Securing Text File on Audio Files using Least Significant Bit (LSB) and Blowfish

Ahmad Rizky Fauzan [a,1], Al Farissi [b,2*], Muhammad Naufal Rachmatullah [b,3]

[a] Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia
09021281722042@student.unsri.ac.id; [2] alfarissi@unsri.ac.id*; [3] naufalrachmatullah@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received -
Revised -
Accepted -

Keywords
Blowfish
Cryptography
Least Significant Bit
Steganography

ABSTRACT

Along with the development of technology, communication can be done in various ways, one of which is digital messages. But often the messages sent do not reach their destination and are obtained by irresponsible parties. This happens because of the lack of security in the file. For this reason, security is needed so that messages cannot be stolen or seen by other parties. There are various ways to secure messages, including Steganography and Cryptography techniques. This study uses a combination of the Least Significant Bit method and the Blowfish algorithm to secure secret messages in audio files. This research will measure encryption and decryption time, analysis of message file size changes after encryption and decryption, and PSNR value of audio files. The result of encryption using blowfish is a change in the size of the message file caused by the size of the message file being less than the block cipher size, so additional bytes are given so that the message size matches the block cipher size. The speed of the encryption and decryption process using the blowfish algorithm results in an average time for encryption of 547.98ms while the average time for decryption is 538.19ms. The longest time for the encryption process is 557.30ms and the fastest is 534.50ms, while the longest time for the decryption process is 548.74ms and the fastest is 531.46ms. Hiding messages in audio files using LSB produces PSNR values above 30dB.

1. Introduction

As technology advances, there are more and more conveniences in communicating, one of which is digital messages. Often the message that you want to send does not reach its destination and falls into the hands of irresponsible people because of the lack of security in the message. For that we need security that can access the message in order to get to the destination[1].

There are many methods to secure messages, including steganography and cryptography. This research will protect messages in WAV format audio files using the Least Significant Bit (LSB) and blowfish methods. This research is useful for knowing the time required for the encryption and decryption process using blowfish and also knowing the difference in the quality of the audio file (audio carrier) before and after the message is hidden in it (stego audio) based on the PSNR value.

Steganography is the art of inserting a secret message into a medium, the secret message will be hidden without changing the contents of the file, so that unauthorized parties cannot see the hidden message. The steganography process is a continuation of the cryptographic process which in practice is a way of hiding messages in other media so that third parties cannot see where the message is[2].

Cryptography comes from two Greek words, namely "cryptó" which means secret and "graphy" which means writing. To maintain data security, cryptography encrypts messages (plaintext) in the form of ciphertext that cannot be understood. Cryptography is a science that studies mathematical techniques related to information security aspects, such as data confidentiality, data validity, data integrity, and data authentication [4].

2. Literature Study

A. Theoritical Basic

1. Least Significant Bit (LSB)

LSB is a steganographic method that is widely used for the purpose of inserting data into other digital media. In addition, LSB is easy to implement in applications [3]. The LSB method in audio files works by modifying the last few bytes of the audio file to hide the sequence of bytes that contain secret data.

The following illustrates the process of inserting a message into an audio file using LSB. If it is known that the bytes of the message file to be inserted are 11010011 and the bits of the audio file are

10101110 10101100 11100111 11100001
10010111 11100011 11011011 00101110

Then the result of the LSB insertion is:

10101111 10101101 11100110 11100001
10010110 11100010 11011011 00101111

2. Peak Signal-to-Noise Ratio (PSNR)

PSNR is an indicator to find out the changes that occur between the original file and the stego file by calculating the bit changes that occur between the two. Before calculating the PSNR, you must find the MSE (Mean Square Error) value. With the formula that can be seen in the following equation [6]:

$$MSE\ (x,y) = \frac{1}{N}\sum_{1=1}^{N}(xi - yi)^2$$

N = Number of data
Xi = audio carrier
Yi = stego audio

After the MSE value has been obtained, it is continued with the search for the PSNR value by calculating the maximum signal value divided by the MSE value. With the formula that can be seen in the equation below

$$PSNR = 10\ \log_{10}\left(\frac{255}{MSE}\right)$$

3. Blowfish

The blowfish algorithm was designed by a cryptanalyst named Bruce Schneier to replace the DES algorithm in 1993. blowfish is a symmetric key cryptography algorithm. blowfish uses the same key mechanism both in the encryption and decryption processes that use input and output data in the form of data blocks with a size of 64 bits.[5]

