A Method of Cover Audio Selection for Embedding Based on Various Criteria

Muhammad Harith Noor Azam [1], *Farida* Ridzuan [1,2,*], *M Norazizi Sham* Mohd Sayuti [3], *AH* Azni [1,2], *Sakinah* Ali Pitchay [1,2], and *Najwa Hayaati* Mohd Alwi [1,2]

[1] Faculty of Science and Technology, Universiti Sains Islam Malaysia, 71800 Nilai, Malaysia
[2] Cyber Security and Systems Research Unit, Faculty of Science and Technology, Universiti Sains Islam Malaysia, 71800 Nilai, Malaysia
[3] Faculty of Engineering and Built Environment, Universiti Sains Islam Malaysia, Malaysia

Abstract. The main goal of an audio steganography method is to improve the capacity, imperceptibility, and robustness. Several methods of audio steganography have been proposed to enhance its capabilities. To further optimize the efficiency of the audio steganography method, this paper proposes to select the appropriate audio cover for concealing the message. The selection of an appropriate cover can significantly improve the quality of the output and facilitate the development of innovative audio steganography techniques. This paper proposed a new audio cover selection method, which can further enhance the characteristics of the resulting output. This method ranked each cover according to the qualities that need to be boosted using various measurements. These measurements are the Maximum Available Space (MAS), Peak Signal to Noise Ratio (PSNR) and Bit Error Rate (BER) for capacity, imperceptibility, and robustness respectively. Based on the experiment conducted using five cover audios stored in the self-created database, each audio is ranked differently based on the measurement used to determine its characteristic level. In conclusion, the proposed cover selection method can be used to select the most proper cover, hence improving the characteristics of audio steganography.

1 Introduction

Audio steganography is a method to hide secret message by using certain method to modify the audio hence, ensuring that only the sender and the intended receiver know about the existence of the secret message [1]. To achieve covert communication, an audio steganography method is utilized by the sender to embed a confidential message within user selected cover audio. The method embeds the message into the audio to create the stego-file. The sender then delivers to the receiver over an unsecured channel [2]. The recipient decodes the stego file using the same audio steganography method to extract the message.

It is essential to satisfy three characteristics of audio steganography which are capacity, imperceptibility, and robustness to ensure audio steganography to be effectively implemented. Capacity refers to the potential amount of the secret data embedded in the cover audio [4]. Next, imperceptibility describes as the measure of the ability to embed the secret message without affecting the audio signal that human perceive [5]. Lastly, robustness refers to the ability of stego-file to withstand the attacks [6]. However, there are three trade-off exist between them which are: 1) capacity-imperceptibility, 2) imperceptibility-robustness, and 3) robustness-capacity [7].

*Corresponding author: farida@usim.edu.my

Generally, achieving optimal performance in any steganography system requires making certain trade-offs that may hinder the simultaneous maximization of all desirable characteristics. For instance, the capacity-imperceptibility trade-off highlights that increasing the message size results in a reduction in the quality of the cover. This is due to the fact that larger messages require more modifications, thereby reducing the overall quality of the cover[8], [9].

Examining certain characteristic and parameters, proposing a new method, and evaluating that method are typical steps in developing a new audio steganography method [10] thus enhancing the targeted characteristics. Despite various methods proposed by researchers, the importance of using quality cover audio for the audio steganography is often disregarded. Selecting poor quality cover audio may lead to a downgrade in certain characteristics. As an example, cover A and cover B are embedded with similar secret message. If cover A has a greater binary difference from the secret message compared to cover B, it will require more binary flipping to embed the secret message. This reduces the imperceptibility of the cover beyond what is necessary, which can be avoided by selecting cover B. Hence, the research on the cover audio selection method has the potential to serve as a complement to the audio steganography ecosystem by safeguarding and ensuring the intended high level of characteristics are maintained despite used by unknowledgable user.

Several cover selection methods were proposed for image steganography [11]–[14]. [11] proposed cover selection based on processing distortion and embedding distortion. The flow from original image to become stego-file is shown in Fig. 1.

