Digital Text Security with Steganography Least Significant Bit and Audio Feature Extraction

M Delina¹, H Nasbey¹, A Gunawan¹ and S Muhasyah²
¹Physics Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta. Jl. Rawamangun Muka No1, Jakarta Timur 13220, Indonesia.
²Ministry of Foreign Affairs of the Republic of Indonesia. Jl. Taman Pejambon No 6. Jakarta Pusat 10110, Indonesia.
Email: mutia_delina@unj.ac.id

Abstract. The study developed a digital text security system by inserting the classified message or text into digital audio (stego audio) with the least significant bit (LSB) steganography and audio feature extraction. The text bits were inserted into each cover audio frame based on the value of feature extraction energy and zero-crossing rate (ZCR). The study has demonstrated that the least significant bit and audio feature extraction can cover the digital text in stego audio and then be extracted back to the original digital text. There is a small difference in feature extraction energy value between stego audio and cover audio. While there is no difference in feature extraction zero-crossing rate between stego audio and cover audio. The study also measured signal-to-noise ratio (SNR); the detected noise in stego audio because of the change in the least significant bit. The SNR is 111.97 – 130.00 dB.

1. Introduction
Nowadays, digital transmission through the internet increases rapidly [1]. Therefore, the information exchange easily worldwide every second. However, for a classified file that is only accessed by an authorized person, security and authenticity are still the primary concern [2]. During transmission, the classified information has risked being attacked by an eavesdropper [3] such as interruption, interception, modification, and fabrication[4]. As prevention, this classified information can be protected by steganography [5] which covers the information with various media[6], e.g., text [7], images [8], audio [9], and video [10]. The cover is called the cover image. Meanwhile, the text with the cover is a stego image. With the aid of cover image, the classified information is not visible directly and looks natural. Human vision cannot recognize the different between cover image and Stego image [11].

Steganography has been applied in various fields. In pharmacy, steganography present to secure the electronic prescription and patient data from the eavesdropper [12]. In medical, steganography employs to protect the medical image. Therefore, no theft can steal the data nor modified it. An illegal modification of medical images will lead to a false diagnose[13]. In politics, the steganography identified cheaters in an election[14]. In transportation, steganography is practiced for a ticket barcode to protect the passenger data such as transaction, bank account, and traveling history[15]. In Information and Communication Technology (ICT), steganography secures the Internet of Things (IoT) data [16].

This study employed the steganography least significant bit [17] with audio feature extraction [18] to protect the classified text. Here, the information is covered by a 16-bit double cannal digital audio with various genres. By combining audio feature extraction and steganography least significant bit, the study obtained a better security level result.
Steganography employs a cover to hide digital information. This cover is like a wrapping box for the information during the transmission from sender to receiver. The digital information may in bits or a block. The cover itself has various types like text, image, audio, video, sound, or other media [19].

Least Significant Bits (LSB) is a steganography method in the spatial domain. This technique is simple because the cover LBS’s can be replaced by message bits directly. For example, in steganography with cover image, the LSB in each image pixel is changed by the message bits.

Feature extraction is essential for audio analysis, such as pattern recognition and machine learning. With this feature, the audio characteristic such as human voice, music, and explosion can be recognized [20]. This feature can also distinguish music genres such as pop, rock, jazz, and dangdut. The feature extraction consists of two main parameters, time and frequency based. The time-based itself composed of energy and Zero-Crossing Rate (ZCR) feature. The energy feature \( E(i) \) can be denoted as follow.

\[
E(i) = \frac{1}{W_L} \sum_{n=1}^{W_L} |x_i(n)|^2, \tag{1}
\]

where \( W_L \) and \( x_i \) are the frame number at the \( i \)-th audio sample, respectively. \( W_L \) is obtained from ratio between number of samples in the audio cover reference with number of text character. Meanwhile, the ZCR to count the noise signal is denoted as follows.

\[
Z(i) = \frac{1}{2W_L} \sum_{n=1}^{W_L} [\text{sgn} \left[ x_i(n) \right] - \text{sgn} \left[ x_i(n-1) \right]] \tag{2}
\]

where

\[
\text{sgn}[x_i(n)] = \begin{cases} 
1, & x_i(n) \geq 0, \\
0, & x_i(n) < 0.
\end{cases}
\tag{3}
\]

The difference expected signal with the noise signal can be obtained by measuring the Signal-to-Noise Ratio (SNR) on the following equation.

\[
SNR = 10 \log_{10} \frac{\sum_{n=1}^{N} x(n)^2}{\sum_{n=1}^{N} [x(n) - y(n)]^2} \tag{4}
\]

2. Method
The research started with the development of audio feature extraction program. Here, there were three audio files as the cover audio in the .wav format (audio 1, audio 2, and audio 3), and a (classified) digital text in the .txt format.

Each text was then inserted into the audio cover by least significant bit steganography technique. The process was started by extracted and inserted the (classified) text into the cover audio, between the initial and final reff (reff). Furthermore, the initial and final reff time was the stego key. The audio file details and time interval of reff are shown in table 1 and table 2, respectively.

