MATEC Web of Conferences **402**, 02002 (2024) https://doi.org/10.1051/matecconf/202440202002
The 4th ICETIR 2022

Analyzing the impact of N-Frame value on audio watermarking

Kholidiyah Masykuroh [1*], *Zein* Hanni Pradana[1], *Indrarini* Dyah Irawati[2], *Gelar* Budiman[2] , and *Arfianto* Fahmi[2]

[1]Institut Teknologi Telkom Purwokerto, Jalan D.I. Panjaitan No.128, Purwokerto, Indonesia
[2]Telkom University, Jalan Telekomunikasi No. 1, Bandung, Indonesia

Abstract. Progress is being made in digitizing data from text, images, sound, and video. However, audio producers must invest in protecting digital rights and documenting authentication. One method for safeguarding message confidentiality is through watermarking. In this study, we employ the Lifting Wavelet Transform (LTW) technique and the Compressive Sensing (CS) technique, specifically Orthogonal Matching Pursuit (OMP), to propose a compressed watermark embedding technique for audio data. The performance analysis of the proposed system is based on the impact of the N-Frame value on SNR, ODG, C, and BER values. The results of the experiments indicate that increasing the number of N-Frames affects the SNR and C values, with higher SNR values and lower C values corresponding to an increased number of N-Frames.

1 Introduction

To combat unauthorized duplication and distribution of digital audio content, the digital audio industry employs audio watermarking as a means of copyright protection. Audio watermarking involves embedding imperceptible and unique identifiers within the audio data itself. These identifiers serve as digital signatures, enabling content owners to prove their ownership and track the authorized distribution of their digital audio files.

The implementation of audio watermarking technology has significantly contributed to the protection of intellectual property rights in the digital audio industry. By embedding watermarks within digital audio files, content creators and rights holders can effectively deter piracy and enforce copyright regulations.

One of the primary advantages of audio watermarking is its transparency. Unlike traditional forms of copyright protection, which may alter the audio quality or user experience, audio watermarking is designed to be imperceptible to the human ear. This ensures that the overall listening experience remains unaffected while still providing a robust mechanism for tracking the source of unauthorized distribution.

Furthermore, audio watermarking offers scalability and adaptability, supporting a wide range of audio formats and playback devices. This versatility ensures that protected content can be accessed and enjoyed by a broad audience while maintaining its copyright integrity.

While audio watermarking provides an effective means of copyright protection, content owners and industry stakeholders need to remain vigilant against emerging piracy techniques. Continual monitoring and advancements in watermarking technology are crucial to staying one step ahead of infringers.

Technological advancements in the digital audio industry have necessitated the implementation of copyright protection measures, such as audio watermarking. By embedding imperceptible and unique identifiers within digital audio files, content creators can safeguard their intellectual property rights and mitigate the impact of unauthorized duplication and distribution. The use of audio watermarking not only serves as a deterrent to piracy but also allows content owners to track and prove ownership of their digital audio content, thus fostering a sustainable and thriving digital audio industry.

Technological developments improve the quality of various forms of digital audio information. Additionally, the ease of access that enables the duplication and transfer of digital content constitutes an infringement of ownership rights, which is detrimental to the digital audio industry. Consequently, there is a need for regulation and supervision of these illegal activities. Therefore, audio watermarking is employed as a copyright protection measure for digital content [1].

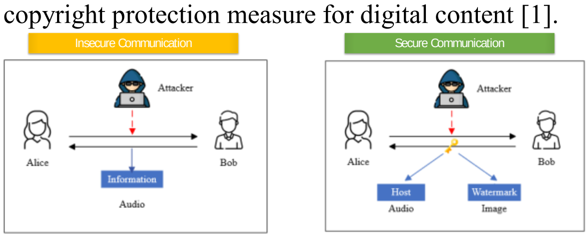

Fig. 1. Secure and Insecure Communication.

Watermarking methods in the context of copyright protection need to satisfy three specific criteria: imperceptibility, robustness, and payload. These criteria ensure that the watermarked content remains

* Corresponding author: kholidiyah@ittelkom-pwt.ac.id

unnoticeable to human perception, possesses resilience against various attacks, and allows for the embedding of additional information.

