The Combination of Bit Matching-Based Steganography and DES Cryptography for Data Security

Budi Prasetiyo 1)
Semarang State University, Semarang
e-mail : prasemath@gmail.com

Rahmat Gernowo 2) Beta Noranita 3)
2) 3) Diponegoro University, Semarang

Abstract - This research discussed the combination of steganography and cryptography to secure data without changing the quality of cover medium. Steganographic method is used to match bit of the message with bit of the MSB image cover. Matching process is done by divide and conquer method. The result will be a bit indexposition, and then it will be encrypted using cryptographic DES (Data Encryption Standard). The input are text message, image, and key. The output is ciphertext bit index which can be used to secure the messages. To read the contents of the message, we require the same image cover and key. Outcomes of proposed method can be used to secure the data. The advantages of this method are the image quality will not change and the capacity of stored messages can be larger than the image. According to the research, both grayscale and colorful images can be used as image cover, except the image contains 100% black and 100% white. Bit matching process on image which have much variety of color takes less time. The damage of messages due to the addition of “salt and pepper” noise starts from 0.0067 of MSE value and gaussian starts from MSE 0.00234.

Keywords: steganography, bit matching, divide and conquer, bit index, MSB, encryption, decryption, DES.

I. INTRODUCTION

These recent years, the human need for information is increasing. In the midst of rapid development of information technology, the internet is no longer providing secure information. The development of search-engine coupled with the development of virus, bugs, spam and hackers who can steal confidential data (Kautzar, 2007). To solve this problem, various ways have been developed to improve data security, such as cryptography and steganography.

Steganography is the art and science of hiding data in other media as a covert(e.g. image) in order to make the data looks sketchy (Provos and Honeyman, 2003). Cryptography is the art and science of maintaining the confidentiality of data (Schneier, 1996). In cryptography, the original data is converted into another form that can not be read. The combination of steganography and cryptography can simultaneously increase the security of the data (Krenn, 2004).

Method for combining steganography and cryptography has been developed. In general, the mostly used technique is message encrypt first (cryptography), then hiding it into media cover (steganography)(Raphael and Sundaram, 2011). However, the embedding process can affect the quality of the cover media.

Efforts to minimize the quality changes of cover image can be done by embedding the data in the least significant bit. Changes in the quality of cover is invisible (Chan and Cheng, 2004), but the embedding of cover into the least bit tends to make the cover prone to robust. Robust resistance can be done by embedding the data in the first bit (most significant bit), but it will change the quality of the cover and it will look suspicious.

Other studies conducted by (Challita and Farhat, 2011) developed a new way of merger steganography and cryptography without changing the media cover. The technique is performed by matching the message bits on the cover, and then continue the process of encryption (cryptographic). One well-known cryptographic algorithms since 1977 and became a worldwide standard is the Data Encryption Standard (DES).

This research will combine steganography and cryptography without changing the media cover. The steganography method used is a method based on bit matching in the first bit (most significant bit) and the cryptographic method used is the DES algorithm.

II. REVIEW OF LITERATURE

Various methods have been developed for data security. In general, the techniques used are encrypting the message first (cryptography process), and then embedding it into the media cover (steganography process) (Raphael and Sundaram, 2011).

Combination techniques are not only limited as shown in Figure 1. Research (Narayana and Prasad, 2010) examined two approaches to secure steganography media cover (image). Securing steganography images done by encrypting. The first method, steganography image is directly encrypted with S-DES, the result is a ciphertext. The second method, the image is encrypted then ciphertext from encryption will be embedded on another image.
splitting it into two smaller parts (divide), then match each section (conquer). The results of each part of the solution then combined into a total solution (combine).

### III. BASIC THEORY

#### 3.1 Steganography

Steganography comes from the Greek, meaning Steganos (meaning hide) and Graphos (meaning writing), so that steganography is defined as “hidden writing (covered writing)”. Steganography is the science and art of hiding a secret message (hiding message) so that the existence of the message is not detected by human senses. The data hiding process into media is called embedding, whereas the reverse process is called extraction. In general, the process is shown in Figure 4.

One of the easiest methods of steganography is LSB (Least Significant Bit). The procedure to perform this method is to embed the least bit at each pixel with the message bits. Terminology LSB is reviewed by Sharp (2001). The LSB embedding will change the bit value, but it will be invisible, so that the third party does not know the existence of the secret message behind the media cover (Chan and Cheng, 2004).

