A Secure Approach to Audio Steganography

Nisha Kundu and Dr. Amandeep Kaur

M.Tech Student , CST Department, Central University of Punjab, Bathinda, India
Associate Proff.CST Department, Central University of Punjab, Bathinda, India

Abstract
As the internet growing day by day secure transmission of information is crucial. Steganography and cryptography both techniques provide data confidentiality and help us to protect data from day by day attacks. Digital communication has become an essential part of infrastructure now a days and also lots of applications are internet based. So the communication made must be secret. Steganography is a process of hiding the information of secret messages or one can say that it is an art of sending the hidden data or message over a public network. In the proposed approach message is first encrypted through using The Vigenere Square Encryption Algorithm and then ASCII conversion of data takes place. After those characters of information or data is embedded into deeper layers through modified LSB method. After that further audio transposition encryption technique is used for audio.

Keywords:
Audio Steganography, Encryption, Audio Encryption, Modified LSB, Vigenere Square Encryption Algorithm, LSB Coding , PSNR-Peak signal to noise ratio, MSE-mean square error.

1. Introduction

Internet communication is essential part of communication now a day. Data needs to be protected and secure when it is transmitted over the public network. So we have to secure data by using security techniques and make it confidential [1] .and steganography is proposed to provide more security to the data .Audio steganography is a technique in which the secret data is embedded into the cover audio by using various techniques like LSB, with the secret message hidden from the unauthorized party and to protect against the attacks. The main requirements for any steganographic system are - transparency, hiding capacity and robustness [2]. After embedding of message is done , the stegno audio is obtained by use of various steganographic methods[8] At the receiver's side the hidden data can be extracted from the stegno signal using the reverse algorithm as that of used for embedding. Encryption is done for more security of data and information travelled through a public network. Encryption is a part of cryptography .Cryptography provides data confidentiality, data integrity and In the proposed approach combination of encryption. [4]. The proposed method works on modified vigenere cipher technique. The Vigenere cipher encrypts its inputs by using the key and the plaintext as indices into a fixed lookup table: the Vigenere square [3]. After that the message characters are converted to ASCII values and then embedding is done into the deeper layers using modified LSB method. For more security further audio encryption is done by using transposition method [6] [7]. Finally MSE and PSNR are calculated so that we can compare the results with other methods.

2. Related Work

Authors in [9] told about LSB basic method that how embedding is done and also the Capacity of LSB-Based Audio Steganography .To extract a secret message from an LSB encoded sound file, the receiver needs to use same approach as used in the embedding process. Normally, the length of the secret message to be encoded is smaller than the total number of samples in a sound file. One must decide then on how to choose the subset of samples that will contain the secret message.

In [13], authors have suggested an algorithm in which the data is first subjected to encryption using Data Encryption Standard (DES). The encrypted message is then passed to embedding phase. In embedding phase the encrypted message will embedded into the cover medium which is either image or audio or video resulting in a stego medium. The embedded stego medium contains the encrypted text message which is extracted at the receiver side. The extracted text is decrypted using decryption module.

Authors in 2010 presented a scheme that uses an image file as the secret data to be hidden in the audio file taken as cover object because the size of the image is generally quite small as compared to audio [10] in which it must be hidden. The proposed method is based on the basic technique for hiding secret data in audio. According to the proposed method, least significant bits up to three LSB positions are used as a stego-key to encode the image bits in stego-object. Thus, the image data can be embedded according to the embedding algorithm taking various MSB positions and same process is used for receiver end.

Authors in 2011 purposed Discrete Wavelet Transform (DWT) Audio steganography method so that

Manuscript received February 5, 2026
Manuscript revised February 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.2.17

data is embedded with signal in the LSBs of the wavelet coefficients of audio signal [11]. To improve the inaudibility of embedded data employed a hearing threshold when hiding data in the integer wavelet coefficients, while avoided data hiding in the silent parts of audio signal. In this method robustness is achieved and here due the presence of hearing threshold the inaudibility is also improved , so it is better than that of HAS and that of above one in term of security too.

Authors in, 2011 purposed an algorithm in which the secret data is first encrypted using AES algorithm [12]. This encrypted data is then embedded into an audio file. The authors then encrypt the audio file using Spread Spectrum technique before transmitting it over the network. So this will help enhance the security.
Authors in 2011 purposed an enhanced LSB technique which is tested on .wav file [13] and it choose random positions for embedding the secrete bits into audio file. So security and robustness is increased by some level by using this technique.