Fig. 1. Flowchart of proposed cover selection method.

Based on Fig.1, the processing distortion is caused by the processing such as contrast enhancement or image denoising. The idea of conducting the image preprocessing is to fulfill certain requirements or criteria needed to pass the communication channel. On the other hand, embedding distortion is caused by the embedding process due to the modification of the image pixel. All the cover image in the database is calculated their processing distortion and embedding distortion. The images are arranged in ascending order after combining these two distortions. The image that has lowest distortion are selected for the embedding process. [12] proposed cover selection method that straight forward. During the cover selection process, the percentage of number '1' from the array of cover image pixels' binary value in the database is computed. Then the percentage of '1' from the binary value of secret message is also retrieved. Then both percentage values are compared and the cover image that has the most similar percentage value of '1' with secret message are selected. [13] proposed cover selection that used four evaluation metrics which are PSNR, match bits, Different Image Histogram (DIH) and Revisited Weighted Stego (RWS). Then it produces four distinct ranking lists based on these four metrics. Each one of them is independent with the others. [14] proposed a cover selection based on structural similarity index measure (SSIM). The cover selection method is trained by machine learning and embedded with 0.1 bit per pixel with the increment ranging from 0.1 to 1.0. After each image has been embedded with a different size of secret message, an array of SSIM values is collected. It simply determined by how many round SSIM of these images are managed to achieve above 0.99. Images that achieve the

highest scores are regarded as the highest quality of cover and can be utilized for actual embedding.

Based on the author's knowledge, there is only one cover selection method proposed for audio steganography while the rest of cover selection method are mainly proposed for image steganography and video steganography. [15] proposed a cover audio selection method that focus on the trade-off between capacity and the imperceptibility. The method uses the Signal-to-Noise Ratio (SNR) to measure imperceptibility and audio sample size to measure capacity. It employs NSGA-II as a searching mechanism to obtain the optimal set of solutions.

Thus, there is a lack of research conducted on cover audio selection methods. Although there are cover image selection methods focusing on robustness, the measurement cannot be directly implemented in audio. On top of that, the previous method on audio focuses only on the capacity and the imperceptibility characteristics [15]. Therefore, this research proposes a new cover selection method for audio steganography focusing on three different characteristics which are capacity, imperceptibility, and robustness. The research on audio steganography has far-reaching implications for users seeking to protect their sensitive information. By providing the best cover audio to be selected by the user, this study empowers users to communicate more securely and safeguard their data from unauthorized party. Overall, this research is a valuable tool for anyone seeking to safeguard their privacy and protect their digital assets.

The next section discusses the proposed work, followed by Result and Analysis. Last section concludes the work.

2 Proposed Work

The proposed work aims to give rank to each audio in the database based on different characteristics which are capacity, imperceptibility, and robustness. A general illustration of the proposed method is presented in Fig. 2.

Based on Fig. 2, the user must first enter the secret message. Following that, the method begins by evaluating the database's cover audio. During the evaluation process, the hidden message is inserted into the cover audio based on the number of embedded bits per sample (bps). This evaluation process as highlighted in Figure 1 uses three metrics to measure imperceptibility, capacity, and robustness. The evaluation process is explained in detail in subsection 2.1. There are two loops used in this method. The outer loop is used to manipulate the index audio variable while the inner loop is used to manipulate the bps variable. As this research intends to use audio that has 16-bit audio sample, therefore, the maximum bps used is 8, which is the maximum reported in [15] for enhancing robustness. It is because these 16-bit is divided equally into least significant bit (lsb) and most significant bit (msb). Any alteration to the bits under the MSB can lead to a significant distortion of the data, which in turn can compromise its imperceptibility [16]. Therefore, it is not advisable to use the bits under the MSB for any kind of alteration. Doing so can lead to a loss of data quality and make it unsuitable and unfeasible for use in steganography. On the other hand, the cover audio does not change much and hard to be detected if it is altered at the lsb [17]. After the evaluation is complete, each cover in the database is assigned to their own distinct rankings based on capacity, imperceptibility, and robustness.

