

# Audio Steganography using Echo Hiding in Wavelet Domain with Pseudorandom Sequence

Sounak Lahiri

B.Tech

Department of Electronics and  
Communication Engineering  
Institute of Engineering and  
Management, Kolkata

## ABSTRACT

This paper deals with the encoding of message bits in a cover audio carrier signal. It applies the basic concepts of audio steganography in transform domain to achieve higher efficiency in data transmission, while preserving the secrecy of the information being transmitted. The proposed model in this paper deals with the application of echo hiding of binary message bits in the carrier signal, in the transform coefficients obtained by applying 2D- discrete Haar Wavelet Transform on the cover signal. Moreover, this algorithm applies pseudorandom sequence to encode the data, which gives the method more efficiency and prevents unauthorized decoding at any moment. The application of echo hiding technique makes the stego signal more immune to noise and disturbances, during transmission through the channel network. The performance of this method is analyzed on the basis of the output SNR and PSNR values calculated for several test cases, that has been discussed later in the paper.

## General Terms

Audio Steganography, Discrete Wavelet Transform, Pseudorandom sequence, Echo Hiding.

## Keywords

2D-Discrete Haar Wavelet Transform, Pseudorandom sequence, Echo Hiding.

## 1. INTRODUCTION

Steganography is a method of communicating a secret message from one point to another in such an efficient manner, so as to preserve the secrecy of the message without any hindrance to its transmission. It uses a cover signal to carry the payload signal. In this scenario, the message is concealed or embedded within the cover object to form the stego object, which is then transmitted through the network. The embedding process can be done in several ways both in temporal as well as in transform domain.

For an efficient steganographic technique, it should have [1]:-

- High perceptibility with SNR values greater than 20dB
- Reliable and robust.
- High capacity to carry more information.
- High speed of embedding.
- Immune to noise.

The proposed algorithm in this paper deals with embedding in the transform domain. To be specific, wavelet domain is used here, to encode the data.

A wavelet is a waveform of limited duration that has an average value of zero. E.g. Haar Wavelet as shown in the Figure 1.



Figure 1- Haar Wavelet

The Haar wavelet transform decomposes a discrete signal into two sub-signals of half its length. One is a running or average trend and the other one is a running difference or fluctuation. The advantage of using discrete haar wavelet transform is to obtain data that is more accurate by providing different resolutions at different parts of the time-frequency plane [2].

Haar wavelet is defined as in equation (1),

For  $\psi: \mathbb{R} \rightarrow \mathbb{R}$

$$\Psi(t) = \begin{cases} 1, & \text{for } t \in [0, \frac{1}{2}), \\ -1, & \text{for } t \in [\frac{1}{2}, 1), \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The concept of pseudorandom generators is that, it is a mathematical function that generates a random number sequence but with a specific period of number of values. After completion of the certain period number of values, the sequence repeats itself. The general generator function of a pseudorandom generator can be given as,

$$X_{n+1} = \text{mod}(P1 * X_n + P2, \text{max}) \quad (2)$$

Where, X is the sequence, P1 and P2 are arbitrary constants and max defines the maximum value till which the pseudorandom sequence might run.

The advantage of using pseudorandom sequence is that, it is always unique to the encoder and decoder. Hence, for any coding system, it highly increases the efficiency of the system and the secrecy of the information being sent is preserved.

This paper deals with the related works in this field and a proposed algorithm that might help to improve the efficiency of the system, by hiding the data in the wavelet domain using echo hiding technique on a cover audio signal.

## 2. RELATED WORK

Several research works have been done in the field of audio steganography and its application in hiding information during transmission. A few related works has been described in the following lines, which have aimed to achieve increased efficiency and robustness.

Ali M. Meligy, Mohammed M. Nasef, Fatma T.Eid,2015,[3], proposed a method of transforming the cover signal using LWT and applying LSB coding using random keys to achieve successful embedding. This method uses three random keys – one to embed the message, one to store the length and the third to randomly skip bits to increase robustness.

