Contents lists available online at TALENTA Publisher

DATA SCIENCE: JOURNAL OF COMPUTING AND APPLIED INFORMATICS (JoCAI)

Journal homepage: https://jocai.usu.ac.id

Comparative Analysis of Ciphertext Enlargement on Generalization of the ElGamal and Multi-factor RSA

Imanuel Zega[1], Mohammad Andri Budiman[2], and Syahril Efendi[3]

[1,2]Master of Informatics Program, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia
3Doctor of Computer Science Program, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia

ARTICLE INFO

Article history:
Received 30 November 2022
Revised 4 December 2022
Accepted 27 January 2023
Published online 31 January 2023

Keywords:
Generalization of the ElGamal
Multi-factor RSA
Cryptography
Chipertext Comparison

Email:
[1]nuelzega49@gmail.com
[2]mandrib@usu.ac.id
[3]syahril1@usu.ac.id

Corresponding Author:
Mohammad Andri Budiman

ABSTRACT

Information and communication security has become more crucial and has become a new problem in relation to security, accessibility, data management, and other information policy challenges as a result of how easy it is for all users to use communication media. One of the fields of science that has a technique or art for disguising the data sent by the sender to the recipient with the aim of maintaining the confidentiality of the data is called cryptography. In determining better cryptographic algorithms for data security systems, in addition to considering strength, key length and ciphertext enlargement are also important factors to consider. Therefore, in this study, we attempted to compare the ciphertext magnification of the generalization of the ElGamal and multi-factor RSA algorithms by utilizing the same key length. Generalization of the ElGamal and Multi-factor RSA are both asymmetric algorithms that have public and private key pairs for encryption and decryption. However, at the level of security the RSA algorithm is based on the difficulty of finding large integer factors into two prime factors. In contrast to the ElGamal algorithm, security is based on the difficulty of calculating the discrete logarithm of a large prime modulus. The results of the comparison algorithm carried out are represented in the form of a table containing the plaintext, key length, and size of the data.

IEEE style in citing this article:

I. Zega, M.A. Budiman, and M. Syahril, " Comparative Analysis of Ciphertext Enlargement on Generalization of the ElGamal and Multi-factor RSA." *Data Science: Journal of Computing and Applied Informatics (JoCAI)*, Vol. 7, No. 1, pp. 44-50, 2023.

1. Introduction

The rapid advancement of information technology also encourages communication media as a medium for delivering information from one place to another, making it easier for people to access communication media. With everyone having easy access to communication media, information and communication security is becoming increasingly important and a new challenge in terms of security, accessibility, data management, and other information policy issues [1]. This will have an impact on data security for users of information media or messages using these communication media.

Information becomes very easy to know, manipulate, retrieve, and use by irresponsible people [2]. Thus, it becomes very important to address sensitive issues in providing data security, especially in the current era of modern digitalization, which continues to grow [3]. It must explore the development, perfection, and evolution of ecosystems in the digital world from the perspective of information technology, with the need for a method that can maintain the confidentiality of information known as cryptography.

Cryptography is a field of science that has the technique or art of disguising the data sent by the sender to the recipient to maintain the authenticity and integrity of the data. Budiman and Poltak [4], explained that cryptography

e-ISSN : 2580-829X | p- ISSN: 2580-6769

is a mathematics-based technique and is related to data security, including confidentiality, authentication, and data integrity.

The Generalization of the ElGamal algorithm is the latest derivative of ElGamal, which has security difficulties by relying on discrete logarithms. In contrast to the RSA Multi-factor algorithm, which is one of the RSA variants based on the modification of the modulus structure of ordinary RSA and relies on factoring a very large integer into its prime factors.

Symmetric and asymmetric cryptography are two types of cryptography that are distinguished by the type of key. The symmetric key is the key used in the encryption and decryption processes. Whereas the encryption and decryption operations for the asymmetric key use different keys, the encryption process uses the public key while the decryption process uses the private key. In terms of security, asymmetric cryptographic methods outperform symmetric methods [5]. The following is a cryptographic classification shown in Figure 1.