### Table 1. Cover audio

| Audio Stego | Frequency (Hz) | Data       |
|-------------|----------------|------------|
| Audio 1     | 44,100         | 10,601,328 |
| Audio 2     | 44,100         | 9,532,272  |
| Audio 3     | 48,000         | 12,206,064 |

### Table 2. Time interval of the first reff.

| Cover Audio | Initial Reff (s) | Final Reff (s) |
|-------------|------------------|----------------|
| Audio 1     | 59               | 90             |
| Audio 2     | 96               | 149            |
| Audio 3     | 62               | 89             |
The text character's number determined the energy feature and ZNR frame. Therefore the character's number was similar to the frame length \( W_f \). The initial and final reff was also obtained to extract the energy feature and ZNR. Here, the energy feature influenced the bit arrangement in the text, where an 8 bits character was inserted in each frame. Once the feature extraction in a frame was above the average energy, the 8 bits character was inserted in the initial frame. Meanwhile, if the energy feature extraction was below the average energy, the 8 bits character was inserted into the middle of the frame.

Finally, the SNR was obtained to analyze the stego audio difference before and after the text was inserted. Thus, the SNR determined the success of the steganography method application. As a comparison, SNR from previous research was employed, the details is shown in table 3.

| Research       | SNR (dB)     |
|----------------|-------------|
| Wakiyama [21]  | 45.2 – 50.4 |
| Divya [22]     | 55.37       |
| Meligy [23]    | 78.11 – 86.09 |
| Meligy [24]    | 94.76       |

3. Results and Discussion

This research obtained three audio files as the audio cover and four texts. Yet, to concise the graph presentation, this paper only showed the figure of audio1.wav. The energy and ZNR feature of audio 1 before the text was inserted was shown by figure 1.

**Figure 1.** (a) Energy feature of audio1.wav before inserted by text. (b) ZNR of audio1.wav before inserted by text.

Figure 1(a) showed the energy feature of audio1.wav or stego audio 1 before inserted by (classified) text. Meanwhile figure 1(b) was the ZNR of audio1.wav before the audio was inserted by the text.
The energy feature difference of stego audio $E_S$ with cover audio $E_C$ after the text was inserted to audio1.wav.

Figure 2 showed the energy feature difference of stego audio $E_S$ and cover audio $E_C$ after the text was inserted to the audio1.wav. The graph showed that the energy feature difference was small. Therefore, it was not influenced the text position in the stego audio.

The ZCR feature extraction difference of stego audio $E_S$ with cover audio $E_C$ after the text was inserted to audio1.wav.

Figure 3 showed that the ZCR feature of stego audio and cover audio has similar value. The comparison of stego audio before and after the text was inserted obtained from SNR. The SNR of three audio stego and four text were shown in table 4.

| Digital Information | Audio Stego (dB) | Audio Stego (dB) | Audio Stego (dB) |
|---------------------|-----------------|-----------------|-----------------|
| Text 1              | 116.84          | 125.97          | 121.03          |
| Text 2              | 115.37          | 124.65          | 119.64          |
| Text 3              | 121.00          | 125.77          | 120.56          |
| Text 4              | 111.97          | 121.22          | 116.30          |

The SNR was 111.97 – 130.99 dB. This result was more significant than SNR from previous research in table 3. The SNR value larger than 40 dB was good because the stego audio’s noise was smaller than cover audio.

4. Conclusion
This study show that the combination of steganography least significant bits with the audio extraction feature give a better security result. It shown that this combination method was able to cover the digital information which is shown by the SNR value 119.95 – 136.03 (dB).
References

[1] S. Dhawan and R. Gupta, “Analysis of various data security techniques of steganography: A survey,” *Inf. Secur. J.*, vol. 30, no. 2, pp. 63–87, 2021, doi: 10.1080/19393555.2020.1801911.

[2] M. R. Islam, T. R. Tanni, S. Parvin, M. J. Sultana, and A. Siddiqua, “A modified LSB image steganography method using filtering algorithm and stream of password,” *Inf. Secur. J.*, vol. 00, no. 00, pp. 1–12, 2020, doi: 10.1080/19393555.2020.1854902.

[3] R. Ranjith Kumar, S. Jayasudha, and S. Pradeep, “Efficient and secure data hiding in encrypted images: A new approach using chaos,” *Inf. Secur. J.*, vol. 25, no. 4–6, pp. 235–246, 2016, doi: 10.1080/19393555.2016.1248582.

[4] M. Delina, B. H. Iswanto, H. Permana, and S. Muhasyah, “The simulation of a symmetric quantum key distribution,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 434, no. 1, 2018, doi: 10.1088/1757-899X/434/1/012023.

[5] M. Fateh and M. Rezvani, “An email-based high capacity text steganography using repeating characters,” *Int. J. Comput. Appl.*, vol. 0, no. 0, pp. 1–7, 2018, doi: 10.1080/1206212X.2018.1517713.

[6] H. Dutta, R. K. Das, S. Nandi, and S. R. M. Prasanna, “An Overview of Digital Audio Steganography,” *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, vol. 37, no. 6, pp. 632–650, 2020, doi: 10.1080/02564602.2019.1699454.