2.1 Evaluation Process

Before the evaluation process was conducted, the cover was embedded with a secret message using the traditional LSB technique with variation of *bps*. Next, the evaluation on the cover

audio is conducted to measure capacity, and two evaluations on the stego-file produced are conducted to measure imperceptibility and robustness.

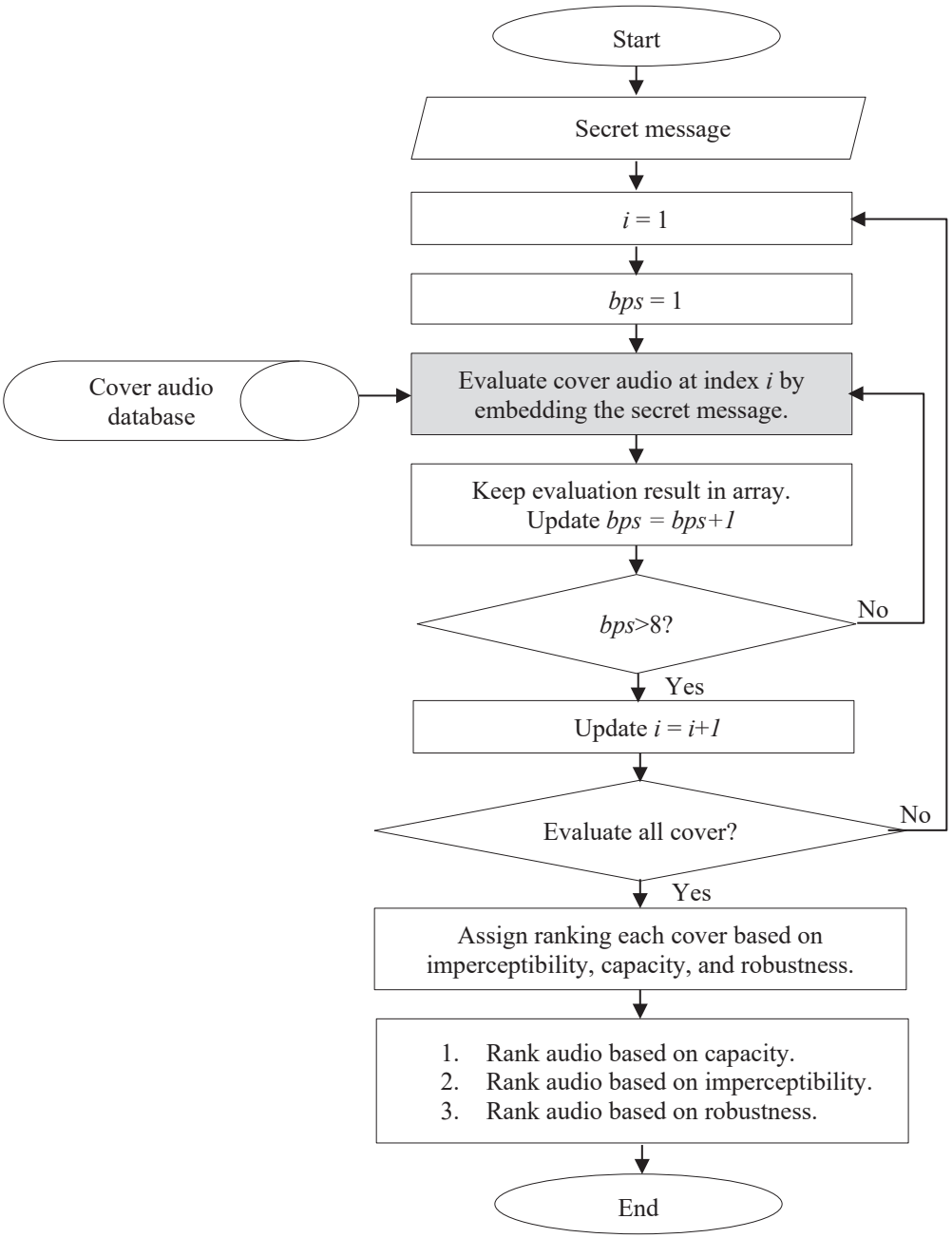

Fig. 2. Flowchart of proposed cover selection method.

