ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 2, Issue 1, January 2013

Audio Steganography – A Review

Mrs. Namita Verma
M.E.(Communication)
S.S.C.E.T. Bhilai (C.G.)

Mr. Vinay Kumar Jain
Associate Professor
S.S.C.E.T. Bhilai (C.G.)

ABSTRACT

With the fast growth of technology & the large role it will play in the future the security of sending confidential information is becoming more& more of an issue. One could say that this is an "Information Age". In present day to day life, effective data hiding methods are needed due to attack made on data communication. The large demand of internet application requires data to be transmitted in a secure manner. Data transmission in public communication system is not secure because of interception & improper manipulation by eavesdropper. The attractive solution for this problem is steganography, which is the art & science of writing hidden messages in such a way that no one, apart from sender & intended receiver, suspect the existence of the message. It is a form of security through obscurity.

1. INTRODUCTION

The term "Security through Obscurity" is the belief that a system of any sort can be secure so long as nobody outside of its implementation group is allowed to find out anything about its internal mechanisms.

Data hiding is considered as security by obscurity system. Number of techniques has been implemented towards improving secure data hiding approaches. Two main considerations in these techniques are the amount of data hidden & the secrecy of the data against the attackers. One of such technique is "Steganography".

2. OVERVIEW OF STEGANOGRAPHY

Steganography has proved to be one of the practical ways of securing data. It is used to hide secret data inside other digital mediums. Audio files & signals make appropriate medium for steganography due to high data transmission rate & high level of redundancy.

Generally, all digital mediums, signals or files can be used in steganography process as cover media. The choice of cover media depends on the level of redundancy. Redundancy can be described as the bits of media, signals or file that offer accuracy more than needed for the object use. Image, video & audio files fulfill this requirement.

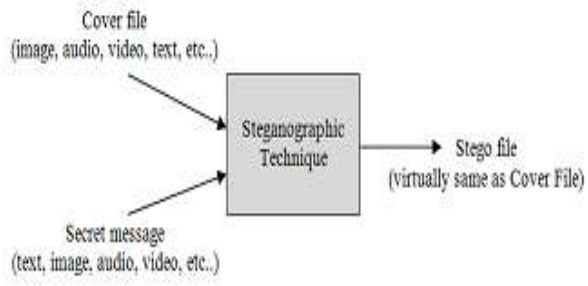

Figure 1. Fundamental scheme of steganography process

Steganography is an art of hiding secret information inside a carrier file so that the representation of carrier file is not altered. The basic process in steganography involves embedding the secret message in the carrier file & thus the stego file is created .This stego file resembles the carrier file & is transmitted in the transmitter side. It is received at receiver side & the reverse process of extracting the secret information from the stego file is performed.

2.1. Data Hiding in Audio Files

Encoding secret messages in audio is the most challenging technique to use when dealing with Steganography. This is because the human auditory system (HAS) has such a dynamic range that it can listen over. The HAS 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. Perturbations in a sound file can be detected as low as one part in ten million. However there are some "holes" available in this perspective range where data may be hidden. While the HAS has a large dynamic range, it often has a fairly small differential range. As a result, loud sounds tend to mask out quiet sounds. There are some environmental distortions so common as to be ignored by the listener in most cases.

There are two concepts to consider before choosing an encoding technique for audio. They are the digital format of the audio and the transmission medium of the audio.

There are three main digital audio formats typically in use. They are Sample Quantization, Temporal Sampling Rate and Perceptual Sampling.

• Sample Quantization which is a 16-bit linear sampling architecture used by popular audio formats such as .WAV and .AIFF.

• Temporal Sampling uses selectable frequencies (8 kHz, 9.6 kHz, 10 kHz, 12 kHz, 16 kHz, 22.05 kHz and 44.1 kHz.) to sample the audio. Sampling rate puts an upper bound on the usable portion of the frequency range. Generally, the higher the sampling rate is, the higher the usable data space gets.

• Perceptual Sampling format changes the statistics of the audio drastically by encoding only the parts the listener perceives, thus maintaining the sound but changing the signal. This format is used by the most popular digital audio on the Internet today in ISO MPEG (MP3).

Transmission medium (path the audio takes from sender to receiver) must also be considered when encoding secret messages in audio. The four transmission mediums are discussed below.

• Digital end-to-end environment: If a sound file is copied directly from machine to machine, but never modified, then it will go through this environment. As a result, the sampling will be exactly the same between the encoder and decoder. Very little constraints are put on data hiding in this environment.

• Increased/decreased resampling environment: In this environment, a signal is resampled to a higher or lowers sampling rate, but remains digital throughout. Although the absolute magnitude and phase of most of the signal are preserved, the temporal characteristics of the signal are changed.