D.M. Ballesteros L, J.M. Moreno A, 2012, [4] proposed a method that utilizes the Frequency Masking concept with sorting of the wavelet coefficients of the payload by performing indirect LSB substitution for embedding speech signals into speech signals. It has four steps: decomposition and pre-scaling by transforming the decimal coefficients of the signals to a binary representation by the Discrete Wavelet Transform (DWT), then sorting the wavelet coefficients of the payload and then indirect LSB substitution for embedding. Finally, the output is reconstructed from the signals by applying the inverse wavelet transform.

K.P.Adhiya, S.A. Patil, 2012, [3, 5], proposed a method in which each audio sample is converted into bits, followed by the embedding of textual information in it. In the embedding process, first the payload is converted into its equivalent binary. The last 4 bits of this binary is taken into consideration and redundancy check of the binary code is performed to check the prefix for either 0 or 1 whichever is used.

H.I. Shahadi, R. Jidin and W.H. Way, 2014, [6] proposed an approach that uses Integer-to-Integer Lifting Wavelet Transform (Int2Int LWT) and Least Significant Bits (LSBs) substitution. It analyzes each cover frame (4-samples) by using 2-levels of Haar Int2Int LWT and generates an adaptive key (stego-key) that is used for embedding data. The first 12-bits from the payload are inserted in the first 6 LSBs after the Starting Depth (SD) from the first and second coefficients of the first detail sub-band, respectively. The reminder 4-bits from the payload are inserted in the first 4 LSBs after the SD from the coefficients of the second detail sub-band. The output stego frame is constructed by performing inverse of 2-levels Haar Int2Int LWT after converting all sub-bands to decimal.

N. Gupta and Ms. N. Sharma, 2014, [7] proposed a system that improves robustness and security by using the concept of Discrete Wavelet Transform and Least Significant Bit. The method emphasizes on the scheme of image hiding in audio. The approach compares it with simple LSB insertion method for data hiding in audio.

M. P. Jain and Prof. V. Trivedi, [8], proposed a method that compares coefficients in the Discrete Cosine Transform domain. Two Coefficients of a segment are compared and payload bits are embedded, accordingly. It segments the cover audio into parts and then performs DCT on each part to give two outputs; one a DC signal having highest power and lowest frequencies and the others are AC signals with decreasing power from lowest to highest frequencies. This method presents high capacity and high stego-signal quality.

## 3. PROPOSED METHOD

The proposed algorithm uses echo hiding technique to hide the payload message signal in the audio signal being used as

carrier. The echo hiding procedure can be explained as follows:-

**Echo hiding:** - Echo hiding technique is an efficient method to embed messages in an audio carrier (cover) signal by introducing an echo at various instances of the discrete signal. The advantages of this method include – high data transmission rate, superior robustness [9] and high noise immunity. Every echo produced from the cover signal can store a single bit of message signal. Hence, initially the original signal is broken down into blocks. After the embedding process, the blocks are joined back together to retrieve the output signal. [9, 10, 11]

Data hiding using echo has three parameters; that is, decay rate, offset and initial amplitude [12]. The initial amplitude is used to determine the cover audio amplitude. Decay rate helps to determine the echo function to be made and offset gives the distance between the echo and the cover signal data points.

There is a representation of the concept of the parameters shown in Figure-2.

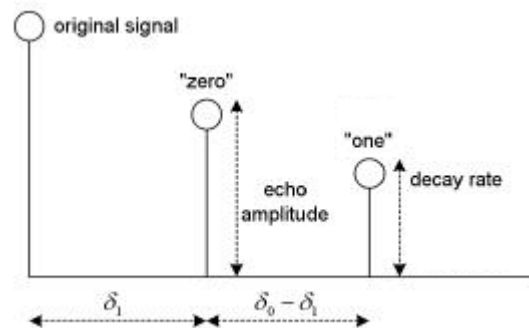


Figure 2 - The Echo hiding parameters.

