Design and Implementation of Rivest Shamir Adleman’s (RSA) Cryptography Algorithm in Text File Data Security

Hengki Tamando Sihotang\textsuperscript{1,}\textsuperscript{*}, Syahril Efendi\textsuperscript{2}, Elvyawati M Zamzami\textsuperscript{3}, Herman Mawengkang\textsuperscript{4}

\textsuperscript{1,2,3,4} Doctoral Program, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Sumatera Utara, Indonesia

E-mail: hengki.tamando@yahoo.com

Abstract. Security of data text file on a computer can be done by utilizing encryption and decryption techniques. One technique is encryption and decryption of data encryption system text file with cryptography. Cryptography is the science or art to randomize and secure the message to avoid the manipulation of data by performing encryption and decryption on the data of the text file. Cryptographic algorithm consists of several types of one example is the algorithm Rivest Shamir Adleman (RSA). The RSA algorithm is one of asymmetric cryptography technique i.e. cryptography modern technique that has two keys of the public key used for encryption and a private key techniques to make the process deskripsi. The purpose of this research is to design and implement RSA cryptography algorithm on text file data security. The results of this design is a desktop based software. This research is expected to provide benefits to the users, individually as well for other organizations.

1. Introduction

Data and information are needed in the development of an institution, be it a company, government and even people’s lives. Information is now a very important commodity for anyone. Along with the development of Science and Technology, especially in the field of information technology, data processing into information has also changed and the way in which information dissemination / transfer has changed is faster with digital systems. Companies or Organizations must of course follow this development, problems that often occur in data transfer paths or in storage media often occur data theft, data damage, data manipulation, this is done by people who have special interests often called hackers, crackers and attackers,\textsuperscript{[1]} and even this often happens when data processing is in progress. If this happens to an organization or institution it will certainly cause a large loss and even lead to the closure of the company or organization. Because, every organization or company certainly has confidential data. Also in the development of science and technology, it should be balanced with the quality of data and information security. There are several theories about data security that have long been discovered. The data security technique is called cryptographic technology. Cryptography is the study of data security methods, namely the art and science of maintaining message security\textsuperscript{[2][3]. While the steps for securing data are called cryptographic algorithms\textsuperscript{[4]. For the file text data security system so that it cannot be read by unauthorized people is called Encryption, then the data
will be returned to what is later called Decryption\cite{5}\cite{6}. There are two types of cryptographic algorithms, symmetric and asymmetric cryptographic algorithms\cite{6}. The symmetric algorithm (conventional algorithm), which is an algorithm that uses only one key and the asymmetric cryptographic algorithm, also called (public key algorithm), is already sophisticated because it distinguishes between public key and private key\cite{6}. Asymmetric cryptographic algorithms consist of EEC, LUC, RSA, EL. The RSA algorithm is a third tier algorithm\cite{7}\cite{5}. In this study the authors will develop the science of cryptography for data.text data security by selecting the Ron Riverst Shamir Leonard and Adleman (RSA) algorithm\cite{8}\cite{9}. This algorithm is abbreviated as RSA because this algorithm was established by three people and abbreviated from the three names Riverst Shamir Adleman. This algorithm is already modern by spreading public keys and making secret keys\cite{9}. By using cryptographic RSA algorithm it can be used to examine the authenticity of a digital data\cite{10}\cite{11}. Cryptography with RSA algorithm has been able to secure data up to the row and column level while maintaining the integrity of the data within the authority of each user. The RSA algorithm is the most advanced Asymmetric randomization data algorithm in the field of public key cryptography\cite{7}\cite{12}\cite{13}.

2. Methods
Problem solving in this study by applying cryptographic algorithms with the RSA Method (Ron Riverst Shamir Leonard and Adleman)\cite{14}. The Completion Flow using this method is in the picture and explanation below:

a. Key Formation Flow
   The flow of the key formation program can be described as follows:

![Key Formation Flow](image)

b. File Encryption Process Flow
   Encryption is called data security to ensure confidentiality. The original data is called plaintext or cleartext which is changed into codes that are not understood\cite{15}. This encryption is interpreted by chipher or code\cite{16}. Data encoding is done using an algorithm that can be used to encode data. In the encryption process there are several steps that must be passed by the user before getting the encrypted file results, following the flowchart image to do the encryption process:
c. Program Decryption Process File

In the decryption process, there are several steps that must be passed before the user gets the decrypted file[17], the following is a flowchart to do the decryption process:

![Flowchart of Program Decryption Process File Algorithm](image)

**Figure 3.** Program Flow Process Description File Algorithm

d. RSA algorithm

The RSA cryptographic system encryption and decryption algorithm rests on the assumption of a one-way function constructed by a modular exponential function in the multiplication
group \( \mathbb{Z}_n^* \) and multiplication group \( \mathbb{Z}_n^* \) with \( n = pxq \), \( p, q \) is a prime number and \( \phi(n) = (p - 1) \times (q - 1) \). \( \mathbb{Z}_n^* \) \( \mathbb{Z}_n^* \) [13].

