A Novel Audio Steganography Scheme using Amplitude Differencing

Shafi K[1], Sankaranarayanan A[2],Prashanth G[3],Akash Mohan[4]
Amrita Vishwa Vidyapeetham
Coimbatore, India
{[1]shaff.k, [2]sankar.narayanan8, [3]prashanth.amrita, [4]akashmhn}@gmail.com

Abstract—**A novel method for embedding a covert message in cover audio for secure communication is proposed. The covert message to be embedded can be of any format. In this process, two cover audio files are taken and difference of the amplitude values is calculated. These differences are then classified into number of ranges and are substituted with a new value to embed the secret message. The embedded secret message can be extracted in its original format from the resulting two stego-audio files without referring the original audio files. The secret message can be encrypted for added security. Experimental results show that the technique achieves imperceptible embedding, large payload, and accurate data retrieval.**

KeyWords- Audio Steganography; Data Hiding; PVD; Steganalysis

I. INTRODUCTION

Any data can be represented as digital data. With the advancement of the technologies and the internet, storage and communication via digital data has gained a lot of significance. As a direct consequence, the need for data security in digital communications, copyright protection of digitized properties and secure communication has also gained equal importance [1].

The term steganography is the technique of embedding secret information in a communication channel in such a manner that the very existence of the information is concealed [2]. The aim is to embed and deliver secret messages in digital data without any suspiciousness. The secret message might be a caption, a plain text, another image, a control signal, or any data that can be represented in bit stream form. The secret message may be compressed and encrypted before the embedding begins.

Steganography techniques have been successfully applied on text files, images, audio and video files. Although, steganography in audio is a field not much explored. It is because of the fact that embedding secret messages in digital sound is usually a more difficult process than embedding messages in other media, such as digital images [3].

Audio Steganography, or information hiding in audio signals, is gaining widespread importance for secure communication of information such as covert battlefield data and banking transactions via open audio channels [1].

A Steganography system, in general, is expected to meet three key requirements, namely, imperceptibility of embedding, correct recovery of embedded information, and large payload. Some degradation in the perceptual quality of the stego-signal from that of the original host signal may be acceptable. Practical audio embedding systems, however, face hard challenges in fulfilling all three requirements due to the large power and dynamic range of hearing, and the large range of audible frequency of the human auditory system [1].

Audio Steganography relies on the imperfection of the human auditory system. In several audio steganography techniques, the secret message is embedded in the frequency range which is in either higher or lower than the frequency range which can be audible to the human ears. Several other Audio steganography techniques take advantage of the psychoacoustic masking phenomenon of the human auditory system [HAS]. Frequency masking occurs when human ear cannot perceive frequencies at lower power level if these frequencies are present in the vicinity frequencies at higher level [4]. But such techniques are often vulnerable to many steganalysis techniques which can detect presence of such frequencies. In this paper we present a scheme that meets all the requirements and is imperceptible to many known Steganalysis techniques. This method is an extension of work initiated by Da-Chun Wu and Wen-Hsaing Tsai who developed and algorithm for embedding covert data in images [5].

In following section, we describe the method of embedding and retrieving covert data in audio signals. In section III the experimental results are discussed. In section IV, the work is summarized, conclusions are drawn and possibilities for further results are indicated.

II. PROPOSED SYSTEM

Hiding data in the LSBs of the values of amplitude of audio signals is a common information hiding method. As we increase the number of LSBs used to embed the data, more distortion is produced in the resulting stego-audio signal. In the proposed system, we use two cover audio signals. There are two advantages of using more than one cover audio signals – it distributes the payload equally among the cover audio signals, hence preventing the addition of noise and it provides additional security as the secret message cannot be extracted with just one of the stego-audio signals.

