Data Hiding in Audio Signal: A Review

Article · January 2009

3 authors, including:

Debnath Bhattacharyya
Vignan's Institute of Information Technology
239 PUBLICATIONS   **1,177** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   software testing View project

Project   A review on bio medical image processing View project

Data Hiding in Audio Signal: A Review

Poulami Dutta[1], Debnath Bhattacharyya[1], and Tai-hoon Kim[2]

[1]*Heritage Institute of Technology, Kolkata-700107, India*
{poulamiduttacse,debnathb}@gmail.com
[2]*Hannam University, Daejeon, Korea*
taihoonn@empal.com

Abstract

Information hiding technique is a new kind of secret communication technology. The majority of today's information hiding systems uses multimedia objects like audio. Embedding secret messages in digital sound is usually a more difficult process. Varieties of techniques for embedding information in digital audio have been established. In this paper we will attend the general principles of hiding secret information using audio technology, and an overview of functions and techniques.

1. Introduction

The fast improvement of the Internet and the digital information revolution caused major changes in the overall culture. Flexible and simple-to-use software and decreasing prices of digital devices (e.g. portable CD and mp3players, DVD players, CD and DVD recorders, laptops, PDAs) have made it feasible for consumers from all over the world to create, edit and exchange multimedia data. Broadband Internet connections almost an errorless transmission of data helps people to distribute large multimedia files and make identical digital copies of them. In modern communication system Data Hiding is most essential for Network Security issue. Sending sensitive messages and files over the Internet are transmitted in an unsecured form but everyone has got something to keep in secret. Audio data hiding method is one of the most effective ways to protect your privacy.

2. Overview

General principles of data hiding technology, as well as terminology adopted at the First International Workshop on Information Hiding, Cambridge, U.K. [1] are illustrated in Figure 1. A data message is hidden within a cover signal (object) in the block called embeddor using a stego key, which is a secret set of parameters of a known hiding algorithm. The output of the embeddor is called stego signal (object). After transmission, recording, and other signal processing which may contaminate and bend the stego signal, the embedded message is retrieved using the appropriate stego key in the block called extractor [2].
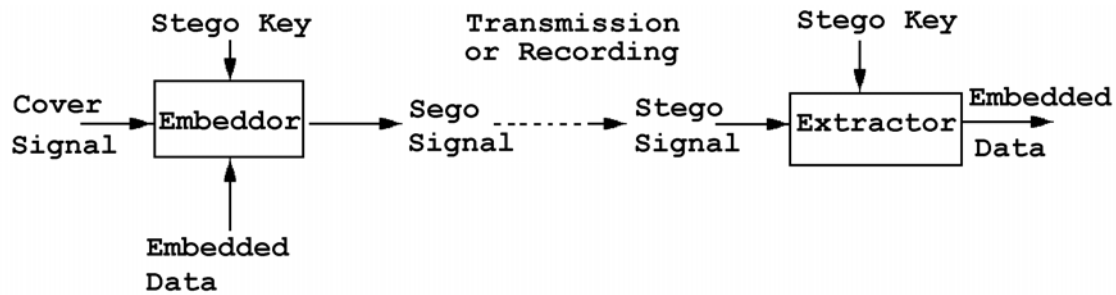
Figure 1. Block diagram of data hiding and retrieval.

A number of different cover objects (signals) can be used to carry hidden messages. Data hiding in audio signals exploits imperfection of human auditory system known as audio masking. In presence of a loud signal (masker), another weaker signal may be inaudible, depending on spectral and temporal characteristics of both masked signal and masker [3] Masking models are extensively studied for perceptual compression of audio signals [2] In the case of perceptual compression the quantization noise is hidden below the masking threshold, while in a data hiding application the embedded signal is hidden there. Data hiding in audio signals is especially challenging, because the human auditory system operates over a wide dynamic range. The human auditory system perceives over a range of power greater than one billion to one and a range of frequencies greater than one thousand to one. Sensitivity to additive random noise is also acute. The perturbations in a sound file can be detected as low as one part in ten million (80 dB below ambient level).However, there are some "holes" available. While the human auditory system has a large dynamic range, it has a fairly small differential range. As a result, loud sounds tend to mask out quiet sounds. Additionally, the human auditory system is unable to perceive absolute phase, only relative phase. Finally, there are some environmental distortions so common as to be ignored by the listener in most cases [4]. Now we will discuss many of these methods of audio data hiding technology.

