HIDING TEXT IN AUDIO USING MULTIPLE LSB STEGANOGRAPHY AND PROVIDE SECURITY USING CRYPTOGRAPHY

S.S. Divya, M. Ram Mohan Reddy

Abstract: : Steganography is an art and science of writing hidden messages in such a way that no one apart from the intended recipient knows the existence of the message.The maximum number of bits that can be used for LSB audio steganography without causing noticeable perceptual distortion to the host audio signal is 4 LSBs, if 16 bits per sample audio sequences are used.We propose two novel approaches of substitution technique of audio steganography that improves the capacity of cover audio for embedding additional data. Using these methods, message bits are embedded into multiple and variable LSBs. These methods utilize upto 7 LSBs for embedding data.Results show that both these methods improve capacity of data hiding of cover audio by 35% to 70% as compared to the standerd LSB algorithm with 4 LSBs used for data embedding. And using encryption and decryption techniques performing cryptography. So for this RSA algorithm used.

Keywords- Information hiding,Audio steganography,Least significant bit(LSB),Most significant bit(MSB)

Introduction:

Stenography is an art of secret communication. It is sub-discipline of information hiding that focuses on concealing the existence of messages [1]. The term hiding refers to the process of making the information imperceptible or keeping the existence of the information secret. The Stenography algorithm were primarily Developed for digital images and video sequences. Interest in research in audio Stenography started. Any Stenography Technique has to satisfy two basic requirements. The first requirement is perceptual transparency that is cover object (object containing any additional data) and stego object (object containing secret messages) must be perceptually indiscernible. The second constraint is high data rate of the embedded data[1]. In a computer based audio Stenography system, secret messages are embedded in digital sound. Used audio file as a cover object to audio Stenography [1]. In audio Stenography the weakness of the Human auditory system is used to hide information in the audio. However embedding secret messages in digital sound is usually a more difficult process then embedding data in other media and can hide data into a host signal is perceptually transparent[1]. Embedding information into audio Stenography seems more secure due to less steganalysis techniques for attacking to audio. Furthermore, natural sensitivity and difficulty of working on audio and improvement in related techniques is needed. All these Stenography techniques deal with a few common types of Stenography procedure depending on the variation of the host media. That means the cover object or carrier object which will be used to hide the data.

- Mtech(VLSI),Department of ECE,PBR VITS, Kavali, A. P. Email: divya.vits10@gmail.com
- M. Ram Mohan Reddy, Assoc. Professor, Department of ECE, PBR VITS, Kavali, A. P.

Different media like images, text, video and audio has been used as a carrier in different times[2]. Audio Stenography has wide range of applications such as Covert communication, Digital water marking, access control, etc [2]. An effective of audio used audio file as a cover object to audio. Stenography scheme should possess the following Characteristics:

Inaudibility of distortion:

It evaluates the audible distortion due to signal modification like messages embedding or attacking. The audio without affecting the perceptual quality of the host audio signal.

Robustness:

It measures the ability of the embedded data to due to signal intention and unintentional attacks. Unintentional attacks generally include common data manipulations such as re-sampling, re-quantization. Intentional attacks include addition of noise, resizing etc [3].

Data rate (capacity):

It refers to the amount of information that a data hiding scheme can successfully embedded without introducing perceptual distortion. In other words the bit rate of the message is the number of the embedded bits with in a unit of time and is usually given in bits per second (bps) [3]. The maximum limit on the number of used LSBs for data embedding that do not cause audible distortion to the host audio signal restricts the amount of data for hiding purpose. As far as Stenography is concerned there is always a need to embed large amount of data in the host audio signal. Thus, in this paper we present 2 novel approaches of LSB coding that increase the capacity of cover audio so as to embed large amount of data. These methods maintain the perceptual transparency of the resulting audio signal, text, image, etc [3].

