Estimating the $\phi(n)$ of Upper/Lower Bound in its RSA Cryptosystem

Chenglian Liu[1] and Ziwei Ye[2]

[1]Department of Electrical Engineering, National Tsing-Hua University, Taiwan
chenglian.liu@gmail.com
[2]Department of Computer Science and Technology, Tsing-Hua University, Beijing, China.
454109135@QQ.COM

Abstract. The RSA-768 (270 decimal digits) was factored by Kleinjung et al. on December 12 2009, and the RSA-704 (212 decimal digits) was factored by Bai et al. on July 2, 2012. And the RSA-200 (663 bits) was factored by Bahr et al. on May 9, 2005. Until right now, there is no body successful to break the RSA-210 (696 bits) currently. In this paper, we would discuss an estimation method to approach lower/upper bound of $\phi(n)$ in the RSA parameters. Our contribution may help researchers lock the $\phi(n)$ and the challenge RSA shortly.

Keywords: RSA cryptosystem; Euler's totient function; Factoring;

1 Introduction

Challenge RSA [12] is an interesting and difficult work. Recently, most scientists and researchers [1, 4, 8] using general number field sieve (GNFS) algorithm to factor RSA modulus n. In practical environment, it looks like if you to want to break the RSA, you may have best choice to choose GNFS when you already factor the modulus n. In theoretical, Wiener [17] first proposed a cryptanalysis of short secret exponents where the $d < N^{0.5}$ in 1990. Boneh [3] presented 'Twenty years of attacks on RSA cryptosystem' in 1999. He classified and described varieties attack. Followed by Boneh and Durfee [2], they suggested the provate key d should be greater than $N^{0.292}$ for the security problem. Even though, some bodies focus on secret key d or factor composite number n. Their purpose are clearly. We can not help but think, does it exist a general estimation way without factor to challenge RSA? In this article, we would introduce a new methodology where approach the lower bound and the upper bound of $\phi(n)$. For this generalize conception, it may match any bit length composite number n.

2 Review of RSA Conception

The signer prepares the prerequisite of a RSA signature: Two distinct large prime p and q, $n = pq$, Let e be a public key so that $\gcd(e, \phi(n)) = 1$, where $\phi(n) = (p-1)(q-1)$, then calculate the private key d such that $ed \equiv 1 \pmod{\phi(n)}$. The signer publishes (e, n) and keeps (p, q, d) secretly. The notation as same in [12].

2.1 RSA Encryption and Decryption

In RSA public-key encryption, Alice encrypts a plaintext M for Bob using Bob's public key (n, e) by computing the ciphertext

$$C \equiv M^e \pmod{n} \tag{1}$$

where n, the modulus, is the product of two or more large primes, and e, the public exponent, is an (odd) integer $e \geq 3$ that is relatively prime to $\phi(n)$, the order of the multiplicative group $\mathbb{Z}_n^*$.

2.2 RSA Digital Signature

$$s \equiv M^d \pmod{n} \tag{2}$$

where (n, d) is the signer's RSA private key. The signature is verified by recovering the message M with the signer's RSA public key (n, e):

$$M \equiv s^e \pmod{n} \tag{3}$$

3 Our Methodology

In this section, we would calculate the upper bound and the lower bound of $\phi(n)$ in RSA scheme. The detail described as below.
Notation:
ℓ: means lower bound.
u: means upper bound.
ε: a decimal expansion number (e.g $99/100 = 0.99\cdots$).

3.1 Approaching $\phi(n)$

If n is composite, hence

$$\phi(n) \leq n - \sqrt{n}, \tag{4}$$

Sierpinski [15] mentioned it in 1964. It is know that if equation (4) is a good upper bound for $\phi(n)$. Is there a good lower bound for $\phi(n)$? This question also be discussed by a newsgroup dialog between Ray Steiner and Bob Silverman in 1999 [16]. For $n > 30$, the $\phi(n) > n^{2/3}$, Kemdall and Osborn proved it [7]; for $n \geq 3$, the $\phi(n) > \frac{\log 2}{2}\frac{n}{\log n}$ given by Hatalova and Salat [6].