1. RSA Encryption and Decryption

In the encryption process, the main requirement that must be met is the formation of a public key. In addition, the requirement that must be met in the encryption process is that the plaintext value must be greater than 0 and smaller than the modulus \( n \) value of the public key. In the encryption process also if the plaintext comes from the letters or symbols, it is necessary to convert it to ASCII code. This is to support multiplication of the RSA encryption process[10]. Before doing the multiplication using the above formula, then in RSA encryption, the plaintext must first be converted into decimal numbers. Changing the plaintext into this decimal number must refer to the ASCII code value. The length of the encryption process at RSA is dependent on the amount of text to be encrypted and the length of the key used[16].

In the decryption process of the RSA algorithm, it has similarities with the encryption process. The difference lies in the use of values \( e \) and \( d \). This decryption algorithm is also a modular \( n \) exponential using a private key[8]. The product of the formula is a decimal number. Because in the encryption process, plaintext is converted to decimal numbers using ASCII code values, so in the decryption process this also applies. The plaintext value in the form of decimal numbers must be converted to char using the ASCII code value[2].

Rivest Shamir Adleman’s cryptography (RSA is one of modern cryptography. RSA is an algorithm that involves expressions with exponential functions) [12]. Plain is encrypted in a block, where each block has a binary value less than a certain number \( (n) \). The encryption and decryption process for plain \( P \) block and \( C \) block cipher can be described as follows[7]:

a. Encryption Process

\[ C = P^e \mod n \]

Information:

\( C = \) Chiper
\( P = \) Plain
\( e = \) Public Key (\( e \))
\( n = \) Product of 2 prime numbers

b. Decryption process

\[ P = C^d \mod n \]

Information:

\( P = \) Plain
\( C = \) Chiper
\( d = \) Private Key (\( d \))
\( n = \) Product of 2 prime numbers

Both parties (sender and receiver) must know the value of \( n \). The sender knows the value of \( e \) and only the recipient knows the value of \( d \). So, it can be concluded that the public key of this algorithm is \( KU = e, n \) and the private key is \( KR = d, n \). Not free, must go through certain formulas[10][14].

2. RSA Key Formation

The establishment or generation of RSA keys in the proposed system is a process for establishing public and private keys that will be used in the encryption and decryption process. Simply put the RSA key formation algorithm in the proposed system is as follows[8][18]:

1) Find and determine prime numbers randomly and store them in variables \( p \) and \( q \),
provided the number of bits for this number is the same.

2) Calculate the value \( n = pq \), where this value will be used to modulus the public key and private key.

3) Calculate the value of \( pq = (p - 1) \times (q - 1) \) to use in searching for a private key value.

4) Choose the value of \( e \) for the public key with the conditions \( 1 < e < pq \) and \( \gcd(e, pq) = 1 \), where the value of \( e \) is usually a relatively small value. If the \( e \) criteria do not match these conditions, then the corresponding \( e \) value must be sought, then what must be sought next are the \( p, q, n \) and \( pq \) values. The \( \gcd \) (Greatest Common Divisor) function used in determining the value of \( e \) is a function for calculating the largest factorization value of a number.

5) Then the next step is to choose the value of \( d \), with the condition that the value of \( d \) meets \((dx)e \mod pq = 1\).

\[
\begin{align*}
n &= p \times q \\
\phi(n) &= (p - 1) \times (q - 1) \\
e &\leftarrow Z \text{ with } \gcd(e, \phi(n)) = \ldots 1 \\
d &= e^{-1} \text{ on } \mathbb{Z} \\
K_{publik} &= (e, n), K_{privat} = d = \ldots 2
\end{align*}
\]

3. Results and Discussion

Implementation of information security with cryptographic techniques on the RSA algorithm is presented in Figure 4.

![Cryptographic interface design of tex data security with RSA method](image)

**Figure 4.** Cryptographic interface design of tex data security with RSA method

Testing Scenarios:

1. Testing the Key Formation Process RSA security lies in the difficulty of factoring large numbers into prime factors. Factoring is done to obtain a private key. As long as factoring large numbers into prime factors, a good algorithm has not yet been found, so long as the security of the RSA algorithm is also guaranteed.

