Dual Information Audio Watermarking with Modified Wavelet Based LSB Technique

Nikhila Arava, A. Bhuvaneshwari*, Hajira Fathima

***Abstract*— Audio steganography has gained significant importance in the recent years for transmitting secret information embedded in the audio files. The existing algorithms in audio steganography aim to improve the security of the embedded data and preserve the authenticity of the audio signal with focus on hiding a single type of information. Limited work has been done to embed varied information in audio files. The highlight of the paper is the proposed Modified DWT based LSB (MDWT-LSB) algorithm that conceals dual information namely, text and image in the same audio file. Initially, three major audio steganography algorithms, namely Least Significant Bit (LSB), Echo Hiding (EH), and Spread Spectrum (SS), are implemented to hide only text information. The performances are compared in terms of Mean Square Error, Signal to Noise ratio, Peak Signal to Noise ratio and Speech Power. The subjective evaluations are done on stego audio signals by plotting histograms, Power Spectral Density and Correlation Plots. The results show that LSB algorithm gives good results compared to the other techniques. As an improvement, a new Modified DWT based LSB (MDWT-LSB) algorithm is proposed to hide dual information, namely text and image, in the audio file. Analyzing the results, it is inferred that the MDWT-LSB algorithm gives the maximum PSNR (105.1001 dB) compared to the existing LSB technique with the advantage of embedding text and image in the audio. The subjective evaluations and computed metrics justify the improved performance of the proposed MDWT-LSB technique to efficiently hide varied secret data.**

***Index Terms*— Audio Steganography, Least Square Bit, Spread spectrum, Modified DWT based LSB.**

I. Introduction

THE technological developments have resulted in an enormous amount of multimedia information, such as text, image, audio, and video files, shared in real-time. Audio steganography is a method of transmitting hidden information in the digital audio. The secure transmission of the hidden message has become a primary concern in the recent years. Basically, two main methods are being used to secure the data, namely steganography and cryptography. Steganography secures the information, whereas cryptography transforms the message into an undetectable format. Steganography conceals the secret information in a cover file in such a way that only the sender and receiver are aware of its presence [1]. In audio Steganography, the secret messages are hidden in .wav and MP3 sound files causing the binary sequence of associated audio files to change [2]. It initiates the identity of information by embedding it into the digital file in such a way that its removal is hardly possible, thereby preventing its unauthorized use [3]. Extensive steganographic methods are devised to simultaneously enhance security, quality of the stego file, and embedding capacity [4]. The data was hidden in smooth blocks of different sizes employing mean and additive modulus [5]. All these methods describe the embedding of a single type of information in different forms, while hiding varied data in the audio cover has been less frequently explored. This aspect is focused in the paper and a new Modified DWT based LSB (MDWT-LSB) algorithm is proposed for embedding text and images in the audio file. Initially, the basic algorithms namely Echo Hiding (EH), Spread Spectrum (SS), Improved SS and Least Significant Bit (LSB) are implemented to embed text data. This is followed by the proposed method to embed dual information in the audio file. The generalized methodology of audio steganography is given in section II and steganographic algorithms, along with the proposed MDWT-LSB technique, are elaborated in section III. The implementation results and the evaluated metrics are analyzed in section IV, and section V gives the conclusion.

Manuscript received January 10, 2023; revised May 02, 2023, Nikhila Arava is student of BE, Department of Instrumentation, Dayananda Sagar College of Engineering, Bangalore-560078, Karnataka, India. (e-mail: aravanikhila@gmail.com).

*Corresponding author: Dr A. Bhuvaneshwari is an Associate Professor in the Department of ECE, ME (Digital Systems) at Deccan College of Engineering and Technology, Hyderabad -500001, India. (e-mail: sudha_bhuvana@yahoo.com)

Hajira Fathima is student in the Department of ECE, ME (Digital Systems) at Deccan College of Engineering and Technology, Hyderabad -500001, India. (e-mail: meherarif07@gmail.com)

II. Audio Watermarking

Steganographic methods are used in both steganography and digital watermarking to insert data in a concealed manner. The top priority in steganography is the imperceptibility to human senses and concealing the information, while digital watermarking prioritizes robustness. The basic audio watermarking model is shown in Fig 1 [6].

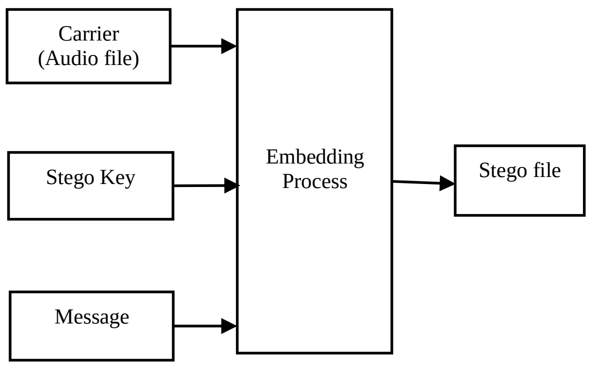

Fig. 1. Basic model of Audio Watermarking

The general model consists of a carrier file, a secret message to be embedded, and a password referred to as the stego-key. The carrier is the cover-file since it conceals

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

confidential information [7]. The secret message can be plain text, a picture, an audio file, or any other file type. A stego-key is a password that assures that only the recipient who knows the appropriate decoding key can extract the message from a cover-file. A stego-file is obtained after embedding the secret message in the carrier file. In audio steganography the hidden message is carried with the audio with a minimum detectable change. Since the small changes in amplitude of the signal can store a huge amount of information, and as audio files are larger than images, audio steganography has more potential to conceal information.

III. AUDIO WATERMARKING ALGORITHMS

The audio watermarking mainly consists of two main functions; integrating the watermark information directly in the audio and its extraction in the same form. It implies that text, image or any data can be embedded within the audio file so that the watermarked audio cannot be differentiated from the host signal. The embedding watermark process is done at the source end while the extraction from the watermarked audio is done by reversing the embedding algorithm. Three major watermarking algorithms namely Echo hiding, Spread spectrum and Least Significant Bit techniques are implemented for embedding text in the audio. The chosen methods best preserve the quality of the host audio signal. They are robust to any manipulation and effective against lossy audio compression with MP3. The LSB method is modified to increase the embedding capacity to conceal dual data; image along with the text information.