3. Previous works

This section presents some common methods used for hiding secret information in audio. Many software implementations of these methods are available on the Web and are listed in the relatives section. Some of the latter methods require previous knowledge of signal processing techniques, Fourier analysis, and other areas of high level mathematics. When developing a data-hiding method for audio, one of the first considerations is the likely environments the sound signal will travel between encoding and decoding. There are two main areas of modification which we will consider. First, the storage environment, or digital representation of the signal that will be used, and second the transmission pathway the signal might travel [4].

3.1. Parity coding

One of the prior works in audio data hiding technique is parity coding technique. Instead of breaking a signal down into individual samples, the parity coding method breaks a signal down into separate regions of samples and encodes each bit from the secret message in a sample region's parity bit. If the parity bit of a selected region does not match the secret bit to be encoded, the process flips the LSB of one of the samples in the region. Thus, the sender has more of a choice in encoding the secret bit, and the signal can be changed in a more unobtrusive fashion [5]. Figure 2, shows the parity coding procedure.

3.2. Phase Coding

The phase coding method works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments. Phase coding, when it can be used, is one of the most effective coding methods in terms of the signal-to perceived noise ratio. When the phase relation between each frequency component is dramatically changed, noticeable phase dispersion will occur. However, as long as the modification of the phase is sufficiently small (sufficiently small depends on the observer; professionals in broadcast radio can detect modifications that are imperceivable to an average observer), an inaudible coding can be achieved [4]. . Phase coding relies on the fact that the phase components of sound are not as perceptible to the human ear as noise is. Rather than introducing perturbations, the technique encodes the message bits as phase shifts in the phase spectrum of a digital signal, achieving an inaudible encoding in terms of signal-to-perceived noise ratio [5].
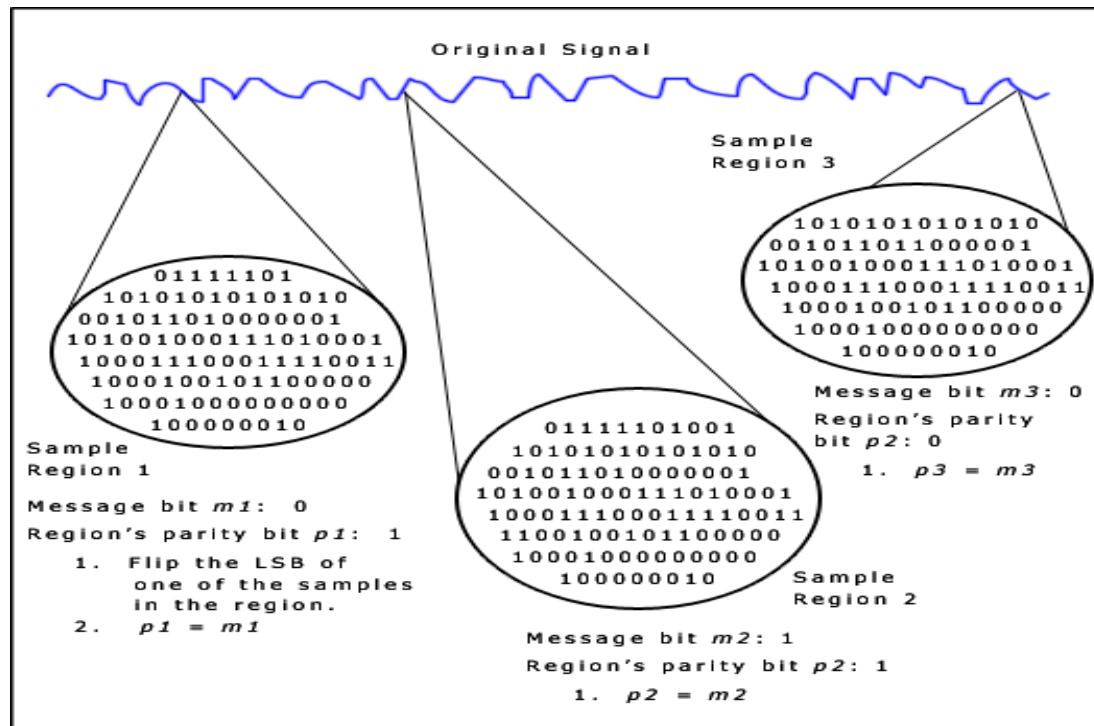

Figure 2. Parity Coding Procedure.

