and b are two distinct positive integers such that $a < b < 2a$. Equation (25) can also be written as

$$\left\lfloor \sqrt{(a^m+1)(b^m+1)} \right\rfloor - \frac{m}{2} \le (ab)^{m/2}$$

or

$$\left\lfloor \sqrt{N} \right\rfloor - \frac{m}{2} \le \phi(N)^{1/2}$$

as shown in equations (23) and (24). Then

$$\left(\left\lfloor \sqrt{N} \right\rfloor - \frac{m}{2} \right)^2 \le \phi(N).$$

If N has the size of u-bit, then

$$N = (a^m+1)(b^m+1) \quad < \quad 2^{u+1}. \tag{26}$$

The dominating term in N is $(ab)^m$. Thus (26) can be rewritten as

$$(ab)^m \quad < \quad 2^{u+1}$$

and

$$m \quad < \quad \log_{ab} 2^{u+1} < (u+1)\log_{2^{2n}} 2 < \frac{1}{2n}(u+1) \tag{27}$$

for some positive integer u (RSA-2048 has $u = 2048$). That is we can obtain $\phi(N)$ only by knowing N and m where N can be obtained publicly and m is a small positive integer. We know that by knowing the value of $\phi(N)$, we can factor N. Hence, N can be factored in polynomial time. ■

The possibility for a^m+1 and b^m+1 to be not under Pollard's class of weak primes is shown in the following remark.

Remark 3.1. *Primes $p = a^m+1$ and $q = b^m+1$ where $m \ge 2$ is a power of 2 and a, b are n-bits numbers and can be represented as*

$$\begin{aligned} p &= a^m+1 = (2 \cdot s_1 \cdot s_2)^m + 1 \\ q &= b^m+1 = (2 \cdot t_1 \cdot t_2)^m + 1 \end{aligned}$$

If all $s_1 < s_2$ and $t_1 < t_2$ are $\left(\frac{n}{m} - 1\right)$-bits primes, then a^m and b^m will be s_2-smooth and t_2-smooth numbers respectively where s_2 and t_2 are large primes and cannot be viewed publicly. Thus, $N = pq$ is susceptible to Pollard $p-1$ integer factorization algorithm.

The next remark justifies our selection criteria on parameter m.

Remark 3.2. *We only conduct our attack when m is a power of 2. The reason is when m is a power of 2, $(a^m + 1)$ can only be factored as complex numbers. Specifically,*

$$a^m + 1 = (a^{m/2} + i)(a^{m/2} - i) \tag{28}$$

where i is a complex number. This means $(a^m + 1)$ will not produce more than one factor over integers which is the properties of a prime number that we need in p and q.

4 PARTICULAR CASE: $m = 2$

In this section, we focus on a particular case of our previous attack where we define $m = 2$. That is we have $p = a^2 + 1$ and $q = b^2 + 1$ as the RSA primes. We have this interesting theorem that help us to justify our focus on this particular case.

Theorem 4.1 (Friedlander-Iwaniec Theorem)**.** *For $x, y \in \mathbb{Z}$, there are infinitely many prime numbers of the form $x^2 + y^4$.*

Proof. The proof is elaborated in (Friedlander and Iwaniec, 1997). ■

The next remark explains our justification.

Remark 4.1. *Primes in the form of $a^2 + 1$ is the special case of Theorem 4.1. This implies that there are infinitely many primes in the form of $a^2 + 1$ and $b^2 + 1$ where $a, b \in \mathbb{Z}^+$, thus there are also infinitely many RSA modulus to take the form of $N = pq = (a^2 + 1)(b^2 + 1)$*

As now there is a motivation to attack RSA modulus in the form of $N = pq = (a^2 + 1)(b^2 + 1)$, the next corollaries show on how the attack works in this particular case.

Corollary 4.1. *Let $a < b < 2a$ where a, b are n-bit numbers. If $\epsilon_1 < \frac{1}{2}\left(\frac{a-1}{a^2}\right)$ and $\epsilon_2 < \frac{1}{2}\left(\frac{b-1}{b^2}\right)$ then $1 < \sqrt{(a^2+1)(b^2+1)} - ab < 1 + \delta$ for some $0 < \delta < 1$.*

Proof. The proof is a direct application of Theorem 3.1 where $m = 2$. ■

Corollary 4.2. *If a and b be two distinct integers such that $a < b < 2a$ then* $\left\lfloor \sqrt{(a^2+1)(b^2+1)} \right\rfloor - ab = 1$.

Proof. The proof is a direct application of Proposition 3.1 where $m = 2$. ■
Using information from the previous corollaries, we conduct an attack against RSA in the next theorem.

