Priyanka Joshi, 2022, 10:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

LSB Modification Techniques of Audio Steganography for Secure Communication

M. Tech. Scholar Priyanka Joshi, Asst. Prof. Ashwini Bade
Department of E & TC,
Siddhant College of Engineering,
Pune, India.
priyanka06vyas@gmail.com, ashwinibade.ceo.44@gmail.com.

Abstract- This paper presents the application of steganography techniques for data hiding in host audio file. Audio steganography is more secured way of techniques than Text and Image Steganography due to its vast size of audio files, it can store more information than other techniques. The main aim of this paper is to present a method of embedding text-based data into a host audio file using the method of least significant bit modification for data hiding without change in quality i.e. of audio file. A method of embedding text in a host audio file through steganography is presented. In it, the audio file is sampled first and then a specific bit of each alternate sample is changed to embed the textual information in an audio file. The audio file can be considering of any type of music styles (pop, rock, techno, jazz). We propose secure communication through Audio with Text based Information using Steganography. Steganography is a method which works by changing a few bits of secret message; we will use specific bit values to represent characters. The resulting audio file will look mostly like the original. We can then send the secret message at the receiver using emails end where the message can be retrieved by knowing which bits are to be decode. In this assignment we will be writing a MATLAB application that will enable us to encode and decode secret messages with another person. The Major goal of this paper is to provide secure communication between authorized people (Sender and receiver).

Keywords- Audio steganography, secure communication, LSB modification.

I. INTRODUCTION

At this time of competition between countries or say between humans, there are many things which are needed to be kept secret or hide from the third party for this purpose there is a novel approach to audio steganography in which embedding is done without making explicit modifications to the host audio file Steganography is the specialty of concealing data in manners that forecast the recognition of stored away messages.

Steganography, got from Greek, in a real sense signifies "covered composition." So while embedding text into an audio file LSB modification creates an undetectable change in the host audio file.

In this LSB modification technique, LSB of binary equivalent of each sample of selected audio file is replaced with a binary equivalent secret message. A program has been developed which can read the audio file bit by bit and stores them in a different output audio file.

For example, if the word "Secret" has to be embedded into an audio file one has to embed the audio binary values of this word into the audio file. For this I have developed algorithm where multiple bits of each sample of the file have been changed or modified to insert text data in it. Also it is observed that the degradation of the host audio file after modification of the bits. The bit modification is done by different ways such as 1,2,3,4 bits were changed in

turn. But after checking through all modification bits, it has noticed that the best will get through 1-bit change in LSB. These methods check the MSBs of the samples, and then number of LSBs for data hiding is decided.

In this way, multiple and variable LSBs are used for embedding secret data in host audio file. This process starts from a suitable position of the data bytes. Then the editing of the least significant bit of the data has been done that have to be embedded.

Taking every alternate sample and changing the least significant bit embeds the whole message.

II. DESIGN METHODOLOGY

The initial model could be developed into an iterative model, with feedback from each phase influencing previous phases; this can be done by using "Waterfall Model". Hence, we are implementing Waterfall Stages for Designing Steganography and cryptography are cousins in the spy craft family.

Cryptography scrambles a message so it cannot be understood and generates cipher text. Steganography is the art of hiding the occurrence of data in another transmission medium to achieve secret communication. A text message in encrypted code, for instance, might arouse disbelief on the part of the recipient while an "invisible" message created with steganographic methods will not understand by third person.

We have selected an audio file with ".wav" extension has been selected as host file. It is assumed that the least significant bits of that audio file should be modified without degrading the sound quality. An MP3 is lossy and compressed. A WAV is lossless and uncompressed. An MP3 format of audio file will never sound better than a Wav, no matter what kbps it's at as it is all still lossy.

WAV files have two basic parts, the header and data. Normally in wav files, the header is situated in the first 44 bytes of file. Leaving the first 44 bytes, the rest of the bytes of the file are all about the data. The data is one chunk of samples that represents the whole audio. While embedding data, one can't deal with the header sections because a minimal change in the header section leads to a corrupted audio file.

A program has been developed which can read the audio file and stores them in a different file. The first 44 bytes should be left unchanged because these are the data of the header section. Then the remaining data field is modified for embeding textual information. For example, if the word "Secret" has to be embedded into an audio file one has to embed the binary values of the word "Secret" into the audio data field.

Table 1. Letters with ASCII values and corresponding binary values.

Letters	Ascii values	Binary values
S	82	1010010
E	68	1000100
C	66	1000010
R	81	1010001
E	68	1000100
T	83	1010011

From the table, one can come to a point that for embedding the word "Secret" into the host audio file actually the corresponding eight bit binary values have to be embedded into the data field of that audio file.