### The encoding process:-

1. At first, the payload message vector and the carrier audio is taken as input.
2. The message vector is converted to its equivalent binary form. For any image payload, the pixel values are individually converted to its binary form. For text messages, the ASCII values of the letters are converted to their equivalent binary form.
3. The carrier audio is first sampled at around 44 KHz (to reduce distortion due to echo addition) and then fragmented using 2D- Discrete Haar wavelet transform. The coefficient matrices of cA, cD, cH and cV are obtained.
4. The length of the carrier coefficients after wavelet transform and the length of the payload are used to generate a pseudorandom sequence which has a maximum value equal to the length of the carrier coefficients and a period value greater than the length of the binary sequence. The initial value of the sequence is taken to start from 3000 (obtained experimentally).
5. **Embedding process**
  - The length of the message is placed at the 100<sup>th</sup> position of the coefficient vector (e.g. – cD) in use.
  - The values of the pseudorandom sequence give the positions for embedding on the carrier audio. For every position, the value at that position and its

corresponding position, from any of the coefficient vectors (cA, cD, cH, cV) is taken to form a vector A.

- The message vector is now checked. For zero the vector A is passed through filter H0 and for one, it is passed through filter H1, to form a vector B. The values of the vector B are placed in the coefficient vector (cA, cD, cH, cV) at positions of two greater than the corresponding position given by the pn sequence.
- The filter functions are given as

$$H0(z) = 1 + 0.5z^{-1} \quad (3)$$

$$H1(z) = 1 + 0.25z^{-2} \quad (4)$$

- The process is repeated for all values of the message vector.

6. The final coefficient vectors are now taken and inverse 2D- Discrete Haar wavelet transform is performed to obtain the final stego signal which contains the embedded payload signal.

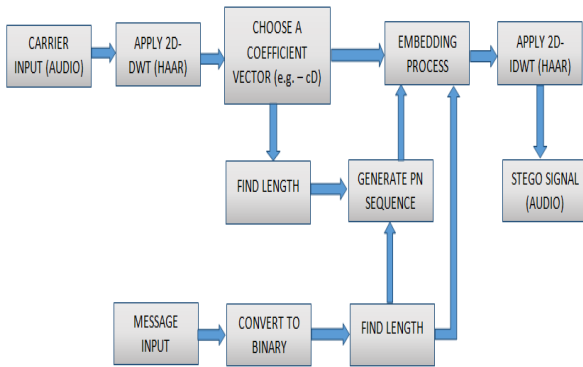


Figure 3 – Block Diagram of the Encoding Process

**The decoding process:**

1. The stego signal is taken as input to the system. It is filtered to get rid of any disturbances or noise from the channel.
2. The stego audio is sampled at 44 KHz and then fragmented using 2D- Discrete Haar wavelet transform. The coefficient matrices of cA, cD, cH and cV are obtained.
3. The 100<sup>th</sup> position of the coefficient vector (e.g. – cD) used is checked to find the length of the message vector under consideration.
4. For this length of the message vector obtained from the stego audio signal, the pseudorandom sequence is generated again.
5. **Extraction process:**
  - For every position given by the values of the pn sequence, the value at that position and its corresponding position is taken to form vector S.
  - Vector S is passed through both filters H0 (as in equation (1)) and H1 (as in equation (2)), to generate HS0 and HS1, respectively.
  - The corresponding positions are checked for values. If it matches with HS0, then a 0 was sent and a zero is placed in the output message signal. Similarly, if it matches with

HS1 then a 1 was sent and hence, a one is placed in the output message vector.

- The process it repeated for the entire length of the message vector, to obtain the output.
- 6. Finally, the output is now converted back to the form in which it was sent (image, text, etc.).