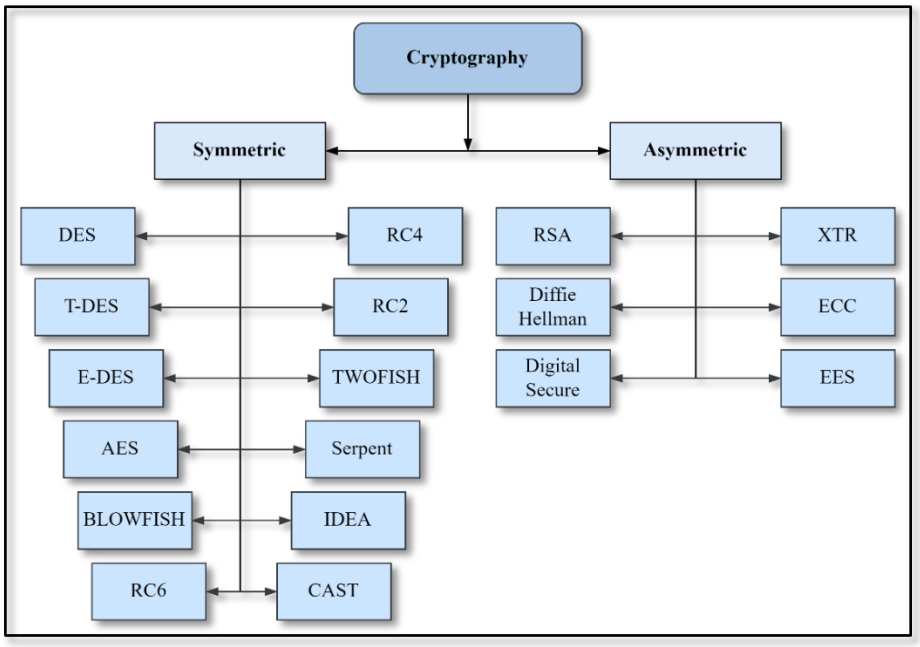

Figure 1. The classification of cryptography

A. Encryption and Decryption

Encryption turns the original data into an incomprehensible form, it is intended to keep the message safe without interference from outside parties. Decryption presents the opposite process of encryption, in which the cipher text is returned to plain text without losing a word in the original text.

B. Cryptographic Purposes

In the data and information security system, there are several things that are the goals and objectives of cryptography, namely:

- Confidentiality: is a service that encrypts data sent or received with the aim of protecting the user's identity and data privacy from being read by others.
- Integrity: a process that protects data from unauthorized insertion, deletion, and updating of data.
- Authentication: identifying the truth on the part of the sender of the message or data.
- Non-repudiation: prevents certain parties from rejecting the sending of messages made with denial.

Based on the foregoing, it is explained that cryptographic algorithms based on encryption key techniques are classified into symmetric and asymmetric key cryptography. However, asymmetric algorithms are superior to symmetrical ones. In the study, it will present a comparison of the Generalization of the ElGamal and Multi-factor RSA algorithms so that it can find out the development of both algorithms in terms of ciphertext magnification using the same prime number length. Thus this research can benefit knowledge as a reference in comparing ciphertex magnification.

2. Methods

In this section, the researcher describes the general stages of the two algorithms used as well as provides a simple example to clarify the steps of a method. The following is the architecture of the message encryption and decryption processes used.