• Analog transmission and resampling: This occurs when a signal is converted to an analog state, played on a relatively clean analog line, and resampled. Absolute signal magnitude, sample quantisation and temporal sampling rate are not preserved. In general, phase will be preserved.

• ‘‘Over the air" environment: This occurs when the signal is ``played into the air" and ``resampled with a microphone". The signal will be subjected to possible unknown nonlinear modifications causing phase changes, amplitude changes, drifting of different frequency components, echoes, etc.

The signal representation and transmission environment both need to be considered when choosing a data-hiding method

2.2. How data is hidden in sounds

Sound samples are, by their very nature, inaccurate estimates of the correct value of the sound wave at a particular moment in time. The sound samples in Windows WAV files are stored as either 8 or 16 bit values that eventually get passed to the DA converter in your soundboard. For 8 bit samples this means that the values can range between 0 and 255. 16 bit samples range between 0 and 65535.

All S-Tools does is to distribute the bit-pattern that corresponds to the file that you want to hide across the least significant bits of the sound sample. For example, suppose that a sound sample had the following eight bytes of information in it somewhere:

132 134 137 141 121 101 74 38

In binary, this is:

1000010010000110 1000100110001101
01111001011001010100101000100110

(LSB of each byte shown in italics)

Suppose that we want to hide the binary byte 11010101 (213) inside this sequence. We simply replace the LSB (Least Significant bit) of each sample byte with the corresponding bit from the byte we are trying to hide. So the above sequence will change to:

133 135 136 141 120 101 74 39

In binary, this is:

10000101 10000111 10001000 10001101
01111000 01100101 01001010 00100111

As you can clearly see, the values of the sound samples have changed by, at most, one value either way. This will be inaudible to the human ear, yet we have concealed 8 bits of information within the sample. This is the theory behind how S-Tools do its job.

2.3. Methods of Audio Steganography

This section presents some common methods used in audio Steganography:

2.3.1. LSB CODING:

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. The following diagram illustrates how the message 'HEY' is encoded in a 16-bit CD quality sample using the LSB method:

In LSB coding, the ideal data transmission rate is 1 kbps per 1 kHz. In some implementations of

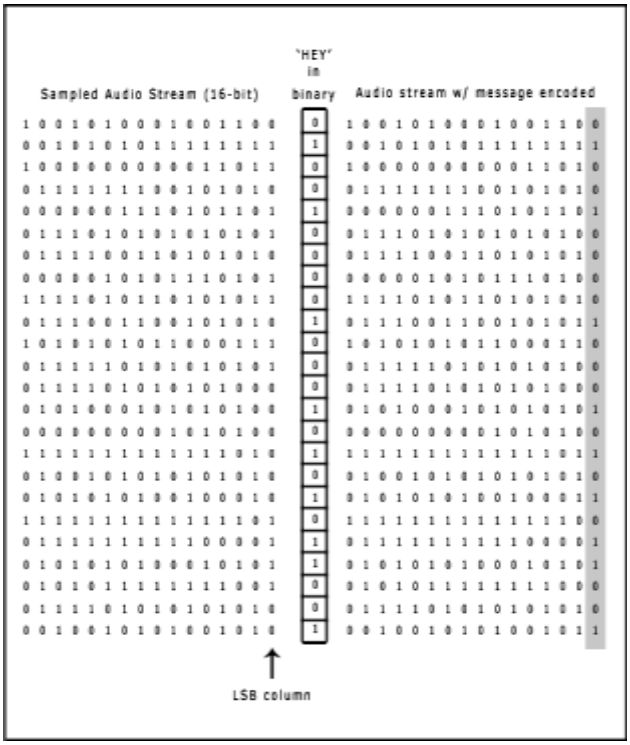

Figure 2. LSB coding

LSB coding, however, the two least significant bits of a sample are replaced with two message bits. This increases the amount of data that can be encoded but also increases the amount of resulting noise in the audio file as well. Thus, one should consider the signal content before deciding on the LSB operation to use. For example, a sound file that was recorded in a bustling subway station would mask low-bit encoding noise. On the other hand, the same noise would be audible in a sound file containing a piano solo.

The main advantage of the LSB coding method is low computational complexity of the algorithm while its major disadvantage : As the number of used LSBs during LSB coding increases or, equivalently, depth of the modified LSB layer becomes larger, probability of making the embedded message statistically detectable increases and perceptual transparency of stego objects is decreased. Low Bit Encoding is therefore an undesirable method, mainly due to its failure to meet the Steganography requirement of being undetectable.