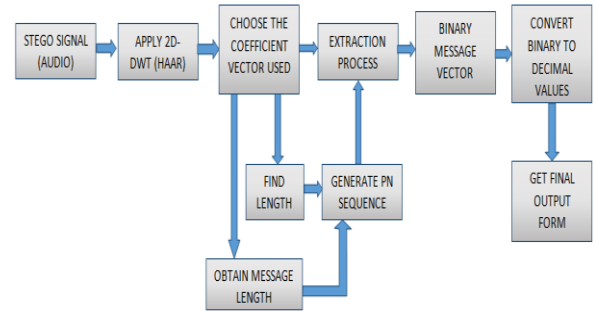


Figure 4 – Block Diagram of the Decoding Process

**4. RESULTS**

The proposed algorithm was coded and tested using MATLAB (R2014a). A few audio signals were used to act as the cover signal. For the payloads, first a random binary sequence was used to check the correctness of the design, followed by a text message and a grayscale image message. The SNR and PSNR values are calculated using the equations given below. The comparison is done on the cover audio signal and the final output stego signal.

The SNR is calculated using the formula [3]:-

$$SNR = 10 * \log_{10} \frac{\sum_{i=1}^n (Carrier(i))^2}{\sum_{i=1}^n [Carrier(i) - Stego(i)]^2} \quad (5)$$

The PSNR is calculated by the formula [3]:-

$$PSNR = 10 * \log_{10} \frac{(Peak Value)^2}{Mean Square Error} \quad (6)$$

Where,

$$Mean Square Error = \frac{\sum_{i=1}^n [Carrier(i) - Stego(i)]^2}{M * N} \quad (7)$$

In the equations given above, Carrier represents the original carrier signal taken as input and Stego represents the stego signal obtained after the encoding process. M and N represents the dimension of the input signal vector and R gives the maximum value of the signal.

The proposed method is now tested for several inputs. An example is shown below in Figure 5. Figure 6 shows a close up look at points in the plot, where embedding takes place for a particular message bit.

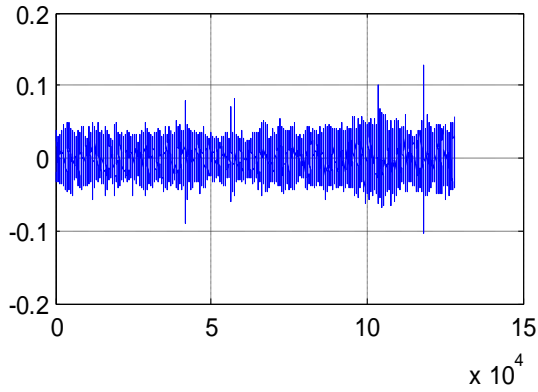


Figure 5- Plot of cD coefficients of Test Audio 3

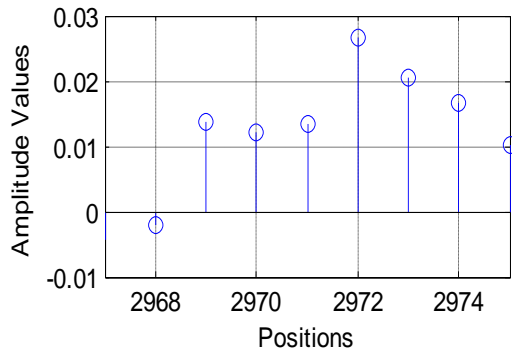


Figure 6- A segment of the plot of cD coefficients of TestAudio 3 before embedding

A part of a message sequence can be shown as follows, in Figure 7.

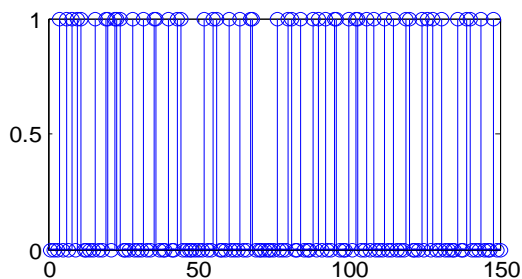


Figure 7- Section of a message sequence showing the first 150 bits (e.g. - Random Binary Sequence)

For every test carrier audio and test message sequence, a coefficient is chosen and the spectrum of both the input and the output signals are checked. Very low variation of the signal is observed, which suggests that the method is efficient and immune to variations in channel noise.