3.1.1 Estimate upper bound Does the equation (4) a good upper bound? In follows, we would estimate a new value where its smaller than previous and optimize.

Theorem 1. *Assume p, q are large prime numbers, where $n = pq$, then $\phi(n) = 4k$, $k \in \mathbb{Z}$ where $1 \leq k \leq \lfloor \frac{n-2\lceil\sqrt{n}\rceil+1}{4} \rfloor$.*

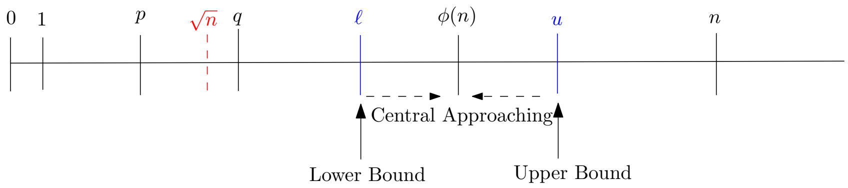

Fig. 1. The lower/upper bound of $\phi(n)$ in RSA scheme.

Proof. As known two variant p and q are large prime numbers, and $p, q > 2$, since $2 \nmid p,\ 2 \nmid q$,
therefore $2 \mid p-1,\ 2 \mid q-1$.
$4 \mid (p-1)(q-1),\ 4 \mid \phi(n)$.
$\phi(n) = 4k,\ k \in \mathbb{Z}^+$.
We will discuss how calculate the range of value k.

$$\begin{aligned} \phi(n) &= (p-1)(q-1) \\ &= pq - (p+q) + 1 \\ &= n - (p+q) + 1. \end{aligned} \tag{5}$$

And

$$\begin{aligned} p+q &\geqslant 2\sqrt{n}, \\ p+q &\in \mathbb{Z}^+, \\ 2 &\mid p+q. \\ p+q &\geqslant 2\lceil\sqrt{n}\rceil. \\ \phi(n) &\leq n+1-2\lceil\sqrt{n}\rceil. \\ \phi(n) &= 4k,\ k \in \mathbb{Z}^+. \\ \phi(n) &\leq 4 \cdot \lfloor \frac{n+1-2\lceil\sqrt{n}\rceil}{4} \rfloor. \end{aligned} \tag{6}$$

Here, we know the maximum value (limit superior) for $k \leq \lfloor \frac{n-2\lceil\sqrt{n}\rceil+1}{4} \rfloor$, we call boundary value.
Consequently, according to above inference, we obtain a best upper bound u of $\phi(n)$ where $\phi(n) \leq 4\lfloor \frac{n-2\lceil\sqrt{n}\rceil+1}{4} \rfloor$.

Theorem 2. *Assume p, q are large prime numbers, and $p, q > 2$, $n = pq$, where $\phi(n) \leq t \leq n, t \in \mathbb{Z}$. Then $t = \phi(n) \iff x^2 - (n+1-t)x + n = 0$ have two positive integer solutions.*

Proof. We describe two properties, necessary and sufficient conditions as follow:
Necessary condition:
If $t = \phi(n) = (p-1)(q-1) = pq-(p+q)+1 = n-(p+q)+1$, then $n+1-t = p+q$,

in the same time, the formula be $x^2 - (p+q)x + pq = 0$. It is clearly, the equation of the two roots p and q are needed to set up.
Sufficient condition:
If equation $x^2 - (n+1-t)x + n = 0$ have two integer solutions.
Assume x_1, x_2 be the equation of the two roots, where $x_1, x_2 \in \mathbb{Z}^+$.
The equation could be transformed to $(x - x_1)(x - x_2) = 0$.
Promptly, $x^2 - (x_1 + x_2)x + x_1 x_2 = 0$.
Then $n = x_1 x_2$, according to n's structure, there are two choice:
1) $x_1 = 1$ and $x_2 = n$ (or $x_1 = n$ and $x_2 = 1$).
2) $x_1 = p$ and $x_2 = q$ (or $x_2 = q$ and $x_2 = p$).
If x_1, x_2 one for 1 and n.
Now, $x_1 + x_2 = n + 1$, because $x_1 + x_2 = n + 1 - t$, hence $t = 0$.
However, in our assumption, the condition is $t > 0$, so this inference contradiction.
The equation have two integers p and q where

$$x_1 + x_2 = p + q, \tag{7}$$

then

$$p + q = n + 1 - t. \tag{8}$$

Promptly,

$$t = n + 1 - (p + q), \tag{9}$$

and

$$t = \phi(n). \tag{10}$$

Thus, for the sufficient condition is setting up.