As can be seen in Fig. 1, Blowfish is a Feistel network consisting of 16 rounds. The input is a 64-bit data element

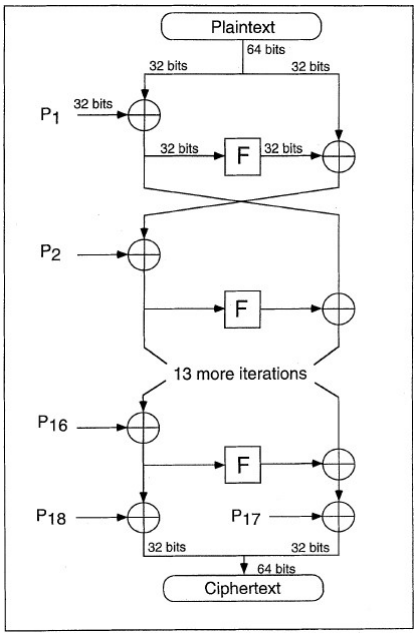


Fig. 1. Block Diagram of Blowfish

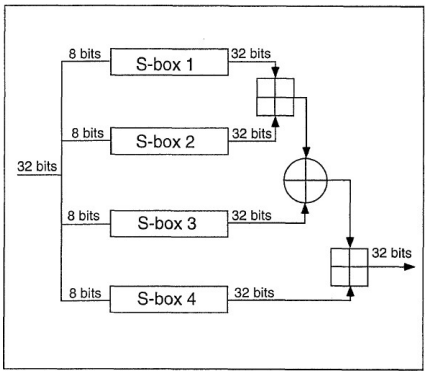


Fig. 2. Function F.

From Fig. 2 it can be simplified that

$$F(XL) = ((S_{1,\,a} + S_{2,\,b} \bmod 2^{32}) \text{ XOR } S_{3,c}) + S_{4,\,d} \bmod 2^{32}$$

4. WAV

A WAVE file (.WAV) is one of several encoding file formats in digital audio. audio files with the .wav format are uncompressed and have good sound quality, so when changes occur to the audio file it will not cause noise so there is no suspicion

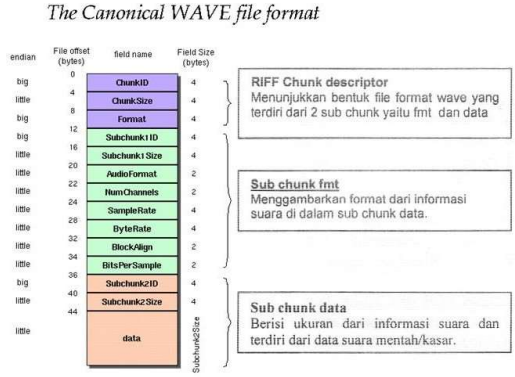


Fig. 3. WAV Format.

Figure II-1 shows the format of the WAVE file. It was explained that in the WAVE file there are 3 parts, RIFF, fmt, and data. the header is in the first 44 bytes of the WAV file, and the 45th to the last bytes are the data from the WAVE file. In this study, the last bit in the header will not be changed because it can make significant changes to the audio file after the message is inserted.

B. Previous Research

Research [6] by using a combination of blowfish and LSB algorithms on audio files. the results of hiding secret messages in 5 audio files show a very good PSNR value, which is between 80.01dB to 93.31dB.

Research [8] using LSB method for image and audio steganography. the results of this study are The main advantage of the method is a very high watermark channel capacity, the use of only one LSB of the host audio sample gives capacity of 44.1 kbps. The obvious disadvantage is the extremely low robustness of the method, due to fact that random changes of the LSBs destroy the coded watermark

Research [9], Steganography on the WAV format audio cover media using the methods of LSB, LSB + 1, LSB + 2, LSB + 3, LSB + 4, LSB + 5, LSB + 6, and LSB + 7. The PSNR evaluation shows that the larger the cover media size used, the smaller the risk of quality degradation, and the bigger the secret message, the greater the noise that will be generated. the most recommended technique is the LSB + 5 technique with a larger steganographic capacity and minimal risk of quality loss.

Research [10] conducted an analysis on the blowfish algorithm, the result is It can be concluded that blowfish presents good avalanche text from the second round. However, blowfish has a good non-linear relation between plaintext and cipher text.