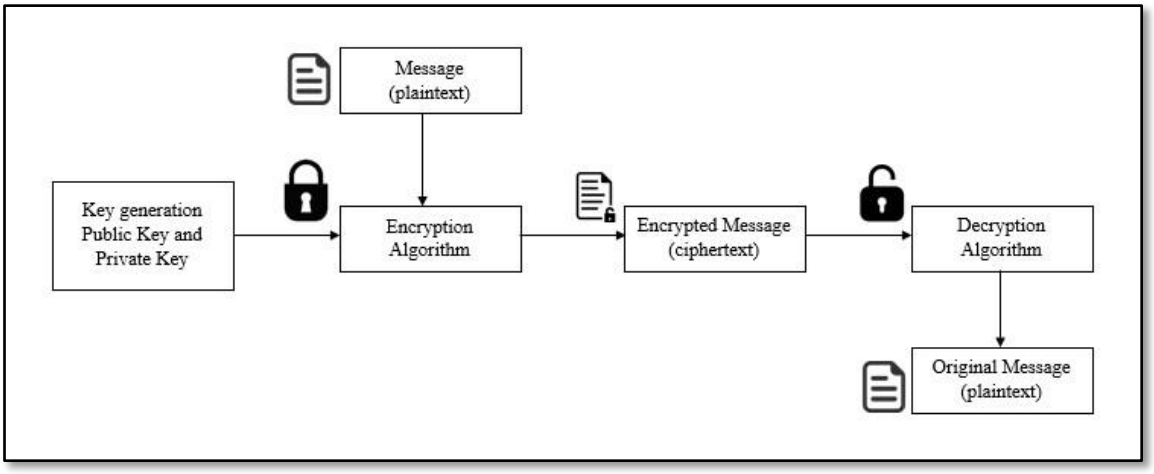

Figure 2. Architecture of the encryption and decryption processes used

2.1. Prime Number Generator

A prime number is a positive integer that has exactly two factors, namely 1 and the number itself. Prime numbers are used in cryptography to calculate public and private keys. Its power depends heavily on the difficulty of decomposing large integers into its factors [6].

One of the techniques that can be used to test the primacy of a number is to use Fermat's theory. Pierre De Fermat was a French mathematician in 1640. Formally, Fermat's theorem is expressed when a prime number p can be ascertained primacy, if the number p is a prime number and a is not a multiple of the number p, and 1 < a < p. Thus, it can be annotated in t with the equation a p−1 ≡ 1 mod p ; 1 < a < p, or it can also be written as a p−1mod p ≡ 1 ; 1 < a < p.

2.2. Generalization of the ElGamal Algorithm

The generalization of the ElGamal algorithm is a variant of ElGamal that involves the encryption process depending on the prime factorization of the plaintext other than the exponential modulus. The exponential modulus is the primitive root taken twice during encryption with respect to the number of different prime factors in the plaintext and the secret encryption key [7].

The generalization of the ElGamal algorithm consists of 3 processes, namely the key generation process, the encryption process, and the decryption process.

The first process is key generation. Here are the steps:

1. Choose a prime number that is greater than $p > 255$.
2. Choose the primitive root number g modulo p, provided that $g > p$.
3. Choose a random number x, with conditions 2 to p-2.
4. Calculate $a \equiv g^x \ (mod\ p)$.

The second process is the encryption process. Here are the steps:

1. Express the message into plaintext blocks: $m = p_1^{a1}.p_2^{a2}...p_i^{ai}$.
2. Change the message block value to an ASCII value.
3. Choose a random number y, provided that it is within 2 to $p - 2$
4. Choose a random number, provided that $1 < iy < p\text{-}1$.
5. Calculate the value of $d \equiv g^i (mod\ p)$.
6. Calculate the value of $b \equiv d^y (mod\ p)$.
7. Calculate the value of $c \equiv m.a^{iy} \ (mod\ p)$.
8. Send b and c to the recipient of the message.

The third process is message decryption. Here are the steps:

1. Calculate plaintext with the equation $b^x = (d^y)^x \equiv (g^i)^{xy}(mod\ p)$.
2. The b^x value obtained in ASCII is then converted into plaintext.
3. Decrypt the ciphertext with $C = [M[a^{i.y}\ mod\ p]]\ mod\ p$.

A. Key Generation of Generalization of the ElGamal (Example)

1. Determine the value of a prime number as $p > 255$.
   $p = 313$
   $g = 296$
   $x = 309$
2. Calculate the value of a with the equation $a = g^x(mod\ p)$.
   $a = 296^{309}\ mod\ 313$
   $= 257$
   Then,
   Public key $=(p, g, a) = (313, 296, 309)$
   Private key$= (x) = (257)$

B. Encryption from the Generalization of the ElGamal (Example)

1. The sender chooses a message (M)
   $M = Z$ →"90"
   Messages are converted into ASCII table values.
2. Determine the value of $i\ and\ y$.
   $i = 7\ and\ y = 12$
3. Calculate the value of d
   $d = 296^7\ mod\ 313$
   $d = 258$
4. Calculate the value of b
   $b = 258^{12}\ mod\ 313$
   $b = 265$
5. Calculate the value of C
   $C = [90[257^{7.12}\ mod\ 313]]mod\ 313$
   $C = 224$
   So, we get the value of $b = 265$ and $c = 224$