The key manufacturing process in the RSA algorithm consists of several steps, namely:

a) Choose two arbitrary primes, \( p \) and \( q \).

b) Calculate \( n = p \times q \) (preferably \( p^1 q \), because if \( p = q \) then \( n = p^2 \) so that \( p \) can be obtained by pulling the square root of \( n \)).

c) Calculate \( \phi(n) = (p - 1) \times (q - 1) \).

d) Choose the public key, \( e \), which is relatively prime to \( \phi(n) \).

e) Generate a private key using the equation \( e \times d = 1 \mod \phi(n) \).

Note that \( e \times d = 1 \mod \phi(n) \) is equivalent to \( e \times d = 1 + k \phi(n) \), so \( d \) can be calculated with \( d = \frac{1+k\phi(n)}{e} \).
There will be an integer $k$ that gives the number round $d$. Results from the algorithm above:

a. The public key is a pair $(e, n)$

b. The private key is a pair $(d, n)$

Note: $n$ is not confidential, but it is required for calculations encryption / decryption.

Sample case:

a) The user chooses $p = 11$ and $q = 17$ (both are prime).
b) Next the user counts: $n = pxq = 187$
c) Calculate the value of $m = (p-1) (q-1) = (11-1) (17-1) = 160$
d) The user chooses public key $e = 3$, because 3 is relatively prime with 160. Because $e = 3 \rightarrow \text{gcd}(3,160) = 1$.
e) Calculate the decryption key $d$ with the following equation:

$$d = \frac{1 + k \cdot \phi(160)}{3} = 3 \cdot 107 \mod 160 = 321 \mod 160 = 1$$

By trying the values $k = 1, 2, 3, \ldots$, we get the value of round $d$ which is 107. This is the private key to decrypt the message. This key must be kept secret by the sender.

In order to obtain public keys and private keys, namely:

Public Key: $(e = 3, n = 187)$

Private Key: $(d = 107, n = 187)$

2. Testing the Encryption Process and Description The cryptographic system encryption process using the RSA algorithm can be carried out as follows:

a) Take the recipient’s public key, $e$, and modulus $n$.
b) Express plaintext $m$ to be blocks $m_1, m_2, \ldots$, so that each block represents a value in interval $[0, n - 1]$.
c) Each $m_i$ block is encrypted into a $c_i$ block with a formula $c_i = m_i^e \mod n$.

Whereas the decryption process can be done by: "Each ciphertext block $c_i$ is decrypted back to the $m_i$ block with the formula $m_i = c_i^d \mod n$”

The following will be explained An example of the Encryption process: Suppose the User will send a word file i.e. try.txt to the Recipient. The try.txt file contains data that is "I AM A STUDENT IN STITA PELITA NUSANTARA". First the User will convert the plain to ASCII Format. The file was changed in the ASCII Format as follows:

| CHARACTER | ASCII | CHARACTER | ASCII | CHARACTER | ASCII |
|-----------|-------|-----------|-------|-----------|-------|
| S         | 83    | SPACE     | 32    | SPACE     | 32    |
| A         | 65    | S         | 83    | N         | 78    |
| Y         | 89    | T         | 84    | U         | 85    |
| A         | 65    | M         | 77    | S         | 83    |
| SPACE     | 32    | I         | 75    | A         | 65    |
| K         | 75    | K         | 75    | N         | 78    |
| U         | 85    | SPACE     | 32    | T         | 84    |
| L         | 76    | P         | 80    | A         | 65    |
| I         | 73    | E         | 69    | R         | 82    |
| A         | 65    | L         | 76    | A         | 65    |
| H         | 72    | I         | 73    |
| SPACE     | 32    | T         | 84    |
| D         | 68    | A         | 65    |
| I         | 73    |           |       |           |       |

Then the User will encrypt the plaintext with the public key that has been obtained previously, i.e. $e = 3$ and $n = 187$. The user encrypts each plaintext block as follows:

