Scientific Journal of Informatics
Vol. 7, No. 1, May 2020

p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

A Combination of Hill Cipher and LSB for Image Security

Ajib Susanto[1], Ibnu Utomo Wahyu Mulyono[2], Muhamad Rizky Fajar Febrian[3], Ghaitsa Ardelia Rosyida[4]

[1,2,3,4]Informatics Engineering Department, Faculty of Computer Sciences, Universitas Dian Nuswantoro, Indonesia
Email: [1]ajib.susanto@dsn.dinus.ac.id, [2]ibnu.utomo.wm@dsn.dinus.ac.id, [3]muhamadrizkyff@gmail.com, [4]ghaitsatha@gmail.com

Abstract

The maximum value obtained to test the encrypt hill cipher uses the avalanche effect with modification of one, two, three, and four key matrices 35.71%, therefore an additional security technique is needed. The Least Significant Bit (LSB) method is used to insert the ciphertext that has been generated from the hill cipher algorithm, has been testing using RGB, CMY, CMYK and YUV shapes with 6142 characters in 128 x 128 character images producing the highest PSNR value of 51.2826 dB in CMYK images. Steganography technique is applied because it has advantages in terms of imperceptibility, for example the results of a stego image are so similar to the original image that it is difficult to be distinguished by the human senses. Tests were carried out with 10 images, five images measuring 512 x 512 and five images measuring 16 x 16. While the messages to be inserted were 240, 480, and 960 characters for images measuring 512 x 512 and 24, 48 and 88 characters for images measuring 16 x16. Test results that I have done are calculated using Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) with a minimum PSNR of 51.2907 dB which means the resulting image is quite good. Computation time is calculated using tic toc on Matlab by encrypting 2000 and 6000 characters, and also computation time.

Keywords: Hill Cipher, LSB, Avalanche Effect, PSNR, computation time.

1. INTRODUCTION

In its application, the internet has several impacts, for example a positive impact on everyone in finding data or information quickly and freely. The negative impact is that crime on the internet or cybercrime is increasing, so security needs to be held [1]. Therefore, to prevent the hacking of information or data from irresponsible parties, a method is needed to hide messages or information, for example by watermarking, steganography, cryptography and digital signatures [2].

Steganography is a technique for hiding messages into other digital forms, such as pictures, videos or audio so that they are not seen concretely by others. According to the domain, steganography can be divided into two types, namely the frequency domain and spatial domain [3]. Cosine Discrete Transformation (DCT) and Wavelet Discrete Transformation (DWT) are examples of frequency domains that have advantages over image manipulation [4]. The Least Significant Bit (LSB) method uses spatial domains that are easy to manipulate and easily lose information. The steganography method in this study uses the Least Significant Bit (LSB) method. The Least Significant Bit (LSB) method is the lightest and simplest algorithm because it only changes the last bit with the message bit. The use of steganography in this study has advantages in imperceptibility [5], for example the resulting image is so similar to the original image that it is difficult to be distinguished by the human senses.

Cryptography is a technique for encrypting messages so that messages cannot be read directly. However, the growing development of knowledge and technology allows very powerful parties to know the message from the cover image so that additional coding is needed on the message or information, by combining cryptographic and steganographic methods in the journal [3] [6]. The encryption process is done using certain cryptographic algorithms. Cryptography is a technique for hiding messages or randomizing a message that is easily understood into a message that is not understood by others.

The algorithm carried out by researchers is the Hill Cipher algorithm. This Hill Cipher algorithm is one of the same algorithm keys and has several advantages. Researchers use the cryptographic method that is the Hill Cipher algorithm to produce ciphertext by encrypting a message. Hill Cipher is a cryptographic technique that is competent and difficult to solve if it does not know the key of the matrix, but is easily resolved when knowing the key of the matrix, so as to cover the shortcomings of this method and to protect multiple layers of data the ciphertext insertion technique is used in the digital image of the image using the method steganography is Least Significant Bit (LSB).

