An Approach to Factorization and Attack Against the Asymmetric Cryptographic Algorithm RSA

Petar Ts. Antonov[1] and Valentina R. Antonova[2]

***Abstract* – Problems of the security of the most popular asymmetric cryptographic algorithm RSA are considered. A new approach to factorization of large numbers and attack against RSA is proposed. A recommendation about secure generation of public and secret keys of RSA is formulated.**

***Keywords* – RSA, factorization, crypto-attack.**

I. INTRODUCTION

RSA (Rivest-Shamir-Adleman) is the most often used asymmetric cryptographic algorithm in security schemes of computing and communication systems. It is one of the few known cryptographic algorithms of this class that are distinguished for their universal application – in ciphering of messages, digital signatures and exchange of session secret keys in hybrid ciphering schemes. It has been perceived by ISO, ITU-T, ANSI, banking and financial spheres, military units, etc. It is used in crypto-systems to protect e-mail, in electronic trade, credit card systems, Internet-browsers, etc. Actually RSA is de facto world–wide standard for asymmetric cryptography, well known also as cryptography with public keys.

The description of RSA is very simple [1, 3, 4, etc.]. Initially every user chooses two sufficiently large prime numbers a and b, whereupon is calculated

$$n = a.b \quad \text{and} \tag{1}$$

$$\Phi(n) = (a-1).(b-1) \tag{2}$$

The public key Kp and the secret key Ks of the pair "public/secret" keys of that user are determined from the correlations:

$$\text{НОД}(K_p, \Phi(n)) = 1 \quad \text{and} \tag{3}$$

$$(K_s.K_p) \bmod \Phi(n) = 1, \tag{4}$$

where НОД is the greatest common divisor of the numbers in square brackets [the number Kp must be accidentally chosen and mutually simple to the number Φ(n)].

Further on in the procedures of ciphering and deciphering respectively the pairs of numbers (Kp,n) and (Ks,n) are used, as the first pair is made public attribute, Ks is kept in secret, and the initial numbers a and b are destroyed.

Any accidental source can address a protected message to that user, ciphering the open text M of the message with the public key K_P and send the resultant ciphered text E. Only that user is able to restore M after deciphering E with his own secret key K_S.

Any eventual offender, received somehow the ciphered text E, will know K_P and n. To restore the open text M however, this offender will have to factorize n to its unknown prime factors a and b (this procedure is known as factorization of n), whereupon to calculate Φ(n) and K_S. But this factorization is very labour-consuming and with sufficiently large values of a and b becomes practically beyond the potentialities of contemporary level of technologies.

It is possible that the offender makes a try to find out Φ(n) directly, without factorization of n, but this is not simpler than the factorization itself.

The third possibility for the offender is to immediately calculate K_S without factorization of n and determination of Φ(n), but with sufficiently large K_S, this possibility is not easier than the factorization itself either.

In essence, the above three crypto-attacks to RSA can be considered as attack of the brute force. In the end, however, the real practical attack to RSA comes to solving the problem with the factorization of n.

II. AN APPROACH TO FACTORIZATION

A number of methods to solve this problem are known (see [2, 3, 4, etc.]), which differ in various labour-consumption. A new method of factorization is presented below, which in a number of cases requires less steps, compared to the other similar methods known. The idea of this method is based on the popular Theorem of Euler, saying that any even number, greater than 2, can be represented as a sum of two prime numbers.

If a successful factorization of n is made and the prime factors **a** and **b** are found, then further on Φ(n) will be calculated and after that the secret key K_S from the correlation:

$$(K_s.K_p) \bmod \Phi(n) = 1. \tag{5}$$

To expose the main point of the proposed method for factorization, we shall use the following example:

[1]Peter Ts. Antonov is with the Department of Computer Science and Engineering, Technical University of Varna, 1, Studentska Str., 9010 Varna, Bulgaria, E-mail: peter.antonov@ieee.org

[2]Valentina R. Antonova is with the Department of Computer Science and Engineering, Technical University of Varna, 1, Studentska Str., 9010 Varna, Bulgaria,E-mail:valyvarna@yahoo.com

a=13, b=29, n=377, $\lfloor n^{0.5} \rfloor$=19.