A. *Echo Hiding Algorithm*

In the Echo hiding technique a short echo is introduced in the host signal without leading to any change in statistical and perception qualities of the stego audio. The initial amplitude, delay and decay rate of the echo signal are altered for data hiding [8]. The echo hiding is simple, easy to implement and generates low noise but is costly in steganalysis [9]. The flowchart for embedding the text in the audio file using the Echo Hiding embedding algorithm is shown in Figure 2 and the steps of EH method is summarized [10][11].

Step 1. The cover audio is divided into segments. Each segment is echoed with the delay corresponding to the data bit to be encoded.

Step 2. The number of bits to be hidden are 'N' and 'L' represents the length of segments. The length is chosen such that is not greater than length of the audio signal.

$$(N \times L) < \text{length of audio signal} \quad (1)$$

Step 3. A mixer signal is generated using data bits and a smoothed signal ensures the distortion between adjacent segments is reduced.

Step 4. The echoed signals are filtered with mixer signal before adding up onto cover audio signal.

Step 5. The delayed signal samples are represented as k_1 and k_2, the cover audio is 's' and 'x' is the stego audio signal. The embedding process is written as

$$x = s + k_1 . mixer + k_2(1 - mixer) \quad (2)$$

Step 6. The smoothed mixer signal is generated using the data bits to embed.

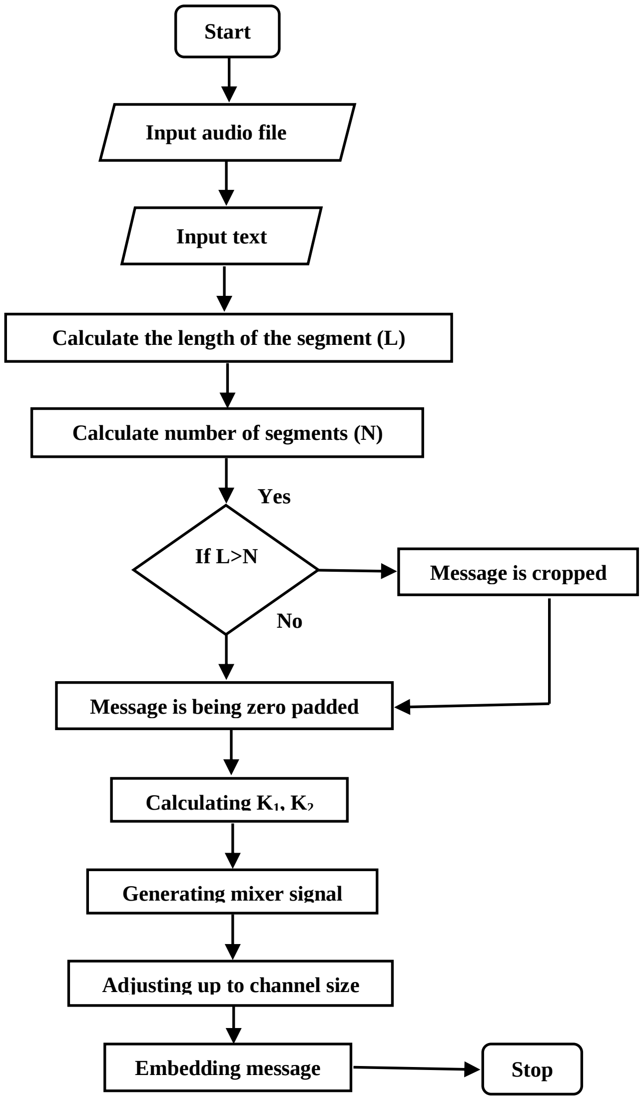

Fig. 2. Echo Hiding Embedding Algorithm

Step 7. The echoes are thus added onto the cover audio after filtering with mixer signal.

In the decoding process, Cepstrum analysis is used in Echo Hiding methods [12]. The cepstrum function detects the presence of echo signal and then offsets the delay if any echoes are added [12]

$$Cn[i] = ifft(\log(\text{abs}(\text{fft}(sn[i])))) \quad (3)$$

The steps for extracting the text from the stego audio using Echo Hiding algorithm are summarized.

Step 1. The stego audio signal is the input used for retrieving the hidden message.

Step 2. The number of frames 'N' is computed by dividing the length of the signal with the length of the segment.

$$N = length(signal)/L \quad (4)$$

Step 3. Reshape the stego audio signal by dividing the signal into frames.

Step 4. Real Cepstrum is performed which detects the existence of echo signal and then offset delay of echo is added.

Step 5. The recovered message is scaled and the watermark text is retrieved.

Echo hiding methods are robust that even when extra information is embedded, hidden data is not lost. The data losses are observed for segments with silent points. In order to drop the error rate, silent segments can be forced by increasing their energy or these segments can be skipped by detecting before embedding. Due to reduced embedding rate and low security, the conventional Echo hiding method has limited application.

B. Spread Spectrum Algorithm

The spread spectrum technique uses the principle of Direct Sequence Spread spectrum (DSSS), to spread the coded secret data across the audio signal thereby increasing the bandwidth. The watermark technique devised on images is applicable for audio data [13]. The flowchart for embedding the text in the audio file using Spread Spectrum is in Figure 3 and steps are summarized.

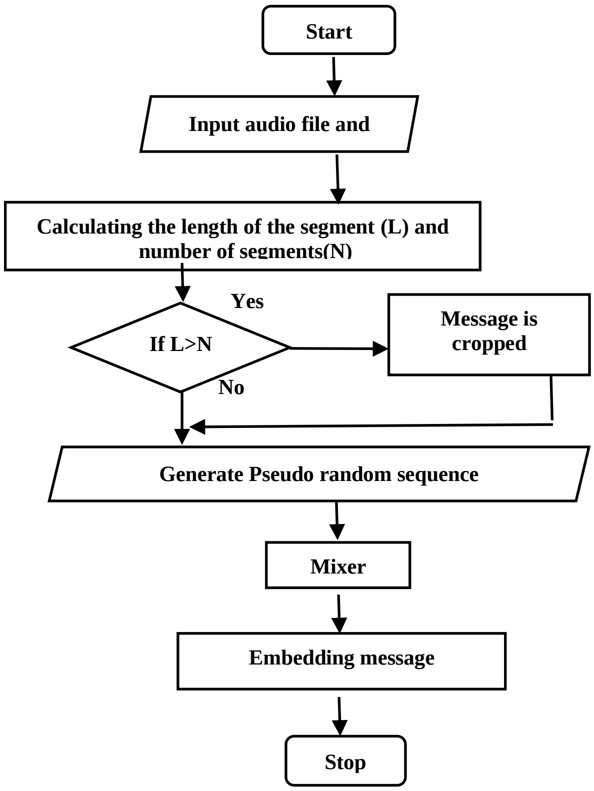

Fig. 3. Spread Spectrum Embedding Algorithm

Step 1. The inputs are; the cover audio file divided into segments, text to be embedded, and the key signal of Pseudo-random Sequence (PN). The PN sequence selected for hiding the text has the same data rate as the original signal.
Step 2. Convert the encrypted text message into bit-sequence information. Represent this sequence using -1 and 1, with 0s mapped to - 1 and 1s mapped to 1 itself.