2.3.2. PHASE CODING:

Phase coding addresses the disadvantages of the noise-inducing methods of audio Steganography. 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.

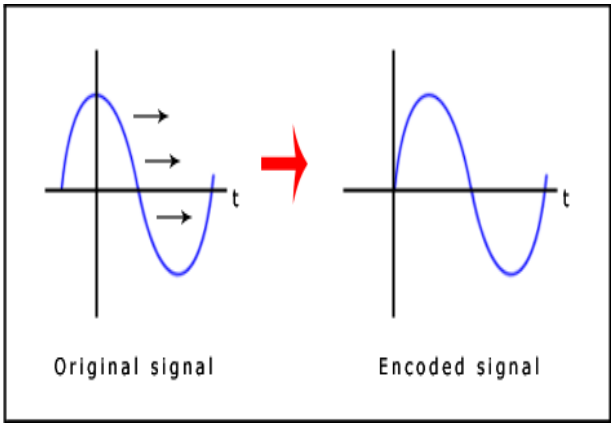

Figure 3. Phase Coding

The phase coding method breaks down the sound file into a series of N segments. A Discrete Fourier Transform (DFT) is applied to each segment to create a matrix of the phase and magnitude. The phase difference between each segment is calculated, the first segment (s0) has an artificial absolute phase of p0 created, and all other segments have newly created phase frames. The new phase and original magnitude are combined to get the new segment, Sn. These new segments are then concatenated to create the encoded output and the frequency remains preserved. In order to decode the hidden information the receiver must know the length of the segments and the data interval used. The first segment is detected as a 0 or a 1 and this indicates where the message starts.

This method has many advantages over Low Bit Encoding, the most important being that it is undetectable to the human ear. Like all of the techniques described so far though, its weakness is still in its lack of robustness to changes in the audio data. Any single sound operation or change to the data would distort the information and prevent its retrieval.

2.3.3 ECHO HIDING:

Echo hiding embeds its data by creating an echo to the source audio. Three parameters of this artificial echo are used to hide the embedded data, the delay, the decay rate and the initial amplitude. As the delay between the original source audio and the echo decrease it becomes harder for the human ear to distinguish between the two signals until eventually a created carrier sound's echo is just heard as extra resonance.

In addition, offset is varied to represent the binary message to be encoded. One offset value represents a binary one, and a second offset value represents a binary zero. 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.

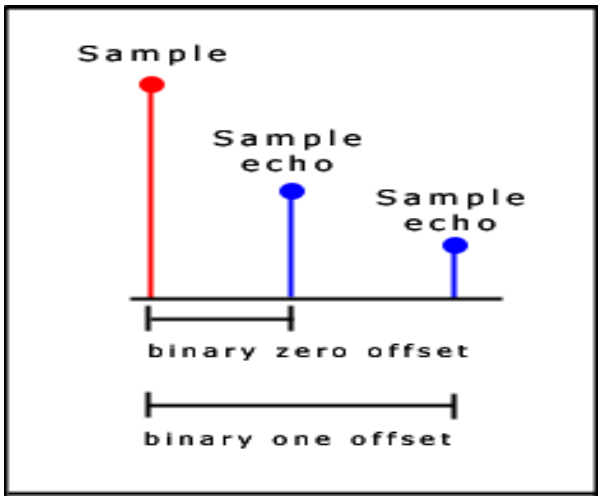

Figure 4. Echo Hiding

The blocks are recombined to produce the final signal.

The "one" echo signal is then multiplied by the "one" mixer signal and the "zero" echo signal is multiplied by the "zero" mixer signal. Then the two results are added together to get the final signal. The final signal is less abrupt than the one obtained using the first echo hiding implementation. This is because the two mixer echoes are complements of each other and that ramp transitions are used within each signal. These two characteristics of the mixer signals produce smoother transitions between echoes.

To extract the secret message from the stego-signal, the receiver must be able to break up the signal into the same block sequence used during the encoding process. Then the autocorrelation function of the signal's spectrum (the spectrum is the Forward Fourier Transform of the signal's frequency spectrum) can be used to decode the message because it reveals a spike at each echo time offset, allowing the message to be reconstructed.

The following diagram summarizes the second implementation of the echo hiding process.