Phase coding is explained in the following procedure:

    a. The original sound signal is broken up into smaller segments whose lengths equal the size of the message to be encoded.

    b. A Discrete Fourier Transform (DFT) is applied to each segment to create a matrix of the phases and Fourier transform magnitudes.

    c. Phase differences between adjacent segments are calculated.

    d. Phase shifts between consecutive segments are easily detected. In other words, the absolute phases of the segments can be changed but the relative phase differences

between adjacent segments must be preserved. Therefore the secret message is only inserted in the phase vector of the first signal segment as follows:

$$phase\_new = \begin{cases} \pi/2 & if \ \ message \ \ bit = 0 \\ -\pi/2 & if \ \ message \ \ bit = 1 \end{cases}$$

e. A new phase matrix is created using the new phase of the first segment and the original phase differences.
f. Using the new phase matrix and original magnitude matrix, the sound signal is reconstructed by applying the inverse DFT and then concatenating the sound segments back together.

To extract the secret message from the sound file, the receiver must know the segment length. The receiver can then use the DFT to get the phases and extract the information (consider Figure 3 for phase cosing procedure).
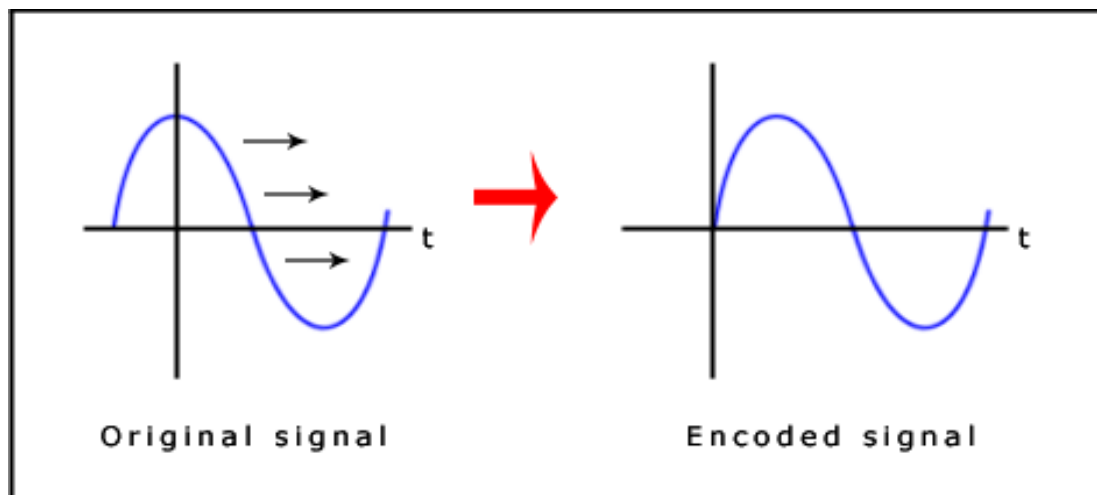

Figure 3. The signals before and after Phase coding procedure.

3.3. Spread Spectrum

In a normal communication channel, it is often desirable to concentrate the information in as narrow a region of the frequency spectrum as possible in order to conserve available bandwidth and to reduce power. The basic spread spectrum technique, on the other hand, is designed to encode a stream of information by spreading the encoded data across as much of the frequency spectrum as possible. This allows the signal reception, even if there is interference on some frequencies. While there are many variations on spread spectrum communication, we concentrated on Direct Sequence Spread Spectrum encoding (DSSS). The DSSS method spreads the signal by multiplying it by a chip, a maximal length pseudorandom sequence modulated at a known rate. Since the host signals are in discrete-time format, we can use the sampling rate as the chip rate for coding. The result is that the most difficult problem in DSSS receiving, that of establishing the correct start and end of the chip quanta for phase locking purposes, is taken care of by the discrete nature of the signal. Consequently,

a much higher chip rate, and therefore a higher associated data rate, is possible. Without this, a variety of signal locking algorithms may be used, but these are computationally expensive [4].