Authors in 2012 proposed a three layered model for audio steganography based on least significant bit replacement [14]. In this purposed research the secret message to be transmitted is passed through two layers before it is embedded within the cover message in the third layer. The stego(secret message after applying the steganography) message is transmitted over the network to the receiver side. The objective of the paper is to make sure the confidentiality of the secret message. As we know confidentiality is equivalent to privacy. Measures are used to ensure confidentiality to prevent sensitive information from reaching to the wrong people. Access must be restricted to the authorized people and take care of capacity, transparency and robustness.

Authors in 2012 proposed method that adapts the Frequency Masking concept using an efficient sorting of the wavelet coefficients of the secret messages and use an indirect LSB substitution for hiding speech signals into speech signals [15]. The proposed method contains 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 using an efficient sorting of the wavelet coefficients of the secret messages, then using an indirect LSB substitution for hiding speech signals into speech signals, and final reconstruct the signals and post-scaling it by applying the inverse wavelet transform.

Authors in 2012 presented a novel approach of submission technique of audio Steganography [16]. Using genetic algorithm, message bits are embedded into multiple and higher LSB layer values based on algorithm that embed data into deeper LSB bits which is resulting in increased robustness. By using this purposed method the robustness would be increased against those intentional attacks which try to reveal the hidden message and also some unintentional attacks like noise addition as well. So overall the integrity and security and robustness are maintained by using this approach because the hidden message is not under the insecurity of attacks to change it.

Authors in [17] proposed technique, the sound is divided into samples where each sample is of 24bit, 8 bits are to be hidden in each sample by distributing the bit pattern [17] that corresponds to the secret gray scale image across the LSBs of the preprocessed sound samples (i.e. the preprocessed sound waves take the shape of a RGB colored image). So the embedding capacity is 8 bits per audio sample which results in large embedding capacity. Additionally, hiding the secret bit pattern by distributing it in the layers of the colored image, add more secrecy to the hidden data

Authors in[18] proposed a method in which consecutive LSB's in each sample of cover audio is replaced with secret message bit.LSB method is very easy to implement but have low robustness. The paper also compares the spectra of original audio signal before embedding and audio signal after embedding and compare the results by using various audio specifications.

Authors in [19] discussed a method of hiding text or secret in audio using multiple LSB steganography and provide security using cryptography techniques. The research has proposed two approaches of substitution technique of audio steganography that improves the capacity of cover audio for embedding data. Here message bits are embedded into multiple and variable LSBs. From the results these methods improves the capacity of data hiding of cover audio.

Authors [20] studied a detailed look of audio steganography techniques using LSB and genetic algorithm approach .This research has study of various techniques of audio steganography using different algorithms like genetic algorithm approach and LSB approach. It has tried some approaches that help in audio steganography. It has the art and science of writing hidden messages in such a way that only the sender and intended recipient suspects the existence of the message.

Authors in [21] proposed, dual layer randomization approach. First layer of randomization is achieved by randomly selecting the byte number or samples from all. An additional layer of security is provided by randomly selecting the bit position at which embedding is done in the selected samples. Using this proposed algorithm the

transparency and robustness of the steganographic technique is increased. Due to less robustness and more vulnerability to be attacked LSB method is not preferred much. Instead two bits (2nd and 3rd LSB's) are used for hiding message. This method is good in terms of integrity and robustness also .A filter is designed to minimize the changes occurred in stego file using a unique key. By this unique key the security of data hiding is increased. The key will derive to extract the correct message at receiver's end. So this method provides more security than that of the LSB method used earlier.

Authors in [25] proposed a new approach that overcomes the problems of substitution techniques in audio steganography one problem is that they are less robust against intentional attacks that try to reveal hidden message and second problem is having low robustness against unintentional attacks. The algorithm will hide the message as per the proposed solution (in deeper layers of audio sample and will modify other bits to minimize the error). The method currently uses 2 bits per byte of audio sample. This will progress towards achieving higher capacity and robustness.