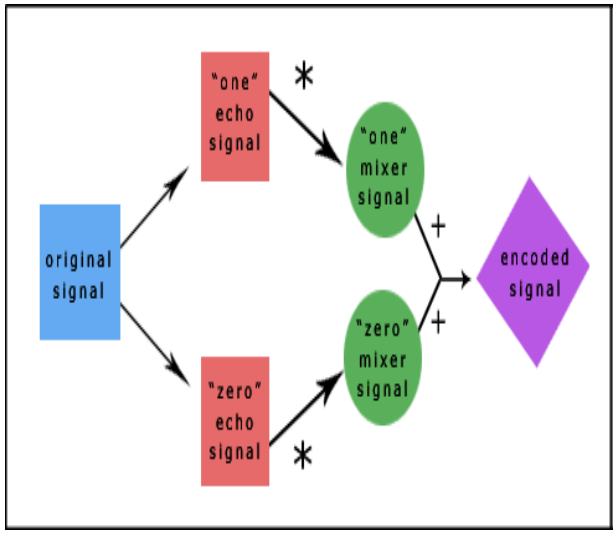

Figure 5. Echo Hiding

Much like phase encoding this has considerably better results than Low Bit Encoding and makes good use of research done so far in psychoacoustics. As with all sound file encoding, we find that working in audio formats such as WAV is very costly, more so than with bitmap images in terms of the "file size to storage capacity" ratio. The transmission of audio files via e-mail or over the web is much less prolific than image files and so is much more suspicious in comparison. It allows for a high data transmission rate and provides superior robustness when compared to the noise inducing methods.

2.3.4. SPREAD SPECTRUM:

Spread spectrum systems encode data as a binary sequence which sounds like noise but which can be recognised by a receiver with the correct key. The technique has been used by the military since the 1940s because the signals are hard to jam or intercept as they are lost in the background noise. Spread spectrum techniques can be used for watermarking by matching the narrow bandwidth of the embedded data to the large bandwidth of the medium.

Two versions of SS can be used in audio Steganography: the direct-sequence and frequency-hopping schemes. In direct-sequence SS, the secret message is spread out by a constant called the chip rate and then modulated with a pseudorandom signal. It is then interleaved with the cover-signal. In frequency-hopping SS, the audio file's frequency spectrum is altered so that it hops rapidly between frequencies.

The following procedural diagram illustrates the design:

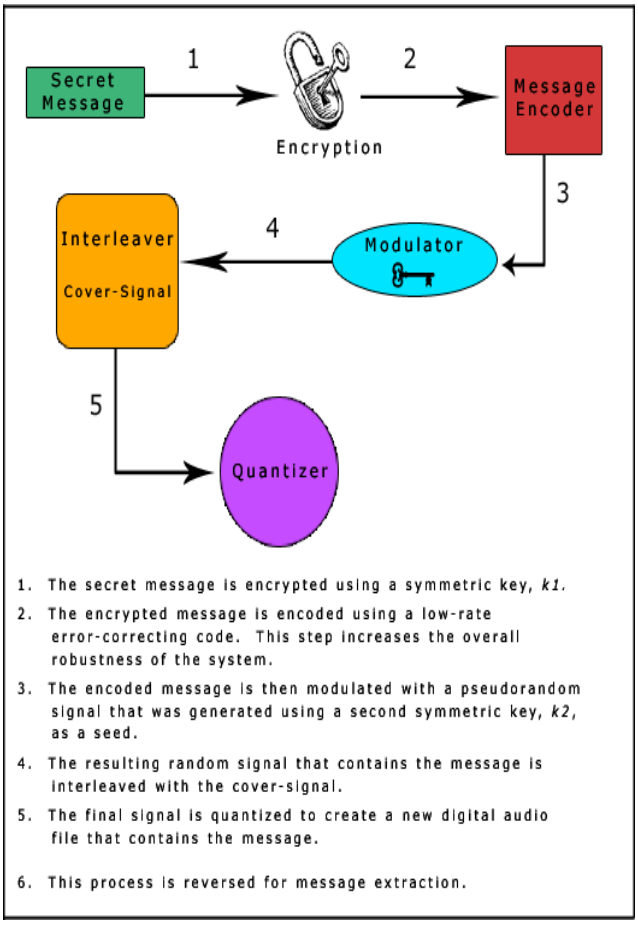

Figure 6. Spread Spectrum (SS)

Spread Spectrum Steganography has significant potential in secure communications – commercial and military. Audio Steganography in conjunction with Spread Spectrum may provide added layers of security.

Spread spectrum encoding techniques are the most secure means by which to send hidden messages in audio, but it can introduce random noise to the audio thus creating the chance of data loss. They have the potential to perform better in some areas than LSB coding, parity coding, and phase coding techniques in that it offers a moderate data transmission rate while also maintaining a high level of robustness against removal techniques.

3. Evaluation of Audio Steganography

3.1. ADVANTAGES:

- Audio based Steganography has the potential to conceal more information:
  - Audio files are generally larger than images.
  - Our hearing can be easily fooled.
  - Slight changes in amplitude can store vast amounts of information.