Least Significant Method:

In this technique, LSB of binary equivalent of each sample of digitized audio file is replaced with a binary equivalent secret message[4]. A program has been developed which can read the audio file bit by bit and stores them in a different file [4]. For example if the word “Audi” has to be embedded into an audio file one has to embedded the audio binary values of this word into the audio file as shown in below[5].

Table.1 Letters with ASCII values and corresponding Binary values

Letter	ASCII Value	Binary Value
A	065	01000001
U	117	01110101
D	100	01100100
I	105	01101001

From the table, one can come o point that to embed the word" Audio" into the host file actually the corresponding eight bit binary values have to be embedded into the digitized audio file where each sample is represented with 16 bits.

PROPOSED METHOD:

We use 2 novel approaches of LSBs of audio samples for data hiding. 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. These proposed methods remarkably increase the capacity for data hiding as compared to standard LSB without causing any noticeable distortion to the data [5].

Using MSB algorithm:

This method considers the value of the MSB of the digitized samples of cover audio for data hiding. Representation of the embedding procedure. The steps for data embedding and data retrieval as follow [4].

MSB bit	No of LSBs needed for data embedding
0	**6**
1	**7**

Steps for Data Embedding:

1. Read the cover audio signal.
2. Write the text in an in file to be embedded. Convert it into a sequence of binary bits.
3. Every message bit from step 2 is embedded into the variable and multiple LSBs of the samples of the digitized cover audio cover.
4. For embedding purpose, the MSB of the cover sample is checked. As shown in above table.

If MSB is '0' then use 6 LSBs for data embedding.
If MSB is '1' then use 7 LSBs for data embedding.

5. The modified cover audio samples are then written to the file forming the stego object.

Steps for Data Retrieval:

1. Read the stego object.
2. Retrieval of message bits is done by checking the MSB of the samples.

If MSB is '0' then use 6 LSBs for data retrieve.
If MSB is '1' then use 7 LSBs for data retrieve.

3. After every such 16 messages bits retrieved, they are converted into their decimal equivalents and finally the secret audio signal reconstructed.

The Capacity by the proposed method is estimated by using [5]
Cp=p1*7+p2*6
Where p1 and p2 are the probabilities of the samples with value as '1' and '0' respectively.
The Percentage increase in Capacity is given by ECP= (cp/c)*100
Where 4 bits per Sample.

Applications:-

Usage in modern printers:
Steganography is used by some modern printers, including HP and Xerox brand color laser printers. Tiny yellow dots are added to each page. The dots are barely visible and contain encoded printer serial numbers, as well as date and time stamps.
Defense Applications: Steganography is mostly used in Defense Applications.

Alleged use by intelligence services: In 2010, the Federal Bureau of Investigation revealed that the Russian foreign intelligence service uses customized steganography software for embedding encrypted text messages inside image files for certain communications with "illegal agents" (agents under non-diplomatic cover) stationed abroad.

Public key Cryptography:
In this performing encryption and decryption of data by using RSA an algorithm. The sender sign a message with its private key exchange that is two sides cooperates to exchange message[6]. In this algorithm one user uses a public key and other user uses a secret (private) key.

The arithmetic calculations of the algorithm:

1. Choose one number and divide it to two prime numbers.
2. Select p and q.
3. Calculate n=p*q.
4. Calculate Z=(p-1)(q-1)
5. Select integer 'e' which is relatively prime to 2. (Z=Φ(n))
6. Calculate e*d=1 Mod(Φ(n)) and determine 'd'.
7. Public key = {e, n} and private key={d, n}
8. For Data Encryption:
   c=M^e Mod n.
9. For Data Decryption :
   M=c^d Mod n.

Results:

Using multiple LSBs were tested with audio sequences were with sampling frequency 44100 Hz audio file represented by 16 bits per sample. Duration of the clips ranged from 2 to 8 seconds. This method analyzed in terms of PSNR (peak signal to noise ratio), incr_cap (Increased Capacity) and MSE. The results of the proposed method that consider only MSB for increasing the capacity of the cover audio, as shown in below. The estimated increase in capacity is either close or more than this estimated value and the average value is 6.774.And is they simple technique , the embedded text into audio more secure and the hidden information is recovered without any errors .