Capacity Evaluation. Higher capacity indicates a bigger size of secret message can be embedded into that specific cover audio [18]. To measure the capacity effectively, this method

calculates the maximum capacity based on the embedding technique mentioned earlier using maximum available space. Maximum available space (MAS) is formulated as in Equation 1:

$$MAS = bps * d * sr \quad (1)$$

where *bps* refers to the bit embedded per sample, *d* is the duration of the audio in second and *sr* is the sample rate of the audio.

Imperceptibility Evaluation. Higher imperceptibility indicates a lower distortion produced from the embedding process [19]. This method measures the cover audio imperceptibility by measuring the PSNR of the stego-file produced. PSNR can be calculated using Equation 2 [15]:

$$PSNR = 10\, log_{10} \frac{R^2}{MSE} \quad (2)$$

where R is the peak signal value in the original cover audio and Mean Squared Error (MSE) is defined using Equation 3 [15]:

$$MSE = \frac{1}{n}\sum_{1}^{n}(x(i) - y(i))^2 \quad (3)$$

where x and y are original cover audio and stego-signals respectively, while i is denoted as sample index and n is denoted as the total audio sample.

Robustness Evaluation. Higher robustness indicates the stego-file has higher resistance towards the attacks [20]. Although there are many attacks implemented to investigate the robustness level, this method uses Additive White Gaussian Noise (AWGN) attack to measure the robustness level of the stego-file. A single point low level of AWGN attack at 100dB is introduced to the stego-file to ensure there is a differentiation between low and high robust stego-file. The robustness of the stego-file can be tested using the Bit error rate (BER), which calculates the error from the secret message retrieval and evaluates the ratio of the number of embedded message bits that result in an error during the retrieval process to the overall size of the secret message. BER was calculated using Equation 4 [21].

$$BER = 1 - \left(\frac{number\ or\ retrieval\ error}{total\ number\ of\ message\ bits}\right) \quad (4)$$

3 Result and Analysis

This section presents the result and analysis of the output of the proposed method. The objective of the experiment is to demonstrate the impact of audio quality on different bps settings. A database contains five audios with different durations which have 44100 sample rate were used. A specific sample rate was used, as it has a huge influence on the capacity performance. It determines the number of samples available per second. Higher sample rates increase the capacity performance; therefore, audio needs to be standardized. The size of secret message is set to 4KB for the evaluation purpose. MAS, PSNR and BER for each cover are recorded. In this experiment, the recorded results were ranked to determine the best and worst performing ones. Table 1 shows the result of cover audio index with bps used with their respective MAS, PSNR and BER while Table 2 shows the top 10 results of index audio with respective bps used, sorted from best to worst for each characteristic.

Based on Table 2, it was observed that the sorted selected covers with its bps based on capacity and robustness yielded similar results. This is because embedding using higher bps

leads to embedding at a higher level, which can improve robustness. Hence, these two ranking lists are equivalent. On the other hand, the imperceptibility ranking is different because it depends on the number of modifications made. It can vary depending on the number of audio samples which require flipping. Depending on the choice of audio and bits per sample (bps), the performance of audio steganography can either improve or deteriorate. The research showed that the suggested method for selecting cover audio can significantly affect the performance of audio steganography hence this method can complement the strength of other audio steganography methods used.

Table 1. Results of MAS, PSNR and BER of five audios with their respective *bps*.