Authors in [27] proposed method for Embedding and Extraction Process using Random keys. In this paper, they have presented an audio steganography technique based on LWT(Lifting Wavelet Transformation) and modified LSB technique by using three random keys. They used three random keys to increase the robustness of the LSB. The first key is used for embedding the type of the secret message "text or audio or image", the second key is used for embedding the length of the secret message and the third is used for skip some of bits randomly to increase the robustness of the LSB. Authors used CD coefficient because of high frequencies. Also, the SNR values of proposed method are better than other known methods. Where x is the original signal, y is the stego signal, M and N are the numbers of rows and columns of the input signal and R is maximum value of the signal [27].This method also provides the comparison for signal to noise ratio.

$$PSNR = 10*\log_{10}\frac{R^2}{MSE}$$

$$MSE = \frac{\sum_1^n [x-y]^2}{M*N}$$

Fig. 1. PSNR and MSE formula [27]

Authors in [28] purposed method information hiding using encryption in audio Steganography. They firstly encrypt the message using cipher techniques in this method; the original plain text is subjected to classical Vigenère cipher followed by double columnar transposition and then also encrypts the audio the audio file is subjected to encryption. Encryption is carried out by transposition in which the orders of the audio frames are changed generating a scrambled audio file. This scrambling is carried out by the random numbers generated by Blum Blum Shub pseudo random number generation algorithm. The random numbers are generated dynamically using BBS algorithm. The first step of encryption is to restrict the magnitude of each random number to the number of audio frames present in the audio by making use of modulo operation. The audio frames are then traversed starting from the very 1st audio frame. Each audio frame is swapped with another; the latter"s frame number being picked by the random number generated at that instant. The procedure is repeated till all the audio frames are exhausted. Subsequently, the encrypted audio file is obtained.

3. Proposed Technique

Phase 1:

The first phase is related to encrypting the secret data by using modified Vigenere cipher algorithm that is called The Vigenere Square Encryption Algorithm. The primary weakness of classical Vigenere cipher is the repeating nature of its key. Firstly Vigenere cipher encrypts its inputs by using the key and the plaintext as indices into a fixed lookup table called as the Vigenere square. For easy computation, the algorithm first maps letters to numbers: A=1, B=2, ... SPACE=27. As shown in the matrix below, each row in the square is derived from the row above by circularly shifting it one place to the left:

1 2 3 4 5 ... 26 27
2 3 4 5 6 ... 27 1
3 4 5 6 7 ... 1 2
...
27 1 2 3 4 ...25 26

To encrypt, replicate the letters in the key so the key and plaintext are the same length. Then, derive each cipher text letter by lookup in the Vigenere square: use the key letter as the row index and the plaintext letter as the column index. If the key K and the plaintext P are n letters long, form the cipher text result C by indexing into the Vigenere square V, as follows:

$$C(n) = V(K(n), P(n))$$

Decryption simply reverses the process, using the key letter to determine the row index and the cipher text letter to

determine the column index, which is the plaintext letter. For example:

$$P(n) = FIND(V(K(n,:)) == C(n))$$

Phase 2:

In this phase, each character obtained by encrypted message (C) is extracted and the ASCII value of the character is used to generate bit pattern. Each bit of the pattern replaces the last bits of an audio frame by using modified LSB. Each character is then represented by an 8 bit binary number. Thereby, for each character, here we will have 8 consecutive audio frames.
ASCII conversion is as follow:

Character to be embedded – A

ASCII value of A - 65

8 bit binary representation of the ASCII value: 01000001

8 consecutive audio frames in binary format (consider the 8 bits)

10010010
01010101
10010101
11110011
10100000
11010101

In this phase we have modified Least Significant Bit (LSB) encoding technique in which the encrypted message (C) is embedded into the cover audio making use of deeper layers to embed message by the modified LSB technique. As shown in table:

Table 1: Modified LSB method

1st MSB	2nd MSB	Secret Message Bit
0	0	3rd LSB
0	1	2nd LSB
1	0	1st LSB
1	1	1st LSB

In this way algorithm works and embedding is to be done.

Phase 3:

Last phase is to encrypt the audio file further for more security. This is done by using transposition method. Here the performance of transposition technique is measured using MSE. In this test value of original and encrypted file are compared. The encryption process will begin with preprocessing segmentation of audio data. Each audio chunks are separated by intervals of bits in which the form of chunks of audio data array will be used as an input in the transposition. Then transposition process will swap the index array for audio encrypt of chunks. The key is going to affect the outcome of exchange index. After index interchangeable, the audio chunks of data will be redeveloped into new audio file as a result of encryption. Then we will decrypt the audio file at another end by use of key.