C. Decryption of the Generalization of the ElGamal (Example)

1. Receive ciphertext from the sender, i.e., $b = 265$ and $c = 224$
2. Compute z
   $z = 265^{309}\ mod\ 313$
   $= 79$
3. Compute w
   $w = 79^{-1}\ mod\ 313$
   $= 210$
4. Compute m
   $m = 224 \times 210\ mod\ 313$
   $= 90$
   So, from the decrypted value, 90 is the character "Z".

2.3. Multi-factor RSA Algorithm

The multi-factor RSA algorithm is a modification of the RSA algorithm with a modulus RSA structure, where $n = pqr$ atau $n = p^2r$. In the formation of keys in the multi-factor algorithm, there is an added parameter b [8]. Here are the stages in the multi-factor RSA algorithm.

The first process is key generation. Here are the steps:

1. Generate different prime numbers $b > 3$
2. Calculate the value of $n \leftarrow \prod_{i=1}^{b} P_i$
3. Find the value of $\varphi(n) \leftarrow \prod_{i=1}^{b}(P_i - 1)$
4. Use the same value of e as in standard RSA keys, which is $e = 65537$

5. Calculate the value of $d = e^{-1} \, mod \, \varphi(n)$, so that we get the public key $= (n, e)$ and the private key $= (d, p1, p2, \cdots pb)$.

The second process is the encryption process. Here are the steps:

1. Take the n and e as public key
2. Calculate the ciphertext with the equation $C = m^e \, (mod \, n)$

The third process is message decryption. Here are the steps:

1. Prepare the generated private keys and the previously encrypted message, which is c $(d, P_1, P_2, P_3, \cdots, P_n)$
2. Apply the Chinese Remainder Theorem (CRT) to calculate the value of d with the formula $d_i = d \, mod \, (p_i - 1)$, where $1 \leq i \leq b$.
3. Obtain the value for $B = p_1 \times \, p_2 \times p_3 \times \cdots p_n$
4. Calculate $B_i = B/p_i$
5. Find the value $S_i = B_i - 1 (mod \, p_i)$
6. Get value of $m = \sum_{i=1}^{k} ai. B_i. S_i (mod \, B)$.

A. Key Generation of Multi-factor RSA (Example)

1. Generate prime numbers as much as b.
   $b = 3$
   $p_1 = 103, p_2 = 113, p_3 = 127$
2. Calculate the value of n
   $n = p_1 \cdot \, p_1 \cdot p_1 = 103 \cdot 113 \cdot 127$
   $n = 1.478.153$
3. Calculate the value of $\varphi(n)$
   $\varphi(n) = (p_1 - 1)(p_2 - 1)(p_3 - 1)$
   $= 102 \cdot 112 \cdot 126$
   $= 1.439.424$
4. Use the RSA algorithm's standard value e, which is $e = 65.537$.
5. Calculate the value of $d = e^{-1} \, mod \, \varphi(n)$
   $d = e^{-1} mod \, \varphi(n)$
   $= 65537^{-1} \, mod \, 1.439.424$
   $= 1.106.369$
   Then,
   Public key $= (n, e) = (1.478.153, 65.537)$
   Private key$= (p_1, p_2, p_3, d) = (103, 113, 127, 1.106.369)$

B. Multi-factor RSA Encryption (Example)

1. The sender selects the message (M)
   $M = Z \rightarrow$"90"
   Messages are converted into ASCII table values.
2. Encrypt messages with equations $C = M^e mod \, n$
   $C = 90^{65537} \, mod \, 1.478.153$
   $= 1.428.647$

C. Multi-factor RSA Decryption (Example)

1. Receive ciphertext from the sender
   $C = 1.428.647$
2. Decrypt the ciphertext with the equation $M = C^d \, mod \, n$
   $M = 1.428.647^{1.106.369} \, mod \, 1.478.153$
   $= 90$
   So, from the decrypted value, 90 is the character "Z".