In this research, we will describe the implementation of hiding messages or information in digital images by combining the Hill Cipher algorithm and the Least Significant Bit (LSB) method. Will describe the quality of the results of this method into various things, such as imperceptibility, stego image size and encryption quality. The quality of imperceptibility can be measured by MSE, PSNR, and histogram. While the quality of encryption is measured by entropy [10]. The calculation of the time required by the system to perform a process can be used computation time in Matlab with the tic toc process.

2. METHODS

2.1. LSB Steganography

Steganography is a technique for disguising messages or information into other digital content such as in the form of images, video or audio so as not to be seen from clay. Steganography comes from Greek namely Steganos which means "hidden or veiled" and the word graphein which means "writing" so that steganography has the meaning of "writing (writing) veiled" [6]. This method saves better security than pure steganography. Problems that will occur if sharing a key. If the other party knows the key, it will be easier to describe and access the original information. Steganography is applied to text, images, video clips, music and sound media.

The Least Significant Bit (LSB) method is a method for securing messages or information in the form of media, by directly entering messages or information into pixels from the cover image [1]. By changing a small part of each pixel bit, where the position of the bit will be replaced by a message that will be inserted in the selected main media [3], then the changes that occur in the color value does not affect the image quality, this method has a good imperceptibility so that people others cannot see a significant change in the image [2]. Bit changes are made sequentially starting from the first bit until the last bit in accordance with the length of the message to be hidden.

2.2. Hill Cipher Cryptography

Cryptography is a technique used to encrypt a message so that it cannot be read directly. Cryptos is derived from the Greek word meaning "secret" while graphe which means "writing". Components of the cryptographic algorithm [7]:

1. Input: Plaintext, the message (data or information) that will be sent (containing data or information in the original language). Plaintext contains the original text and is used in the encryption process.
2. Ouput: Ciphertext, the result of encryption of the plaintext in the form of a code or letter that has no meaning and is not recognized as a message or information or data.
3. Encryption: a process for converting plain text into ciphertext.
4. Description: a process to convert back from plaintext to ciphertext.
5. Key: the encryption and description process requires a key, which can be either a public key or a private key.

The Hill cipher was discovered by Lester S. Hill in 1929. This method is one of the symmetric key cryptographic algorithms. Basic matrix theory uses multiplication between matrix and inverse of the matrix . Hill code is a polyalphabeth code, using matrix multiplication calculation using the substitution method. The hill code includes polyalphabeth which means that each alphabet can be mapped to more than one kind of character [8]. Hill cipher is a cryptographic algorithm that is difficult to solve if the person only has the ciphertext without the key. But it is very easy to solve if the person has the key.

2.3. Avalanche Effect

Calculation of bit testing that occurs on the media to see the number of bits that change in messages with long characters using the avalanche effect as show in Eq. (1).

$$Avalanche\ effect = \left(\frac{amount\ of\ flip\ bit}{amount\ of\ bit}\right) x\ 100\% \tag{1}$$

Where:
Flip bit = the bit that changes after encryption
Total bits = total bits in the text

2.4. Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR)

A study needs to be tested to determine the advantages and disadvantages of the research. An image is tested to determine the quality of the image with the help of MSE (Mean Square Error) and PSNR (Peak Signal Noise Ratio) measuring devices [3]. Mean Square Error is a measuring tool used to measure the quality of an image, measuring the average squared cumulative error between the stego and the original image. Error shows irregularities in the image. Peak Signal to Noise Ratio is the ratio of the maximum value of the signal measured by noise that affects the signal, used to determine the maximum value of the signal measured with noise that affects the signal and can also be used to express image quality. This formula can be calculated by Eq. (2) and Eq. (3).