Theorem 4.2. *Let $N = pq$ be a valid RSA modulus where $p = a^2 + 1$ and $q = b^2 + 1$ for any positive integer a, b such that $a < b < 2a$. Then N can be factored in polynomial time.*

Proof. We can see that

$$N = pq = (a^2+1)(b^2+1) \tag{29}$$

and

$$\begin{aligned} \phi(N) &= (p-1)(q-1) = a^2b^2. \\ \phi(N)^{0.5} &= ((p-1)(q-1))^{0.5} = ab. \end{aligned} \tag{30}$$

From Corollary 4.2 we know that

$$\left\lfloor \sqrt{(a^2+1)(b^2+1)} \right\rfloor - ab = 1 \tag{31}$$

if a and b are two distinct positive integers such that $a < b < 2a$. Equation (31) can also be written as

$$\left\lfloor \sqrt{(a^2+1)(b^2+1)} \right\rfloor - 1 = ab$$

or

$$\left\lfloor \sqrt{N} \right\rfloor - 1 = \phi(N)^{0.5}$$

as shown in equations (29) and (30). Then

$$\left(\left\lfloor \sqrt{N} \right\rfloor - 1 \right)^2 = \phi(N).$$

That is we can obtain $\phi(N)$ only by knowing N which can be obtained publicly. We know that by knowing the value of $\phi(N)$, we can factor N. Hence, N can be factored in polynomial time. ■

The next remark shows how primes $p = a^2 + 1$ and $q = b^2 + 1$ where a, b are n-bits numbers can be the extension of Pollard's class of weak primes.

Remark 4.2. *Primes $p = a^2 + 1$ and $q = b^2 + 1$ where a, b are n-bits numbers can be represented as*

$$\begin{aligned} p &= a^2 + 1 = (2 \cdot s_1 \cdot s_2)^2 + 1 \\ q &= b^2 + 1 = (2 \cdot t_1 \cdot t_2)^2 + 1 \end{aligned}$$

where $s_1 < s_2$ and $t_1 < t_2$ are $\left(\frac{n}{2} - 1\right)$-bits primes. This implies that a^2 and b^2 will be s_2-smooth and t_2-smooth numbers respectively where s_2 and t_2 are large primes and cannot be viewed publicly. Thus, $N = pq$ is susceptible to Pollard $p - 1$ integer factorization algorithm.

Now we present a numerical example for this attack.

Example 4.1. *We use RSA-2048 modulus in this example. Specifically, we are given*

$$\begin{aligned} N = \; & 18524330360427191997567078229332509778314586480819926 90 \\ & 8976269255355807274483548362398715225859598689173174975 \\ & 2515118828296111313684792338625763511075584966752062520 \\ & 2111431594826103656523797329009654709621127229289321801 \\ & 0739417690130966164073953718995247536385878164622553388 \\ & 0958760640287223935334192281332072318636466559983784094 \\ & 8607572684415507458478207398284507082349398498183282187 \\ & 6637602673821309629555765759977907353909606661952170897 \\ & 8265888589190206621430585530459833315616661830804394060 \\ & 5164627730656281139218996459251822176848240037024856662 \\ & 1199938201035456776405174550331448285108959236714001358 \\ & 736121500089. \end{aligned}$$

Then we calculate

$$\begin{aligned}\tilde{\phi} &= \left(\left\lfloor\sqrt{N}\right\rfloor - 1\right)^2 \\ &= 18524330360427191997567078229332509778314586480819926908 9\\ &\quad 7626925535580727448354836239871522585959868917317497525 15\\ &\quad 1188282961113136847923386257635110755849667520625202111 43\\ &\quad 1594826103656523797329009654709621127229289321801073941 76\\ &\quad 9013096616407395371899524753638587816462255338809587606 40\\ &\quad 2872239353341922813320695442580414127644904255105101050 94\\ &\quad 9091849027996018004407761126783900494654125491541031125 40\\ &\quad 2881251247642706429451373125112572759594193717541907454 12\\ &\quad 3528375539719547721782645744854394195420510175604047348 81\\ &\quad 8551281266760879137161360137976569118626666783011745084 14\\ &\quad 705161440267209521134964353769432331608269958 56.\end{aligned}$$

Given N and $\tilde{\phi}$, we can calculate

$$p,q = \frac{(N-\tilde{\phi}+1) \pm \sqrt{(N-\tilde{\phi}+1)^2 - 4(1)(N)}}{(2)(1)}.$$

as $(X-p)(X-q) = X^2 - (N-\tilde{\phi}+1)X + N$ *as* $N - \tilde{\phi} + 1 = p + q$. *This is true because* $\tilde{\phi} = \phi$. *That is,*

$$\begin{aligned}p &= 11191204239430195832293236986233960997659031066669180013 72\\ &\quad 0115079630389483400905162580944819966558368679306567968 740\\ &\quad 2614407298383430996251818183800725419827061070204375549 495\\ &\quad 4321291650448263165774990071224742177406013099298892347 062\\ &\quad 5432697889873191744934658873033501733060897836466604049 284\\ &\quad 035941786611573315 7\end{aligned}$$

and

$$
\begin{aligned}
q \;=\;& 16552580012041997104400265485399505417950419144711100008560\\
& 11075349672991446612670566553119081151459602539847933152 77\\
& 57411289261915248931616216606787014394760947803017381857 13\\
& 30240733240702022794376708220073245941132359950050763409 27\\
& 36155896536201352921571417204599641075695748968329422308 49\\
& 6698707709178771077.
\end{aligned}
$$

