Text Steganography on Digital Video Using Discrete Wavelet Transform and Cryptographic Advanced Encryption Standard Algorithm

Megah MULYA[1], Osvari ARSALAN[2], Latifah ALHAURA[3], Rendy WIJAYA[4], Syahrul Ramadhan A.S.[5], Christofer YEREMIA[6]

[1,2,3,4,5,6]*Informatics Engineering, Computer Science Faculty, Sriwijaya University, Indonesia*

ABSTRACT

The rapid development of the internet has increased the ease of sharing information to people around the world. However, this advancement also raises a distress about data manipulation when the data are sent by the sender to the recipient. Therefore, information security is a major problem in data communication. Steganography and cryptography play important roles in the field of information security. Steganography can be applied though various digital media such as video, images, and audio to hide information in such a way that no one else knows that there is a hidden information, except for the relevant parties. Cryptography refers to the art of converting a plaintext (message) into an unreadable format. Both steganography and cryptography techniques are robust.

The purpose of this study is to prove that the quality of the results of steganography used to hide secret messages in a video (stego video) using Discrete Wavelet Transform (DWT) and Advanced Encryption Standard (AES) is good. Good quality is measured with Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). The message is encrypted using a cryptography algorithm, AES, to be more secure. The message is then placed into the one frame frequency domain in the video using the DWT method. Video quality before and after the message is inserted is measured based on the PSNR and SSIM values.

From the experiments, it is known that the PSNR value ranges between 39 to 40 dB and the SSIM value is close to 1, which means the change of the video frame due to insertion does not cause large noise, thus, the quality of the resulting video stego is satisfactory.

Keywords: *steganography, cryptography, DWT, AES*

INTRODUCTION

Steganography is one of the information security techniques that is often used in communication through digital media. One of the main objectives of steganography is to prevent information to not be taken by unauthorized parties by hiding the existence of information that is to be exchanged on media [1].

Several studies have suggested that a Discrete Wavelet Transform (DWT) can be applied in the steganography process [2] - [4]. DWT is a technique that uses the frequency domain on digital media which is the container in conducting steganography. The use of the frequency domain has a number of obstacles in the steganography process seeing that the media must be transformed first to get the media frequency.

In addition to steganography, cryptography is also often used to secure information. Cryptography aims to transform original information into meaningless forms by applying the coding method [5]. Various coding methods have been widely introduced. One of the most famous and widely used in the world is the Advanced Encryption Standard algorithm. This algorithm has been recognized as one of the powerful cryptographic algorithms.

Some studies even combine the two data security techniques [6], [7]. By combining both techniques, confidential information can be sent safely through the internet. Both steganography and cryptography have important roles that can improve data security : steganography can make information untraceable while cryptography makes information incomprehensible. Both techniques can be implemented to various digital media such as text, audio, and video.

Based on the considerations above, this study will apply a combination of steganography and cryptography techniques using DWT and AES. The media used is a digital video with the AVI extension.

METHOD PROPOSED

This study applies a combination of steganography and cryptographic techniques in order to improve data security. The method used in this research is the Discrete Wavelet Transform and Advanced Encryption Standard.

Discrete Wavelet Transform (DWT). DWT is used to convert spatial domains into frequency domains. This conversion process is called decomposition. The decomposition process is executed with the help of a highpass filter and lowpass filter. The first filtering process is done by bottom sampling each row in the image. For example, we have a picture. Bottom-sample will produce two images Px (Q/ 2). Next, do the bottom sampling for the columns in the image to get the four parsed images (P/ 2) x

(Q/ 2). This decomposition process will produce four subbands namely LL, LH, HL, and HH. These four subbands can be executed using insertion process by modifying the most significant bits of each pixel in the subband.

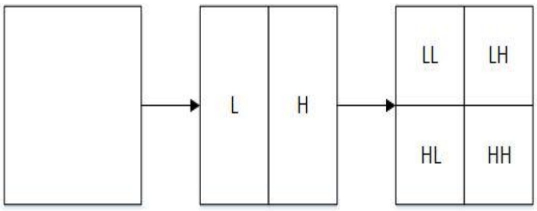

Figure 1. Decomposition Process, (a) Original Image, (b) First Down-sample, (c) Second Down-sample (Kumar, dkk)

Advanced Encryption Standard (AES). AES is a symmetric-key and a block cipher type. AES consists of three block chipers, AES-128, AES-192, and AES-256. Each cipher encrypts and decrypts data in blocks of 128 bits using cryptography keys of 128-,192- and 256-bits, respectively. Symmetric or secret-key ciphers use the same key for encrypting and decrypting. Figure II-6 outlines the process of the AES algorithm[3],[8].