$$w = \{wi\}, wi\ \{-1,1) \quad (5)$$

Step 3. The number of bits to be hidden (N) and length of the segment (L) are chosen such that it does not exceed the length of the audio signal.

$$N \times L < length\ of\ the\ audio\ signal \quad (6)$$

Step 4. If the length of the bits is greater than the number of segments, the message gets cropped. A smoothed mixer signal is generated to reduce the distortion between the segments.

Step 5. In the embedding process the amplitudes of the cover audio signal are multiplied by the PN sequence and message bits on a sample-by-sample basis. The embedded cover audio results in the stego file.

Improved Spread Spectrum

The Improved Spread Spectrum (ISS) technique is redefined by a slight modification in embedding process by authors Malvar, H.S. and Florêncio, D.A [14]. The steps 1 to 4 remain the same as in Spread Spectrum (SS). But the ISS varies the amplitude of the inserted chip sequence by a function $(x, b)u$ as per Equation 7. The secret text message is embedded and the stego audio signal 's' is given as [14]

$$s = x + (x, b)u \quad (7)$$

Where 'x' is the host audio signal, 'b' is the number of bits and 'u' is the chip sequence.

The steps for extraction of text from the audio file for SS and ISS algorithms are summarized.
Step 1. Load the stego file.

$$y = {\sim}x + w{\sim} \quad (8)$$

Where y = stego signal, $\sim x$ = host signal, $w\sim$ = watermark
Step 2. Length of the segment 'L' and no. of segments 'N i' are calculated.
Step 3. The same pseudo-random sequence is generated as done in the embedding process.
Step 4. The recovered watermark is scaled by converting the message into double.
Step 5 . The recovered message is displayed after reshaping the data.
The ISS has limited distortion compared to the original method. The detection method remains the same for ISS and is applicable for most watermarking schemes.

C. Least Significant Bit Algorithm

The Least significant Bit (LSB) is the simplest and preferred method for embedding data in the audio cover. In this, the least significant bits of the cover samples are used for inserting the watermark [15]. The basic LSB method to embed text message in the audio cover is shown in the Figure 4 and the steps are described.

Step 1. The audio file and the message (".txt") to be hidden are taken as inputs.
Step 2. The size of the audio signal and message text are determined for embedding. The length of the message to be hidden must be smaller than the total samples in the audio file.
Step 3. The string value of the message to be hidden is converted to double. The binary equivalents of the decimal samples are obtained for carrier and message.
Step 4. In the embedding process the Least Significant Bits of the cover audio file (i,j) is replaced by the LSB of the watermark.
Step 5. The secret message is encoded and the stego audio is obtained.
The watermarked audio shows unnoticeable degradation. The LSB method is more suitable in noise free scenarios with the simple embedding of the message bits in the audio samples.

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

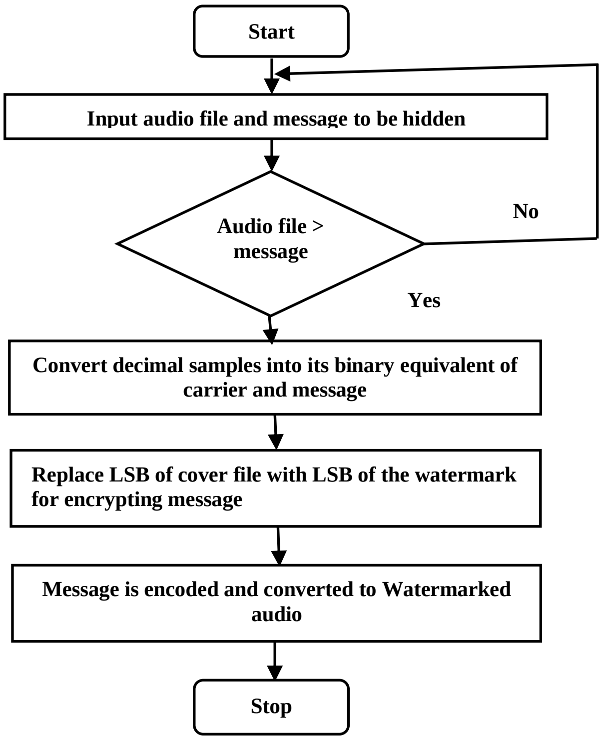

Fig 4. LSB Text Embedding Algorithm

In the extraction process the watermark message is recovered exactly using the LSB extraction algorithm. The steps are summarized below.

Step 1. Read the Stego audio signal to extract the hidden message from the LSB encoded sound file.

Step 2. The receiver needs to access the sequence of sample indices employed in the embedding process.

Step 3. The encrypted message characters are decrypted from the LSB of the stego audio.

Step 4. The recovered information is scaled by converting the message from binary to decimal.

Step 5. The extracted message is displayed after reshaping the data.

The main advantage of LSB technique is that it provides high perceptual transparency and is a safe way for the transmission of text [16]. There is no change in the size of the audio file after encoding and the scheme is well suited for different audio file formats.

D. Modified DWT based LSB (MDWT-LSB) algorithm

The section describes a modified LSB algorithm that embeds dual information namely text and image in the audio file. In the proposed Modified DWT- LSB (MDWT-LSB) algorithm, the image to be hidden is transformed from spatial to frequency domain by applying Discrete Wavelet Transform (DWT). The transformation decomposes the image into four non-overlapping sub-bands. The secret text is embedded in the sub band coefficients of the image data. The image coefficients are reconstructed, and dual information (text and image) are embedded into the audio signal to generate the watermarked host. The Discrete Wavelet Transform is used in the watermarking process since it provides a compact representation of the signal in time and frequency domain in addition to its computation efficiency [17]. The flowchart for the Modified DWT-based LSB algorithm is shown in Figure 5 and steps are summarized.