[index audio, bps]	MAS	PSNR	BER	[index audio, bps]	MAS	PSNR	BER
[1,1]	132260	99.306	0.5	[3,5]	661300	81.319	0.78
[1,2]	264520	95.772	0.63	[3,6]	793560	76.744	0.82
[1,3]	396780	91.373	0.69	[3,7]	925820	70.945	0.84
[1,4]	529040	86.572	0.77	[3,8]	1058080	66.021	0.86
[1,5]	661300	81.261	0.78	[4,1]	132260	99.359	0.5
[1,6]	793560	76.541	0.82	[4,2]	264520	95.750	0.63
[1,7]	925820	70.899	0.84	[4,3]	396780	91.308	0.69
[1,8]	1058080	66.591	0.86	[4,4]	529040	86.674	0.77
[2,1]	132260	99.343	0.5	[4,5]	661300	81.383	0.78
[2,2]	264520	95.742	0.63	[4,6]	793560	76.688	0.82
[2,3]	396780	91.382	0.69	[4,7]	925820	70.767	0.84
[2,4]	529040	86.628	0.77	[4,8]	1058080	66.154	0.86
[2,5]	661300	81.319	0.78	[5,1]	132260	99.359	0.5
[2,6]	793560	76.744	0.82	[5,2]	264520	95.750	0.63
[2,7]	925820	70.945	0.84	[5,3]	396780	91.308	0.69
[2,8]	1058080	66.021	0.86	[5,4]	529040	86.674	0.77
[3,1]	132260	99.343	0.5	[5,5]	661300	81.383	0.78
[3,2]	264520	95.742	0.63	[5,6]	793560	76.688	0.82
[3,3]	396780	91.382	0.69	[5,7]	925820	70.767	0.84
[3,4]	529040	86.628	0.77	[5,8]	1058080	66.154	0.86

Table 2. Sorted Result.

Characteristic	**Sorted selected index audio with bps used (Best to Worst)**									
Capacity	[1,8]	[2,8]	[3,8]	[4,8]	[5,8]	[1,7]	[2,7]	[3,7]	[4,7]	[5,7]
Imperceptibility	[4,1]	[5,1]	[2,1]	[3,1]	[1,1]	[1,2]	[4,2]	[5,2]	[2,2]	3,2]
Robustness	[1,8]	[2,8]	[3,8]	[4,8]	[5,8]	[1,7]	[2,7]	[3,7]	[4,7]	[5,7]

4 Conclusion

In this paper, a new cover selection method was proposed for audio steganography. The proposed method used three evaluation metrics which are MAS, PSNR and BER to measure capacity, imperceptibility, and robustness. The experimental results indicate that different

cover audio and bps used affect capacity, imperceptibility, and robustness. The method generated a list of cover audios with its ranking for the user to select from, emphasizing the importance of audio quality when using it as a cover. As a conclusion, this method provides a solution in ensuring that an audio steganography method could perform to its optimal performance and allowing unknowledgeable users to perform steganography efficiently.

This work was funded by the Ministry of Higher Education (MOHE) of Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/USIM/02/1). The authors would like to express their gratitude to Universiti Sains Islam Malaysia (USIM) and MoHE for the support and facilities provided.