Theorem 3. *Assume p, q are large prime numbers, and $p, q > 2$, $n = pq$, $t = 4k$ where $\phi(n) \le t \le n, t \in \mathbb{Z}$. If $\sqrt{(n+1-t)^2 - 4n}$ is an integer number, the equation $x^2 - (n+1-t)x + n = 0$ has two integer solutions.*

Proof. Since p, q are both prime numbers where $n = pq$, $2 \nmid n$, but $2 \mid n + 1$.
Suppose $t = 4k$, and $2 \mid t$. Thus $2 \mid n + 1 - t$.
If $\sqrt{(n+1-t)^2 - 4n} \in Z$, so $2 \mid \sqrt{(n+1-t)^2 - 4n}$.
The equation $x^2 - (n+1-t)x = 0$ of the two solutions are:
$x = \frac{n+1-t \pm \sqrt{(n+1-t)^2 - 4n}}{2}$.
Because $2 \mid n + 1 - t$, and while $\sqrt{(n+1-t)^2 - 4n} \in \mathbb{Z}$,
it exists $2 \mid \sqrt{(n+1-t)^2 - 4n}$.
When $\sqrt{(n+1-t)^2 - 4n} \in \mathbb{Z}$, $2 \mid n + 1 - t \pm \sqrt{(n+1-t)^2 - 4n}$.
Due to $\sqrt{(n+1-t)^2 - 4n} \in \mathbb{Z}$, the x is also $\in \mathbb{Z}$.
Here, the two solutions of equation $x^2 - (n+1-t)x + n = 0$ are both integers.

3.1.2 Estimate lower bound Loomis et al. [10] thought the Shapiro's [13] lower bound $\phi(n) > n^{(\log 2)}/(\log 3)$ as a (naive) lower bound for E_n, they can determine when all members of a given E_n have been found. Powell [11] pointed out that Konyagin and Shparlinksi's lower bound $N_1(n, p) > (p-1)/2 - p^{3/2}/n$ where $n > 1$ is a positive integer and that p is an odd prime number with $p \equiv 1 \pmod{n}$; that is a good bound if p is a small compared to n, and establishes that $N_1(n,p) \geq (\sqrt{\phi(n)}(\prod_{\substack{q\ prime \\ q|n}} q^{1/(q-1)})/n)p^{1-1/\phi(n)}$. Powell also discussed an improvement the upper and lower bounds in [11]. What is the optimal lower bound? The other discuss in following:

Theorem 4. *For all $n \geq 3$ we have $\phi(n) \geq \frac{n}{e^{\gamma}\log\log n} + \text{Ø}(\frac{n}{(\log\log n)^2})$, where γ is the Euler-Mascherone Constant, and the above holds with equality infinitely often.* ***Remark:****note in particular that since $loglogn \rightarrow \infty$ as n grows large, we see that the result $\frac{n}{m} < \phi(n)$ can not hold for any fixed integer m.*

Proof. Consider *R*, set of all n such that $m < n$ implies $\frac{\phi(n)}{n} < \frac{\phi(m)}{m}$. This set is then all of the 'record breaking' n. If $n \in R$ has k prime factors, let n^* be the product of the first k prime factors. If $n \neq n^*$ and $\frac{\phi(n)^*}{n^*} \leq \frac{\phi(n)}{n}$, which is impossible. Hence, *R* consist entirely of n of the form $n = \prod_{p \leq y} p$ for some y. Now for $n \in R$, we can choose y so that $\log n = \sum_{p \leq y} \log p = \theta(y)$. Then using one of Mertens estimates we see that $\frac{\phi(n)}{n} = \prod_{p \leq y}(1 - \frac{1}{p}) = \frac{e^{-\gamma}}{\log y} + O(\frac{1}{(\log y)^2})$. Since $\log\log n = \log(\theta(y)) = \log y + \text{Ø}(1)$ by Mertens estimates again, we have for $n \in R$, $\phi(n) = \frac{ne^{-r}}{\log\log n} + O(\frac{1}{(\log\log n)^2}$.