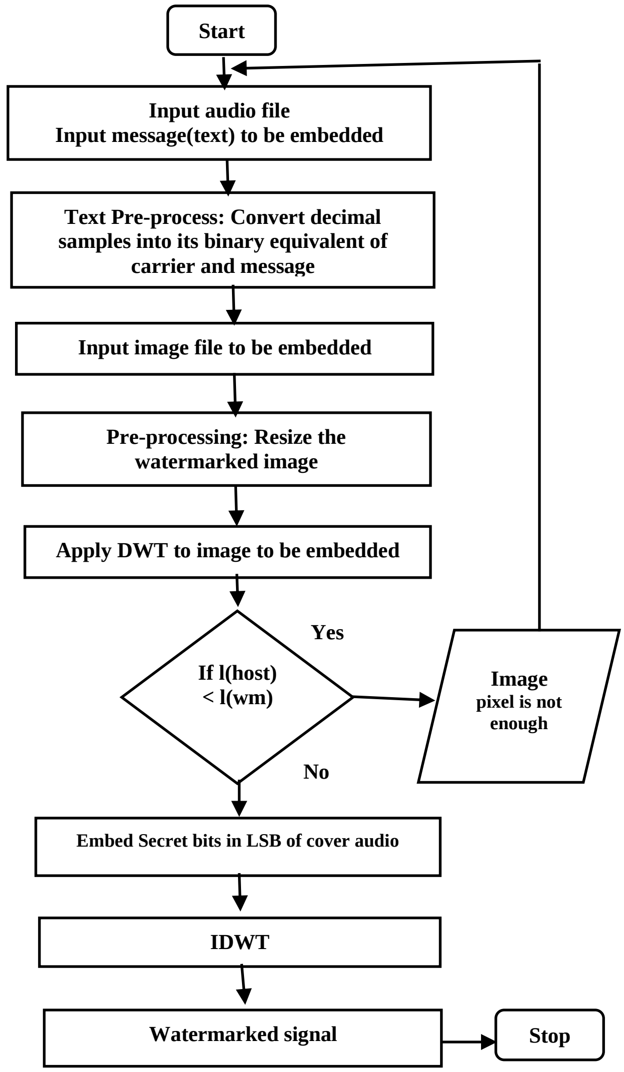

Fig. 5. Modified DWT- LSB Embedding Algorithm

Step 1. The input is the audio file of any format which is used to hide both message and image data.

Step 2. In the next step suitable preprocessing is done for the message to be hidden by converting decimal samples into binary equivalent.

Step 3. The image to be embedded in the audio file can be of any format (".jpg", ".png", .tif").

Step 4. The audio signal and watermark image are resized prior to embedding and the length of the watermark is determined.

Step 5. Apply Discrete Wavelet Transform (DWT) to decompose the image into four sub bands. Embed the LSB of the watermark (i,j) to the value of the LSB of the cover signal (i,j).

Step 6. The embedding is done in the transform domain. This is followed by applying Inverse Discrete Wavelet Transform (IDWT) to get the stego audio signal which hides image and text simultaneously.

The steps for extracting the watermark image from the audio file using the LSB algorithm are summarized.

Step 1. The watermarked stego audio is read. The post processing is done and the size of the watermarked image is determined.

Step 2. The encrypted message characters are decrypted using the extracted message size.

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Step 3. The LSB of the audio signal is used to recover the watermark from the first plane of the host image.
Step 4. The size is represented in binary and the number of ones in it is gives the count. The recovered watermark image is scaled by converting it from binary to decimal.
Step 5. The watermark image is reshaped and the extracted image is displayed.

E. Performance Metrics

The audio steganography evaluation includes various indexes such as steganography capacity, undetectability, imperceptibility, robustness and security. The performances of the implemented algorithms are compared with subjective and objective evaluations.

1) Objective evaluation

In the objective evaluation the quality of the watermarked stego audio is measured using the metrics Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR) and Speech Power (SP)[18].

a) Mean Square Error

In audio steganography, the Mean Square Error (MSE) is computed as the average squared difference between the host audio and the stego audio. It is given by the Equation 9.

$$MSE = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}(P(x,y) - q(x,y))^2 \quad (9)$$

where P(x,y) is the host signal and q(x,y) is the watermarked signal. A smaller value of MSE indicates reduced error.

b) Signal-to-Noise Ratio

The Signal to Noise Ratio (SNR) refers to the level of the audio signal as compared to the noise level present in that signal. It is given by the Equation 10.

$$SNR = 10\log_{10}\left\{\frac{\sum_{n=0}^{N-1}x^2(n)}{\sum_{n=0}^{N-1}[\tilde{x}(n) - x(n)^2]}\right. \quad (10)$$

where x(n) is the host signal and watermarked signal is $\tilde{x}(n)$ .The larger SNR value indicates better quality of the signal which suggests there is more useful information than unwanted data.

c) Peak Signal-to-Noise Ratio

The Peak Signal-to-Noise Ratio (PSNR) is a ratio used as a quality measurement between the host and stego signals. It is measured in decibels and is given by the Equation 11.

$$PSNR = 20\log_{10}[max/(MSE)^{1/2} \quad (11)$$

There is an inverse relation between MSE and PSNR. The quality of signal is better if the PSNR ratio is high.

d) Speech Power

The standard way to analyse the quality of audio signal is the perception capability which is expressed in terms of speech power. It is the value of mean opinion score calculated after normalization. The values range in between 1.0 and 4.5.

$$Speech\ power = Sum(x(n))^2 \quad (12)$$
$$Speech\ Power = Sum(\tilde{x}(n))^2 \quad (13)$$

where x(n) is the original signal and $\tilde{x}(n)$ is the watermarked signal. The perceptual evaluation of audio quality is used to measure the decline of audio, and its output objective difference grade ranges from very annoying speech (-4) to Imperceptible (0).

2) Subjective Evaluation

In the subjective evaluation, Power Spectral Density (PSD), Histogram and Cross correlation plots are obtained for the watermarked audio signal. The plots are shown for the modified DWT based LSB (MDWT-LSB) technique since it gives the best performance compared to other methods. The cross correlation is plotted between the host and stego audio signal. The host audio and overlapped watermarked audio signals are depicted for Echo Hiding, Spread Spectrum, Improved Spread Spectrum, LSB and MDWT-LSB algorithms.

IV. RESULTS AND DISCUSSION

Audio watermarking offers an efficient method to hide information in a more secure way. The work aims to design an effective audio watermarking technique to hide dual information with minimized risk. The goal is achieved when a maximum PSNR and minimum error are achieved. The original host audio signal is a .wav file in which the information is embedded. The host audio signal used as a cover to embed the secret message is shown in Figure 6. The size of the audio signal is 1293kb and the specifications are summarized in Table I.