III. DESIGN ALGORITHM

To develop this algorithm multiple bits of each sample of the file have been changed or modified to insert text data in it. It has also been observed the degradation of the host audio file after modification of the bits. The bit modification was done by various ways, like 1, 2, 3, 4 bits were changed in turn. But after going through all the modification it has been observed that 1-bit change in LSB gave the best result. Thus, data can be embedded according to the following algorithm.

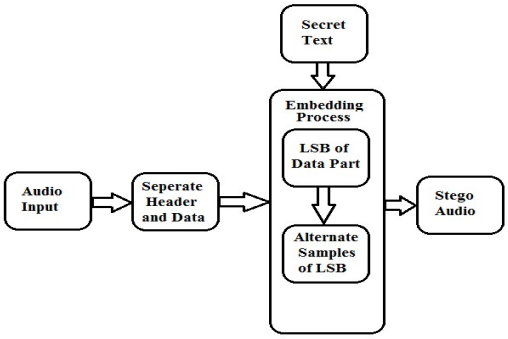

Fig 1. Algorithm (For Embedding of Data).

First audio input is to be given then the header and data of the host audio file are to be separated. Leaving the header part (that is first 50 bytes) untouched; the embedding process is done by taking every alternate sample and changing the Least Significant bit in a queue to embed the whole secret message.

The data retrieving algorithm at the receiver's end follows the same logic as the embedding algorithm with slight change. Algorithm (For Extracting of Data):

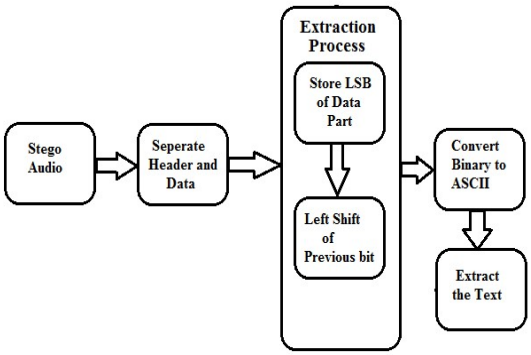

Fig 2. Algorithm (For Extracting of Data).

After the completion of the embedding process, stegno audio is obtained as the output. Now this Stegno output is the input for the extraction process. Leaving the first 50 bytes and starting from the 51st byte and store the least significant bit in a queue.

Check every alternate sample and store the least significant bit (LSB) in the previous queue with a left shift of previous bit. Convert the binary values to corresponding decimal to get the ASCII values of the secret message. From the ASCII the secret message can be found.

IV. INTERPRETATION AND RESULT

Sample No. Binary Values of corresponding sample Binary value to be embedded Binary values after Modification.

An audio file named "secret.wav" has been selected. After checking the binary values of each sample, first 50 samples were left without any changes. If the binary value of the corresponding sample is "01010010" then "0" should be modified. From Table I it can be observed that to embed the letter "S", the sender has to embed the binary value "01010010". That is why according to the embedding algorithm "S" should be embedded according to Table II.

Table 2. Sample of Audio Files with Binary Before and After Embedding.

Sample No.	Binary Values of corresponding sample	Binary value to be embedded	Binary values after Modification
51	01110100	0	01110100
53	01011110	1	01011111
55	10001011	0	10001010
57	01111011	1	01111011
59	10100010	0	10100010
61	00110010	0	00110010
63	11101110	1	11101111
65	01011100	0	01011100

Sample No. Binary value with embedded secret data Bits that are stored in Queue.

From Table II the embedding process of the letter "S" was discussed hence, in Table III, the extraction process of "S" is depicted. We have started from the 51st sample, every alternate sample is checked and the least significant bit is stored into a queue with a left shift of previous bit.

After getting all the bits in the queue, start from the left hand side, take 8 bits and convert them into corresponding decimal to get the ASCII, from the ASCII retrieve the embedded textual message. From the table, it is clearly observed that after getting 01010010 in the queue it is converted into the corresponding decimal that is 82, the ASCII of "S". Thus "S" is retrieved. Similarly, the next letters are also extracted and hence the complete word "SECRET'.

The below figure shows the graph of time and frequency domain plot of original audio file and encoded audio file. It shows graph remains same for both before and after embedding text , the overall size of the audio file remains the same. Thus, the

audio steganography is successful as there are undetectable or non-audible differences between the original and the stego audio files.