RSA-200	Same digits length
n	27997833911221327870829467638722601621070446786955428537560009929326128400107609345671052955360856061822351910951365788637105954482006576775098580557613579098734950144178863178946295187237869221823983
$\phi(n)$	27997833911221327870829467638722601621070446786955428537560009929326128400107609345671052955360856050364020022070262634017415134803482520365925322995768594715101139912289736681370959747280607953550168

Fig. 2. The same digits of $\phi(n)$ and modulus n parameters in RSA.

From above, it seems so complexity. Does it exist a simple computation method? We observed the modulus n with $\phi(n)$, there have some characteristics. An example for RSA-200, the modulus n and the $\phi(n)$ are 200 decimal digits. We compared n and $\phi(n)$ each other, there are the same digits from left to right **110** digits. The example showed in Figure 2. Discuss on RSA modulus number with haft of the bit prescribed, be introduced some literatures in [5, 9, 14]. To RSA-704, the n and $\phi(n)$ had same

Table 1. Comparison of some types in RSA parameters. Unit: decimal digits

type	Modulus n length	$\phi(n)$ length	p length	q length	same digits n & $\phi(n)$	same digits $\phi(n)$ & u
RSA-200	200	200	100	100	110	101
RSA-210	200	200	105	105	?	?
RSA-704	212	212	106	106	106	108
RSA-768	232	232	116	116	115	120

digits 106, it amounts same length with p or q. We computed the upper bound value according to Theorem 1, this upper bound had same 108 digits with its $\phi(n)$. And analyzed the RSA-768, the n had 115 digits same with $\phi(n)$; the $\phi(n)$ had 120 digits same with its upper bound u. Please see Table 1 We observed the relationship of $\phi(n)$ and its boundary value k, when $\phi(n)$ divided by k, we obtained this quotient are very approaching to number 4, these lower bounders are very close to multiples of number 4. We say $3.\overline{999}$, and have 106's 9 after decimal point for case of RSA-210 type. The lower bound approximation figure diagram be shown in figure 4 and in Table 2. As

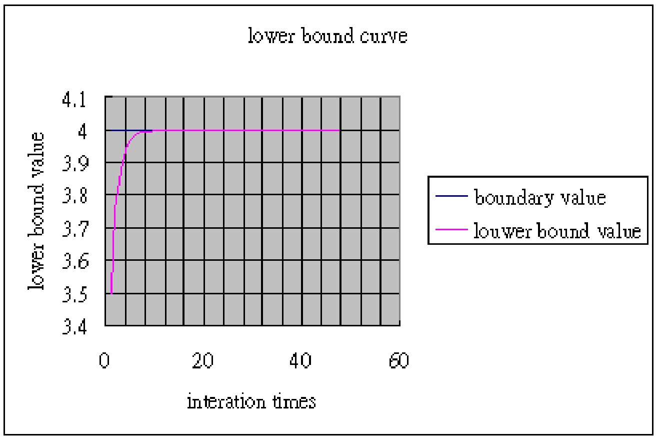

Fig. 3. The lower bound approximation curve status.

Table 2. The relationship of $\phi(n)$ and its boundary value k.

type	$\phi(n)/k$	statement
RSA-200	$3.\overbrace{99999}^{99's\ 9}8$	there have 99's 9 after the decimal point
RSA-210	$3.\overbrace{99999}^{106's\ 9}2$	Estimating have 106's 9 after decimal point
RSA-704	$3.\overbrace{99999}^{107's\ 9}8$	there have 107's 9 after decimal point
RSA-768	$3.\overbrace{99999}^{117's\ 9}7$	there have 117's 9 after decimal point

known as the modulus number n of RSA-210, we re-estimated its lower/upper bounds. We assume:

$$(3+\varepsilon)\lfloor\frac{n-2\lceil\sqrt{n}\,\rceil+1}{4}\rfloor \leq \phi(n) \leq 4\lfloor\frac{n-2\lceil\sqrt{n}\,\rceil+1}{4}\rfloor, \tag{11}$$

where $\varepsilon = 0.\overbrace{99999}^{106's\ 9}$. We therefore compute the upper bound u and lower bound ℓ; those shown result in Figure 4.