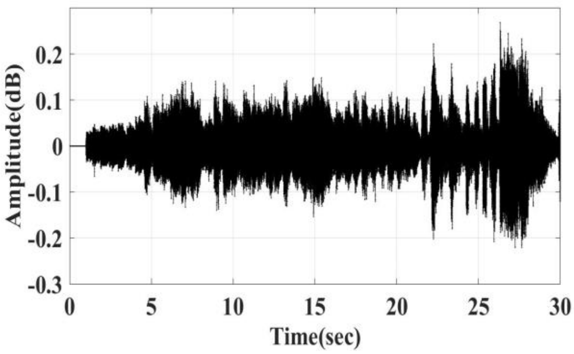

Fig 6. Host audio signal

Table I: Specifications

Parameters	Values
Number of Channels	1
Sample Rate	22050
Total Samples	661500
Duration	30
Bits per Sample	16

The efficiency of the algorithms is evaluated with parameters MSE, SNR, PSNR and speech power. The metrics obtained by EH, SS, ISS, and LSB algorithms for hiding the hidden are compared.

A. Audio watermarking using Echo Hiding

The Figure 7 shows the stego audio signal after embedding the watermark text-1 into the host audio signal. The embedding performance of EH algorithm is evaluated on two different text files (text-1 and text-2). The first file has the message (Text) and second file has the information

(Message). The size of the text file is 1kB. The Stego audio signal is generated after embedding the text information as in Figure 7. It is observed that the original host audio signal represented by the markers and the stego audio are almost similar; with a slight difference between the two signals.

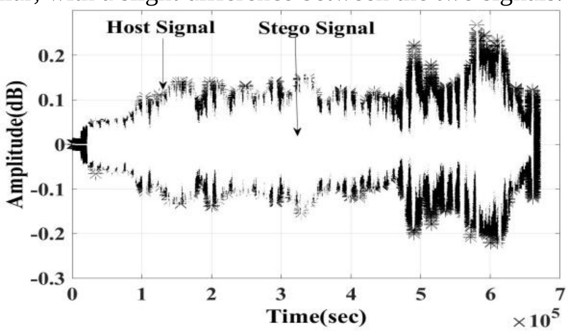

Fig. 7 Overlapped Host and Stego audio signal using Echo Hiding

The evaluated metrics of the stego signal with text-1(text) and text-2 (message) hidden using EH algorithm is given in Table II.

Table II: Metrics of stego audio using EH algorithm

Metrics (dB)	Hidden text-1 (text)	Hidden text-2 (message)
MSE	0.01863	0.018628
SNR	5.5252	5.6289
PSNR	82.7605	82.7617
Speech Power	1181.301	1175.207

As the length of the secret message increases more bytes must be embedded in the cover audio resulting in the decrease of the speech power. Comparing the speech power results of hidden text-1 and text-2, it is observed that it is lesser for text-2 due to increased number of bits in it.

B. Audio Watermarking using SS and ISS Algorithm

The Spread spectrum spreads the input data signal over the channel by multiplying it with a pseudo-random sequence. Figure 8 shows the graph of Spread Spectrum with data bits for text-1. The Figure 8 shows the host audio signal depicted with a marker super imposed on the watermarked audio signal . A slight difference between the two signals can be observed at the edges of the waveform.

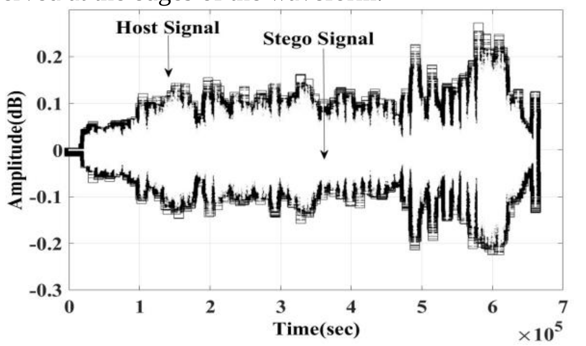

Fig. 8. Overlapped Host and Stego audio signal using SS algorithm

The evaluated metrics of the stego signal with text-1 and text-2 hidden using SS algorithm are given in Table III.

Table III: Metrics of stego audio using SS algorithm

Metrics	Hidden text-1 (text)	Hidden text-2 (message)
MSE	0.00499	0.00499
SNR	17.3746	17.7283
PSNR	94.1896	94.2004
Speech Power	939.6414	938.3203

The Improved Spread Spectrum (ISS) embedding is implemented and the results are summarised. The Figure 9 shows the overlapped host signal and the stego signal when text-1 is hidden in the cover audio. A slight variation between the two signals is observed at the envelope of the waveform.

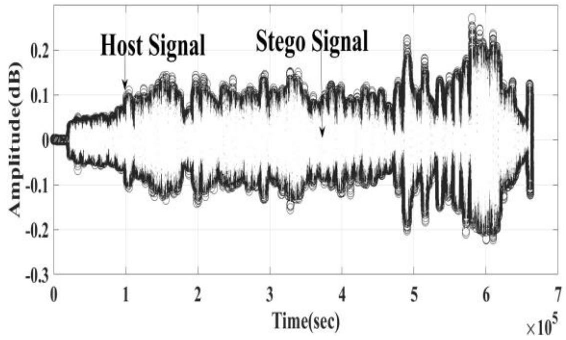

Fig. 9. Overlapped Host and Stego audio signal using ISS algorithm

The evaluated metrics of the stego signal with text-1 and text-2 hidden using the ISS algorithm is given in Table IV.

Table IV: Metrics of stego audio using ISS Algorithm

Metrics (dB)	Hidden text-1 (text)	Hidden text-2 (message)
MSE	0.0048	0.004967
SNR	17.7541	18.0171
PSNR	94.5191	94.2431
Speech Power	938.2279	937.3185

The comparison metrics for Spread Spectrum and Improved Spread Spectrum algorithms are shown in Table V. It is observed that a higher PSNR is obtained with ISS. This can be attributed to the fact that Improved SS embedding is not sensitive to amplitude scaling as in conventional Spread spectrum.