The use of LSB on the combination of steganography and cryptography was done in a research conducted by Sharp (2001). The process consists of three stages, namely encryption, steganography, and decryption. Encryption and decryption is done with DES algorithm (Data Encryption Standard). The use of LSB can minimize the image quality changes, but the capacity of messages that can be accommodated is due to the size of the image. (Kekre et al., 2012) conducted a LSB steganography study to increase the message capacity with PVD approach (Pixel value differencing). LSB insertion is then compared with the MSBit value. Most Significant Bit. If the value of the first 4 MSB bits is “1”, then embed it on the last 4 bits. If the first 3 bits MSB is “1”, then embed it on the last 3 bits. If the first 2 MSB bit is “1”, then embed it on the last 2 bits. If the value is outside the criteria, then the embedding is done on the last bit (least).

Image quality is an important component in steganography. (Challita and Farhat, 2011) developed another way to combine both steganography and cryptography without changing the image quality. The technique is performed by matching the message bits on the cover, the results is in the form of bit position index. Index is then encrypted. The output is bit index ciphertext. Bit matching is done by divide and conquer (Cormen, 2009) that consist of three processes, namely divide, conquer, and combine. Arrangement of long bits is

#### 3.2 Cryptography

Cryptography comes from two Greek words, Crypto which meansthe secretand Grapho which means writing. Cryptography is the study of mathematical techniques related to aspects of information security, such as data confidentiality, data authenticity, data integrity, and authentication of data (Menezes et al., 1996).

Cryptography basically consists of two processes, namely the encryption and decryption process. In general, the encryption and decryption process can be seen in Figure 5.

#### 3.3 Data Encryption Standard (DES) Algorithm

DES is a block cipher algorithm that operates on 64-bit input and block key sizes of 128 bits (Munir, 2006). DES algorithm in general scheme is shown in Figure 6 (Munir, 2006).
as follows:

a. Broke the 64-bit plaintext into L(32bit) R(32bit)
b. Perform initial permutation (IP)
c. Encrypt in 16 rounds (enchipering). Internal locks on each different lap
d. Invert the initial permutation (IP-1).

3.4 Combination of Steganography and Cryptography

Combination of steganography and crytography in general is performed with the cryptographic process first and then steganography, which encrypts the message first and then embeds the encrypted ciphertext into a cover media (Raphael dan Sundaram, 2011). The merging concept of cryptography and steganography is shown in Figure 8.

![Figure 7. Combination of Steganography and Cryptography (Raphael dan Sundaram, 2011)](image)

3.5 Bit Matching in Divide and Conquer

Message bit matching in the bit image is done by dividing and conquering in which consists of four processes: divide, conquer, and combine. It means to break down and divide smaller parts (divide), and then recursively solve any smaller issues (conquer). Then the solutions of every minor problem are merged into one main solution (combine) (Cormen et al., 2009).

The use of divide and conquer method in steganography is done by Chalita and Farhat (2011) in his research to match the location of messaged bits and 0s of the image. Suppose given sequence $S_1$ and $S_2$, and denoted in LCS($S_1$, $S_2$), an algorithm for long substring search (longest common subsequence) of $S_1$ that appears in $S_2$. Next it will yield the true value if the entire subsequence in $S_2$. Algorithm illustration (LCS) is given in Figure 9.

![Figure 8. Image matching algorithm on the message](image)

3.6 Methodology

The method used in this study is the integration of steganography and cryptography. Cryptographic algorithms used are the DES. There are two processes involved in steganography, the embedding and extraction. This study constructed a stego-crypto software with the waterfall model. Waterfall method is shown in Figure 10.

![Figure 9. Waterfall Method (Pressman, 2001)](image)

Waterfall methodology divides the research into four phases which are interrelated and influence. The four stages are needs analysis, design, code, and test (Pressman, 2001).

The combination of cryptography and steganography in this study required 4 processes, they are bit matching, encryption, decryption and reconstruction. The details are as follows.

3.1. Matching Bit

In this study, the method of matching is done with divide and conquer (Cormen, 2009). Input to this process is the message and image.