Transposition key column-table encryption

The process begins with inserting the indices of audio chunks of data into a horizontal table. The sorted key is managed and exchanged position of table column. Next, the indices of audio chunks of data is vertically read and formed into new order indices that can be seen in Fig:

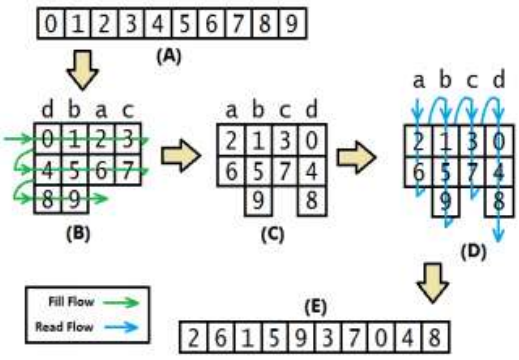

Figure 3. Indices of audio chunks of data (A), Inserting Table Column (B), Ordering Column Based-on key character (C), Reading Table Column (D), and The results of Column Encryption (E).

Same process is beginning when we do row-table encryption only row plays a major role instead of column. When we want to decrypt the message again then procedure is same as that of encryption.

4. Results

As per proposed technique we have taken the message to hide , input audio and then efficiency of this method is analyzed. Result portion contains the about original message, key value , cipher data , audio embedding and its results , PSNR, MSE values and computational time. This part highlights on the waveforms obtained as per message and audio file specifications.

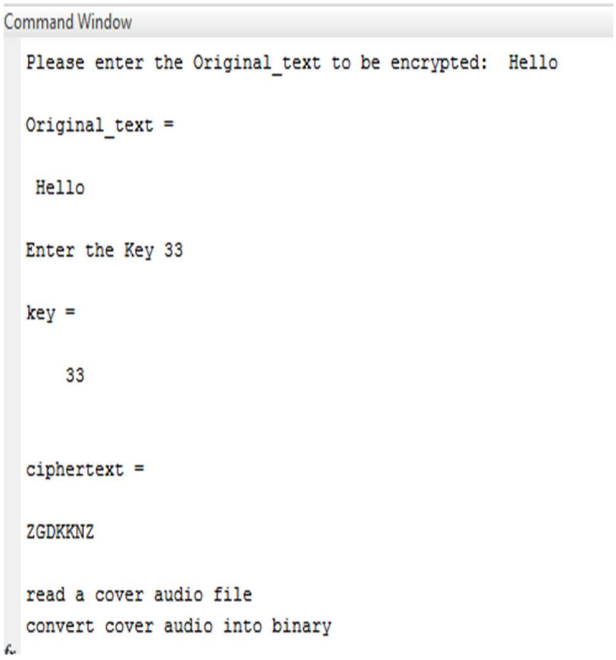

Figure.5. Results of message by using The Vigenere Square Encryption Algorithm.