Table V: Comparison of SS and ISS

Metrics (dB)	Hidden text-1 (text)		Hidden text-2 (message)	
	SS	ISS	SS	ISS
MSE	0.004997	0.0048115	0.0049914	0.004967
SNR	17.3746	17.7541	17.7283	18.0171
PSNR	94.1896	94.5191	94.2004	94.2431
Speech Power	939.6414	938.2279	938.3203	937.3185

C. Audio Watermarking using LSB Algorithm

The LSB is a spatial domain technique and is preferred due to its increased hiding capacity compared to the transform domain methods. The Figure 10 depicts the audio signal obtained after embedding text-1 using LSB algorithm.

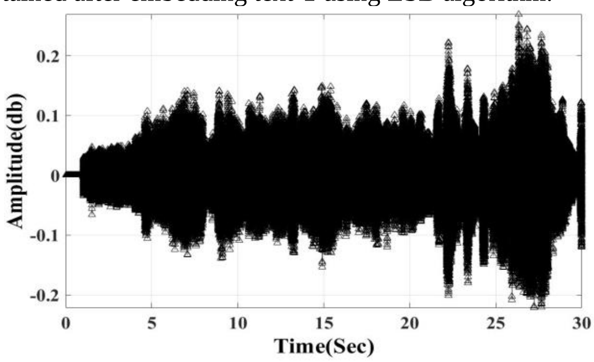

Fig. 10. Stego audio signal using LSB algorithm

The Figure 11 shows overlapped host signal and stego signal using LSB algorithm with text-1 embedded in it.

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

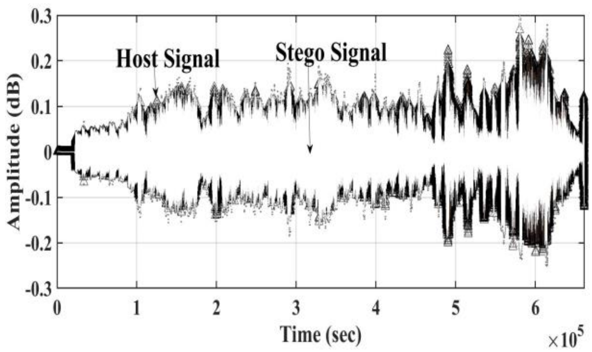

Fig.11. Overlapped Host and Stego audio signal using LSB algorithm

The evaluated metrics of the stego audio signal with text message hidden using LSB algorithm are given in Table VI.

Table VI: Metrics of stego audio using LSB algorithm

Metrics (dB)	Hidden text-1 (text)	Hidden text-2 (message)
MSE	2.517e-07	2.7057e-07
SNR	92.7085	93.5250
PSNR	180.1471	179.5191
Speech Power	922.7514	922.7514

D. Comparison of audio steganography algorithms

Digital audio watermarking techniques are implemented using different algorithms. The performances are compared with MSE, SNR and PSNR on two different data sets. The metrics with text-1 and text-2 hidden are shown in Tables VII and VIII.

Table VII: Comparison of steganography algorithms with text-1 hidden

Metrics (dB)	EH	SS	ISS	LSB
MSE	0.0186	0.004997	0.004811	2.517e-07
SNR	5.5252	17.3746	17.7541	92.7085
PSNR	82.760	94.1896	94.5191	180.147

Table VIII: Comparison of steganography algorithms with text-2 hidden

Metrics (dB)	EH	SS	ISS	LSB
MSE	0.01862	0.00499	0.00496	2.7057e-07
SNR	5.6289	17.7283	18.0171	93.5250
PSNR	82.7617	94.2004	94.243	179.519

It is observed that the watermarking of the audio signal using the Least Significant Bit method gives less MSE, greater SNR, and PSNR values compared to other techniques. Hence, LSB audio watermarking is considered the best method compared to Echo Hiding, Spread Spectrum, and Improved Spread Spectrum. The comparison is shown in Figure 12.

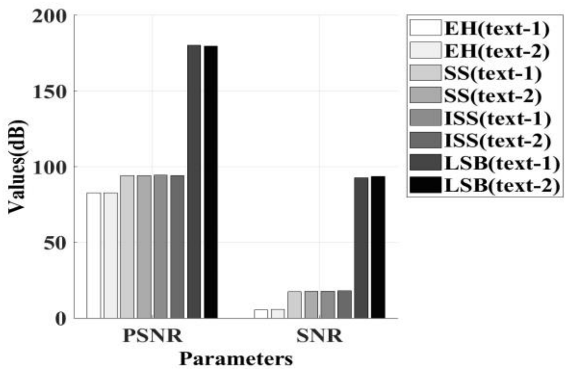

Fig. 12. Comparison of PSNR and SNR for Host and Stego File

Since the LSB algorithm is found to be more efficient, this method is selected to hide dual information namely text and image in the same audio file. This is achieved by performing the embedding in transform domain using DWT and a modified LSB algorithm is proposed to hide the text and image data.

Initially, the performance of LSB to embed image is evaluated. The image (Watermark.png) is hidden and the watermarked audio is generated. The stego audio signal obtained after embedding the watermark image into the cover audio using LSB algorithm is given in Figure 13.

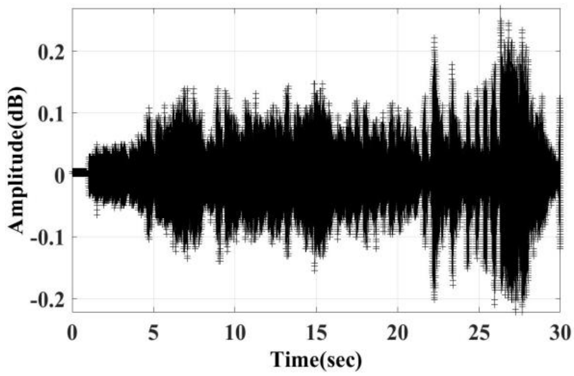

Fig. 13. Stego audio signal with image hidden using LSB algorithm

Figure 14 shows the over-lapped host and stego audio signal

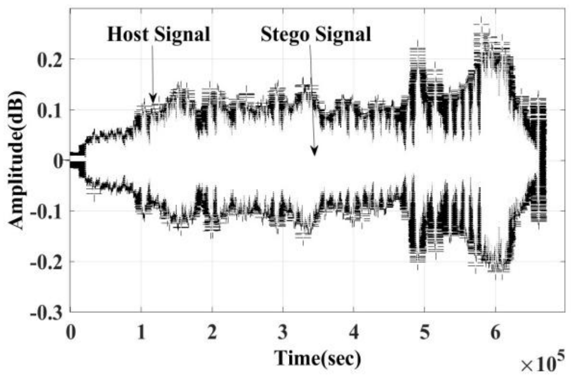

Fig. 14. Overlapped Host and Stego signal with image hidden using LSB

The image (Watermark.png) is embedded into the host audio file using the LSB algorithm. The embedded image or the watermark image is shown in Figure 15.