The example works because we can find

$$
\begin{aligned}
p &= a^2 + 1 = (2 \cdot s_1 \cdot s_2)^2 + 1\\
q &= b^2 + 1 = (2 \cdot t_1 \cdot t_2)^2 + 1
\end{aligned}
$$

where

$$
\begin{aligned}
s_1 \;=\;& 782131431622224301962022328989869164506945802441560328216 3\\
& 5946564531700694291\\
s_2 \;=\;& 676283319789936746425439336668710558435486783683510815108 4\\
& 7653815623424568287
\end{aligned}
$$

and

$$
\begin{aligned}
t_1 \;=\;& 824708936865457005879911675319324735088130619530026046247 8\\
& 2578893369979398737\\
t_2 \;=\;& 780013557889123248900640831895670759539274699178210045161 1\\
& 0572909688369715201
\end{aligned}
$$

where s_1, s_2, t_1, t_2 are primes. ■

Remark 4.3. *Both p and q in example 4.1 satisfy the conditions stated in FIPS PUB 186-4.*

5 THE ALGORITHM

We propose an algorithm to conduct the attack based on our results. The algorithm takes RSA modulus, N as its input and only outputs p and q if N is susceptible to our attack.

Algorithm 1 Factoring RSA modulus N with u-bits size

```
Input: $N, u$
Output: $p, q$ or $\perp$
Set $m = 2$.
Set $n = \left\lfloor \frac{u}{m} \right\rfloor$.
while $m < \frac{1}{2n}(u+1)$ do
    calculate $\phi = \left(\left\lfloor \sqrt{N} \right\rfloor - i\right)^2$
    calculate $p = \frac{(N-\phi) \pm \sqrt{(N-\phi)^2 - 4(1)(N)}}{(2)(1)}$.
    if $p \in \mathbb{Z}$ then calculate $q = N/p$.
        Output $p$ and $q$
    else set $m = m \cdot 2$ and set $n = \left\lfloor \frac{u}{m} \right\rfloor$.
    end if
end while
Output $\perp$
```

We define $m = 2$ as the initial point of the algorithm and terminates the algorithm when $m < \frac{1}{2n}(u+1)$ as shown in (27).

6 COUNTERMEASURE OF THE ATTACK

We have shown the devastating effect our attack can do on RSA modulus which falls under this new class of weak primes. In this section, we propose a countermeasure against the attack. The countermeasure is as follows:

Countermeasure of the attack

Given p and q, if

$$\sqrt{p-1} \quad \text{or} \quad \sqrt{q-1} \tag{32}$$

is an integer, the RSA standard must find the respected new p or q.

Our proposed countermeasure is simple and can be easily implemented in

RSA key generation standard practices.

7 CONCLUSION

We present a new method to factor RSA modulus $N = pq$ where the RSA primes have special structures in general case that may hinder current factorization algorithm to solve it. We also present the attack in a particular case where we argue that there are infinitely many RSA modulus can be formed in such case. Both attacks are presented in an algorithm that can run in polynomial time. To avoid the leakage, we propose a countermeasure which identify the new class of these weak primes.

REFERENCES

Asbullah, M. A. and Ariffin, M. R. K. (2016). Design of Rabin-like Cryptosystem without Decryption Failure. *Malaysian Journal of Mathematical Sciences*, 10(S):1–18.

FIPS, P. (2013). 186-4: Federal information processing standards publication. Digital Signature Standard (DSS). *Information Technology Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, MD*, pages 20899–8900.

Friedlander, J. and Iwaniec, H. (1997). Using a parity-sensitive sieve to count prime values of a polynomial. *Proceedings of the National Academy of Sciences*, 94(4):1054–1058.

Gordon, J. (1984). Strong primes are easy to find. In *Workshop on the Theory and Application of of Cryptographic Techniques*, pages 216–223. Springer.

Lenstra, A. K., Lenstra, H. W., Manasse, M. S., and Pollard, J. M. (1993). The number field sieve. In *The development of the number field sieve*, pages 11–42. Springer.

Lenstra Jr, H. W. (1987). Factoring integers with elliptic curves. *Annals of mathematics*, pages 649–673.

Pollard, J. M. (1974). Theorems on factorization and primality testing. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 76, pages 521–528. Cambridge University Press.

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.

Rivest, R. L. and Silverman, R. D. (1997). Are Strong primes needed for RSA? In *The 1997 RSA Laboratories Seminar Series, Seminars Proceedings*.

Williams, H. C. and Schmid, B. (1979). Some remarks concerning the MIT public-key cryptosystem. *BIT Numerical Mathematics*, 19(4):525–538.