Type	Wave Sound (.wav)
Size	80KB
Length	3 seconds
Bit Rate	352 kbps

Figure .6 . Audio file specifications

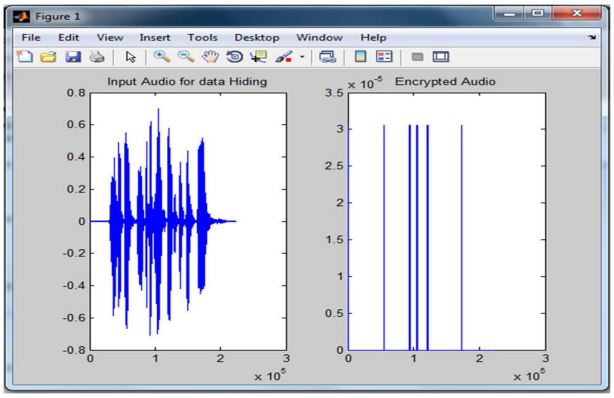

Figure. 7. input audio and encypted audio using modified LSB method

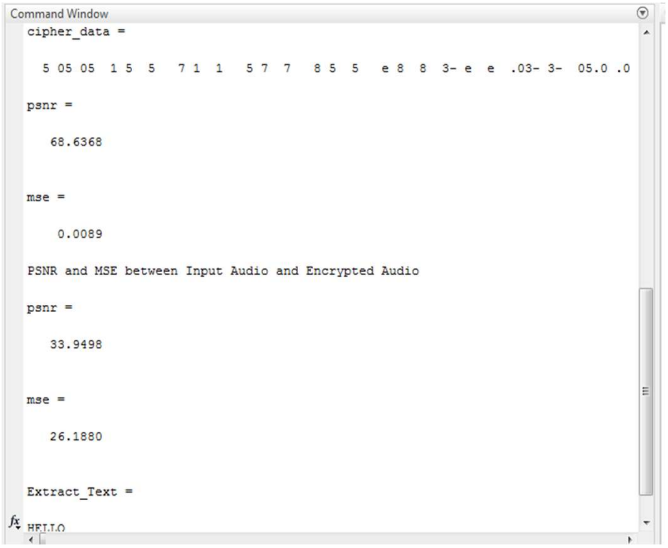

Figure. 8. results of cipher data using transpositional audio encryption , MSE , PSNR values , extract text and computation time

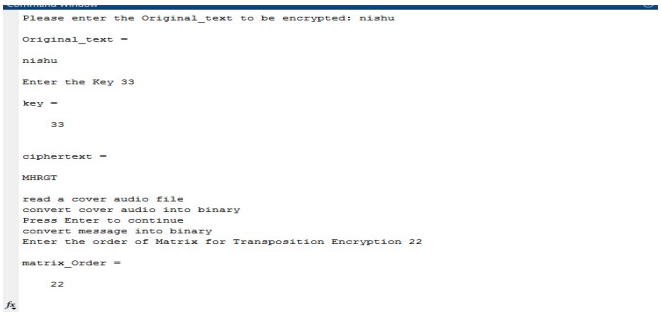

Figure.9. Results of message by using The Vigenere Square Encryption Algorithm.

Type	Wave Sound (.wav)
Size	1.20MB
Length	57seconds
Bit Rate	176 kbps

Figure.10 .Audio file specifications

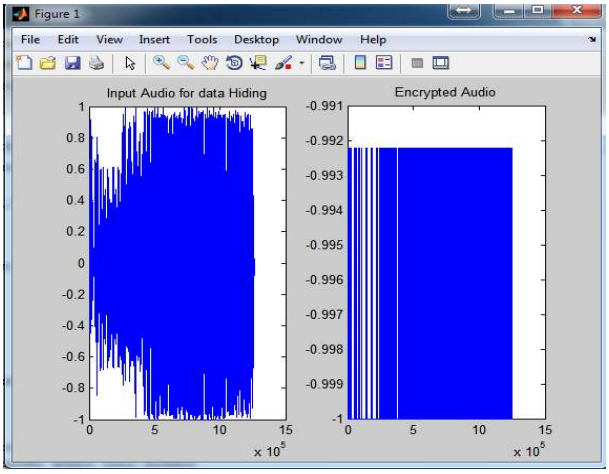

Figure. 11. input audio and encypted audio using modified LSB method

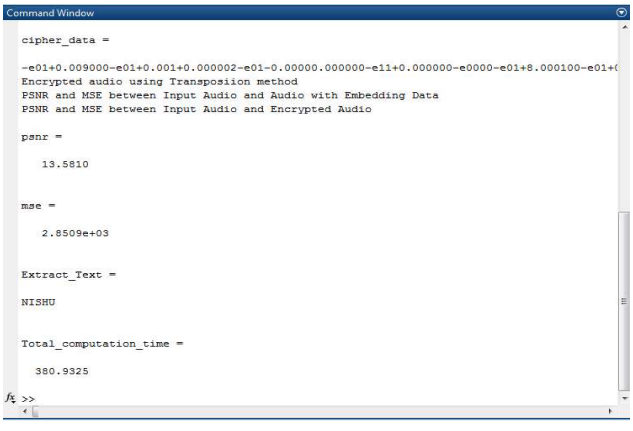

Figure. 12 .results of cipher data using transpositional audio encryption , MSE , PSNR values , extract text and computation time

AUDIO	CAPACITY(BYTES)	BIT RATE	SNR
Gtr-nylon22.wav	1.00 MB	352kbps	79.94
Hello.wav	350 KB	352kbps	91.15
input.wav	438 KB	706kbps	90.89
Position.wav	750KB	188kbps	87.76
Sheep.wav	980 KB	178kbps	85.31
Fonky.wav	1.20 MB	176kbps	79.73
Guitar.wav	890 KB	352kbps	88.67

Figure. 13 .results of cipher SNR after embedding

5. Conclusion

As discussed earlier, security of information over the internet is becoming a major concern. In this paper, a security for information from intruders using a secure approach of cryptography as well as steganography is there. The secret data is first encrypted which is embedded into an audio file and then this audio file is encrypted before being transmitted over the network. This combination of cryptography and steganography ensures that even if the audio file is intercepted by an unauthorized person, the person doesn't discover the secret information. As a part of future work, the authors recommend more secure encryption algorithms to be utilized for text encryption. Further, different steganographic techniques can also be used.