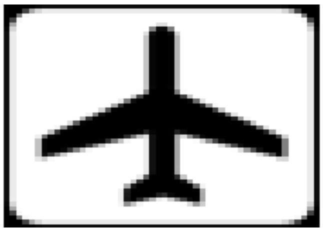

Fig. 15. Watermark Image

The parameters of the image are given in Tables IX.

Table IX: Parameters of an Image

Parameters	Values
Size of an image	255x255(pixels)
Format	.png

The evaluated metrics of the stego signal using LSB algorithm with image hidden are given in Table X.

Table X: Metrics of stego audio using LSB algorithm with image hidden

Metrics (dB)	Values
MSE	0.001423
SNR	99.0455
PSNR	29.2618
Speech Power	923.8451

E. Audio Watermarking using MDWT-LSB Algorithm

The Modified DWT based LSB (MDWT-LSB) algorithm is implemented to hide text and image in the audio file. The image to be hidden is transformed into its sub bands by applying DWT. The Haar wavelet is used for transformation and embedding is performed on the transform coefficients. The Figure 16 shows the image obtained after applying DWT.

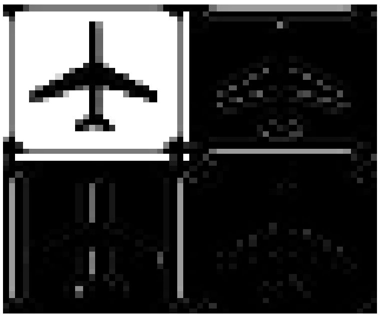

Fig 16. Watermark image after applying DWT

The secret image and text are embedded in the audio file and the stego audio is obtained as shown in Figure 17.

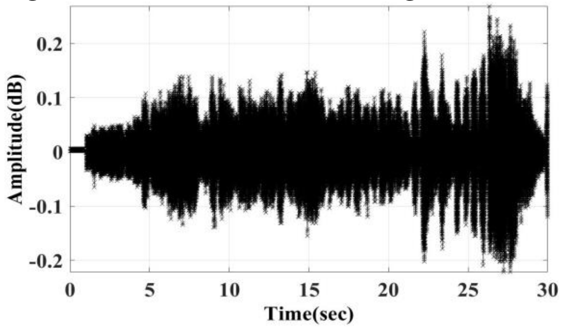

Fig 17. Stego audio signal using MDWT-LSB algorithm

To show the difference between the original and watermarked signal, the host and stego audio signals are overlapped as shown in Figure 18.

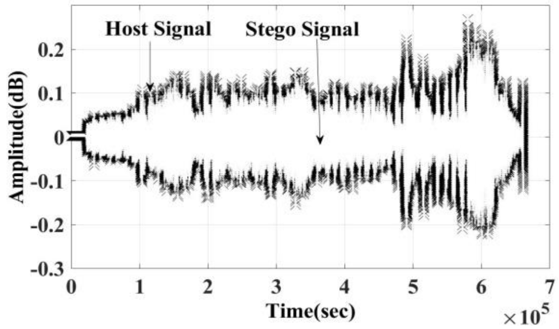

Figure 18. Overlapped Host and Stego audio using MDWT-LSB

The Power Spectral Density (PSD) is the measure of signal's power content and frequency. The amplitude of the PSD is normalized by the spectral resolution employed to digitize the signal. The Figure 19 shows the PSD of the stego audio signal.

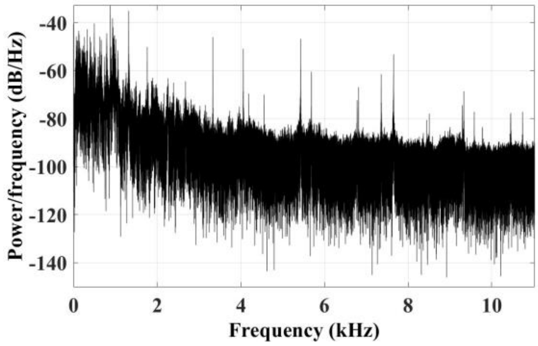

Fig 19. Power Spectral Density of stego audio signal

The Figure 20 shows the histogram of the stego audio using the modified LSB algorithm.

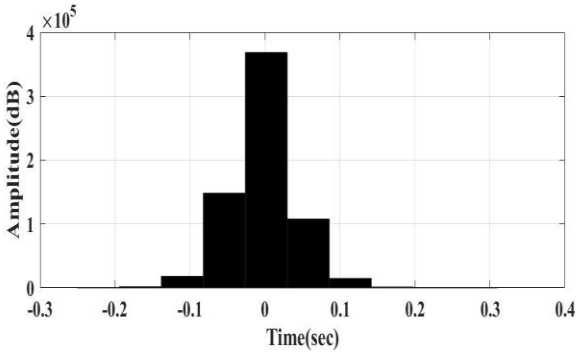

Fig. 20. Histogram of stego audio signal

The cross-correlation between two signals indicates how identical the signals are. If there is a correlation between the signals, it suggests that they are more or less dependent on each other. The Figure 21 shows the cross-correlation of stego audio with its watermark sequence for a modified DWT-based LSB algorithm.

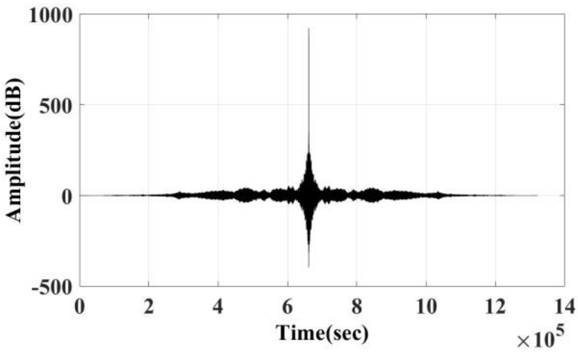

Fig. 21. Cross-Correlation of stego audio with its watermark sequence

The evaluated metrics of the stego signal using modified DWT-based LSB algorithm with dual hidden information is given in Table XI. It gives least Mean Square Error and maximum PSNR values.