$$MSE = \sum_{u=0}^{U-1} \sum_{i=0}^{I-1} \sum_{o=0}^{O-1} \|H_i(u,i,o) - S_i(u,i,o)\|^2 \quad (2)$$

$$PSNR_{dB} = 10\, log_{10} \left(\frac{2^8-1}{MSE}\right) \quad (3)$$

Where:
u,i,o = Size of image (width*height*number of layer)
Hi(u,i,o) = Host image with size u*i*o
Si(u,i,o) = Stego image with size u*i*o

2.5. Entropy

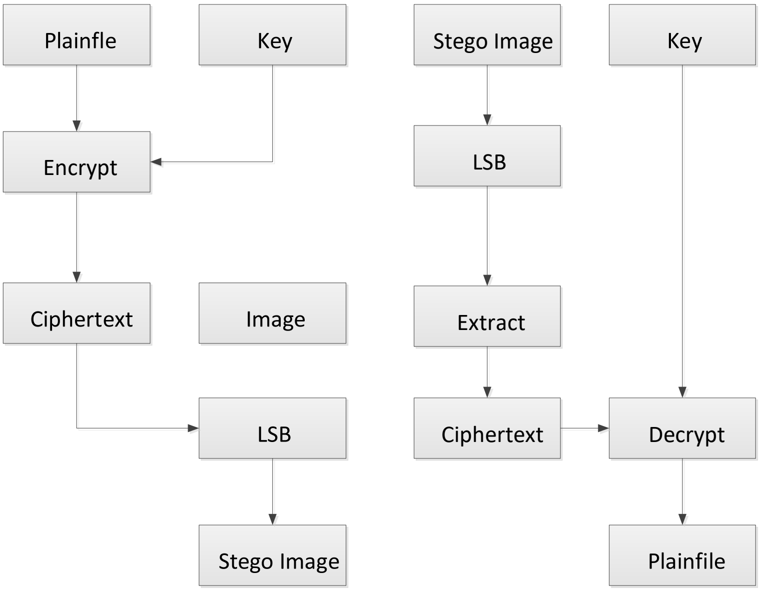
Entropy is an unordered theory and there are circumstances that may be uncertain. The definition of entropy related to information theory is the amount of measure of information contained in a message. Can be expressed in units of bits. Useful for encoding messages, can be calculated by Eq. (4).

$$\text{He} = -\sum P(k)\ {}^{log}\ 2n\ k = 0\ (P(k)) \quad (4)$$

Where:
He = Entropy
n = Gray value of the image
Pk = The probability of the symbol occurrence k

3. RESULT AND DISCUSSION

Here, we proposed Hill Cipher-LSB. Cover image used in this study uses two sizes, namely 512 x 512 pixels and 16 x 16 pixels. While the number of characters used in inserting messages in a 512 x 512 pixel image is 240 characters, 480 characters, and 960 characters. And for the image size of 16 x 16 pixels using a number of characters as many as 88 characters. The maximum character that can be accommodated by 16 x 16 pixel images is only 90 characters, so testing is only done up to 88 characters to avoid errors. The proposed method has been explain in Figure 1.



a) (b)

Figure 1. (a) Encrypt and embed, (b) Decrypt and extract

3.1 Hill Cipher Testing

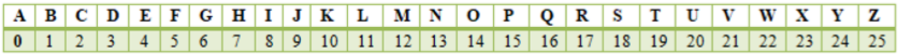
In testing in this study the researchers used a matrix key which has the order of 2x2. The message to be encoded does not specify how many characters and each character must be between the letters A - Z and does not distinguish between uppercase and lowercase letters. The sequence of letters used for example A = 0, B = 1, C = 3 and so on until the letter Z = 25 and spaces are not counted in this encoding. P is plaintext, C is ciphertext, K is key as shown in Eq (5).

$$C = P \; x \; K \tag{5}$$

The encryption process steps are as follows:

P = KAMU and $K = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$

1. The space is removed from between words.
2. If the number of letters in the message is odd then a dummy is added, so the number of letters becomes even.
3. The plaintext is converted into a 1 x 2 matrix format into P = $\begin{bmatrix} K \\ A \end{bmatrix} \begin{bmatrix} M \\ U \end{bmatrix}$
4. The plaintext matrix is converted into numeric shapes according to the following:



A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Thus becoming : $P = \begin{bmatrix} 10 \\ 0 \end{bmatrix} \begin{bmatrix} 12 \\ 20 \end{bmatrix}$

5. P (plaintext) multiplied by K (key) and produce C = $\begin{bmatrix} 84 \\ 92 \end{bmatrix}$
6. Look for the remainder of the quotient with modulo 26 based on the results of the K and P,

   $\begin{bmatrix} 10 \\ 0 \end{bmatrix} mod\ 26 = \begin{bmatrix} 20 \\ 10 \end{bmatrix}$

   $\begin{bmatrix} 84 \\ 92 \end{bmatrix} mod\ 26 = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$

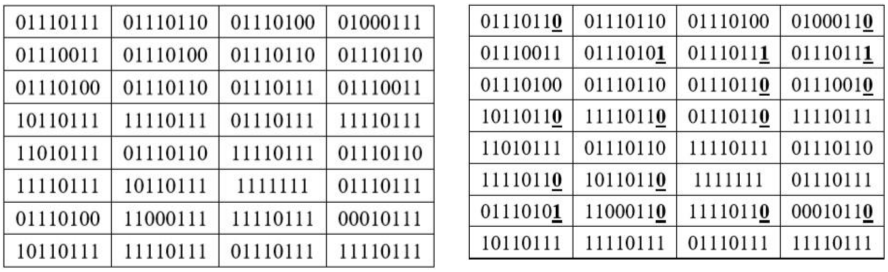
   The modulo result matrix is converted into letters based on the conversion table in point 4.

   $\begin{bmatrix} 20 \\ 10 \end{bmatrix} mod\ 26 = \begin{bmatrix} U \\ K \end{bmatrix}$

   $\begin{bmatrix} 6 \\ 14 \end{bmatrix} mod\ 26 = \begin{bmatrix} G \\ O \end{bmatrix}$
7. Result of ciphertext, c = $\begin{bmatrix} U \\ K \end{bmatrix} \begin{bmatrix} G \\ O \end{bmatrix}$

3.2 LSB Testing

The images used are 24-bit color images, images consisting of (R, G, B) each color has a color depth of 8 bits, messages or information will be hidden into R bits, G bits, and B bits in each every pixel. Steganography steps to hide messages into image images using the LSB algorithm with the results of the calculation of the UKGO matrix as a ciphertext which has been calculated using the Hill Cipher algorithm. K = 01001011, A = 01000001, M = 01001101, U = 01010101. The first step is to change the ciphertext to binary. After the message is converted into binary form then the message will be hidden into an image by the LSB method with the binary value. The last step, merging the RGB layer after the message is converted into binary form and inserted into a stego image then the image is saved as shown in Figure 2.



01110111	01110110	01110100	01000111
01110011	01110100	01110110	01110110
01110100	01110110	01110111	01110011
10110111	11110111	01110111	11110111
11010111	01110110	11110111	01110110
11110111	10110111	1111111	01110111
01110100	11000111	11110111	00010111
10110111	11110111	01110111	11110111

(a)

0111011**0**	01110110	01110100	0100011**0**
01110011	0111010**1**	0111011**1**	0111011**1**
01110100	01110110	0111011**0**	0111001**0**
1011011**0**	1111011**0**	0111011**0**	11110111
11010111	01110110	11110111	01110110
1111011**0**	1011011**0**	1111111	01110111
0111010**1**	1100011**0**	1111011**0**	0001011**0**
10110111	11110111	01110111	11110111

(b)