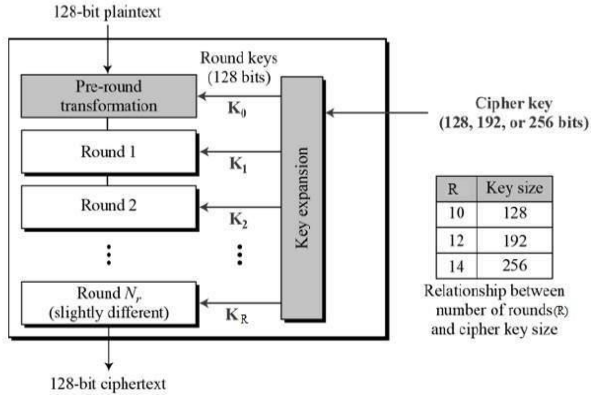

Figure 2 Global Scheme of AES Algorithm (S. Rawal, 2016)

1. Do permutation of plaintext with an initial permutation (IP) matrix.
2. The permutation result is enciphered 16 times (16 rounds) where each round is a Feistel Network (Figure 2) which is mathematically expressed as:
3. First block of plaintext is divided into two parts, right ( ) and left ( ). These two parts will be calculated for 16 rounds. In each round, a block is combined with the internal key in function. This function produces an output that will be XORed with block.
4. The enciphering result is the permuted with inverse initial permutation (IP^{-1}) into ciphertext. From the elaboration of DWT and AES, we get the steps of research work described in figures 3 and 4 below.

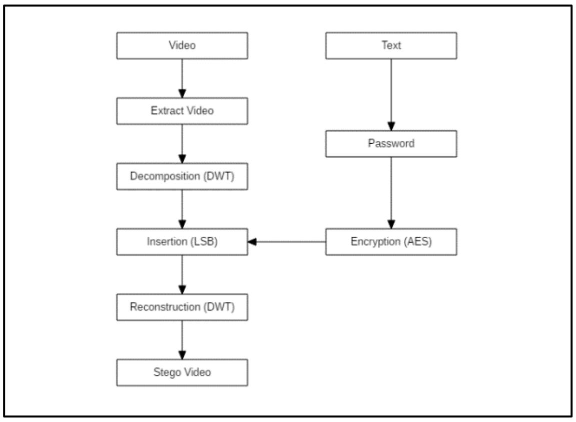

Figure 3. Embedding Process

1. Input a AVI video, text, and passcode.
2. Extract frames from a video and selects a frame to be inserted text.
3. Perform the encryption process using AES on the text.
4. The encryption results in a ciphertext.
5. Perform the decomposition process on the selected frame using DWT which divides a frame into four subbands (frequency domain).
6. The ciphertext will be inserted to frame.
7. Reconstruct the frequency domain of the frame into the spatial domain using IDWT.
8. Last, reassemble the selected frame with an entire frame to make a new stego video.

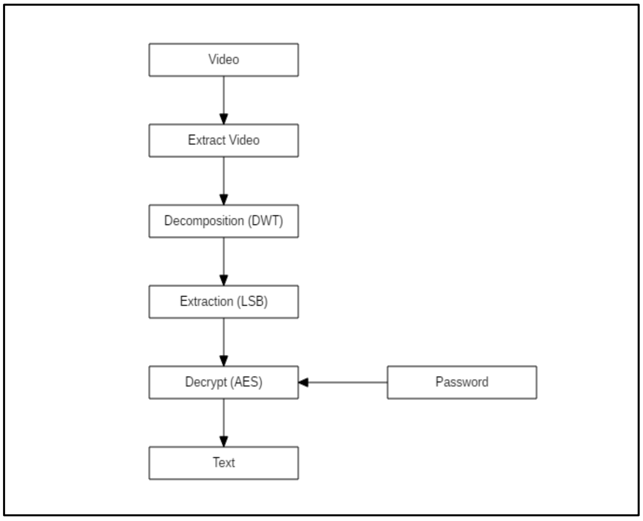

Figure 4. Extraction Process

1. Input stego video and passcode used in the embedding process.
2. Extract frames from the video and gets a selected frame that contains a message.
3. Perform the decomposition process on the selected frame using DWT to get the frequency domain.
4. Perform the message extraction process using the LSB method. The result is ciphertext.
5. Decrypt ciphertext using AES algorithm to get the original text.

RESULTS AND ANALYSIS

The data used in the test consist of 4 video files with a resolution of 320x240, 640x480, 960x540, 1920x1080. Whereas the inserted text are 50 characters, 500 characters, and 5000 characters in size for each original and encrypted text.

Results and Video Quality Measurement. Measurement of video quality before and after insertion aims to determine video quality changes that occur after the insertion process. Video quality testing is accomplished by calculating the value of the Structural Similarity Index (SSIM) to determine the level of video similarity and Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE) to determine changes in messages that are inserted frames into. In this study, the data set of the videos are collected from the stored videos in the researcher's computer. Test results for measuring MSE, PSNR, and SSIM values can be seen in Table 1 below.