The plot of the output signal after embedding can be shown as follows, in Figure 8.

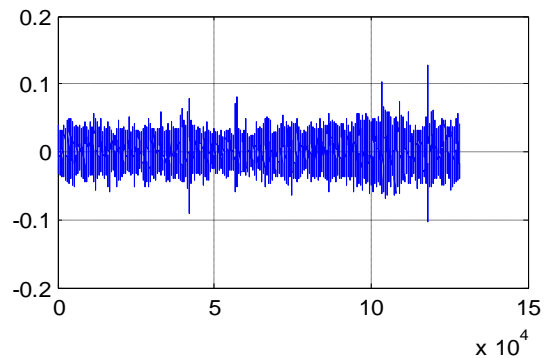


Figure 8- Plot of cD coefficients of output after embedding message

Similarly, for the output signal, the close up look of the signal coefficients in the same region can be given as shown in Figure 9. It can be seen that, the coefficient values of cD matrix changes after the embedding process. The same segment of the matrix has been shown in both Figure 6 and Figure 9. After embedding the message signal, the amplitude at the first two positions remains the same and the consecutive position amplitudes change because of embedding the filtered values. This illustrates the embedding process for a single bit of message. The process is exactly the same, for the whole message vector.

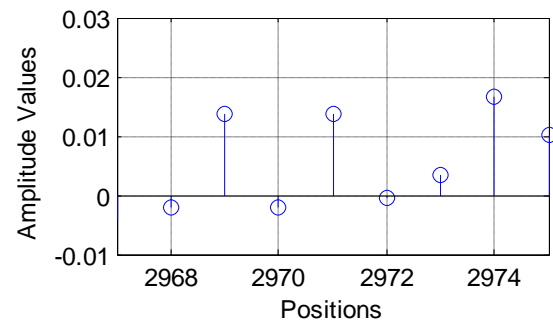


Figure 9- A segment of the plot of cD coefficients of stego signal after embedding

Comparing Figure 5 and Figure 8, it can be analyzed that there is very small difference between the coefficient vectors before and after the encoding process. This is because the echo effect does not become prominent when the audio signal is sampled at frequencies greater than 1 KHz. Now, the SNR and PSNR values are observed for both the approximation coefficients (cA) or low frequencies and the detail coefficient (cD) or the high frequencies. The output SNR and PSNR values are given as follows in Table.1.

**Table 1 – SNR and PSNR values for embedding the data in the test audio signals for both cA and cD coefficients**

MESSAGE	CAPACITY(bits)	CARRIER	CAPACITY(bits)	cA		cD	
				SNR	PSNR	SNR	PSNR
RANDOM BINARY SEQUENCE	1320	TESTAUDIO1	453450	48.6113	80.484	83.4294	115.3021
		TESTAUDIO2	300339	41.4006	74.2471	89.8194	122.6659
		TESTAUDIO3	256709	45.715	64.1617	116.6133	135.0601
		TESTAUDIO4	158291	34.6642	71.5561	65.9103	102.8021
<b>AVERAGE</b>				42.59778	72.61223	88.9431	118.9576
TEXT MESSAGE	1216	TESTAUDIO1	453450	50.0696	81.9422	84.7162	116.5888
		TESTAUDIO2	300339	42.7899	75.6365	90.4850	123.3316
		TESTAUDIO3	256709	46.5743	65.0210	117.6441	136.0908
		TESTAUDIO4	158291	35.5668	72.4586	66.6987	103.5905
<b>AVERAGE</b>				43.75015	73.76458	89.886	119.9004
IMAGE	1200	TESTAUDIO1	453450	50.6927	82.5654	84.8689	116.7916
		TESTAUDIO2	300339	43.3266	76.1731	90.6293	123.4759
		TESTAUDIO3	256709	47.5488	65.9956	117.9608	136.4075
		TESTAUDIO4	158291	36.0733	72.9651	66.8953	103.7871
<b>AVERAGE</b>				44.41035	74.4248	90.08858	120.1155

Table 2 represents the average SNR and PSNR values for each audio taken during the experiment.