3. Methodology

A. Data

The type of data used as the object of research is secondary data. There are 2 types of data needed in this study, audio data in WAVE format and message data in pdf and txt formats.

Table 1. Message File.

No.	File Name	File Size (Byte)
1.	PesanTXT1.txt	19,772 bytes
2.	PesanTXT2.txt	21,114 bytes
3.	PesanTXT3.txt	23,590 bytes

4.	PesanTXT4.txt	26,883 bytes
5.	PesanTXT5.txt	30,189 bytes
6.	PesanPDF1.pdf	38,017 bytes
7.	PesanPDF2.pdf	107,056 bytes
8.	PesanPDF3.pdf	109,642 bytes
9.	PesanPDF4.pdf	116,959 bytes
10.	PesanPDF5.pdf	124,672 bytes

The message file in Table. 1. are files that will be encrypted and hidden in the audio file

Table 2. Audio Files.

No.	File Name	File Size (Byte)
1.	Carrier 1.wav	1,073,218 bytes
2.	Carrier 2.wav	2,146,166 bytes
3.	Carrier 3.wav	5,226,766 bytes
4.	Carrier 4.wav	10,406,738 bytes

The audio files in **Table. 2.** are files that will be used as a container audio for hidden messages to be inserted in it.

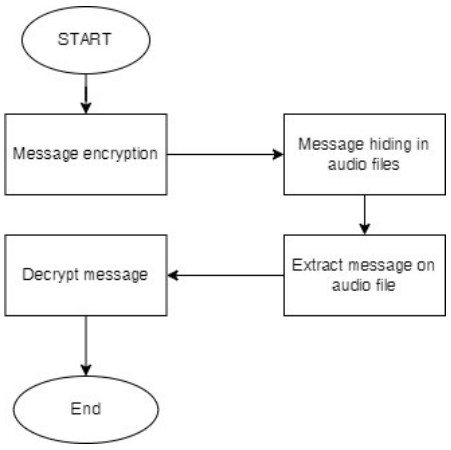


Fig. 4. Research Stages Diagram

Based on **Fig. 4.** the research will be carried out by encrypting the message first, then the encrypted message will be inserted in the audio file and provide output in the form of an audio file that has been inserted a message (stego audio). then to retrieve the hidden message file, extract the message and will provide output in the form of an encrypted message file, then decrypt the message file so that the message file can be opened again as before.

4. Result

A. Message File Encryption and Decryption Test

Table. 3. Encryption and Decription Test Result

No.	Message Files	Message file size	Message file size after encryption (Byte)	Time (ms)	
				Encrypt	Decrypt
1.	PesanTXT1.txt	19,772 bytes	19,776 bytes	543.00	548.74
2.	PesanTXT2.txt	21,114 bytes	21,120 bytes	539.32	531.46
3.	PesanTXT3.txt	23,590 bytes	23,592 bytes	544.33	536.40

4.	PesanTXT4.txt	26,883 bytes	26,888 bytes	554.06	541.11
5.	PesanTXT5.txt	30,189 bytes	30,192 bytes	557.30	544.22
6.	PesanPDF1.pdf	38,017 bytes	38,024 bytes	555.94	534.77
7.	PesanPDF2.pdf	107,056 bytes	107,064 bytes	561.28	533.03
8.	PesanPDF3.pdf	109,642 bytes	109,648 bytes	542.36	538.40
9.	PesanPDF4.pdf	116,959 bytes	116,960 bytes	534.50	539.13
10.	PesanPDF5.pdf	124,672 bytes	124,680 bytes	547.77	534.71

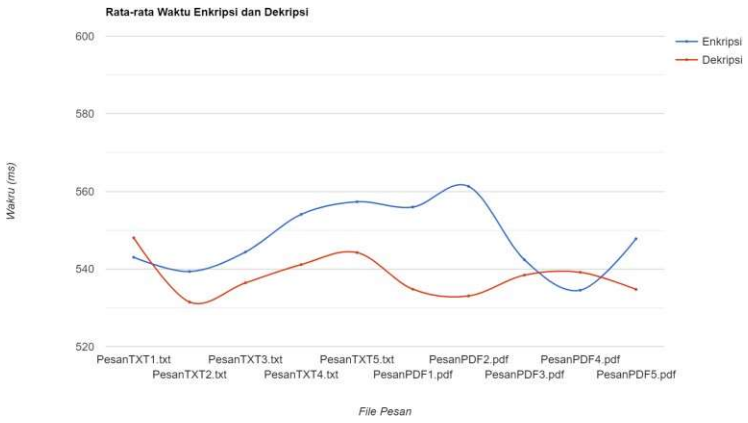


Fig.5. Encryption and Decryption Average Time

Message encryption and decryption testing can be seen in **Table. 5.** The result of testing the message file using blowfish is the difference in the size of the message file after the encryption process is carried out. This is because the length of the encrypted file is less than the predefined cipher block size, which is 64bit, so Blowfish will add padding or layers so that the cipher block size is met.