Audio watermarking involves inserting a watermark into an audio signal to safeguard it against unauthorized use. To achieve this, the audio watermarking process can be broken down into two distinct steps: embedding and extraction.

In the embedding step, additional content is inserted into the original audio. This additional content typically consists of the watermark itself, which contains unique identifiers or information related to copyright protection. The embedding process must be performed in a manner that minimizes any perceptible changes to the audio, ensuring that the watermark remains imperceptible to listeners.

Once the watermark has been successfully embedded, the extraction step comes into play. This step involves recovering the watermark from the watermarked audio signal. By extracting the watermark, it becomes possible to verify the authenticity of the audio and assert copyright ownership. The extraction process should be designed to reliably retrieve the watermark, even in the presence of various attacks, such as compression, filtering, or signal alteration.

Ensuring imperceptibility, robustness, and payload capacity is crucial for designing effective audio watermarking methods for copyright protection. By meeting these criteria, watermarking techniques can provide a powerful means of safeguarding valuable audio content from unauthorized use and piracy, thereby preserving the rights and interests of the content creators.

Watermarking methods for copyright protection must meet three criteria: imperceptibility, robustness, and payload. Audio watermarking involves inserting a watermark into the audio, making it difficult for different attacks to remove the watermark. The process of audio watermarking can be divided into two steps: (a) embedding, which involves inserting additional content into the original audio, and (b) extraction, which aims to recover the watermark [2].

Figure 1 illustrates how secure and insecure communication occurs. The information sent through communication media is essential for keeping the information safe during delivery. An example is Alice and Bob exchanging information in the form of audio through a telecommunications network. Data transmitted through communication channels becomes vulnerable if a third party seeks to eavesdrop on the audio conveying instructions from Alice to Bob. This underscores the importance of implementing authentication for the audio, providing peace of mind for Alice and Bob regarding the security and integrity of the exchanged information. This problem can be solved by adding a watermark to the audio that Alice sends to Bob or vice versa.

Several criteria determine the effectiveness of an audio watermarking system. The first criterion is imperceptibility, which refers to ensuring that there is no noticeable change in perception after inserting the watermark image into the audio host. The second criterion is robustness, which means that the watermarked audio can still be extracted reliably even after an intentional or unintentional attack. The third criterion is security, which ensures that only the intended party can extract the designed watermarked audio. The fourth criterion is capacity, which indicates the number of watermarks that can be inserted into the audio host. Lastly, the criterion of computation is designed to ensure that the computations involved are not complex. Figure 3 shows a general block diagram of audio watermarking [3]. The development of the audio watermarking technique utilizing the transformation domain method is carried out in this study, employing the Lifting Wavelet Transform (LWT).

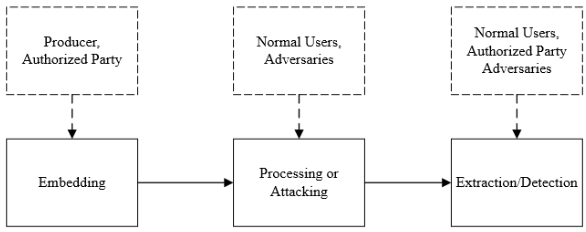

Fig. 2. A General Audio Watermarking Block Diagram.

Due to the wide range of power perception, research, research [4] shows that no other four human auditory systems (HAS) are more sensitive than the human visual system (HVS). Consequently, numerous audio watermarking techniques have been developed. These techniques can be classified into four types: time domain, transformation domain, compressed domain, and combined domain.

The study focuses on analysis the effects of the number of N-Frames used on parameters such as SNR, ODG, C, and BER. This paper consists of several sections including the introduction, literature review, research methods, results and discussion, and conclusion.

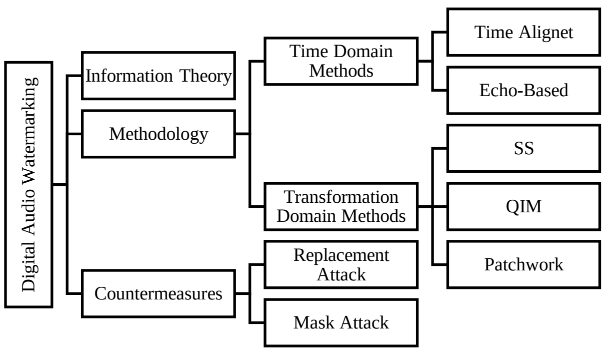

Fig. 3. Digital Audio Watermarking Structure.