[7] Z. Jalil and A. M. Mirza, “Innovations in Computing Sciences and Software Engineering,” *Innov. Comput. Softw. Eng.*, vol. 10, 2010, doi: 10.1016/j.proeng.2012.01.917.

[8] T. H. Park, J. G. Han, Y. H. Moon, and I. K. Eom, “Performance improvement of LSB-based steganalysis using bit-plane decomposition of images,” *Imaging Sci. J.*, vol. 64, no. 5, pp. 262–266, 2016, doi: 10.1080/13682199.2016.1171452.

[9] Y. Xiang, I. Natgunanathan, D. Peng, W. Zhou, and S. Yu, “A dual-channel time-spread echo method for audio watermarking,” *IEEE Trans. Forensics Secur.*, vol. 7, no. 2, pp. 383–392, 2012, doi: 10.1109/TIFS.2011.2173678.

[10] K. Srinivasan, T. Gowthaman, and J. Kanakaraj, “A novel copyright marking approach using steganography and robust RSA asymmetric-key cryptographic technique in audio files,” *J. Discret. Math. Sci. Cryptogr.*, vol. 20, no. 8, pp. 1563–1571, 2017, doi: 10.1080/09720529.2017.1402575.

[11] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Information Hiding — A Survey,” vol. 87, no. 7, 1999.

[12] A. Omotosho, J. Emuoyibofarhe, and C. Meinel, “Securing E-Prescription from Medical Identity Theft Using Steganography and Antiphishing Techniques,” *J. Appl. Secur. Res.*, vol. 12, no. 3, pp. 447–461, 2017, doi: 10.1080/19361610.2017.1315788.

[13] S. K. Patel, C. Saravanan, and V. K. Patel, “Cloud-based Reversible Dynamic Secure Steganography Model for embedding pathological report in medical images,” *J. Inf. Optim. Sci.*, vol. 0, no. 0, pp. 1–9, 2019, doi: 10.1080/19393555.2016.1641273.

[14] B. Lakshmi Sirisha, S. Srinivas Kumar, and B. Chandra Mohan, “Identification of cheaters in elections using steganography,” *J. Inf. Optim. Sci.*, vol. 37, no. 2, pp. 271–278, 2016, doi: 10.1080/02564602.2015.1130891.

[15] O. E. K. Amh, S. Muhasyah, and M. Delina, “Least significant bit steganography method for the digital data protection in the barcode,” *AIP Conf. Proc.*, vol. 2169, no. November, 2019, doi: 10.1063/1.5132672.

[16] A. K. Bairagi, R. Khondoker, and R. Islam, “An efficient steganographic approach for protecting communication in the Internet of Things (IoT) critical infrastructures,” *Inf. Secur. J.*, vol. 25, no. 4–6, pp. 197–212, 2016, doi: 10.1080/19393555.2016.1206640.

[17] M. Baritha Begum and Y. Venkataraman, “LSB based audio steganography based on text compression,” *Procedia Eng.*, vol. 30, no. 2011, pp. 703–710, 2012, doi: 10.1016/j.proeng.2012.01.917.

[18] A. G. Chiţu, L. J. M. Rothkrantz, P. Wiggers, and J. C. Wojdel, “Comparison between different feature extraction techniques for audio-visual speech recognition,” *J. Multimodal User Interfaces*, vol. 1, no. 1, pp. 7–20, 2007, doi: 10.1007/bf02884428.

[19] S. Khan, N. Ahmad, and M. Wahid, “Varying index varying bits substitution algorithm for the implementation of VLSB steganography,” *J. Chinese Inst. Eng. Trans. Chinese Inst. Eng. A*, vol. 39, no. 1, pp. 101–109, 2016, doi: 10.1080/02538389.2015.1082933.

[20] E. Keogh, “Automatic Feature Extraction for Classifying,” *Artif. Intell.*, pp. 127–149, 2005.
[21] M. Wakiyama, Y. Hidaka, and K. Nozaki, “An audio steganography by a low-bit coding method with wave files,” Proc. - 2010 6th Int. Conf. Intell. Inf. Hiding Multimed. Signal Process. IHMSP 2010, pp. 530–533, 2010, doi: 10.1109/IHMSP.2010.135.

[22] S. S. Divya and M. R. M. Reddy, “Hiding Text In Audio Using Multiple LSB Steganography And Provide Security Using Cryptography,” Int. J. Sci. Technol. Res., vol. 1, no. 6, pp. 68–70, 2012.

[23] A. M. Meligy, M. M. Nasef, and F. T. Eid, “An Efficient Method to Audio Steganography based on Modification of Least Significant Bit Technique using Random Keys,” Int. J. Comput. Netw. Inf. Secur., vol. 7, no. 7, pp. 24–29, 2015, doi: 10.5815/ijcnis.2015.07.03.

[24] A. M. Meligy, M. M. Nasef, and F. T. Eid, “A Hybrid Technique for Enhancing the Efficiency of Audio Steganography,” Int. J. Image, Graph. Signal Process., vol. 8, no. 1, pp. 36–42, 2016, doi: 10.5815/ijigsp.2016.01.04.