978-1-4244-9008-0/10/$26.00 ©2010 IEEE

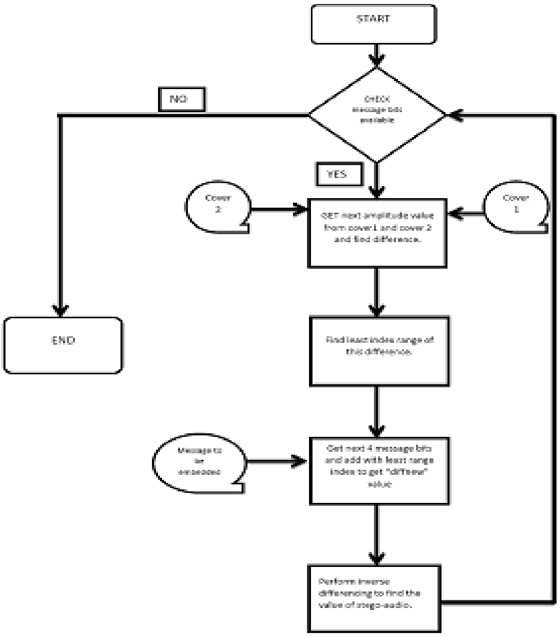
START
CHECK message bits available
NO
YES
END
Cover 2
Cover 1
GET next amplitude value from cover1 and cover 2 and find difference.
Find least index range of this difference.
Message to be embedded
Get next 4 message bits and add with least range index to get "diffnew" value
Perform inverse differencing to find the value of stego-audio.

Figure 1.Flow chart for embedding algorithm

The flowchart of the proposed embedding method is shown in “Fig.1”. The process of calculation of range indices and data embedding are described subsequently.

A. Calculation of range indices

When the data is embedded in the last two bits of the audio signal, the distortion produced is very minimal. In this system, the data is embedded in the last two bits of both of the cover audio signals. Therefore, totally 4 bits are available for embedding the secret message. We consider the secret message as a long sequence of 8 bit streams. Since all files of all formats are saved in the memory as bytes, all the files can be represented as a long sequence of bit streams. Hence the secret message to be embedded may be of any format.

The 8 bit stream of the secret message is divided into two 4 bit streams and embedded in the cover audio signals. Since 4 bits are available, a total of 16 values ranging from 0 to 15 can be embedded. Hence the least range index of range R from 0 through 255 varies as {0, 15, 31, 47...255}.

B. Data Embedding

In the proposed system, the audio signals are considered as byte streams rather than continuous wave signal which is logical since the audio signals are saved as bytes in memory. First, two cover audio signals are taken, preferably of almost equal size. Let the amplitude values of the audio signals be denoted by $a1_i$ and $a2_i$. The difference between each amplitude value D_i is calculated. The difference value D_i is compared to each of the max range index R_m of range R. When D_i is less than R_m, the value the least range index R_l of range R is assigned to ND_i. The 8 bit stream of the secret message is divided into two 4 bit streams. The value of the 4 bit stream S_j of secret message is added to ND_i to obtain the new amplitude difference value.

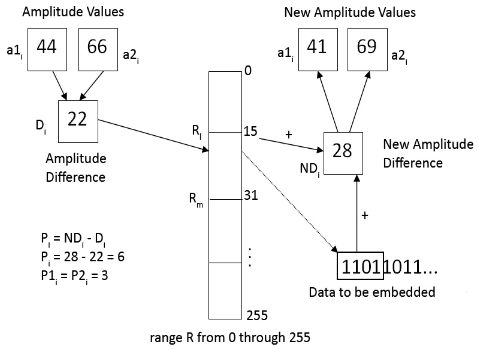

Figure 2.Embedding process

The difference P_i between the values ND_i and D_i is calculated. Since the range varies as multiples of 16, the maximum value of P_i can be 15. Hence P_i can be represented in 4 bits. These 4 bits P_i is represented as sum of two 2 bit values namely $P1_i$ and $P2_i$. $P1_i$ is then subtracted from the amplitude value $a1_i$ and $P2_i$ is added to the amplitude value $a2_i$ to get the desired stego-audio signal. This process is termed as Inverse Differencing (Fig.2).

C. Data Extraction

For extraction of the embedded data, the two stego-audio signals are taken. The difference D_i between each amplitude values $a1_i$ and $a2_i$ of the two stego-signals is calculated. The difference value D_i is compared to each of the max range index R_m of range R. When D_i is less than R_m, the value the least range index R_l of range R is assigned to ND_i. The difference between ND_i and D_i gives the value of the hidden data.

III. EXPERIMENTAL RESULTS

This section introduces the test results achieved. To evaluate the performance of the proposed Steganalysis algorithm, we randomly picked 300 wav files from the wav surfer database [6]. These entire wav files have different audio characteristics. The sampling rate is 44:1 kHz with 16 bits per sample. The size audio files vary from 100KB to 5MB. The covert data to be embedded is represented in bits.