Procedure: In DSSS, a key is needed to encode the information and the same key is needed to decode it. The key is pseudorandom noise that ideally has flat frequency response over the frequency range, i.e., white noise. The key is applied to the coded information to modulate the sequence into a spread spectrum sequence.

The DSSS method: The code is multiplied by the carrier wave and the pseudorandom noise sequence, which has a wide frequency spectrum. As a consequence, the spectrum of the data is spread over the available band. Then, the spread data sequence is attenuated and added to the original file as additive random noise (see Figure 4). DSSS employs bi-phase shift keying since the phase of the signal alternates each time the modulated code alternates (see Figure 5). For decoding, phase values f0 and f0 + p are interpreted as a "0" or a "1," which is a coded binary string [4]. Spread Spectrum is shown in Figure 5.

In the decoding stage, the following is assumed:
    a. The pseudorandom key is maximal (it has as many combinations as possible and does not repeat for as long as possible). Consequently it has a relatively flat frequency spectrum.
    b. The key stream for the encoding is known by the receiver. Signal synchronization is done, and the start/stop point of the spread data is known.
    c. The following parameters are known by the receiver: chip rate, data rate, and carrier frequency.
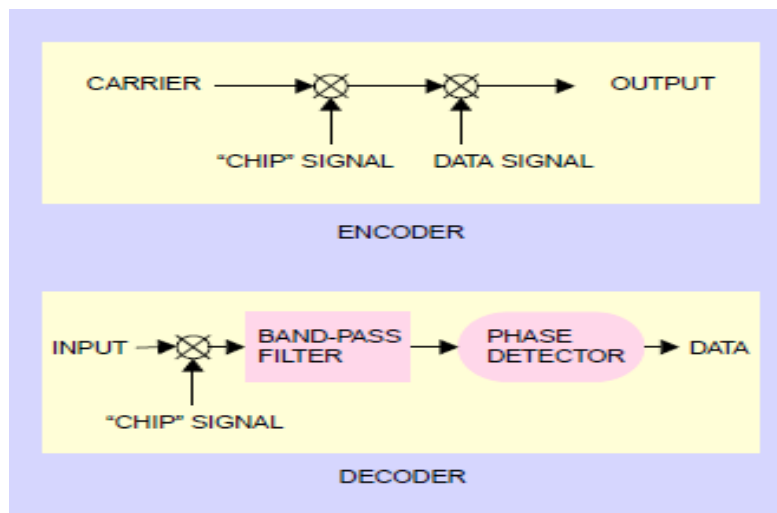

Figure 4. Spread spectrum encoding.

3.4. Echo Hiding

In echo hiding, information is embedded in a sound file by introducing an echo into the discrete signal. Like the spread spectrum method, it too provides advantages in that it allows for a high data transmission rate and provides superior robustness when compared to the noise inducing methods. If only one echo was produced from the original signal, only one bit of information could be encoded. Therefore, the original signal is broken down into blocks

before the encoding process begins. Once the encoding process is completed, the blocks are concatenated back together to create the final signal [5]. Echo Hiding is shown in Figure 6.

Also, a message can be encoded using musical tones with a substitution scheme. For example, a Fistone will represent a 0 and a C tone represents a 1. A normal musical piece can now be composed around the secret message or an existing piece can be selected together with an encoding scheme that will represent a message [7, 8].
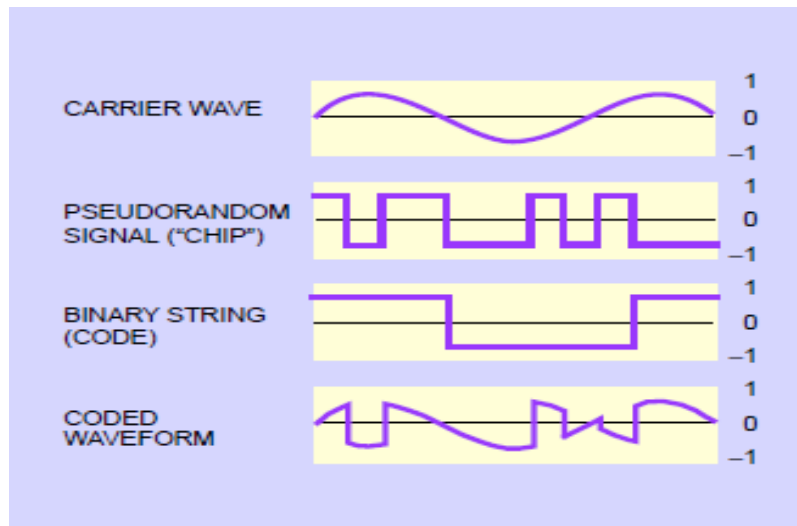