Table XI: Metrics of stego audio using MDWT-LSB algorithm

Metrics (dB)	Values
MSE	0.001233
SNR	105.1001
PSNR	30.8961
Speech Power	923.5021

F. Comparison of LSB and MDWT-LSB Algorithm

The metrics obtained for LSB and MDWT-LSB audio watermarking algorithms are compared in Table XII.

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table XII: Metrics for LSB and MDWT-LSB algorithms

Metrics (dB)	LSB	MDWT-LSB
MSE	0.0014231	0.0012335
SNR	99.0455	105.1001
PSNR	29.2618	30.8961

It is observed that the Modified LSB algorithm has the least Mean Square Error and higher SNR, PSNR values compared to the LSB technique. The MDWT-LSB technique hides a 256X256, 8 bit gray scale image of capacity 524288 bits in addition to a text file of 1KB. The hiding capacity of the MDWT-LSB algorithm has increased compared to the conventional LSB method. Additionally varied dual data; such as text and image are embedded in MDWT-LSB with an improved performance. The Figure 22 shows the comparison of MSE and PSNR for stego signals of LSB algorithm and MDWT-LSB algorithm.

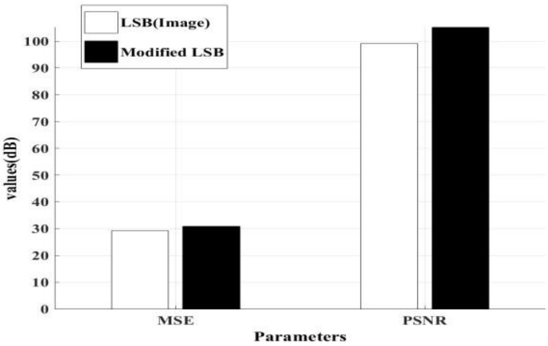

Fig. 22. Comparison of LSB and MDWT-LSB

Analysing the results, it is inferred that the proposed MDWT-LSB watermarking algorithm gives the best performance compared to the implemented watermarking techniques.

V. CONCLUSIONS

In this work, EH, SS, ISS, and LSB audio watermarking algorithms are evaluated to embed text into the audio signal. As LSB gives high PSNR, additional information is embedded into it and a Modified DWT-based LSB audio watermarking algorithm (MDWT-LSB) is proposed to embed the image and text information. Various parameters are calculated for the stego signal. The graphs are plotted to evaluate the performance of all algorithms. Analysing the results, it is observed that the proposed Modified DWT-based LSB (MDWT-LSB) algorithm gives improved results compared to other algorithms. It gives maximum PSNR and minimum MSE value compared to other techniques. Therefore, it can be concluded that MDWT-LSB is a more efficient algorithm for audio watermarking in the context of hiding dual varied data with improved capacity.

REFERENCES

[1] D. Tan, Y. Lu, X. Yan and X. Wang, "A Simple Review of Audio Steganography," 2019 *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2019, pp. 1409-1413, doi: 10.1109/ITNEC.2019.8729476.

[2] Dutta, Hrishikesh, Rohan Kumar Das, Sukumar Nandi, and SR Mahadeva Prasanna. "An overview of digital audio steganography." IETE Technical Review 37, no. 6 (2020): 632-650.

[3] Than, M., and S. Sin. "Secure Data Transmission in MP3 file using LSB and Echo Hiding." *International Journal of Advanced Research in Computer Science* 10, no. 4 (2019): 45.

[4] Maniriho, Pascal, and Tohari Ahmad. "Enhancing the Capability of Data Hiding Method Based on Reduced Difference Expansion." *Engineering Letters* 26, no. 1 (2018).

[5] Agrawal, S. and Kumar, M., 2017. Mean value based reversible data hiding in encrypted images. *Optik, International Journal for Light and Electron Optics*, vol 130, pp.922-934, (2017)..

[6] Khalil, Mohammed, and Abdellah Adib. "Audio watermarking system based on frequency hopping modulation." *In 2014 International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 211-216. IEEE, 2014.

[7] Nababan, Erna Budhiarti, Goklas Tomu Simbolon, and Opim Salim Sitompul. "Multi-LSB and Modified Vernam Cipher to Enhance Document File Security." *IAENG International Journal of Computer Science* 47, no. 4 (2020).

[8] D. Gruhl A. Lu, and W. Bender. Echo hiding In In R. J. Anderson, editor, *Information Hiding, First International Workshop, Cambridge*, U.K., May 30 - June 1, 1996, Proceedings, volume 1174 of *Lecture Notes in Computer Science*, pages 293{315. Springer, 1996.

[9] K. Bansal, A. Agrawal and N. Bansal, "A Survey on Steganography using Least Significant bit (LSB) Embedding Approach," *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*,2020,pp64-69..

[10] Tekeli, Kadir, and Rifat Asliyan. "A comparison of echo hiding methods." *The Eurasia Proceedings of Science Technology Engineering and Mathematics* 1 (2017): 397-403.

[11] H. S. Malvar and D. A. F. Florencio, "Improved spread spectrum: a new modulation technique for robust watermarking," in *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 898-905, April 2003, doi: 10.1109/TSP.2003.809385

[12] Khalil, Mohammed, and Abdellah Adib. "Audio watermarking system based on frequency hopping modulation." *In 2014 International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 211-216. IEEE, 2014.

[13] I. J. Cox, J. Kilian, F. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. Technical report 95-10, NEC Research Institute, Princeton, NJ, USA, 1995.

[14] Malvar, Henrique S., and Dinei A. Florêncio. "An improved spread spectrum technique for robust watermarking." In 2002 *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. IV-3301. IEEE, 2002.

[15] Pradeep Kumar Singh, and R.K.Aggrawal, "Enhancement of LSB based Steganography for Hiding Image in Audio*", International Journal on Computer Science and Engineering*, Vol. 02, No. 05, 2010, P 1652-P1658.

[16] J. D. Gordy and L. T. Bruton, "Performance evaluation of digital audio watermarking algorithms," Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems (Cat.No.CH37144), 2000, pp. 456-459 vol.1, doi: 10.1109/MWSCAS.2000.951682

[17] Z. Haque, A. Saha, T. S. Das, A. Basu and A. Chattopadhyay, "Performance Evaluation of Audio Watermarking in EMD Framework," 2018 *Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 2018*, pp. 269-274..

[18] Al-Hooti, Mohammed Hatem Ali, Tohari Ahmad, and Supeno Djanali. "Improving the Capability of Reduced Difference Expansion based Digital Image Data Hiding." *IAENG International Journal of Computer Science* 46, no. 4 (2019).