**Table 2 – Average SNR and PSNR values for both cA and cD coefficients for the test audio signals**

CARRIER SIGNAL AUDIO	CAPACITY(bits)	cA		cD	
		AVERAGE SNR	AVERAGE PSNR	AVERAGE SNR	AVERAGE PSNR
TESTAUDIO1	453450	49.7917	81.6639	84.3382	116.2275
TESTAUDIO2	300339	42.5057	75.3522	90.3112	123.1578
TESTAUDIO3	256709	46.6127	65.0594	117.4061	135.8528
TESTAUDIO4	158291	35.4348	72.3266	66.4014	103.3932

The variation of SNR and PSNR, for cA coefficients can be illustrated with the help of the following graph as shown in Figure 10.

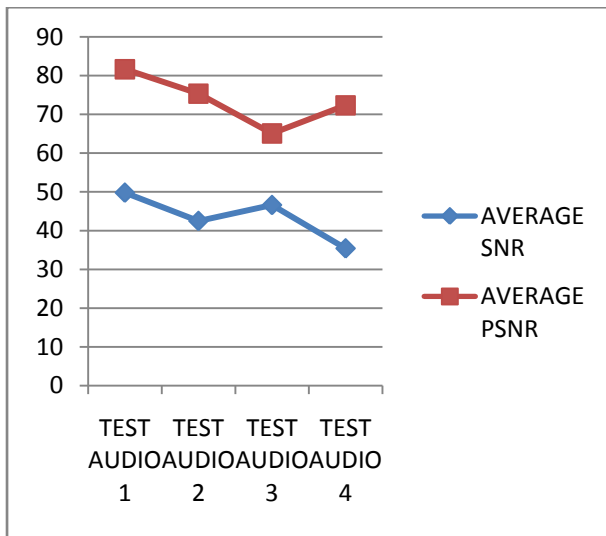


Figure 10 – Variation of SNR and PSNR for cA coefficients

Similarly, the variations for cD coefficients are shown in Figure 11, below.

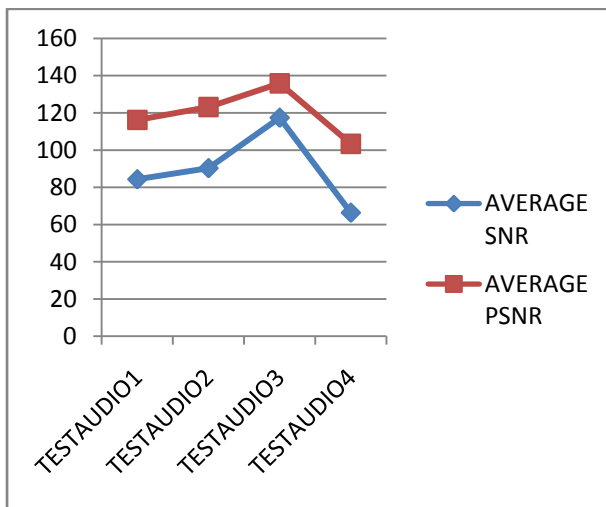


Figure 11 – Variation of SNR and PSNR for cD coefficients

## 5. CONCLUSION

This paper provides a modified method of encoding messages in a cover audio signal. In contrast to the LSB addition embedding process as suggested in the other related works, the proposed method uses echo hiding technique. The advantage lies in the fact that this process provides the required immunity to channel noise during transmission. Moreover, embedding in the transform domain increases robustness and reliability. The application of Discrete Wavelet Transform makes the system more efficient and dynamic as signal reconstruction becomes easier, in comparison to other transforms that may be used for the purpose. Furthermore, wavelet transform allows analysis in both time and frequency domain, which has more advantage over the other transforms. The use of pseudorandom sequence to encode data at different

positions increases efficiency and preserves the secrecy of the information being transmitted from one point to another, because the pseudorandom sequence that is generated is unique to the encoder - decoder pair. Therefore, the information in the stego signal cannot be easily decoded without the equivalent matching decoder. Hence, the system as a whole, has more efficiency and is immune to noise, while preserving secrecy of message during transmission.