The steps are performed in the bit matching are:

a) Convert the message and image in binary form
b) Taking the value of MSB image
c) Perform the matching messages on MSB image. If the bit message is contained in MSB image, then proceed to save the position of the bit index. Index saving consists of index position of the first (start) and the position index of the last bit (end). If the matching process does not occur, continue the process d) as follows.
d) Divide the message into two parts of equal length of the left (L[i]) and right (R[i])
e) Repeat the same steps as in number b), with L[i] and R[i] as input. If all the bit message are contained in the image, then the matching process is completed and continue to f). If not, repeat step c) with L[i] and R[i] as an input.
f) Keep all bit index from matching results
g) The output is a vector that contains the index structure of bit position.

For example, suppose bit message and bit image are known as follows:

Message (M) : 10110111
Image (I) : 10010001010101010011

Since M is not contained in I then M was split into two parts left (L) and right (R), namely:

1. L[1]: 1011 which is located at the position index “11 14”, which is 10010001010101010011.
2. R[1]: 0111, not present in the image, then divide R[1] into two parts, namely:
   a. L[2]: 01, located at index position “3 4”
   b. R[2]: 11, located at index position “8 9”
3. Because of all the bit positions found, then the matching processes are completed and proceed to step 4.
4. Combine all solutions from step 1, step 2a, and 2b. Retrieved whole bit index position “11 14 3 4 8 9”.

International Conference on Information Systems for Business Competitiveness (ICISBC 2013)
3.2. Encryption

Position vector of bit is obtained at 3.1 section later in encryption. Encryption process is done with the DES algorithm.

3.3. Decryption

Input to this process is the ciphertext and the key. Decryption of the ciphertext is the inverse of the encryption process. DES uses the same algorithm for encryption and decryption. In the decryption process, the key sequence used is the inverse one namely K16, K15, ..., K1. For each round of 16, 15, ..., 1, the output at each round of decryption is obtained as follows:

\[ L_i = R_{i-1} \]
\[ R_i = L_{i-1} \oplus f(R_{i-1}, K) \]

3.4. Reconstruction

Reconstruction aims to restore the message to its original form. Input at this stage consists of bit and image index location. Process that is carried out is taking the composition of the image bit based on vector of index bit location. The process output is in the format of bit message composition.

The steps are performed in the reconstruction process are:

a) Convert the image in binary form and take the bit of MSB image.
b) Read the contents of two index vectors. The first index is a bit’s early position (start) and the second index is the bit’s end position (end).
c) Taking the value of bit image based on step b),
d) Repeating the process b) and c) until the last index position.
e) The composition of the bits will create an output in the form of bit message.

For example, suppose the unknown vector and image as follows:

Vector: 11 14 3 4 8 9
Image: 1001000110101101010011

Extraction steps is done by taking the value of the bit image based on the location of the vector. Putting all the bit value from matches result of the vector. In that case a match is obtained:

* Vector11,14, produced in 1011.
* Vector34, generating 01.
* Vector89, yielding 11.

All the above results are combined, resulting in outputs 10,11,111.

3.5. Combination Steganography and Cryptography

3.5.1. Overview

Combination of steganography and cryptography in this study consists of the main processes, namely the process of embedding and extraction which is generally shown in Figure 11.

3.5.2. Embedding Process

Embedding process (Figure 12) consists of bit matching and encryption, the result is ciphertext. Extraction process (Figure 13) consists of decryption and reconstruction; the results are the image of the message.

3.5.3. Extraction Process

Extraction process aims to store the message to its original form in order to maintain the original content. The inputs of extraction process are ciphertext vectors, a key, and imagery.

---

Figure 10: General description of the combination of steganography and cryptography on the study

Embedding process (Figure 12) consists of bit matching and encryption, the result is ciphertext. Extraction process (Figure 13) consists of decryption and reconstruction; the results are the image of the message.

Figure 11: Embedding Process

Embedding steps are as follows:

a) Integrate the input in the form of images, messages, and key.
b) Convert the message and image in binary form.
c) Match the bit message with the bit of MSB image. The same bit positions are stored in the bit vector index.
d) Encrypt the bit vector index with DES algorithm.
e) The output is ciphertext. The ciphertext contains a bit vector that has been encrypted.
f) Finish.

Figure 12: Extraction Process

Extraction process aims to store the message to its original form in order to maintain the original content. The inputs of extraction process are ciphertext vectors, a key, and imagery.
Figure 12. Extraction process

Extraction process steps are as follows:

a) Input key, ciphertext vector, andimagery.

b) Decrypt the vector with the key, the decryptedplaintext is the output bit index.

c) Do a message reconstruction by matching bit of MSB image based on bit index vector.

d) The output is the message.

e) Finish.