Figure 5. Synthesized spread spectrum information encoded by the direct sequence method
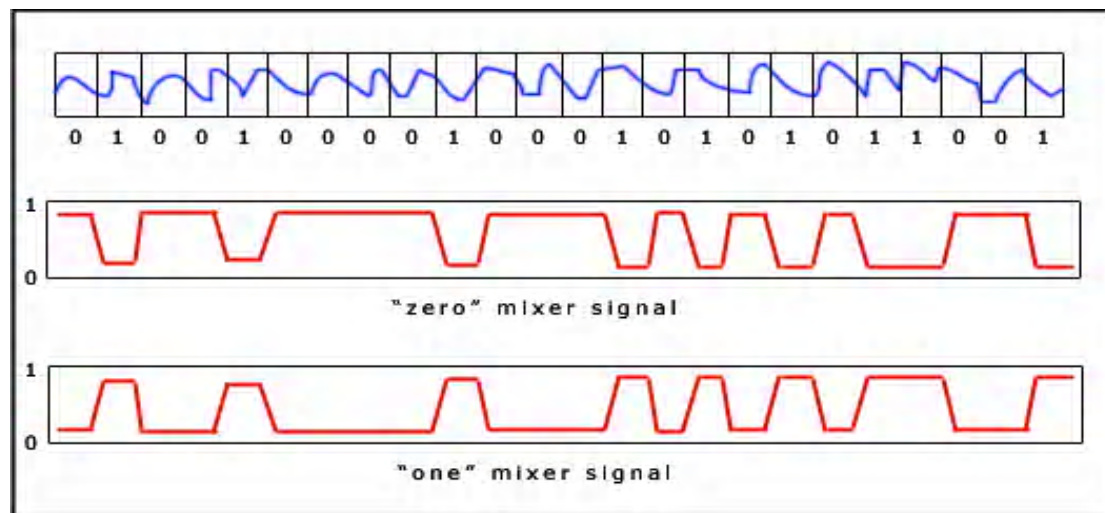

Figure 6. Echo hiding

4. Proposed work

Here we will discuss the disadvantages of the previous procedure and how those are different with present method. There are two main disadvantages associated with the use of methods like parity coding. The human ear is very sensitive and can often detect even the slightest bit of noise introduced into a sound file, although the parity coding method does

come much closer to making the introduced noise inaudible. Another problem is robustness. One disadvantage associated with phase coding is a low data transmission rate due to the fact that the secret message is encoded in the first signal segment only. Phase coding method is used when only a small amount of data needs to be considered.

Least significant bit (LSB) coding is the simplest way to embed information in a digital audio file. By substituting the least significant bit of each sampling point with a binary message, LSB coding allows for a large amount of data to be encoded. Among many different data hiding techniques proposed to embed secret message within audio file, the LSB data hiding technique is one of the simplest methods for inserting data into digital signals in noise free environments, which merely embeds secret message-bits in a subset of the LSB planes of the audio stream.

The following steps are:

    a.   Receives the audio file in the form of bytes and converted in to bit pattern.
    b.   Each character in the message is converted in bit pattern.
    c.   Replaces the LSB bit from audio with LSB bit from character in the message.

This proposed system is to provide a good, efficient method for hiding the data from hackers and sent to the destination in a safe manner. This proposed system will not change the size of the file even after encoding and also suitable for any type of audio file format. Encryption and Decryption techniques have been used to make the security system robust.