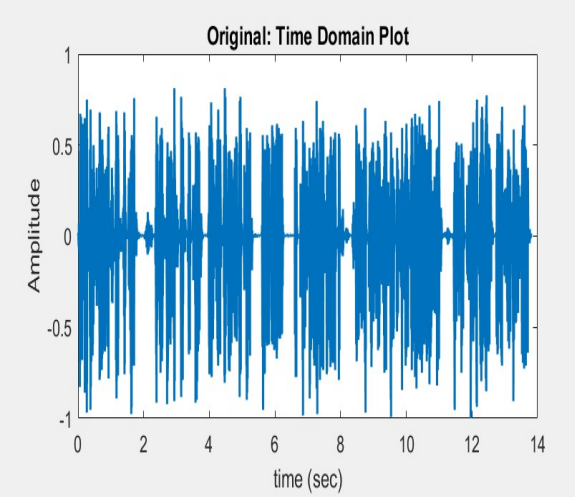

Fig 3. Chart 1: Time Domain Plot of Original Audio.

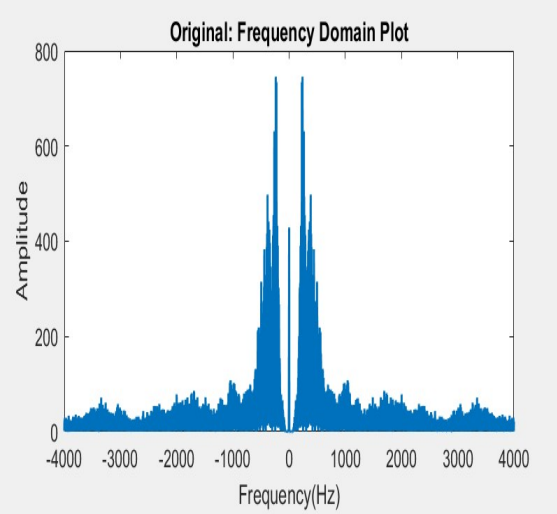

Fig 4. Chart 2: Frequency Domain Plot of Original Audio.

Table 3. Extraction of Data from Audio File.

Sample No.	Binary value with embedded secret data	Bits that are stored in Queue
51	01110100	0
53	01011111	01
55	10001010	010
57	01111011	0101
59	10100010	01010
61	00110010	010100
63	11101111	0101001
65	01011100	01010010

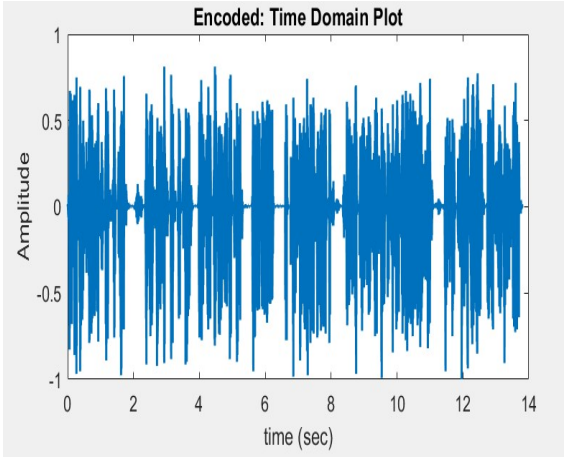

Fig 5. Chart 3: Time Domain Plot of Encoded Audio.

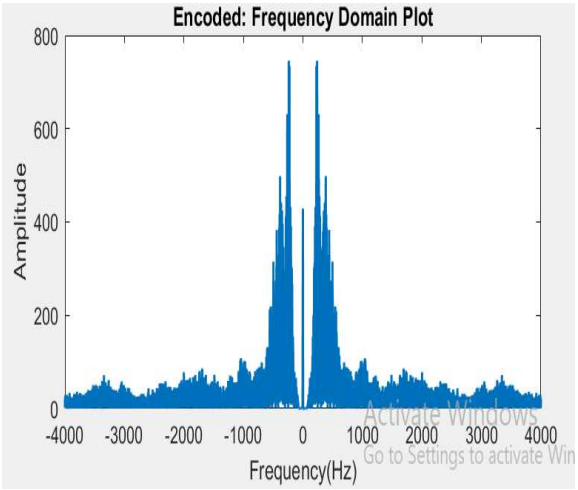

Fig 6. Chart 4: Frequency Domain Plot of Encoded Audio.

V. CONCLUSION

A method of embedding text-based data into a host audio file using the method of bit modification has been presented in this paper. A procedure has been developed in which the data field is edited to embed intended data into the audio file. To proceed with this, the header section of the audio need to check perfectly because a minimal change in the header section may leads to a corruption of whole audio file.