IV. RESULTS AND DISCUSSION

The result of this research is in the form of application programs which is developed using the programming language MATLAB R2009b then it was tested for date security.

4.1. Testing the image color

1) Embedding Process

The step-to-step embedding process are as follows:

a) Selecting the message file which will be embedded. In this case, the message file was selected.

b) Choosing an image file, for example Baboon.bmp (Figure 14).

c) Typing the password as the encryption key for DES (for example key: 1234567).

d) The output of the embedding process is after a bit index for DES (Figure 15), and the bit index of ciphertext (Figure 15).

<table>
<thead>
<tr>
<th>No</th>
<th>Citra</th>
<th>Embedding (left)</th>
<th>Ekstraksi (right)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Block bmp</td>
<td>1.327</td>
<td>24.285</td>
<td>25.612</td>
</tr>
<tr>
<td>2</td>
<td>Grayscale bmp</td>
<td>0.374</td>
<td>31.53</td>
<td>32.904</td>
</tr>
</tbody>
</table>

Rata rata: 0.850 | 27.897 | 28.740 | 0.309 | 30.315 |

Figure 13. Citra cover

2) Extraction Process

In testing the extraction process, the author will return the message from the vector that has been encrypted by bit extraction.

The step-to-step extraction process are as follows:

a) Choosea vector file, vektor_Baboon.txt

b) Chooseagcovimage file, i.e.Baboon.bmp (Figure 14).

c) Key input. Key must be the same as the one when performing embedding, which is "1234567".

d) Perform the extraction.

After the extraction process, the message successfully returned normal, with the output:

Money transfer, 50 millions via ATM
Bank account: 0123456
Password PIN: 9x8d7g
Under the name of Budi Prasetjo

Table 1. Execution time on the black and white image

<table>
<thead>
<tr>
<th>No</th>
<th>Citra</th>
<th>Embedding (detik)</th>
<th>Ekstraksi (detik)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lena.bmp</td>
<td>0.165</td>
<td>12.417</td>
<td>12.582</td>
</tr>
<tr>
<td>2</td>
<td>Peppers.bmp</td>
<td>0.167</td>
<td>12.19</td>
<td>12.357</td>
</tr>
<tr>
<td>3</td>
<td>Jetting.bmp</td>
<td>0.138</td>
<td>12.028</td>
<td>12.166</td>
</tr>
<tr>
<td>4</td>
<td>Baboon.bmp</td>
<td>0.162</td>
<td>10.121</td>
<td>10.283</td>
</tr>
<tr>
<td>5</td>
<td>Forensic.bmp</td>
<td>0.199</td>
<td>15.105</td>
<td>15.304</td>
</tr>
</tbody>
</table>

Rata rata: 0.164 | 11.932 | 12.097 | 0.164 | 12.257 |

The test showed that averagely the embedding process of black and white image took 28.946 sec to 0.850 sec for bit matching. 27.897 sec for encryption, 30.115 sec for extraction, 29.215 sec for decryption and message reconstruction on tool 0.595 sec. Whilethe color image, averagely the embedding process took 12.19 sec, 0.164 sec for matching, and 9.325 sec for encryption.

The extraction process took 12.774 sec, 12.417 sec for decryption and 0.1638 sec for message reconstruction.

4.2. Test Results with Different Size Resolution

The application was also tested with different imagesizes, ranging from 512px, 256px, 128px, to 64px. The test result (Table 3) shows that the larger the image resolution, the longer the bit matching process will take. The shortest bit matching is bit matching of "Baboon" (0.590 sec), while the longest bit matching is the bit matching of "Block" (0.202 sec). Baboon has the mostcolor variation, while the "Block" only has 2color variations (black and white).