2 Literature Review

A research paper [5] suggests an audio watermarking strategy based on singular value decomposition (SVD), exponential operation (EO), and logarithmic operation (LO) in the discrete cosine transform (DCT) domain. The simulation results demonstrate that the proposed method provides strong defense against various types of attacks including noise, cropping, resampling, re-quantization, and mp3 compression. The watermark is embedded at the significant singular value of the DCT sub-most tier, ensuring the fulfillment of this criterion. The objective of this study is to enhance robustness. Additionally, this study recommends the utilization of synchronization and error-correcting codes.

A synchronization code is used in research [6] to apply Singular Value Decomposition (SVD) in the Discrete Wavelet Transform (DWT) domain and create an audio watermarking system. Based on QIM, the watermark data is added to the singular value of the wavelet that represents the original audio signal. According to the experimental findings, it is impossible to distinguish between audio with and without a watermark. The proposed approach is also resistant to mp3 compression, cropping, low-pass filtering, additive noise, resampling, re-quantization, the insertion of echo, and denoising, as indicated by the test results. Therefore, the suggested technique is suitable for copyright protection applications. Moreover, the proposed algorithm is simple and does not require complex computations.

There are two types of digital watermarking: one involves temporal watermarking, and the other involves transformation watermarking. This temporal and transformation domain is typically achieved by modifying the host signal and then inserting a watermark into the coefficients of the new domain. Finally, an inverse transformation is applied to the modified coefficients to obtain a watermarked signal. Several transformations can be used with discrete signals, including discrete Fourier Transforms (DFT), discrete Cosine Transforms (DCT), and discrete Wavelet Transforms (DWT). Furthermore, several researchers have proposed a method of watermarking images based on Discrete Fractional Fourier Transforms (DFRFTs) and Discrete Fractional Order Random Transforms (DFRNTs). However, audio watermarking schemes employing fractional transforms have not been reported until now [7].

The transformation techniques FFT, DCT, and DWT are all applicable. The following text will explain some of these transformations for audio watermarking.

2.1 Discrete Transform for Audio Watermarking

When the lifting scheme is employed to generate integer-to-integer wavelets, the resulting coefficients become integers for all integer signals. This approach eliminates the need for scaling and directly converts the coefficients into binary values. However, this conversion process may introduce errors such as rounding and out-of-range errors. These errors can affect the accuracy and fidelity of the transformed audio signal.

To mitigate these issues, one must consider the handling of these errors during the conversion process. Techniques such as error correction algorithms and precision adjustments can be employed to minimize the impact of rounding errors. Furthermore, the range of the resulting coefficients should be closely monitored to prevent out-of-range errors.

By addressing these conversion errors, the lifting scheme enables the integration of binary values into the transformed audio signals. This is crucial for embedding secret information, such as the secret bit mentioned earlier, into the subsequent coefficients of the discrete wavelet transform. The ability to accurately encode and decode binary data within the transformed audio opens possibilities for various applications, including secure communication, watermarking, and data hiding.

It is important to note that the lifting scheme, while effective in generating integer-to-integer wavelets, must be implemented with precision and attention to detail. As with any data conversion process, the quality and fidelity of the transformed audio greatly relies on the proper handling of these conversions. Thus, it is crucial to ensure robust algorithms and techniques are in place to minimize errors and optimize the performance of the lifting scheme in practice.

The lifting scheme addresses the issue of non-integer coefficients in the discrete wavelet transform and enables the conversion of resulting coefficients to binary values. However, due to the inherent nature of the conversion process, errors, such as rounding and out-of-range errors, may arise. Proper error handling and precision adjustments are essential to mitigate these errors and maintain the accuracy and fidelity of the transformed audio signal.

In the frequency domain process, after transforming the DWT back to the original audio, the secret bit is inserted into the subsequent coefficients [8].