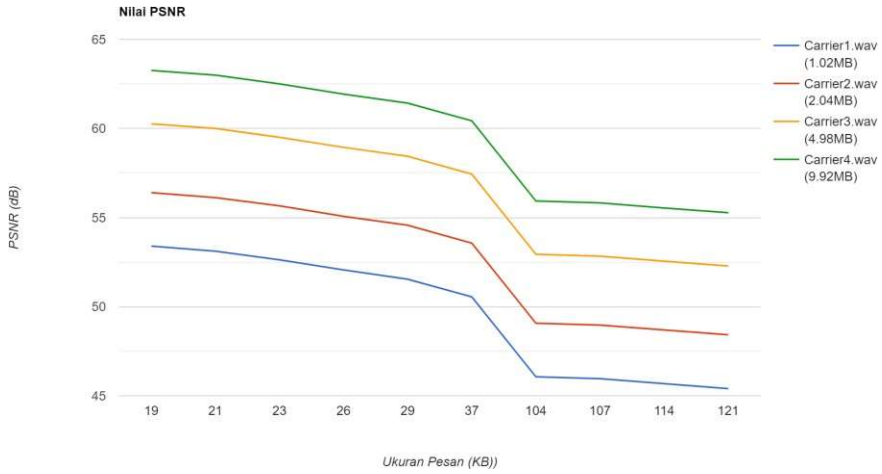
While testing the encryption and decryption time for message files gives an average time for encryption of 547.98ms while the average time for decryption is 538.19ms. The longest time for the encryption process is 557.30ms and the fastest is 534.50ms, while the longest time for the decryption process is 548.74ms and the fastest is 531.46ms.

B. Hiding Messages on Audio Files Test

Table. 4. PSNR of Audio File

No.	**Message File**	***Audio Carrier***	***Stego Audio***	**PSNR (dB)**
1.	PesanTXT1Enc.txt	Carrier 1.wav	StegoTXT1-1.wav	53.40
	PesanTXT2Enc.txt		StegoTXT2-1.wav	53.11
	PesanTXT3Enc.txt		StegoTXT3-1.wav	52.63
	PesanTXT4Enc.txt		StegoTXT4-1.wav	52.06
	PesanTXT5Enc.txt		StegoTXT5-1.wav	51.54
	PesanPDF1Enc.pdf		StegoPDF1-1.wav	50.55
	PesanPDF2Enc.pdf		StegoPDF2-1.wav	46.06
	PesanPDF3Enc.pdf		StegoPDF3-1.wav	45.96
	PesanPDF4Enc.pdf		StegoPDF4-1.wav	45.68
	PesanPDF5Enc.pdf		StegoPDF5-1.wav	45.40
2.	PesanTXT1Enc.txt	Carrier 2.wav	StegoTXT1-2.wav	56.40
	PesanTXT2Enc.txt		StegoTXT2-2.wav	56.12
	PesanTXT3Enc.txt		StegoTXT3-2.wav	55.66
	PesanTXT4Enc.txt		StegoTXT4-2.wav	55.07