Table 3. The test results with different imagesizes

<table>
<thead>
<tr>
<th>Citra</th>
<th>Resolusi (px)</th>
<th>Proses Embedding (detik)</th>
<th>Proses Ekstraksi (detik)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Matching Ekstraksi</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>512 x 512</td>
<td>5.57</td>
<td>6.576</td>
<td>7.576</td>
<td>8.576</td>
</tr>
<tr>
<td>&quot;Block&quot;</td>
<td>256 x 256</td>
<td>7.30</td>
<td>30.165</td>
<td>37.465</td>
</tr>
<tr>
<td></td>
<td>128 x 128</td>
<td>10.34</td>
<td>31.673</td>
<td>31.673</td>
</tr>
<tr>
<td>64 x 64</td>
<td>128 x 128</td>
<td>10.25</td>
<td>24.033</td>
<td>24.288</td>
</tr>
<tr>
<td>Rata rata</td>
<td>2.02</td>
<td>23.112</td>
<td>25.134</td>
<td>22.31</td>
</tr>
<tr>
<td>512 x 512</td>
<td>5.54</td>
<td>24.386</td>
<td>30.108</td>
<td>24.72</td>
</tr>
<tr>
<td>&quot;Gradation&quot;</td>
<td>256 x 256</td>
<td>1.72</td>
<td>31.963</td>
<td>33.685</td>
</tr>
<tr>
<td></td>
<td>128 x 128</td>
<td>10.54</td>
<td>27.441</td>
<td>37.985</td>
</tr>
<tr>
<td>64 x 64</td>
<td>64 x 64</td>
<td>10.25</td>
<td>24.033</td>
<td>24.288</td>
</tr>
<tr>
<td>Rata rata</td>
<td>2.01</td>
<td>27.106</td>
<td>29.116</td>
<td>26.45</td>
</tr>
<tr>
<td>512 x 512</td>
<td>1.86</td>
<td>9.681</td>
<td>11.542</td>
<td>11.78</td>
</tr>
</tbody>
</table>

...
In general it can be concluded that the embedding of black and white images is more robust than the colorful image. The process of bit-matching of images with many color variations is shorter than the image that has little color variation. This is due to the color variations in the image allowing many opportunities for bit-similarity between the arrangement of the bit message and the image, so it takes less time.

4.3. Testing with Giving Noise

Noise ‘salt and pepper’ (Figure 15) and Gaussian (Figure 16) was given to the image in the next embedding process. The images were given noise ‘salt and pepper’ with standard deviation $d=0.001; 0.005; 0.01; 0.05$. The image was given noise ‘Gaussian’ with zero mean and standard deviation $d=0.001; 0.005; 0.01; 0.05$. The images have been given anio the threshold to the extraction process before the encryption

Test results of message reconstruction on a black and white image with the addition of salt and pepper and Gaussian noise remain good. Messages can be read, but there was one image which was damaged. While most of the colorful images were damaged, except the image of “Jet” and “Lenna” which only suffered from a little damage. Both “jet” and “lenna” images have the highest MSE value $0.014$. Damage to the colorful images occurred from $0.0067$ for the image “Photos” which incidentally has a faithful representation. The addition of Gaussian noise causes most of the message content to be corrupted. Damage began to occur when the MSE reached $0.00234$. This is due to the addition of noise effectively affecting the bit message, while matching the bit takes the appropriate bit position index. The result of message reconstruction will produce the changed messages well.
V. CONCLUSION

The process of steganography in this study include the bit matching and reconstruction, while the cryptographic processes include encryption and decryption.

The combination of steganography and cryptography in this study can be used for data security. The input are message, image and key. The output is ciphertext. To be noticed, to see the message content we need the same key and image.

Either grayscale or colorful images can be used as the cover media. The only exception is the colorful image with 100 % black or 100 % white, because the image consists of a homogeneous bit structure. All bit values in the image with 100 % white is 0 (zero) and the image with 100 % black is 1 (one). The bit composition of message varies from 0 to 1, so the bit matching will not find any results.

The addition of noise to the image causes some changes in the message content, the degree of changes vary. In the black and white image, the changes are not significant, while in the colorful image the message content changes a lot. Damage occurred on the addition of salt and pepper noise start from MSE 0.0067 and the damage to the gaussian noise start from MSE 0.00234.

The bit matching process with color variation took shorter time than the image with less color variation. One advantages of this method is there was no change of the image quality. In terms of security, even if the index vector of bit was not encrypted, the data was secure enough. This is due to the need of the right image to reconstruct the index into the original message, otherwise the results will be unreadable.

Other researchers can perform encryption on the image first before matching the bit. They can perform operations on the image with 100 % black or 100 % white, so that both black and white images can be used as covers. Besides, other researchers can modify the output ciphertext into stego image.

References