- The flexibility of audio Steganography is makes it very potentially powerful :
  - The methods discussed provide users with a large amount of choice and makes the technology more accessible to everyone. A party that wishes to communicate can rank the importance of factors such as data transmission rate, bandwidth, robustness, and noise audibility and then select the method that best fits their specifications.
  - For example, two individuals who just want to send the occasional secret message back and forth might use the LSB coding method that is easily implemented. On the other hand, a large corporation wishing to protect its intellectual property from "digital pirates" may consider a more sophisticated method such as phase coding, SS, or echo hiding.

  - Another aspect of audio Steganography that makes it so attractive is its ability to combine with existing cryptography technologies.

- Users no longer have to rely on one method alone. Not only can information be encrypted, it can be hidden altogether.

- Many sources and types makes statistical analysis more difficult :
  - Greater amounts of information can be embedded without audible degradation

- Security :
  - Many attacks that are malicious against image Steganography algorithms (e.g. geometrical distortions, spatial scaling, etc.) cannot be implemented against audio Steganography schemes. Consequently, embedding information into audio seems more secure due to less steganalysis techniques for attacking to audio.
  - As emphasis placed on the areas of copyright protection, privacy protection, and surveillance increases, Steganography will continue to grow in importance as a protection mechanism.
- Audio Steganography in particular addresses key issues brought about by the MP3 format, P2P software, and the need for a secure broadcasting scheme that can maintain the secrecy of the transmitted information, even when passing through insecure channels.

3.2. DISADVANTAGES:

- Embedding additional information into audio sequences is a more tedious task than that of images, due to dynamic supremacy of the HAS over human visual system.
- Robustness: Copyright marks hidden in audio samples using substitution could be easily manipulated or destroyed if a miscreant comes to know that information is hidden this way.
- Commercialized audio Steganography have disadvantages that the existence of hidden messages can be easily recognized visually and only certain sized data can be hidden.
- Compressing an audio file with lossy compression will result in loss of the hidden message as it will change the whole structure of a file. Also, several lossy compression schemes use the limits of the human ear to their advantage by removing all frequencies that cannot be heard. This will also remove any frequencies that are used by a Steganography system which hides information in that part of the spectrum.

4. CONCLUSION:

This report has looked in detail at the major techniques used for data hiding in audio files.

Section I gave an overview of Steganography and in particular the concept of Audio Steganography.

Section II described in detail, various Audio Steganography algorithms namely LSB Coding, Phase Coding, Spread Spectrum and Echo Hiding. At the end, feasibility of Audio Steganography was evaluated by considering it's the pros and cons.

In summary, if implemented correctly and in conjunction with cryptographic methods to secure the embedded data before insertion to a cover medium, many of the data hiding methods described above could become powerful tools for the transmission of undetectable and secure communication.

5. REFERENCES:

[1] Poluami D., Debnath B., and Tai-hoon K., "Data Hiding in audio signal : A review", International Journal of Database Theory and Application, vol.2, No.2, June, 2009.

[2]Stallings, W. " Cryptography and Network Security: Principles and Practice," 3rd edition, Prentice-Hall, 2003.

[3] Gary C. Kessler, "Steganography: Hiding Data Within Data", *http://www.garykessler.net/library/steganography.html*, September 2001.

[4] Dr.H.B.Kekre and A.A.Archana, "Information hiding using LSB technique with increased capacity", *International Journal of Cryptography and Security*, vol. 1, No.2, October 2008.

[5] Johnson, N. F., "Steganography", http:/ /www.jjtc.com/ stegdoc/, George Mason University, 2003.

[6] 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.

[7] 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.

[8] Nedeljko Cvejic, Tapio Seppben "Increasing the capacity of LSB-based audio steganography " FIN-90014 University of Oulu, Finland ,2002.

[9] Sajad Shirali-Shahreza M.T. Manzuri-Shalmani "High capacity error free wavelet domain speech steganography" ICASSP 2008

[10] Neil F.Johnson, Z.Duric and S.Jajodia. "Information Hiding Steganography and Watermarking - Attacks and Countermeasures",Kluwer Academic Publishers, 2001.

[11] F.A.P.Petitcolas, R.J.Anderson, M.G.Kuhn:" Information Hiding- A Survey", Process of IEEE, vol.87, of IEICE, ISEC, vol.106 pp.15-22, September 2006. no.7, pp.1062-1078, July, 1999.

[12] Min Wu, Bede Liu. "Multimedia Data Hiding", Springer- Verlag New York, 2003.