References

[1] Behrouz A. Forouzan, "Cryptography and Network Security" special Indian Edition 2007, Tata McGraw-Hill Publishing Company Limited, New Delhi.

[2] S. Das, B. Bandyopadhyay and S. Sanyal, "Steganography and Steganalysis: different approaches",
Cornell University Library, 2011.

[3] Nigam Sangwan, "Text Encryption with Huffman Compression", International Journal of Computer Applications (0975 – 8887) Volume 54– No.6, September 2012

[4] Ayushi, "A Symmetric Key Cryptographic Algorithm" International Journal of Computer Applications (09758887) Volume 1 – No. 15

[5] Wang H and Wang S, “Cyber warfare: Steganography vs. Steganalysis”, Communications of the ACM, vol. 47, no. 10, 2004

[6] K.P.Adhiya and S.A. Patil, *Hiding Text in Audio Using LSB Based Steganography*.: Information and Knowledge Management, 2012.

[7] J. Gilani, and A. Khalid M. Asad, ""An enhanced least significant bit modification technique for audio steganography”, 2011 International Conference on Computer Networks and Information Technology (ICCNIT)," *IEEE*, 2011.

[8] T Seppanen N Cvejic, "Increasing Robustness of, LSB Audio Steganography Using a Novel Embedding Method," in *, Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC04). vol. 2* , Washington, DC, USA, 2004, p. 553.

[9] Kamred Udham Singh, "A Survey on Audio Steganography Approaches," *International Journal of Computer Applications (0975 – 8887)*, p. 9, 2014.

[10] WBender D Gruhl, "Echo hiding," in *proceeding of the 1st Information Hiding Workshop*, England, 1996, pp. 295-315.

[11] M Shirali-Shahreza S Shirali-Shahreza, "Steganography in Silence Intervals of Speech," in *, proceedings of the Fourth IEEE International Conference on Intelligent Information Hiding and Multimedia Signal (IIH-MSP 2008)*, Harbin , china, 2008.

[12] S Wenndt Gopalan, "Audio Steganography for Covert Data Transmissionby Imperceptible Tone Insertion," in *WOC*, Banff , Canada, july 8-10,2004.

[13] and Tapio Seppben Nedeljko Cvejic, "Increasing The Capacity of LSB-Based Audio Steganography," *IEEE*, 2002.

[14] Kriti, and Pradeep Kumar Singh. Saroha, "A variant of LSB steganography for hiding images in audio.," *International Journal of Computer Applications*, pp. 21-27, 2010.

[15] Haider Ismael Shahadi and Razali Jidin, "High capacity and inaudibility audio steganography scheme," in *7th International Conference n Information Assurance and Security (IAS), IEEE*, 2011.

[16] V.Vijaya Bhasker, V. Shiva Md. Shafakhatullah Khan, "An Optimized Method for Concealing Data using Audio Steganography," *International Journal of Computer Applications*, pp. 0975-8887, 2011.

[17] Junaid Gilani, Adnan Khalid Muhammad Asad, "Three Layered Model," in *International Conference on Emerging*, 2012.

[18] D. M. Ballesteros L and J. M. Moreno, "Highly transparent steganography model of speech signals using Efficient Wavelet Masking," 2012.

[19] Katariya Vrushabh R, Patil Komal K Bankar Priyanka R., "Audio Steganography using LSB," *nternational Journal of Electronics*, pp. 90-92, 2012.

[20] R. A., Al-Anani, A. M., Al-Khalid, R. I., & Massadeh Al-Dallah, "An Efficient technique for data hiding in audio signals.," *American Academic & Scholarly Research Journal*, pp. 10-16, 2012.

[21] Gajanan Galshetwar, Amutha Jeyakumar Ashwini Mane, "Data Hiding Technique: Audio Steganography using LSB Technique," *International Journal of Engineering Research and Applications*, pp. 1123-1125, 2012.

[22] M. Ram Mohan Reddy S.S. Divya, "Hiding Text In Audio Using Multiple LSB Steganography And Provide Security Using Cryptography," *International Journal of Scientific &Technology*, pp. 68-70, 2012.