2.1.1 Discrete Wavelet Transform (DWT)

DWT signals are calculated using low-pass and high-pass filters. In the decomposition process, the primary signal is divided into two parts : a high-frequency component that emerges from the high-pass filter and a low-frequency component that emerges from the low-pass filter. The high-frequency and low-frequency sections are further subdivided into low-frequency sections [9].

2.1.2 Lifting Wavelet Transform (LWT)

The lifting scheme is a method used to create and perform discrete wavelet transforms (DWT). The issue with DWT arises when integer signals are used, as the resulting coefficients are not integers. In order to access the binary values of these coefficients, they need to be scaled and converted into binary. Employing a lifting scheme can resolve this problem by generating integer-to-integer wavelets. Consequently, for all integer signals, the resulting coefficients are integers. This scheme eliminates the need for scaling and enables direct conversion of the coefficients into binary. However, this process can lead to errors during the conversion, including rounding and out-of-range errors [10].

2.2 Analysed Parameters

ODG stands for Objective Difference Grade, which is a metric used to evaluate the perceptual difference between two audio signals. It measures how well a watermarking algorithm preserves the original audio quality. The higher the ODG value, the less perceptible the difference between the original and watermarked audio.

SNR refers to Signal-to-Noise Ratio, a parameter that quantifies the quality of a signal by comparing the power of the desired signal to the power of background noise. In audio watermarking, SNR is important because it represents the level of noise introduced by the watermarking process. A higher SNR indicates higher audio quality in the watermarked audio with less audible distortion.

BER, or Bit Error Rate, is a measure of the transmission errors in a digital communication system. It quantifies the percentage of bits that are incorrectly received compared to the total number of transmitted bits. In the context of audio watermarking, a lower BER indicates a more accurate transmission of the watermarked audio.

Finally, channel capacity refers to the maximum rate at which information can be transmitted through a channel with a negligible probability of error. It represents the upper limit of reliable data transmission and depends on various factors such as bandwidth, signal power, and noise level. For audio watermarking, channel capacity determines the maximum amount of data that can be embedded within the audio signal while maintaining an acceptable level of fidelity.

2.2.1 N-Frame

N-Frame or frame length is one of the input parameters. This value affects the framing process. The formula for determining the N-Frame value is [11] :

$$N - Frame = \frac{Host\ length}{Watermark\ length} \quad (1)$$

2.2.2 Objective Difference Grade (ODG)

A Perceptual Audio Quality Rating (PEAQ) is a measurement method that is part of the International Telecommunication Union-Radiocommunication Sector (ITU-R) standard BS.1387. Table I provides a detailed description of ODG [12].

Table 1. ITU-R Grade and ODG for Audio Quality Evaluation

Grade	ODG	Perceptual Description
5.0	0.0	Imperceptible
4.0	-1.0	Perceptible but not annoying
3.0	-2.0	Slightly annoying
2.0	-3.0	Annoying
1.0	-4.0	Very annoying

2.2.3 Signal to Noise Ratio (SNR)

The SNR parameter is used to maintain similarity between the unwatermarked and watermarked audio. SNR can be calculated using the following formula [8] :

$$\mathrm{SNR} = 10 \log_{10} \frac{\sum_{n=1}^{N} X^2(n)}{\sum_{n=1}^{N} [X(n) - X'(n)]} \quad (2)$$

Where X(n) is the host audio signal and X' (n) is the watermarked audio. A higher SNR value indicates louder audio is untraceable.

2.2.4 Bit Error Rate (BER)

BER is expressed as the number of bits degraded when transmitting digitized data. It represents the interference and noise contribution as a percentage. An estimate of the signal's robustness, implemented by BER, is given by

$$BER(X, X') = \frac{\sum_{n=1}^{N} X(n) \oplus X'(n)}{N} \quad (3)$$

Where the watermark range is denoted by N, the n-th bit of the embedded watermark is specified using X(n), and the n-th bit of the retrieved watermark is denoted by X' (n) increase [8].