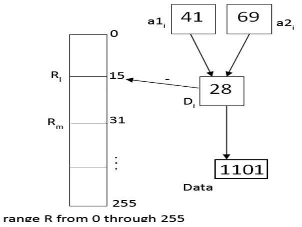

Figure 3.Extraction process

A. *Efficiency of retrieval*

TABLE I. Efficiency of Retrival

Embedded File Format	Efficiency of Retrieval
Plain Text	100%
Rich Text Format	100%
Power point presentation	100%
PDF	100%
Audio Files(.wav,.au,.mp3)	100%
Executable File	100%

The above results were tested and verified. This can be achieved with any file format since all data formats can be represented in byte streams.

B. *Embedding Ratio*

On two 1024 KB audio cover files, the maximum size of image file that was successfully embedded and retrieved was 256 KB. So the maximum embedding ratio was calculated to be 12.5%.

C. *Temporal Centroid (TC)*

The Temporal Centroid is defined as the time average over the energy envelope of the signal. The unit of TC feature is second.

From the table II, it is conspicuous that there is no change in TC in cover and stego-audio signals.

D. *Average Power Spectrum*

Average Power spectrum describes the temporally smoothed instantaneous power of an audio signal. It is a measure of the evolution of the amplitude of the signal as function of time.

The graphs in “Fig.4a” and “Fig.4b” give the power average (over time) of the spectra as a function of frequency, expressed in decibels. We can observe minor variations in the Audio power spectrum of the two audio files.

E. *Standard Deviation*

Standard deviation is used as a measure of dispersion occurring in a given data set. “Fig.5a” and “Fig.5b” show the standard deviation of frequencies with respect to time. It can be perceived that there is negligible difference between the two graphs.

TABLE II

Samples	Temporal Centroid Value
Before Embedding	3.5521
After Embedding	3.5521

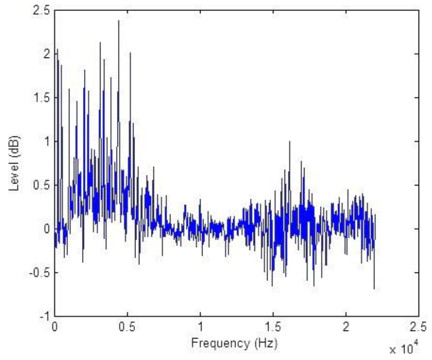

Figure 4a.Average Power Spectrum of a Cover Audio Signal

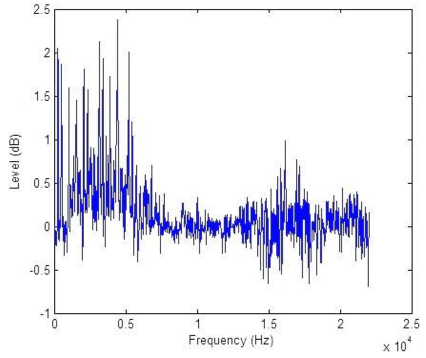

Figure 4b.Average Power Spectrum of a Stego-Audio Signal

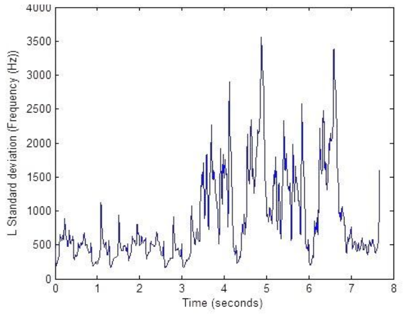

Figure 5a.Standard Deviation of a Cover Audio Signal

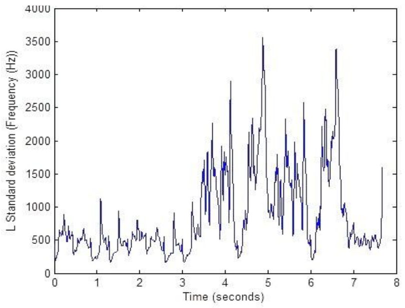

Figure 5b. Standard Deviation of a Stego-Audio Signal

F. Kurtosis

Kurtosis is the fourth moment about the mean divided by square of variance. It is a higher order moment which have been used to classify steganographic audio signals from cover signals.