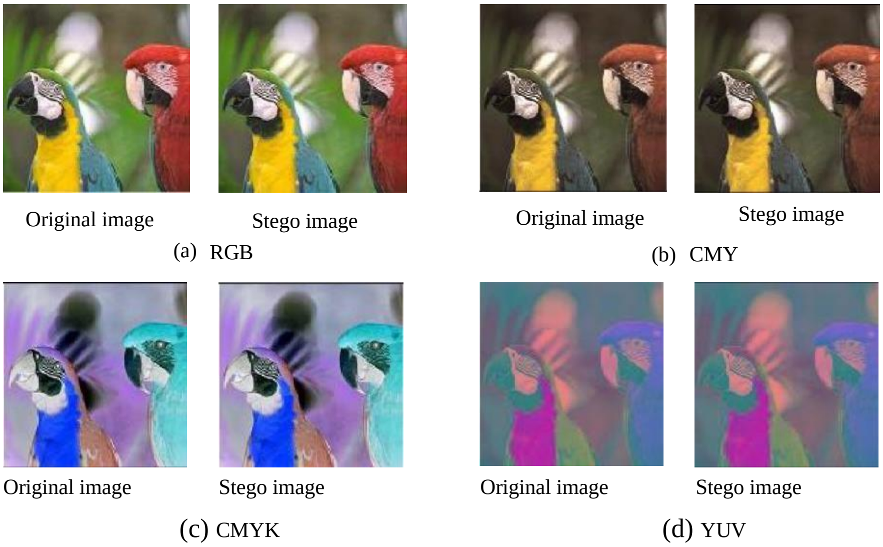
Figure 2. (a) Binary of cover image, (b) Binary after embedding

3.3 Avalanche Effect Testing

Plaintext "udinus" and "polkee" were used for the avalanche effect experiment as explain in Table 1.

Table 1. Avalanche effect testing

No.	Plaintext	Ciphertext	Bit Changed	Avalanche Effect (AE)
1	23·14	EJNAIPYCNACXYV	16	14.29%
2	13.14	BJNAVPGCNAJXHV		
3	15.16	RZNAJDYUNALZJX	8	8.05%
4	17.16	HZNAXDQUNANZLX		
5	23.14	EJNAIPYCNACXYV	16	14.29%
6	23.35	EXNNIJYINNCXYR		
7	01.13	IBNNUVWGNNOJOH	26	23.21%
8	12.13	TBANBVKGANVJTH		
9	01.19	IXNNULWINNOPON	39	34.82%
10	12.49	TGANBYKKANVUTM		
11	71.10	DDNAHNSSNARTDR	37	33.04%
12	74.53	BNANPVGAANHHTX		
13	23.14	EJNAIPYCNACXYV	40	35.71%
14	74.53	BNANPVGAANHHTX		
15	23.55	EXNNIJYINNCXYR	27	24.1%
16	71.10	DDNAHNSSNARTDR		



Original image Stego image

(a) RGB

Original image Stego image

(b) CMY

Original image Stego image

(c) CMYK

Original image Stego image

(d) YUV

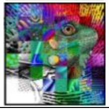
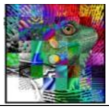
Figure 3. A Comparison of RGB, CMYK, CMY, and YUV between original image and stego image: point a until d

Based on the image results in Figure 3, in Table 2 it can be concluded that the best PSNR value generated is in the YUV color space by hiding messages as much as 6142 characters with a maximum of 6144 characters that can be hidden.

Table 2. Message hiding based on Figure 3 in 128 x128 pixels

Image	Maximum Payload	Message	MSE	PSNR	PSNR
a	9609	6000	0.4869	51.2555	7.7049
b			0.4839	51.2826	7.2258
c			0.4901	51.2280	7.7049
d			0.4912	51.2174	6.7458
	Average		0.3904	51.2458	7.3453

Table 3. A Comparison results of several characters message

Name of Image	Original image	240 characters	480 characters	960 characters
a				
b				
c				
d				
e				

Based on the test results contained in the Tables 3, the difference cannot be seen concretely in the stego image or cover image. This means that the quality of the stego image is very good. However, to determine the quality of stego, it is necessary to measure the Peak Signal Noise Ratio (PSNR) and Mean Square Error (MSE). PSNR is used to assess the quality of the image before and after the message is inserted, while MSE is used to determine the average square error value between the images before and after.