2.2.5 Capacity (C)

Defining channel capacity involves determining the highest possible rate (measured in bits per channel) at which data can be transmitted with an extremely low likelihood of error. According to Shannon's second theorem, the capacity of an information channel is equivalent to the mutual information between two random variables, X and Y. The formula for calculating channel capacity is as follows [13] :

$$C = max_{p(x)} I(X; Y) \quad (4)$$

Two random variables, X and Y, exhibit mutual information as indicated by I (X;Y).

3 Research Methods

The embedding process of hiding the image in the audio involves three key steps: sorting as an insertion point, host domain conversion, and watermarking. These steps ensure that the audio is successfully watermarked with the hidden image. The sorting step determines the appropriate location to insert the image into the audio embedder. Following that, the host domain conversion process converts the audio samples into wavelet domains, which enables the successful embedding of the image.

To maintain the integrity of the audio, it is crucial that the embedded audio remains free of distortion. This is necessary to ensure that the secret data remains concealed within the audio file. Any distortion or alteration to the audio could compromise the integrity of the hidden image.

Once the audio is embedded with the secret data, the extraction process allows for the separation of the audio from the inserted watermark. This step enables the extraction of the hidden image from the audio embedding. By extracting the watermarked audio data, the original audio can be separated from the embedded watermark.

According to a study, watermarked audio has been shown to be resistant to attacks. This indicates that the hidden image remains intact within the audio and is not easily tampered with or compromised. The robustness of the watermarked audio ensures that the hidden image can be reliably extracted during the extraction process.

The algorithm for hiding images in audio consists of the embedding and extraction phases. The embedding process involves sorting, host domain conversion, and watermarking to successfully embed the image into the audio. To maintain the integrity of the hidden data, the audio must be free of distortion. During the extraction process, the watermarked audio data is separated to extract the hidden image. This algorithm, which utilizes wavelet domains, provides a reliable and robust method for hiding images in audio.

The following Figure 3 and Figure 4 explain the embedding and extraction process in audio watermarking [14].

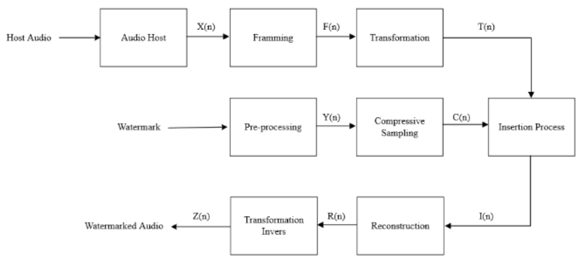

Fig. 4. Embedding Process.

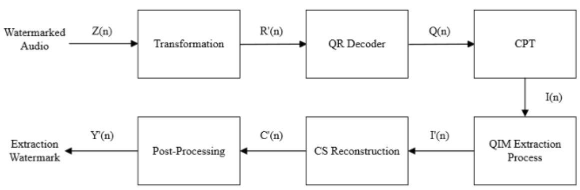

Fig. 5. Extraction Process.

4 Result and Discussion

Table II presents the results of the experiment conducted without any attacks, where the test was repeated seven times for each condition. The experiment considered input variables such as N, N Frame, N bits, Threshold, Scheme, and iscs. On the other hand, the experiment examined output variables including SNR, ODG, C, and BER. The specific value of N-Frame used for these experiments is 32.

Moving on to Table III, it also displays the outcomes of the experiment carried out without any attacks. Similar to the previous table, the test was repeated seven times for each condition. The input variables in this case consist of N, N Frame, N bits, Threshold, Scheme, and iscs. In contrast, the output variables examined are SNR, ODG, C, and BER. The designated N-Frame value used for this set of experiments is 64.

Lastly, Table IV illustrates the results obtained from the experiment conducted without any attacks. Similar to the previous tables, the test was repeated seven times for each condition. The input variables considered are N, N Frame, N bits, Threshold, Scheme, and iscs. Meanwhile, the output variables examined are SNR, ODG, C, and BER. The chosen N-Frame value utilized in these experiments is 128.

The tables provide essential insights into the effect of different N-Frame values on the SNR, ODG, C, and BER variables, enabling a comprehensive understanding of the outcomes of the experiment.

Table 2. The table presents the results of the experiment conducted without any attacks. The test was performed seven times for each condition. The input variables include N, N Frame, N bits, Threshold, Scheme, and iscs. On the other hand, the output variables consist of SNR, ODG, C, and BER. The N-Frame value provided is 32.