So, the ciphertext generated by the User is:
Table 2. Describe Plaintext with a Public Key

| c1 = 833 mod 187 = 128 | c14 = 733 mod 187 = 57 | c27 = 653 mod 187 = 109 |
|------------------------|------------------------|------------------------|
| c2 = 353 mod 187 = 109 | c15 = 323 mod 187 = 43 | c28 = 323 mod 187 = 43 |
| c3 = 893 mod 187 = 166 | c16 = 833 mod 187 = 128 | c29 = 783 mod 187 = 133 |
| c4 = 653 mod 187 = 109 | c17 = 843 mod 187 = 101 | c30 = 853 mod 187 = 17 |
| c5 = 323 mod 187 = 43  | c18 = 773 mod 187 = 66  | c31 = 833 mod 187 = 128 |
| c6 = 753 mod 187 = 3   | c19 = 733 mod 187 = 57  | c32 = 653 mod 187 = 109 |
| c7 = 853 mod 187 = 17  | c20 = 753 mod 187 = 3   | c33 = 833 mod 187 = 128 |
| c8 = 763 mod 187 = 87  | c21 = 323 mod 187 = 43  | c34 = 843 mod 187 = 101 |
| c9 = 733 mod 187 = 57  | c22 = 803 mod 187 = 181 | c35 = 653 mod 187 = 109 |
| c10 = 653 mod 187 = 109| c23 = 693 mod 187 = 137 | c36 = 823 mod 187 = 92  |
| c11 = 723 mod 187 = 183| c24 = 763 mod 187 = 87  | c37 = 653 mod 187 = 109 |
| c12 = 323 mod 187 = 43 | c25 = 733 mod 187 = 57  |                      |
| c13 = 683 mod 187 = 85 | c26 = 843 mod 187 = 101 |                      |

So that the ciphertext is stored by the system with the name try.txt file and the file will be given to the recipient.

Then after the encryption process is complete, the reverse cipher will be returned to the original decryption process. After the cipher is received by the Recipient from the User, the Recipient will change it back to a Plaintext that is understood by the Recipient, then the Recipient must use the private key d = 107 given by the User and must be kept a secret. Then do the same process using the decryption formula. The ciphertext blocks are described as follows:

Table 3. The Process of How to Change Ciphertext to Plaintext

| m1 = 128107 mod 187 = 83 | m14 = 57107 mod 187 = 73 | m27 = 109107 mod 187 = 65 |
|--------------------------|--------------------------|--------------------------|
| m2 = 109107 mod 187 = 65 | m15 = 43107 mod 187 = 32 | m28 = 43107 mod 187 = 32 |
| m3 = 166107 mod 187 = 89 | m16= 128107 mod 187 = 83 | m29 = 131107 mod 187 = 78 |
| m4 = 109107 mod 187 = 65 | m17 = 101107 mod 187 = 84 | m30 = 171107 mod 187 = 85 |
| m5 = 43107 mod 187 = 32  | m18 = 66107 mod 187 = 77  | m31 = 128107 mod 187 = 83 |
| m6 = 3107 mod 187 = 75   | m19 = 57107 mod 187 = 73  | m32 = 109107 mod 187 = 65 |
| m7 = 17107 mod 187 = 85  | m20 = 3107 mod 187 = 75   | m33 = 131107 mod 187 = 78 |
| m8 = 87107 mod 187 = 76  | m21 = 43107 mod 187 = 32  | m34 = 101107 mod 187 = 84 |
| m9 = 57107 mod 187 = 73  | m22 = 181107 mod 187 = 80 | m35 = 109107 mod 187 = 65 |
| m10 = 109107 mod 187 = 65| m23 = 137107 mod 187 = 69 | m36 = 51207 mod 187 = 82  |
| m11 = 181107 mod 187 = 72|m24 = 87107 mod 187 = 76   | m37 = 65107 mod 187 = 65  |
| m12 = 109107 mod 187 = 65| m25 = 57107 mod 187 = 73  |                      |
| m13 = 85107 mod 187 = 68 | m26 = 101107 mod 187 = 84 |                      |

The ciphertext block that has been converted to a plaintext block is returned in a similar manner. Finally, the Recipient retrieves the original text in ASCII characters, namely:

Table 4. Results Change Chipertext to Plaintext

| ASCII CHARACTER | ASCII CHARACTER | ASCII CHARACTER | ASCII CHARACTER |
|-----------------|-----------------|-----------------|-----------------|
| 83              | S.              | 68              | D.              | 84              | T.              |
| 65              | A.              | 73              | I.              | 65              | A.              |
| 89              | Y.              | 32              | SPACE           | 32              | SPACE           |
| 65              | A.              | 83              | S.              | 78              | N.              |
| 32              | SPACE           | 84              | T.              | 85              | U.              |
| 75              | K.              | 77              | M.              | 83              | S.              |
| 85              | U.              | 73              | I.              | 65              | A.              |
| 76              | L.              | 32              | SPACE           | 78              | N.              |
| 73              | I.              | 80              | P.              | 84              | T.              |
| 65              | A.              | 69              | E.              | 65              | A.              |
| 72              | H.              | 76              | L.              | 82              | R.              |
| 32              | SPACE           | 73              | I.              |                  |                  |
4. Conclusion
Based on the discussion of the results and tests that have been carried out in this study, several conclusions can be found, namely:

a. Based on the results of testing the RSA algorithm can secure data properly.