Table1: Results of the proposed Method

Cover Signal	SPEECH1
MSE	3.5e-008
PSNR	170.609
Incr_capa	6.778
Incr_capa%	169.354
SNR	55.37

Fig1: plot of Cover audio signal

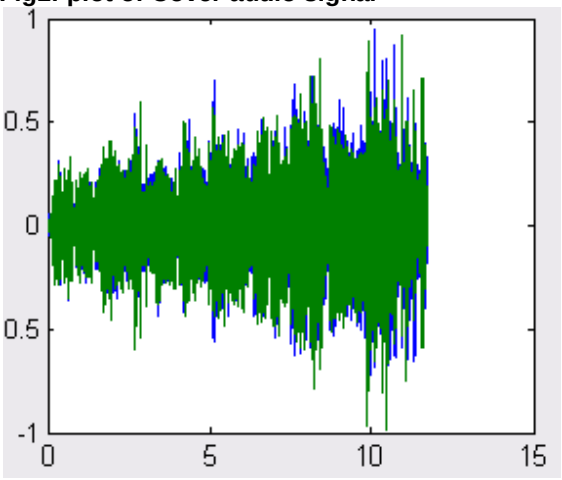
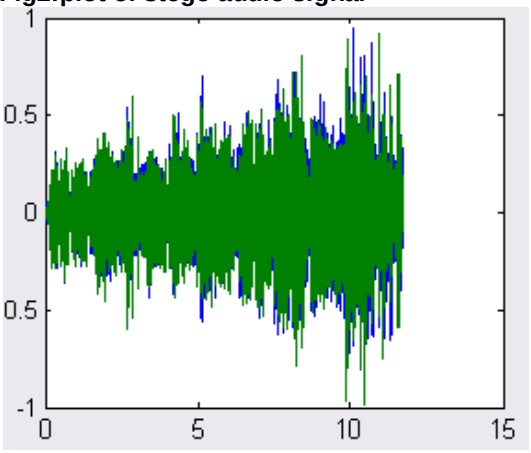


Fig2:plot of stego audio signal



Conclusion:

In this paper, two novel approaches to increase the capacity of the cover audio have been proposed. As compared to standard LSB coding method, these methods embedded data in multiple and variable LSBs depending on the MSBs of the cover audio samples. Here we are checks only the MSB of the cover sample. There is a remarkable increase in capacity of cover audio for hiding additional data and without affecting the perceptual transparency of the Text. And provide the keys concept for secure data.The main advantage of this proposed method is they are simple in logic and the hidden information is recovered without any error. Thus it succeeds in attaining the basic requirement of data hiding.

REFERENCES

[1] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding", IBM Systems Journal, vol. 35, Issues 3&4, 1996, pp. 313-336.

[2] Stefan Katzenbeisser, Fabien A. P. Petitcolas, “Information HidingTechniques for Steganography and Digital Watermarking”. Boston,Artech House, pp. 43 – 82. 2000.

[3] Chen and G.W. Womell, “Quantization index modulation: a class of provably good methods for digital watermarking and information embedding”,IEEE Transactions on Information Theory, Vol. 47, No. 4,pp. 1423-1443, May 2001.

[4] C. C. Chang, T. S. Chen and H. S. Hsia, “An Effective Image Stenographic Scheme Based on Wavelet Transform and Pattern- Based Modification”, IEEE Proceedings of the 2003 International Conference on Computer Networks and Mobile Computing, 2003.

[5] K. Gopalan, “Audio steganography using bit modification”, Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Vol. 2, pp. 421-424, April 2009.

[6] M. M. Amin, M. Salleh, S. Ibrahim, M. R. Katmin and M. Z. I. Shamsuddin, “Information Hiding using Steganography”, 4th National Conference on elecommuni cation Technology Proceedings, Shah Alam, Malaysi.