RSA-210	
u	245246644900278211976517663573088018467026787678332759743414451715061600830038587216952208399332071549102636379525419241883591878719807874925061718037353593039323605526518763037740989017744115767482964632709008
ℓ	245246644900278211976517663573088018467026787678332759743414451715061600830038587216952208399332071549102617986027051721017579734131535066633608723320135703257895405070218987602113186570983810232135299645833216

Fig. 4. The lower/upper bound parameters of $\phi(n)$ in RSA-210.

4 Conclusion

In this paper, we re-estimated a new lower/upper bound values of $\phi(n)$ in RSA-210, our methodology are easily, simply, clearly, no intricately and intuitive. It may be useful to researchers who would quickly reduce the searching ranges. Looking back, more researchers focus on secret d or modulus n, such as well known short exponent attack, side channel attack, or common modulus and cyclic attacks. Our method is differ to previous literatures. Finally, We presented what we claimed actually.

References

1. Bai, S., Thomé, E., Zimmermann, P.: Factorisation of RSA-704 with cado-nfs. IACR Cryptology ePrint Archive p. 369 (2012)
2. Boneh, D., Durfee, G.: Cryptanalysis of RSA with private key d less than $N^{0.292}$. IEEE Transactions on Information Theory 46(4), 1339–1349 (July 2000)
3. Boneh, D.: Twenty years of attacks on the RSA cryptosystem. Notices of the AMS 46, 203–213 (1999)
4. F. Bahr, M. Boehm, J.F., Kleinjung, T.: RSA-200 is factored. Website (2005), `http://www.rsa.com/rsalabs/node.asp?id=2879`
5. Graham, S.W., Shparlinski, I.E.: On RSA moduli with almost half of the bits prescribed. Discrete Applied Mathematics 156(16), 3150–3154 (2008)
6. Hatalova, H., Salat, T.: Remarks on two results in thelementary theory of numbers. Acta. FacRer. Natur Univ Comenian. Math. (20), 113–117 (1969)
7. Kendall, D.G., Osborn, H.B.: Two Simple Lower Bounds for Euler's Function, vol. 17. Texas Journal of Science (1965)
8. Kleinjung, T., Aoki, K., Franke, J., Lenstra, A.K., Thomé, E., Bos, J.W., Gaudry, P., Kruppa, A., Montgomery, P.L., Osvik, D.A., Te Riele, H., Timofeev, A., Zimmermann, P.: Factorization of a 768-bit RSA modulus. In: Proceedings of the 30th annual conference on Advances in cryptology. CRYPTO'10 (2010)

9. Meng, X.: On RSA moduli with half of the bits prescribed. Journal of Number Theory 133(1), 105–109 (2013), http://www.sciencedirect.com/science/article/pii/S0022314X12002302
10. Paul Loomis, M.P., Polhill, J.: Summing up the Euler ϕ function. The College Mathematics Journal 39(1), 34–42 (2008)
11. Powell, C.: Bounds for multiplicative cosets over fields of prime order. Mathematics of Computation 66(218), 807–822 (April 1997)
12. Rivest, R., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM 21, 120–126 (1978)
13. Shapiro, H.: An arithmetic function arising from the ϕ function. The American Mathematical Monthly 50, 18–30 (1943)
14. Shparlinski, I.: On RSA moduli with prescribed bit patterns. Designs, Codes and Cryptography 39 (2006)
15. Sierpinski, W.F.: Elementary Theory of Numbers. North-Holland PWN-Polish Scientific Publishers, Warsawa (1964)
16. Silverman, B.: Euler's phi functions. Website (2012), http://www.math.niu.edu/~rusin/known-math/99/min_phi
17. Wiener, M.J.: Cryptanalysis of short RSA secret exponents. IEEE Transations on Information Theory