Table 1. Video Quality Measurement Results

Video	Number of Characters Inserted	MSE	PSNR (dB)	SSIM
Rhinos.avi (320x240)	50	3,80	42,33	0,9965
	500	3,81	42,32	0,9965
	5000	3,91	42,21	0,9965
Handshake.avi (640x480)	50	7,04	39,66	0,9967
	500	7,04	39,66	0,9967
	5000	7,06	39,64	0,9967
Time changes.avi (960x540)	50	3,31	42,93	0,9983
	500	3,31	42,93	0,9983
	5000	3,33	42,91	0,9983
Morgen.avi (1920x1080)	50	3,18	43,1	0,9964
	500	3,18	43,1	0,9964
	5000	3,19	43,1	0,9964

According to the results shown in Table 1, it can be concluded that the insertion effect using the Discrete Wavelet Transform algorithm and Advanced Encryption Standard on the quality of the stego video is adequate. The SSIM value close to 1 means that the video before and after insertion does not experience significant degradation or deterioration in quality. The PSNR value obtained is around 40 dB, which means the change in pixel value in the frame inserted message does not cause large noise in the frame. The MSE values in Table 1 are quite small. The smaller the MSE value given by PSNR from a larger frame that causes the frame to be categorized under favorable conditions for message insertion.

Results and Analysis Measuring The Percentage of Equality of Extraction Results. Measurements of text similarity are realized to determine whether the text that has been inserted can be extracted correctly and if there is an error in the extraction of the text, what percentage of the similarity of the text extracted with the original text. The test results measuring the similarity of the text can be seen in table 2 below.

Table 2. Execution Testing Results

Video	Text Type	Number of Characters Inserted	Number of Characters Extracted	Number of Characters Error	Percentage of Similarity (%)	Result
Rhinos.avi	Plaintext	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	1	99,98	Not Valid
	Encrypted	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	0	100	Valid
Handshake.avi	Plaintext	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	0	100	Valid
	Encrypted	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	0	100	Valid
Time changes.avi	Plaintext	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	0	100	Valid
	Encrypted	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	0	100	Valid
Morgen.avi	Plaintext	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	5000	0	100	Valid
	Encrypted	50	50	0	100	Valid
		500	500	0	100	Valid
		5000	4998	18	99,63	Not Valid

According to the test results shown in Table 2, most messages entered in the frequency domain (DWT) were extracted successfully. However, there were two failed attempts during message extraction on the Rhinos.avi and Morgen.avi videos. This could be caused by:

- There are some changes in the video pixel, that is, after the insertion process, the pixels in the video ought to be normalized because reconstruction with the Inverse Discrete Wavelet Transform (IDWT) results in several pixels exceeding the range of image values.
- The message is not extracted correctly from the frame, causing changes to the encrypted message. Where, encrypted messages can be recovered or decrypted provided that they meet two conditions, namely the appropriate passcode or key and the correct passcode. The correct ciphertext here is the ciphertext originating from the encoding of the original message. As explained in the first point, the frequency domain will undergo a reconstruction process that causes changes in pixel values. Changes in pixel values also affect the changes in the original passcode. Thus, broken ciphertext causes messages to not be decrypted even if the passcode is entered correctly.

CONCLUSION

This research concludes that Discrete Wavelet Transform and Advanced Encryption Standard can be implemented in the steganography process in AVI video files with the good stego video quality. The video stego quality resulting from the application of the two methods does not cause high degradation with an SSIM value close to-1 and a good PSNR ranging from 39 to 43 dB for each video tested.

Many characters in a message can be entered as many as the number of pixels in a video. However, not all characters in a message can be extracted perfectly. This is because the DWT method can change the video pixel value, which causes some characters in the message to change. Only characters in the LL subband can be extracted correctly.

REFERENCES

[1] E. Ponmani, S. Indhuja, R. Puviarasi, P. Saravanan, and S. Ananthakrishnan, "An Enhanced Least Significant Bit Steganography to Improve the Effectiveness of Graphical Password Authentication," *Int. J. Pure Appl. Math.*, vol. 119, no. 12, pp. 13325–13335, 2018.

[2] S. K. Banik, Barnali Gupta; Bandyopadhyay, "A DWT Method for Image Steganography," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 6, pp. 983–989, 2013.

[3] S. Mahalakshmi, R. Selvarani, J. Thilagam, and N. Tharminie, "Audio Steganography Using AES Algorithm," vol. 4, no. 11, pp. 22–27, 2015.

[4] G. S. N. Kumar, N. Bhavanam, and V. Midasala, "Image Hiding in a Video-based on DWT & LSB Algorithm,". November 2014, 2016.

[5] S. B Sasi, "A General Comparison of Symmetric and Asymmetric Cryptosystems for WSNs and an Overview of Location Based Encryption Technique for Improving Security," *IOSR J. Eng.*, vol. 4, no. 3, pp. 01–04, 2014.

[6] K. Patel and H. Gupta, "Information Hiding using Least Significant Bit Steganography and Blowfish Algorithm," *Int. J. Comput. Appl.*, vol. 63, no. 13, pp. 975–8887, 2013.

A. S. Vaidya, P. N. More, R. K. Fegade, A. Madhuri, and P. V Raut, "Image Steganography using DWT and Blowfish Algorithms," vol. 8, no. 6, pp. 15–19, 2013.

[7] S. Rawal, "Advanced Encryption Standard (AES) and It's Working," *Int. Res. J. Eng. Technol.*, pp. 1165–1169,