	PesanTXT5Enc.txt		StegoTXT5-2.wav	54.57
	PesanPDF1Enc.pdf		StegoPDF1-2.wav	53.56
	PesanPDF2Enc.pdf		StegoPDF2-2.wav	49.07
	PesanPDF3Enc.pdf		StegoPDF3-2.wav	48.96
	PesanPDF4Enc.pdf		StegoPDF4-2.wav	48.69
	PesanPDF5Enc.pdf		StegoPDF5-2.wav	48.42
3.	PesanTXT1Enc.txt	Carrier 3.wav	StegoTXT1-3.wav	60.26
	PesanTXT2Enc.txt		StegoTXT2-3.wav	60.00
	PesanTXT3Enc.txt		StegoTXT3-3.wav	59.50
	PesanTXT4Enc.txt		StegoTXT4-3.wav	58.94
	PesanTXT5Enc.txt		StegoTXT5-3.wav	58.44
	PesanPDF1Enc.pdf		StegoPDF1-3.wav	57.44
	PesanPDF2Enc.pdf		StegoPDF2-3.wav	52.94
	PesanPDF3Enc.pdf		StegoPDF3-3.wav	52.83
	PesanPDF4Enc.pdf		StegoPDF4-3.wav	52.55
	PesanPDF5Enc.pdf		StegoPDF5-3.wav	52.28
4.	PesanTXT1Enc.txt	Carrier 4.wav	StegoTXT1-4.wav	63.26
	PesanTXT2Enc.txt		StegoTXT2-4.wav	62.99
	PesanTXT3Enc.txt		StegoTXT3-4.wav	62.50
	PesanTXT4Enc.txt		StegoTXT4-4.wav	61.93
	PesanTXT5Enc.txt		StegoTXT5-4.wav	61.42
	PesanPDF1Enc.pdf		StegoPDF1-4.wav	60.43
	PesanPDF2Enc.pdf		StegoPDF2-4.wav	55.93
	PesanPDF3Enc.pdf		StegoPDF3-4.wav	55.82
	PesanPDF4Enc.pdf		StegoPDF4-4.wav	55.54
	PesanPDF5Enc.pdf		StegoPDF5-4.wav	55.27



Based on the tests, hiding message files in audio files using the LSB method works well with PSNR values above 30dB.

5. Conclusion

Based on the research that has been done it can be concluded that:

A. The difference in the message file after encryption using blowfish is that there is a slight increase in file size.

B. The speed of encryption and decryption of message files with a size of 20KB to 125 KB using a constant blowfish algorithm with an average length of time for encryption of 547.98ms

while the average time for decryption is 538.19ms. The longest time for the encryption process is 557.30ms and the fastest is 534.50ms, while the longest time for the decryption process is 548.74ms and the fastest is 531.46ms.

C. The results of the study that the audio quality inserted with the message was very good, with a PSNR value above 30dB.

D. The smaller hidden file size in the audio carrier file, the larger the resulting PSNR value.

References

[1] Sari, J. I., & Sihotang, H. T. (2017). Implementasi Penyembunyian Pesan Pada Citra Digital Dengan Menggabungkan Algoritma HILL Cipher Dan Metode Least Significant BIT (LSB). Jurnal Mantik Penusa, 1(2).

[2] R. Munir, Kriptografi. Bandung : Informatika, 2006.

[3] Nur'aini, S. (2019). Steganografi Pada Digital Image Menggunakan Metode Least Significant Bit Insertion. Walisongo Journal of Information Technology, 1(1), 75-90.

[4] Budi Prasetiyo, I. U., Much Aziz Muslim, I. U., & Susanto, H. (2017). Penerapan Kriptografi Algoritma Blowfish pada Pengamanan Pesan Data Teks Implementation Of Blowfish Algorithm Cryptography On Text Message Data Security. Techno. COM, 16(4), 358-366.D. Abdullah, D. N. Saputro, "Implementasi Algoritma Blowfish dan Metode Least Significat Bit Insertion Pada Video Mp3". Jurnal Pseudocode, Vol III. No 2. September 2016. ISSN 2355-5920, 2016.

[5] Hemeida, F., Alexan, W., & Mamdouh, S. (2019, October). Blowfish–secured audio steganography. In 2019 Novel Intelligent and Leading Emerging Sciences Conference (NILES) (Vol. 1, pp. 17-20). IEEE.R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Bhalshankar, S., & Gulve, A. K. (2015). Audio steganography: LSB technique using a pyramid structure and range of bytes. arXiv preprint arXiv:1509.02630.

[7] B. Schneier, Description of a New Variable-Length Key, 64-Bit Block Cipher (Blowfish), Springer-Verlag, 1996.

[8] Chadha, A., & Satam, N. (2013). An efficient method for image and audio steganography using Least Significant Bit (LSB) substitution. arXiv preprint arXiv:1311.1083.

[9] Indrayani, R. (2020, November). Modified LSB on Audio Steganography using WAV Format. In 2020 3rd International Conference on Information and Communications Technology (ICOIACT) (pp. 466-470). IEEE.