Table 4. Message hiding based on able 3 in 512 x 512 pixels and 98112 payload

Image	240 Char			480 Char			960 Char		
	MSE	PSNR	Entropy	MSE	PSNR	Entropy	MSE	PSNR	Entropy
a	0.0012	77.1355	7.6319	0.0024	74.1893	7.6203	0.0049	71.1668	7.6322
b	0.0013	77.0014	7.7432	0.0025	74.0858	7.7432	0.0050	71.1138	7.7431
c	0.0013	76.9633	7.5224	0.0025	74.1164	7.5224	0.0050	71.1413	7.5224
d	0.0012	77.1442	7.5853	0.0024	74.1960	7.5853	0.0050	71.2272	7.5853
e	0.0010	77.6270	5.9969	0.0020	74.9789	6.0032	0.0040	72.0062	6.0150
Average	0.0012	77.1742	7.2959	0.002	74.3132	7.2948	0.0047	71.3310	7.2996

According from Table 4, the PSNR value is determined by the number of messages hidden in the cover image. The smaller the PSNR value the more messages are hidden and the more bits change. In a 512 x 512 image that holds 98112 characters, when a message of 240 characters is inserted it has an average PSNR value of 77.1743 dB, when a message of 480 characters is inserted has an average PSNR value of 74.3133 dB and for a message of 960 characters it has a value of the average PSNR is 71.3311 dB. From the test results obtained PSNR values above 40 dB, which means the quality of the resulting image is very good.

The experimental results above are the results of tests that have been carried out, if viewed concretely there are no significant changes between the stego image and the cover image. This makes the quality of the resulting image very good because there is no change. But of course we need a measuring tool that can determine the quality of stego, namely PSNR (Peak Signal Noise Ratio) which is used to compare the quality of the image before and after insertion. And MSE (Mean Square Error) is used to determine the value of the mean square error between images before and after insertion.

Table 5. Message hiding based on 16 x 16 pixels and 90 payload

Image	24 Char			48 Char			88 Char		
	MSE	PSNR	Entropy	MSE	PSNR	Entropy	MSE	PSNR	Entropy
a	0.1276	57.0722	7.4415	0.2708	53.8038	7.4425	0.4947	51.1866	7.4405
b	0.1523	56.3026	7.4462	0.2890	53.5209	7.5417	0.4973	51.1638	7.4304
c	0.1562	56.1926	7.1271	0.2877	53.5405	7.1145	0.5260	50.9206	7.0933
d	0.1445	56.5312	7.2040	0.2747	53.7416	7.1844	0.4726	51.3853	7.1922
e	0.1510	56.3398	7.3697	0.2747	53.7416	7.8338	0.5221	50.9530	7.4461
Average	0.1463	56.4876	7.3177	0.2793	53.6696	7.4233	0.5025	51.1218	7.3205

According Table 5, the results of tests that have been carried out on images measuring 16 x 16 that can hold a maximum of 90 characters can be seen in table 18. When tested on 24 characters that are inserted it will produce an average PSNR value of 56.4877 dB and when testing is carried out on 48 characters that are inserted will produce an average PSNR value of 53.66697 dB while the 88 characters inserted will produce an average PSNR value of 51.1218. It can be interpreted that the average PSNR value is above 40 dB even though the image used is small and the number of messages inserted is almost reaching the maximum character.

From the results of the tests conducted above, the PSNR value will always decrease, adjusted for the number of characters inserted using the Least Significant Bit (LSB) method. The more messages that are inserted, the more bits will change, but the value of the PSNR will always be above 40 db even though the number of messages that are inserted almost reaches the maximum character that can be contained by the cover image, this can occur because each message that is inserted will only change the last 1 bit of the cover image. This shows that the resulting stego image is very good.