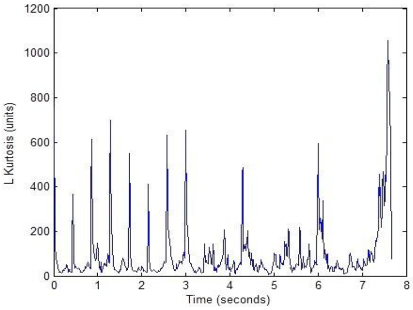

Figure 6a.Kurtosis of a Cover Audio Signal

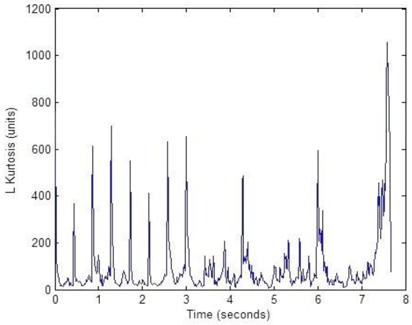

Figure 6b. Kurtosis of a Stego-Audio Signal

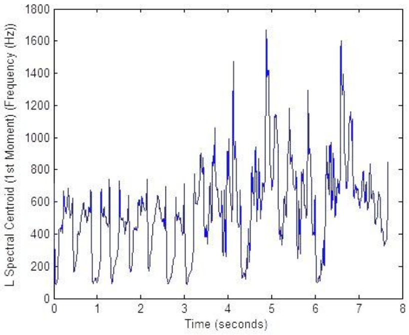

Figure 7a. Audio Spectrum Centroid of a Cover Audio Signal

There is negligible difference between the kurtosis of the cover audio (Fig. 6a) and the embedded audio signals (Fig. 6b). This would make it difficult for classifiers which use kurtosis as a parameter for classification of audio stego signals to identify stego audio signal.

G. Audio Spectrum Centroid (ASC)

Audio Spectrum centroid gives the center of gravity of a log-frequency power spectrum. The log-frequency scaling approximates the perception of frequencies in the human hearing system.

The graphs in "Fig.7a" and "Fig.7b" show negligible differences. Since the ASC is an approximation of the sensitivity of human auditory system to different frequencies, it can be inferred that human auditory system is incapable of perceiving any difference between cover and stego-audio signals.

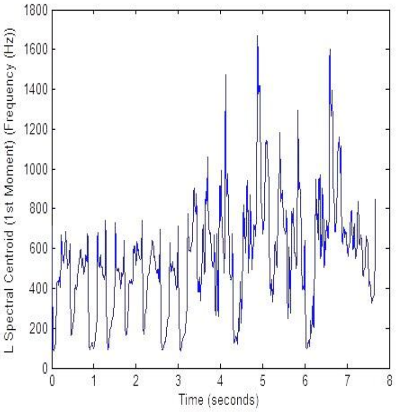

Figure 7b. Audio Spectrum Centroid of a Stego-Audio Signal

IV. Conclusion and Future Works

A new and efficient method for embedding covert messages in audio signals without producing perceivable changes has been proposed. In this method, it is not required to refer to the cover audio signals while extracting the embedded data from stego-audio signals. Covert data is embedded into cover audio signals by replacing the difference value of the amplitude of the two cover audio signals with a new difference in which bits of embedded data are included. The experimental results show that the method meets all requirements for robustness including large payload and imperceptibility. In addition, it provides security as both stego-signals are required to retrieve covert data. In future, this method can be extended to TriPVD [7] algorithm. The same algorithm can also be used to embed covert files in files of other formats.

References

[1] Kaliappan Gopalan, "Audio Steganography by Cepstrum Modification", Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '05), Vol. 5, pp. 421-424, May 2005.

[2] Merriam-Webster Online Dictionary. http://www.merriam-webster.com/dictionary/steganography

[3] Samir K Bandyopadhay, Debnath Bhattacharya, Debashis Gangly, Swarnendu Mukherjee, Poulami Das, "A Tutorial Review on Steganography", Proc. IC 200

[4] Kaliappan Gopalan and Stanley Wenndt, "Audio Steganography For Covert Data Transmission By Imperceptible Tone Insertion", Proc. International Association of Science And Technology For Development (IASTED 2004)

[5] Da-Chun Wu, Wen-Hsaing Tsai, “A steganographic method for images by pixel-value differencing”, Pattern Recognition Letters 2003 -Elsevier

[6] Wave Surfer Database
http://www.wavsurfer.org

[7] Ko-Chin Chang, Chien-Ping Chang, Ping S. Huang, Te-Ming Tu, “A Novel Image Steganographic Method Using Tri-way Pixel-Value Differencing ”, Journal of Multimedia, Vol 3, pp 37-44, Jun 2008