In this algorithm, as an experiment first 50 bytes will be left untouched and starting from the 51th bytes every alternate sample will be modified to embed textual information. How the performance is affected by changing different bit fields has not been reported in this work. However, a rough study was made to see how the changing of a specific bit field creates degradation in the host audio file and in

which point it leads to perceptible change in the audible sound quality to any other third party other than the sender or receiver.

VI. APPLICATION

The host signal should be non-objectionably degraded and the embedded data should be minimally perceptible. The embedded data should be directly encoded into the media, rather than into a header or wrapper, so that the data remain intact across varying data file formats.

Error correction coding should be used to ensure data integrity. The embedded data should be self-clocking or arbitrarily re-entrant. This ensures that the embedded data can be recovered when only fragments of the host signal are available.

Military purpose At this time of competition between countries, there are many things which are needed to be kept secret or hide from the third party, for this purpose there is a novel approach to audio steganography in which embedding is done. For example, if the military of two countries wants to communicate between each other secretly and the existence of secret message have to be unknown to the third country, so they can use this stegnography method for encoding and decoding purpose.

This secure conversation can be saved from wrong hands. An application for data hiding is tamper-proofing. It is used to indicate that the host signal has been modified from its authored state. Modification to the embedded data indicates that the host signal has been changed in some way.

Another application, feature location, requires more data to be embedded. In this application, the embedded data are hidden in specific locations within an image. It enables one to identify individual content features, e.g., the name of the person on the left versus the right side of an image. Typically, feature location data are not subject to intentional removal

VII. FUTURE SCOPE

The main aim of this paper was embedding of text into host audio as a case of steganography. The two basic criteria for successful steganography are that the stegno signal resulting from embedding is indistinguishable from the host audio signal, and the embedded message is extracted correctly at the receiver. In test cases the text-based data is embedded to the audio file to visualize in what extent the target is achieved. However, future scope appears to be endless.

REFERENCES

[1] S Hemalatha; Ramathmika "A Robust MP3 Audio Steganography with Improved Capacity" Published in 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 10.1109/ICCCA49541.2020.9250894

[2] Dingwei Tan; Yuliang Lu; Xuehu Yan; Xiaoping Wang, "A Simple Review of Audio Steganography" Published in 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 10.1109/ITNEC.2019.8729476

[3] C. Sobin; V. M. Manikandan "A Secure Audio Steganography Scheme using Genetic Algorithm" published in 2019 Fifth International Conference on Image Information Processing (ICIIP), 10.1109/ICIIP47207.2019.8985689

[4] P.N.S. Sailaja, B.Chinnakrao, K.Rajeshwari, G.V.Shridhar, "Secret Communication Through Audio For Defense Application", International Conference on Electrical and Electronics Engineering (ICEEE) - 9th Sept, 2012, Guntur-ISBN: 978-93-82208-21-1.

[5] Prof. Samir Kumar, Bandyopadhyay Barnali, Gupta Banik, "LSB Modification and Phase Encoding Technique of Audio Steganography Revisited", Vol. 1, Issue 4, June 2012.

[6] S.S. Divya, M. Ram Mohan Reddy, "Hiding text in Audio using multiple LSB Steganography and provide Security using Cryptography", Vol.1, 6th July, 2012, ISSN 2277-8616 68 IJSTR©2012.

[7] W. Bender d. Gruhl n. Morimotoa. Lu in their "Techniques for data hiding" IBM systems journal, vol 35, nos 3&4, 1996.

[8] Peter H. W. Wong, Oscar C. Au, Justy W. C. Wong in their "Data Hiding and Watermarking in JPEG Compressed Domain By DC Coefficient Modification", Department of Electrical and Electronic Engineering, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.

[9] Budda Lavanya, Yangala Smruthi, Srinivasa Rao Elisala, "Data hiding in audio by using image

International Journal of Science, Engineering and Technology

An Open Access Journal

Steganography technique" Volume 2, Issue 6, November – December 2013 ISSN 2278-6856.

[10] Ravi Saini, Rajkumar Yadav C.M.R.A., GP Sanghi Rohtak, "A New data hiding method using pixel position and logical and operation", et al international journal of computer and electronics research [volume 1, issue 1, June 2012] ISSN: 2778-5795.

[11] Samir Kumar Bandyopadhyay, Indra Kanta Maitra "An Alternative Approach of Steganography using Reference Image", International Journal of Advancements in Technology, ISSN 0976-4860 Vol 1, No 1 (June 2010)