References

1. M. M. Mahmoud, H. T. Elshoush, *Enhancing LSB Using Binary Message Size Encoding for High Capacity, Transparent and Secure Audio Steganography-An Innovative Approach*, IEEE Access, **10**, 29954–29971, (2022).
2. F. Hemeida, W. Alexan, S. Mamdouh, *Blowfish – Secured Audio Steganography*, in 2019 Novel Intelligent and Leading Emerging Sciences Conference (NILES), 2019, **1**, 17–20, (2019).
3. S. B. S.- Smieee, R. S. Mohammed, *Recent Audio Steganography Trails and its Quality Measures*, in 2019 International Conference of Computer and Applied Sciences (CAS2019), 18 December 2019, Baghdad, Iraq, (2019).
4. M. H. N. Azam, F. Ridzuan, M. N. S. M. Sayuti, *A New Method to Estimate Peak Signal to Noise Ratio for Least Significant Bit Modification Audio Steganography*, Pertanika J. Sci. Technol., **30**(1), 497–511, (2022).
5. M. H. Noor Azam, F. Ridzuan, M. N. S. Mohd Sayuti, A. A. Alsabhany, *Balancing the Trade-Off between Capacity and Imperceptibility for Least Significant Bit Audio Steganography Method: A New Parameter*, in 2019 IEEE Conference on Application, Information and Network Security, AINS 2019, 19-21 November 2019, Pulau Pinang, Malaysia, (2019).
6. S. Utama, R. Din, *Performance Review of Feature-Based Method in Implementation Text Steganography Approac*h, J. Adv. Res. Appl. Sci. Eng. Technol., **28**(2), 325–333, (2022).
7. A. A. Alsabhany, A. H. Ali, F. Ridzuan, A. H. Azni, M. R. Mokhtar, *Digital Audio Steganography: Systematic Review, Classification, and Analysis of the Current State of the Art*, Comput. Sci. Rev., **38**, 100316, (2020).
8. S. Ahani, S. Ghaemmaghami, Z. J. Wang, *A Sparse Representation Based Wavelet Domain Speech Steganography Method*, IEEE/ACM Trans. Audio Speech Lang. Process., **23**(1), 80–91, (2015).
9. S. Solak, *High Embedding Capacity Data Hiding Technique Based on EMSD and LSB Substitution Algorithms*, IEEE Access, **8**, 166513–166524, (2020).
10. M. Wakiyama, Y. Hidaka, K. Nozaki, *An Audio Steganography by a Low-Bit Coding Method with Wave File*s, in 2010 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2010, 15-17 October 2010, Darmstadt, Germany, (2010).
11. Z. Wang, X. Zhang, Z. Qian, *Practical Cover Selection for Steganography*, IEEE Signal Process. Lett., **27**(c), 71–75, (2020).
12. P. D. Shah, R. S. Bichkar, *Genetic Algorithm Based Approach to Select Suitable Cover Image for Image Steganography*, in 2020 International Conference for

Emerging Technology, INCET 2020, 05-07 June 2020, Belgaum, India, (2020).
13. R. D. Rashid, *Cover Image Selection for Embedding Based on Different Criteria*, Mob. Multimedia/Image Process. Secur. Appl. 2020, **11399**, 30, (2020).
14. N. Hamid, B. S. Sumait, B. I. Bakri, O. Al-Qershi, *Enhancing Visual Quality of Spatial Image Steganography Using SqueezeNet Deep Learning Network*, Multimed. Tools Appl., **80**(28–29), 36093–36109, (2021).
15. M. H. Noor Azam, F. Ridzuan, M. N. S. Mohd Sayuti, *Optimized Cover Selection for Audio Steganography Using Multi-Objective Evolutionary Algorithm*, J. Inf. Commun. Technol., **22**(2), 255–282, (2023).
16. M. K. Linga Murthy, P. B. Madhavi, S. A. Ahamad, K. Vamsi, Y. Mallikarjuna Rao, *Implementation and Analysis of Image Steganography using Convolution Neural Networks*, Int. Interdiscip. Humanit. Conf. Sustain. IIHC 2022 - Proc., 108–113, (2022).
17. Lindawati, R. Siburian, *Steganography Implementation on Android Smartphone Using the LSB (Least Significant Bit) to MP3 and WAV Audio*, in the 3rd International Conference on Wireless and Telematics 2017, 27-28 July 2017, Palembang, Indonesia, (2017).
18. P. C. Mandal, I. Mukherjee, G. Paul, B. N. Chatterji, *Digital image steganography: A literature survey*, Inf. Sci. (Ny)., **609**, 1451–1488, (2022).
19. A. A. Alsabhany, F. Ridzuan, A. H. Azni, *The Progressive Multilevel Embedding Method for Audio Steganography*, J. Phys. Conf. Ser., **1551**(1), 1-13, (2020).
20. M. Anwar, M. Sarosa, E. Rohadi, *Audio steganography using lifting wavelet transform and dynamic key*, Proceeding - 2019 Int. Conf. Artif. Intell. Inf. Technol. ICAIIT 2019, 133–137, (2019).
21. F. Djebbar, *Securing IoT Data Using Steganography: A Practical Implementation Approach,* Electronics, **10**(21), 2021.