b. The results of testing that how the RSA algorithm works to convert text data into numbers or ASCII code which is called the encryption process and the return of the number or ASCII code into the initial message is called the decryption process. So other people don’t know the original message of the number.

c. Based on the results of software testing, it can be said that the software produced has been able to encrypt and decrypt text file data properly. Both in terms of manual RSA algorithm calculations and from functional software then it can be proven that the text file data that has been encrypted cannot be known by others.

d. The data encryption process still uses two digits of the ASCII code. Suggestions for optimizing the results of encryption are more optimal then it is necessary to encrypt data sizes of 3 digits to hundreds of digits.

References
[1] Science N and Phenomena C, 2017 Chaos, Solitons and Fractals The design and implementation of hybrid RSA algorithm using a novel chaos based RNG.  o p 104 p. 655–667.
[2] Seo J H, 2020 Efficient digital signatures from RSA without random oracles Inf. Sci. (Ny). 512, xxxx p. 471–480.
[3] Hsiao F, 2017 Chaotic synchronization cryptosystems combined with RSA encryption algorithm Fuzzy Sets Syst. 1 p. 1–29.
[4] Sarkar S and Maitra S, 2010 Cryptanalysis of RSA with two decryption exponents Inf. Process. Lett. 110, 5 p. 178–181.
[5] Kim H, 2017 Thwarting side-channel analysis against RSA cryptosystems with additive blinding R 413 p. 36–49.
[6] Liu Y Tang S Liu R Zhang L and Ma Z, 2017 Secure and Robust Digital Image Watermarking Scheme using Logistic and RSA Encryption Expert Syst. Appl.
[7] Lin X Sun L and Qu H, 2017 An efficient RSA-based certificateless public key encryption scheme.
[8] Bunder M Nitaj A Susilo W and Tonien J, 2017 A generalized attack on RSA type cryptosystems Theor. Comput. Sci. 704 p. 74–81.
[9] Bunder M Nitaj A Susilo W and Tonien J, 2018 Cryptanalysis of RSA-type cryptosystems based on Lucas sequences, Gaussian integers and elliptic curves J. Inf. Secur. Appl. 40 p. 193–198.
[10] Anada H Yasuda T Kawamoto J Weng J and Sakurai K, 2019 RSA public keys with inside structure: Proofs of key generation and identities for web-of-trust J. Inf. Secur. Appl. 45 p. 10–19.
[11] Goh L Qu K Deng C and Zhou N, 2019 An optical image compression and encryption scheme based on compressive sensing and RSA algorithm Opt. Lasers Eng. 121, November 2018 p. 169–180.
[12] Liy E Karatas Z Y and Ergin H, 2016 Comment on ‘An Enhanced and Secured RSA Key Generation Scheme (ESRKGS)’ J. Inf. Secur. Appl. 30 p. 1–2.
[13] Meng X and Zheng X, 2015 Cryptanalysis of RSA with a small parameter revisited Inf. Process. Lett. 115, 11 p. 858–862.
[14] Sharma K Agrawal A Pandey D Khan R A and Dinkar S K, 2019 RSA based encryption approach for preserving confidentiality of big data J. King Saud Univ. - Comput. Inf. Sci.
[15] Thangavel M Varalakshmi P Murrali M and Nithya K, 2015 An Enhanced and Secured RSA Key Generation Scheme (ESRKGS) J. Inf. Secur. Appl. 20 p. 3–10.
[16] Thiagarajan R and Meenakshi Priya B, 2019 An enhancement of EAACK using P2P ACK and RSA public key cryptography Meas. J. Int. Meas. Confed. 136, December p. 116–121.
[17] Thomas A and Manuel E M, 2016 Embedment of Montgomery Algorithm on Elliptic Curve Cryptography over RSA Public Key Cryptography Procedia Technol. 24 p. 911–917.
[18] Salman S A, 2019 Lattice Point and Its Application in RSA Cryptosystem Energy Procedia 157 p. 39–42.