Parameter	**Trial I**	**Trial II**	**Trial III**	**Trial IV**	**Trial V**	**Trial VI**	**Trial VII**
N	1,00	2,00	3,00	4,00	5,00	6,00	7,00
N Frame	32,00	32,00	32,00	32,00	32,00	32,00	32,00
N bit	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Threshold	0,10	0,10	0,10	0,10	0,10	0,10	0,10
Skema	7,00	7,00	7,00	7,00	7,00	7,00	7,00
iscs	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Result							
SNR	7,62	10,27	15,30	17,91	21,26	21,49	20,90
ODG	-3,08	-3,58	-2,59	-1,88	-1,08	-2,11	-2,20
C	1435,55	717,77	358,89	179,44	89,72	44,86	22,43
BER	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Table 3. The table presents the results of the experiment conducted without any attacks. The test was repeated seven times under each condition. The input variables include N, N Frame, N bits, Threshold, Scheme, and iscs. On the other hand, the output variables consist of SNR, ODG, C, and BER. The specified value for N-Frame is 64.

Parameter	**Trial I**	**Trial II**	**Trial III**	**Trial IV**	**Trial V**	**Trial VI**	**Trial VII**
N	1,00	2,00	3,00	4,00	5,00	6,00	7,00
N Frame	64,00	64,00	64,00	64,00	64,00	64,00	64,00
N bit	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Threshold	0,10	0,10	0,10	0,10	0,10	0,10	0,10
Skema	7,00	7,00	7,00	7,00	7,00	7,00	7,00
iscs	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Result							
SNR	13,46	15,28	19,13	24,44	27,00	26,59	24,10
ODG	-3,10	-3,10	-1,70	-0,48	-0,16	-1,04	-1,33
C	717,77	358,89	179,44	89,72	44,86	22,43	11,22
BER	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Table 4. The table presents the results of the experiment without any attacks being applied. The test was conducted seven times under each condition. The input variables include N, N Frame, N bits, Threshold, Scheme, and iscs. On the other hand, the output variables consist of SNR, ODG, C, and BER. The value assigned to N-Frame is 128.

Parameter	Trial I	Trial II	Trial III	Trial IV	Trial V	Trial VI	Trial VII
N	1,00	2,00	3,00	4,00	5,00	6,00	7,00
N Frame	128,00	128,00	128,00	128,00	128,00	128,00	128,00
N bit	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Threshold	0,10	0,10	0,10	0,10	0,10	0,10	0,10
Skema	7,00	7,00	7,00	7,00	7,00	7,00	7,00
iscs	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Result							
SNR	18,73	21,42	15,30	27,80	32,55	32,09	25,60
ODG	-2,27	-0,43	-2,59	-0,27	-0,02	-0,41	-1,15
C	358,89	179,44	358,89	44,86	22,43	11,22	5,61
BER	0,00	0,00	0,00	0,00	0,00	0,00	0,00

5 Conclusions

The findings from the experiments indicate that the performance of the proposed LWT and CS-based audio watermarking algorithms is commendable. The influence of the number of N-Frames on the SNR and C values is evident. As the number of N-Frames increases, the SNR values consistently rise. On the other hand, in a series of seven trials, the C value tends to decrease as the number of N-Frames increases.

Further research is warranted to gain a comprehensive understanding of the system's performance and investigate the impact of interference and compression. These factors may significantly affect the effectiveness of the audio watermarking algorithms. By exploring the influence of interference and compression, researchers can enhance and optimize the system's performance in real-world scenarios.

Additionally, another area worth exploring is audio watermarking using fractional transformations. This approach holds promise in terms of providing robust and efficient audio watermarking techniques. Conducting a study on audio watermarking using fractional transformations can potentially lead to innovative methods for embedding and extracting watermarks in audio signals.

By expanding the scope of research to include the impact of interference and compression, as well as exploring new avenues such as audio watermarking using fractional transformations, the field of audio watermarking can continue to evolve, resulting in improved algorithms and techniques. These advancements will ultimately contribute to enhancing security and integrity in various audio applications.