## 6. REFERENCES

- [1] Sumeet Gupta, Dr. Namrata Dhanda,2015, “ Audio Steganography Using Discrete Wavelet Transformation (DWT) & Discrete Cosine Transformation (DCT)”, *IOSR Journal of Computer Engineering (IOSR-JCE)*, e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 2, Ver. V (Mar – Apr. 2015), PP 32-44
- [2] Piotr Porwik, Agnieszka Lisowska,2004, “The Haar-Wavelet Transform in Digital Image Processing:Its Status and Achievements”, *Machine GRAPHICS & VISION* vol . 13, n o. 1/2, 2004, p p .79-98
- [3] Ali M. Meligy, Mohammed M. Nasef, Fatma T.Eid,2015, ” An Efficient Method to Audio Steganography based on Modification of Least Significant Bit Technique using Random Keys”, *I. J. Computer Network and Information Security*, 2015, 7, 24-29
- [4] D. M. Ballesteros L and J. M. Moreno ,” Highly transparent steganography model of speech signals using Efficient Wavelet Masking ” *Expert Systems with Applications*, vol .39, 2012.
- [5] K.P.Adhiya and S.A. Patil,”Hiding Text in Audio UsingLSB Based Steganography”, *Information and KnowledgeManagement*, Vol 2-No.3, 2012.
- [6] H.I. Shahadi, R. Jidin and W.H. Way, “Lossless Audio Steganography based on Lifting Wavelet Transform and Dynamic Stego Key”, *Indian Journal of Science and Technology*, Vol 7-No. 3, March 2014.
- [7] N. Gupta and Ms. N. Sharma, “Dwt and Lsb Based Audio Steganography”, *International Conference on Reliability, Optimization and Information Technology (ICROIT)*, Feb 6-8 2014.
- [8] M. P. Jain, Prof.V. Trivedi “Effective Audio Steganography by using Coefficient Comparison in DCT Domain”, *International Journal of Engineering Research & Technology (IJERT)*, Vol. 2, Issue 8, August – 2013.
- [9] Jayaram P, Ranganatha H R, Anupama H S,2011,” INFORMATION HIDING USING AUDIO STEGANOGRAPHY – A SURVEY”, *The International Journal of Multimedia & Its Applications (IJMA)* Vol.3, No.3, August 2011
- [10] Sajad Shirali-Shahreza M.T. Manzuri-Shalmani “High capacity error free wavelet domain speech steganography” *ICASSP 2008*
- [11] V. Vapnik, “Statistical Learning Theory”, John Wiley, 2008.
- [12] Ary Mazharuddin Shiddiqi, Tirta Priambadha, Baskoro Adi Pratomo,” ECHO DATA HIDING STEGANOGRAPHYAND RSA CRYPTOGRAPHY

ON AUDIO MEDIA”, *INDUSTRI Vol. 11 No. 1 : 1-9*  
*ISSN 1693-0533*

- [13] Jisna Antony, Sobin c. C, Sherly A. P,” Audio Steganography in Wavelet Domain – A Survey”, *International Journal of Computer Applications (0975 – 8887) Volume 52– No.13, August 2012*
- [14] C.P.Sumathi, T.Santanam and G.Umamaheswari,” A Study of Various Steganographic Techniques Used for

Information Hiding”, *International Journal of Computer Science & Engineering Survey (IJCSES) Vol.4, No.6, December 2013*

- [15] Radomir S. Stankovi\_c a, Bogdan J. Falkowski,” The Haar wavelet transform: its status and achievements”, *Elsevier, Computers and Electrical Engineering 29 (2003) 25–44*