Further on we shall write down the two-row sequence of the numbers (1÷$\lfloor n^{0.5} \rfloor$) and ($\lfloor n^{0.5} \rfloor$÷37) so that the sum of the numbers situated one under another is equal to 2.$\lfloor n^{0.5} \rfloor$=2.19=38:

1 2 3 4 5 6 7 8 9 10 11 12 **13** 14 15 16 17 18 19
37 36 35 34 33 32 31 30 **29** 28 27 26 25 24 23 22 21 20 19.

As a<b, then in the first row of the sequence above, the number a=13 is situated closer to 19=$\lfloor n^{0.5} \rfloor$, than b=29 in the second row. The distance from a to $\lfloor n^{0.5} \rfloor$ in this case is La=$\lfloor n^{0.5} \rfloor$-a=19-13=6, and from b to $\lfloor n^{0.5} \rfloor$ - Lb=b-$\lfloor n^{0.5} \rfloor$=29-19=10. If we increase $\lfloor n^{0.5} \rfloor$=19 consecutively by one and two, then the new similar two-row sequences of numbers will look like the ones below:

1 2 3 . . . . . 11 12 **13** 14 . . . . . 19 20
39 38 37 . . . . . . **29** 28 27 26 . . . . . 21 20

1 2 3 . . . . . 11 12 **13** 14 . . . . . 19 20 21
41 40 39 . . . . . 31 30 **29** 28 . . . . . 23 22 21

In the second sequence the numbers a=13 and b=29 are already situated one under another and their sum is equal to 42=2.21, and their product - to n=377. To obtain this correspondence in this case only two iterations came out to be enough.

If in the general case, the necessary number of iterations for factorization of n is denoted by V, it is easy to prove the validity of the following correlation:

$$V = \frac{1}{2}\left(L_a - L_b\right) = \frac{a + b - 2\lfloor \sqrt{n} \rfloor}{2}. \quad (6)$$

It can be seen that the prime factors a and b of the number n we have been after, can be determined from the correlations:

$$a = \lfloor \sqrt{n} \rfloor + V - x \quad (7)$$

$$b = \lfloor \sqrt{n} \rfloor + V + x, \quad (8)$$

which lead to

$$a + b = 2\left(V + \lfloor \sqrt{n} \rfloor\right). \quad (9)$$

In the example, considered above, V=2 and x=8, while a=19+2-8=13 and b=19+2+8=29. Essentially x is the distance between a and b, and the right-hand beginning of the two-row sequence of numbers at the last iteration, which in this case is the second in order. After the second iteration, the numbers a and b proved to be one under another and at the same distance from the beginning of this sequence, which in this example is the number ($\lfloor n^{0.5} \rfloor$+V)=21.

As a.b=n, then the second factor b can be expressed simply by the first factor a, i. e. b=n/a and then:

$$a + n/a = 2\left(V + \lfloor \sqrt{n} \rfloor\right) \quad (10)$$

$$a^2 - 2\left(V + \lfloor \sqrt{n} \rfloor\right)a + n = 0. \quad (11)$$

Solving the quadratic equation above, finally we get:

$$a = V + \lfloor \sqrt{n} \rfloor - \sqrt{\left(V + \lfloor \sqrt{n} \rfloor\right)^2 - n}, \quad (12)$$

$$b = V + \lfloor \sqrt{n} \rfloor + \sqrt{\left(V + \lfloor \sqrt{n} \rfloor\right)^2 - n} \quad (13)$$

These two correlations can be used for factorizing n. For that purpose V is given consecutive integer values, starting with 1. For each consecutive value of V the expression under the radical is calculated

$$\left(V + \lfloor \sqrt{n} \rfloor\right)^2 - n \quad (14)$$

and is checked if it is equal to the square of an integer. If it is not, V is given the next consecutive value and so on. The solution for a and b is reached at that V, for which the expression

$$\sqrt{\left(V + \lfloor \sqrt{n} \rfloor\right)^2 - n} \quad (15)$$

accepts integer value, denoted in this case by x. Then from the expressions above for a and b we calculate their specific needed values.