Low-bit encoding embeds secret data into the least significant bit (LSB) of the audio file. The channel capacity is 1KB per second per kilohertz (44 kbps for a 44 KHz sampled sequence). This method is easy to incorporate.

5. Applications

Audio data hiding can be used anytime you want to hide data. There are many reasons to hide data but most important is to prevent unauthorized persons from becoming aware of the existence of a message. In the business world Audio data hiding can be used to hide a secret chemical formula or plans for a new invention. Audio data hiding can also be used in corporate world.

Audio data hiding can also be used in the noncommercial sector to hide information that someone wants to keep private. Terrorists can also use Audio data hiding to keep their communications secret and to coordinate attacks. In the project ARTUS1 which aims to embed animation parameters into audio and video contents [10]. Data hiding in video and audio, is of interest for the protection of copyrighted digital media, and to the government for information systems security and for covert communications. It can also be used in forensic applications for inserting hidden data into audio files for the authentication of spoken words and other sounds, and in the music business for the monitoring of the songs over broadcast radio.

6. Conclusion

In this paper we have introduced a robust method of imperceptible audio data hiding. This system is to provide a good, efficient method for hiding the data from hackers and sent to the destination in a safe manner. This proposed system will not change the size of the file even after encoding and also suitable for any type of audio file format. Thus we conclude that audio data hiding techniques can be used for a number of purposes other than covert communication or deniable data storage , information tracing and finger printing, tamper detection. As the sky is not limit so is not for the development. Man is now pushing away its own boundaries to make every thought possible. So similarly these operations described above can be further modified as it is in the world of Information Technology. After

designing any operation every developer has a thought in his mind that he could develop it by adding more features to it.

Acknowledgement

References

[1] B. Pfitzmann, "Information Hiding Terminology", First International Workshop on Information Hiding, May 30 – June 1, 1996, Cambridge, UK, pp. 347-350.

[2] Rade Petrovi, Kanaan Jemili, Joseph M. Winograd, Ilija Stojanovi, Eric Metois, "DATA HIDING WITHIN AUDIO SIGNALS", June 15, 1999, MIT Media Lab, Series: Electronics and Energetics vol. 12, No.2, pp. 103-122.
http://pubs.media.mit.edu/?section=docdetail&id=211474&collection=Media+Lab&filtercollection=Media+Lab

[3] J. Johnston and K. Brandenburg, "Wideband Coding Perceptual Consideration for Speech and Music". Advances in Speech Signal Processing, S. Furoi and M. Sondhi, Eds. New York: Marcel Dekker, 1992.

[4] W. Bender, W. Butera, D. Gruhl, R. Hwang, F. J. Paiz, S. Pogreb, "Techniques for data hiding", IBM Systems Journal, Volume 39 , Issue 3-4, July 2000, pp. 547 – 568.

[5] Samir Kumar Bandyopadhyay, Debnath Bhattacharyya, Poulami Das, Debashis Ganguly and Swarnendu Mukherjee, "A tutorial review on Steganography", International Conference on Contemporary Computing (IC3-2008), Noida, India, August 7-9, 2008, pp. 105-114.

[6] Steve Czerwinski, Richard Fromm, Todd Hodes, "Digital Music Distribution and Audio Watermarking".
http://reference.kfupm.edu.sa/content/d/i/digital_music_distribution_and_audio_wat_1045219.pdf
http://http.cs.berkeley.edu/~hodes/watermarking.ps (Source: Computer Science Division, University of California, Berkeley).

[7] Robert Krenn, "Steganography and steganalysis", An Article, January 2004.
http://www.krenn.nl/univ/cry/steg/article.pdf

[8] Francesco Queirolo, "Steganography in Images", Final Communications Report.
http://eric.purpletree.org/file/Steganography%20In%20Images.pdf

[9] Ingemar J. Cox, Ton Kalker, Georg Pakura and Mathias Scheel. "Information Transmission and Steganography", Springer, Vol. 3710/2005, pp. 15-29.

[10] LoboGuerrero, A., Marques, F., Lienard, P.B.J., "Enhanced audio data hiding synchronization using non linear filters", ICASSP '04, 17-21 May 2004, pp. ii- 885-8 vol.2.