3.4 Computation Time

Table 6. A Comparison result between Hill Cipher, LSB and Hill Cipher + LSB

Method	Image Size	Max payload	Message	Time (s)
	-	-	2000	0.016063
Hill Cipher	-	-	4000	0.029069
	-	-	6000	0.048245
	128 x 128	6096	2000	0.603389
LSB	128 x 128	6096	4000	2.099935
	128 x 128	6096	6000	4.263528
	128 x 128	6096	2000	0.648071
Hill Cipher + LSB	128 x 128	6096	4000	2.154402
	128 x 128	6096	6000	4.533753

According Table 6, Computation time is calculated using the tic toc function with a lenna.png image that has a size of 128 x 128 which can hold 6096 characters of messages and the authors test with strings of 2000, 4000 and 6000 characters. With three different methods namely the hill cipher, LSB and Hill Cipher + LSB methods.

4. CONCLUSION

From the results of research that has been done, testing that has been done with the hill cipher by using the avalanche effect only has a range of values of 8.05% to 35.71% with several different key variants, thus requiring additional algorithms to hide the message stronger. So that the cipher text will be inserted in the cover image using the LSB method. And testing is done by using PSNR and MSE measuring instruments, which produces a minimum PSNR value of 51.1218 dB which means that the quality of the resulting image remains of good quality and is difficult to distinguish from the cover image, the hill cipher algorithm is also difficult to solve as long as the key used is unknown by other party. The fastest time to get is 0.016063 seconds while encrypting the hill cipher while the longest time is 4.533753 seconds when encrypting with the hill cipher + hiding the cipher text into the original image with the LSB method. Hill cipher has the fastest time compared to LSB and the combination of hill cipher + LSB. Computational time is affected by the number of encrypted message characters and the number of algorithm methods. The conclusion that can be drawn is that the combination of the hill cipher method and the LSB method produces good stego image quality and messages that are difficult to decrypt if the key is unknown, the computational time of the combination of the two methods is relatively fast.

5. REFERENCES

[1] Ariyanto, A. D. P., Rachmawanto, E. H., & Sari, C. A. (2019, October). Performance Analysis of LSB Image Steganography Combined with Blowfish-RC4 Encryption in Various File Extensions. *In 2019 Fourth International Conference on Informatics and Computing (ICIC)* (pp. 1-6). IEEE.

[2] Cheddad, A., Condell, J., Curran, K., & Mc Kevitt, P. (2010). Digital image steganography: Survey and analysis of current methods. *Signal processing*, *90*(3), 727-752.

[3] Sari, C. A., Rachmawanto, E. H., & Kusuma, E. J. (2019). Good Performance Images Encryption Using Selective Bit T-des On Inverted Lsb Steganography. *Jurnal Ilmu Komputer dan Informasi*, *12*(1), 41-49.

[4] Arrasyid, A. A., Soeleman, M. A., Sari, C. A., & Rachmawanto, E. H. (2018, November). Image Watermarking using Triple Transform (DCT-DWT-SVD) to Improve Copyright Protection Performance. In *2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)* (pp. 522-526). IEEE.

[5] Rajendran, S., & Doraipandian, M. (2017). Chaotic Map Based Random Image Steganography Using LSB Technique. *IJ Network Security*, *19*(4), 593-598.

[6] Chakraborty, S., Jalal, A. S., & Bhatnagar, C. (2017). LSB based non blind predictive edge adaptive image steganography. *Multimedia Tools and Applications*, *76*(6), 7973-7987.

[7] Ilaga, K. R., Sari, C. A., & Rachmawanto, E. H. (2018). A High Result for Image Security Using Crypto-Stegano Based on ECB Mode and LSB Encryption. *Journal of Applied Intelligent System*, *3*(1), 28-38.

[8] Bhowmic, A., & Geetha, M. (2015). Enhancing resistance of hill cipher using columnar and Myszkowski transposition. *International Journal of Computer Sciences and Engineering*, *3*(2), 20-25.