This study is a collaboration between Telkom Institute of Technology Purwokerto and Telkom University. The author expresses gratitude to the institutions for their financial support in conducting this research.

References

1. D. S, R. A, and H. S, "Data Hiding in Audio Signals Using Wavelet Transform with Enhanced Security," in *Computer Science & Information Technology (CS & IT)*, Academy & Industry Research Collaboration Center (AIRCC), Dec. 2013, pp. 137–146. doi: 10.5121/csit.2013.3914.
2. Y. Xiang, I. Natgunanathan, D. Peng, G. Hua, and B. Liu, "Spread Spectrum Audio Watermarking Using Multiple Orthogonal PN Sequences and Variable Embedding Strengths and Polarities," *IEEEACM Trans. Audio Speech Lang. Process.*, vol. 26, no. 3, pp. 529–539, Mar. 2018, doi: 10.1109/TASLP.2017.2782487.
3. G. Hua, J. Huang, Y. Q. Shi, J. Goh, and V. L. L. Thing, "Twenty years of digital audio watermarking—a comprehensive review," *Signal Process.*, vol. 128, pp. 222–242, Nov. 2016, doi: 10.1016/j.sigpro.2016.04.005.
4. S. P. Singh Chauhan and S. A. M. Rizvi, "A survey: Digital audio watermarking techniques and applications," in *2013 4th International Conference on Computer and Communication Technology (ICCCT)*, India: IEEE, Sep. 2013, pp. 185–192. doi: 10.1109/ICCCT.2013.6749625.
5. P. K. Dhar and T. Shimamura, "Blind Audio Watermarking in Transform Domain Based on Singular Value Decomposition and Exponential-Log Operations," *Radioengineering*, vol. 26, no. 2, pp. 552–561, Jun. 2017, doi: 10.13164/re.2017.0552.
6. V. Bhat K, I. Sengupta, and A. Das, "An adaptive audio watermarking based on the singular value decomposition in the wavelet domain," *Digit. Signal Process.*, vol. 20, no. 6, pp. 1547–1558, Dec. 2010, doi: 10.1016/j.dsp.2010.02.006.
7. M. Fan and H. Wang, "Chaos-based discrete fractional Sine transform domain audio watermarking scheme," *Comput. Electr. Eng.*, vol. 35, no. 3, pp. 506–516, May 2009, doi: 10.1016/j.compeleceng.2008.12.004.
8. U. Nair and G. K. Birajdar, "Compressed domain secure, robust and high-capacity audio watermarking," *Iran J. Comput. Sci.*, vol. 3, no. 4, pp. 217–232, Dec. 2020, doi: 10.1007/s42044-020-00059-x.
9. A. Al-Haj, A. Mohammad, and L. Bata, "DWT–Based Audio Watermarking," vol. 8, no. 3, p. 8, 2011.
10. N. Trivedi, V. Kumar, S. Singh, S. Ahuja, and R. Chadha, "Speech Recognition by Wavelet Analysis," *Int. J. Comput. Appl.*, vol. 15, no. 8, pp. 27–32, Feb. 2011, doi: 10.5120/1968-2635.
11. L. Novamizanti, G. Budiman, and B. A. Wibowo, "Optimasi Sistem Penyembunyian Data pada Audio menggunakan Sub-band Stasioner dan Manipulasi Rata-rata Statistik," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 6, no. 2, p. 165, Jul. 2018, doi: 10.26760/elkomika.v6i2.165.
12. M.-J. Hwang, J. Lee, M. Lee, and H.-G. Kang, "SVD-Based Adaptive QIM Watermarking on

Stereo Audio Signals," *IEEE Trans. Multimed.*, vol. 20, no. 1, pp. 45–54, Jan. 2018, doi: 10.1109/TMM.2017.2721642.

13. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc, 2006.
14. I. D. Irawati, G. Budiman, K. Masykuroh, Z. H. Pradana, and A. Fahmi, "High Payload Qr-Based Data Hiding Using Secured Compressed Watermark in Polar Domain," in *Frontiers in Artificial Intelligence and Applications*, A. J. Tallón-Ballesteros, Ed., IOS Press, 2021. doi: 10.3233/FAIA210285.