If in the expression for V we replace b by α.a, i.e. b=αa, then

$$V = \frac{a + \alpha a - 2\lfloor \sqrt{\alpha a^2} \rfloor}{2} = \left\lfloor a\frac{1 + \alpha - 2\sqrt{\alpha}}{2} + 1 \right\rfloor. \quad (16)$$

For approximate calculation of the necessary number of iterations the formula

$$V^* = a\frac{1 + \alpha - 2\sqrt{\alpha}}{2}, \quad (17)$$

can be used, which for the sake of convenience can be presented in the form:

$$\left(\frac{V^*}{a}\right) = 0.5\left(1 + \alpha - 2\sqrt{\alpha}\right) \quad (18)$$

In this formula, for the purpose of simplification of the expression, the requirement for separation the integer part of the result of the square root calculation of αa^2 is removed.

The graph of the changes in (V*/a) depending on α is shown below (fig.1).

If the prime factors a and b have the same decimal rank, i.e. the same number of decimal classes, then

$$a_{\min} = 1000..... \quad (19)$$

$$b_{\max} = 9999..... \text{ and} \quad (20)$$

$$\alpha_{\max} = 9{,}999..... \approx 10\,. \quad (21)$$

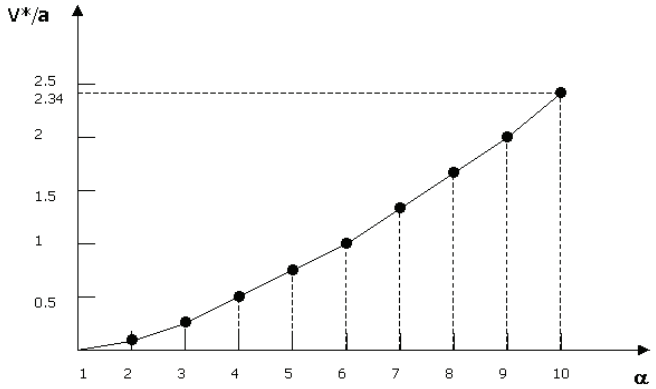

Fig. 1.

In this most unfavourable case of factorization of n, the number of iterations $V^*=2.34a=2.34\cdot 10^h$, where h is the decimal rank of a and b.

III. Conclusion

Considering what has been said above, we can draw the following conclusion: the exposed method for factorization requires smaller number of iterations at smaller difference between the values of a and b. Furthermore, the method is applicable even when a and b are not prime numbers, but odd or even numbers.

The comparative estimations made for the necessary number of iterations between the method presented and the methods for factorization, known from literature, show that in some cases it is with less labour consumption, and in other cases – not. For example, for factorization of n=22317 using the method proposed, only two iterations will be necessary. For comparison, according to the estimations given in [4] for the necessary number of iterations with different methods, for the algorithm of Dixon in this case will be needed:

$$e^{\sqrt{\log n.\log n(\log n)}} \cong 6 \text{ iterations,} \quad (22)$$

And for the Quadratic Sieve Algorithm:

$$e^{\sqrt{\ln n.\ln(\ln n)}} \cong 122 \text{ iterations.} \quad (23)$$

In conclusion we shall annotate that the method proposed can be considered as development of the problem for factorization that has more general nature. Using this method, a quicker crypto-attack can be accomplished to the cryptographic algorithm RSA, provided the numbers a and b have close values.

References

[1] Антонов, П. Ц., С. В. Малчев, *Криптография в компютърните комуникации*, Варна, 2000.

[2] Кнут, Д. *Искусство программирования для ЭВМ*. Т. 2. Получисленные алгоритмы, Пер. с англ. – Москва, Мир, 1977.

[3] Schneier, B. *Applied Cryptography. Protocols, Algorithms and Source Code in C*, John Wiley & Sons Inc., 1996.

[4] Goldwasser, S., M. Belare. *Lecture Notes on Cryptography. – Cambridge*, Massachusetts Institute of Technology, June 1997.