3. Results and Discussions

At this stage, a comparative analysis of the two algorithms is carried out using a 100-digit prime number with the message divided into several blocks containing 3 digits of the decimal value of the message. For example, the message used is "INFORMATICS".

Table 1. The decimal value of each message is based on ASCII codes

I	N	F	O	R	M	A	T	I	C	S
73	78	70	79	82	77	65	84	73	67	83

Then, the message is divided into blocks containing three digits.

$m_1 = 737$ $m_3 = 798$ $m_5 = 658$ $m_7 = 678$

$m_2 = 870$ $m_4 = 277$ $m_6 = 473$ $m_8 = 3$

With the message value that has been divided into several blocks, the enlargement of the ciphertext generated using 100 digits of prime numbers can be seen in the following table.

Table 2. Comparison of ciphertext enlargement

Number of message characters	Original data size	Data Size After Encryption	
		Generalization of the ElGamal	Multi-factor RSA
11	11,1 KB	12,6 KB	13,0 KB
200	11,3 KB	16,0 KB	31,3 KB
500	11,4 KB	20,0 KB	59,8 KB

From the test results shown in Table 2, it is clear that ciphertext magnification occurs in both algorithms. However, the multi-factor RSA algorithm experiences a very large magnification compared to the generalization of the ElGamal. This is because RSA's multi-factor algorithm has a large generator with the result of randomizing public and private keys. By separating several blocks of decimal values from the text and using 100 digits of prime numbers, it can increase the enlargement of the ciphertext and length of key characters, which can affect the security level of the algorithm.

4. Conclusions

This research was conducted to analyze and identify the differences between two algorithms, namely ElGamal and multi-factor RSA. The study showed that the use of the multi-factor RSA algorithm leads to an enlargement of the ciphertext compared to ElGamal. The encryption and decryption of secret text messages revealed that the RSA Multi-factor algorithm employs a large generator, which systematically randomizes public and private keys, resulting in very large ciphertexts. In contrast, the ElGamal algorithm relies on the factorization of the prime number of the plaintext, resulting in relatively smaller ciphertexts. Overall, this research provides a thorough analysis of the two algorithms and their impact on the size of ciphertexts.

References

[1] J. C. Bertot, P. T. Jaeger, and D. Hansen, "The impact of polices on government social media usage: Issues, challenges, and recommendations," *Gov. Inf. Q.*, vol. 29, no. 1, pp. 30–40, 2012, doi: 10.1016/j.giq.2011.04.004.

[2] S. Goel, K. Williams, and E. Dincelli, "Got phished? Internet security and human vulnerability," *J. Assoc. Inf. Syst.*, vol. 18, no. 1, pp. 22–44, 2017, doi: 10.17705/1jais.00447.

[3] Z. Zhang and B. B. Gupta, "Social media security and trustworthiness: Overview and new direction," *Futur. Gener. Comput. Syst.*, vol. 86, pp. 914–925, 2018, doi: 10.1016/j.future.2016.10.007.

[4] M. A. Budiman, P. Sihombing, and I. A. Fikri, "A cryptocompression system with Multi-Factor RSA algorithm and Levenstein code algorithm," *J. Phys. Conf. Ser.*, vol. 1898, no. 1, 2021, doi: 10.1088/1742-6596/1898/1/012040.

[5] O. G. Abood and S. K. Guirguis, "A Survey on Cryptography Algorithms," *Int. J. Sci. Res. Publ.*, vol. 8, no. 7, 2018, doi: 10.29322/ijsrp.8.7.2018.p7978.

[6] Y. G. Narayana and V. Yegnanarayanan, "On Prime number varieties and their applications," *Eng. Appl. Sci. Lett.*, vol. 3, no. 3, pp. 30–36, 2020, doi: 10.30538/psrp-easl2020.0045.

[7] R. Ranasinghe and P. Athukorala, "A generalization of the ElGamal public-key cryptosystem," *J. Discret. Math. Sci. Cryptogr.*, no. May, 2021, doi: 10.1080/09720529.2020.1857902.

[8] D. Boneh and H. Shacham, "Fast variants of RSA," *CryptoBytes*, vol. 5, pp. 1–10, 2002, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.83.7319&rep=rep1&type=pdf.