IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 2, Ver. V (Mar – Apr. 2015), PP 32-44
www.iosrjournals.org

Audio Steganography Using Discrete Wavelet Transformation (DWT) & Discrete Cosine Transformation (DCT)

Sumeet Gupta[1], Dr. Namrata Dhanda[2]
[1]*(Research Scholar/ Department of Computer Science / GITM, Lucknow)*
[2]*(Associate Professor/ Department of Computer Science / GITM, Lucknow)*

Abstract: *Steganography is an art and a science of communicating in a way, which hides the existence of the communication. It is also called as "covered writing", because it uses a "cover" of a message for sending any important secret message. In the steganographic scenario, the secret data is first concealed within another object which is called "cover object", to form "stego-object" and then this new object can be transmitted or saved. Using different techniques, we can send secret data in the form of an image, a music file or even a video file by embedding it into the carrier, forming a stego-signal. At the receiver's end, the secret data can be recovered from the stego-signal using different algorithms. The main goal of Steganography is to communicate securely in a completely undetectable manner and to avoid drawing suspicion to the transmission of a hidden data. It is not only prevents others from knowing the hidden information, but it also prevents others from thinking that the information even exists. If a Steganography method causes someone to suspect there is secret information in a carrier medium, then the method has failed.*

Keywords: *Discrete Wavelet Transformation (DWT), Discrete Cosine Transformation (DCT), Human Auditory System (HAS), Signal to Noise Ratio (SNR), International Federation of the Phonographic Industry (IFPI)*

I. Introduction

Steganography is an art and a science of communicating in a way, which hides the existence of the communication. It is also called as "covered writing", because it uses a "cover" of a message for sending any important secret message. Steganography serves as a means for private, secure and sometimes malicious communication. Steganography is the art to hide the very presence of communication by embedding the secret message into the innocuous looking cover media objects, such as images using the human's visual, aural redundance or media objects' statistical redundance. Steganography is a powerful tool which increases security in data transferring and archiving. In the steganographic scenario, the secret data is first concealed within another object which is called "cover object", to form "stego object" and then this new object can be transmitted or saved. Using different techniques, we can send secret data in the form of an image, a music file or even a video file by embedding it into the carrier, forming a stego-signal. At the receiver's end, the secret data can be recovered from the stego-signal using different algorithms.

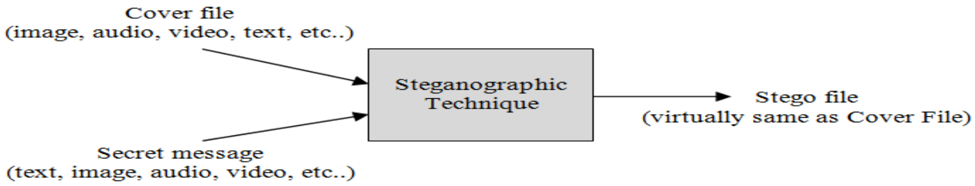

Cover file
(image, audio, video, text, etc..)
Steganographic Technique
Stego file
(virtually same as Cover File)
Secret message
(text, image, audio, video, etc..)

Figure 1 Fundamental process of Steganography

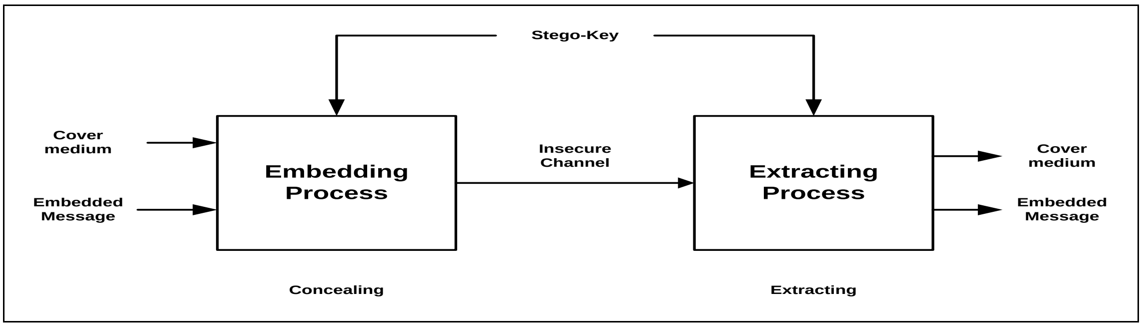


Figure 2 General Steganography System

II. Classification Of Steganography

There are following classification of Steganography

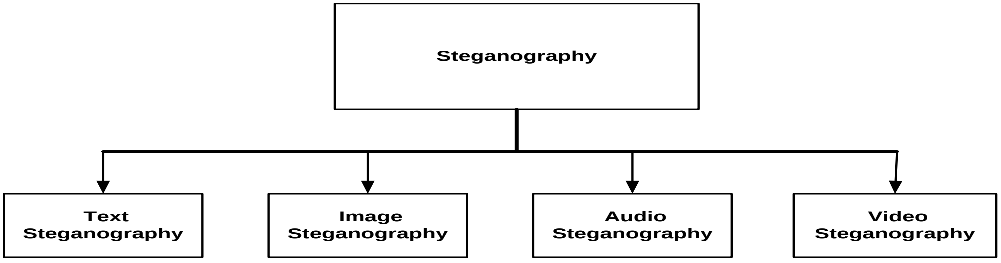


Figure 3 Classification of Information Hiding

2.1 Audio Stegnography

2.1.1 Overview

The main goal of Steganography is to communicate securely in a completely undetectable manner and to avoid drawing suspicion to the transmission of a hidden data. It is not only prevents others from knowing the hidden information, but it also prevents others from thinking that the information even exists. If a Steganography method causes someone to suspect there is secret information in a carrier medium, then the method has failed.

Audio Steganography is focused in hiding secret information in an innocent cover audio file or signal securely and strongly. Communication security and robustness are vital for transmitting important information to authorized entities, while denying access to not permitted ones. By embedding secret information using an audio signal as a cover medium, the very existence of secret information is hidden away during communication. This is a serious and vital issue in some applications such as battlefield communications and banking transactions. In a computer-based audio Steganography system, secret messages are embedded in digital sound. The secret message is embedded by slightly altering the binary sequence of a sound file. Existing audio Steganography software can embed messages in WAV, AU, and even MP3 sound files.

The basic model of Audio Steganography consists of Carrier (Audio file), Message and Password. Carrier is also known as a cover-file, which conceals the secret information. Message is the data that the sender wishes to remain it confidential. Message can be plain text, image, audio or any type of file. Password is known as a stego-key, which ensures that only the recipient who knows the corresponding decoding key will be able to extract the message from a cover-file. The cover-file with the secret information is known as a stego-file.

The information hiding process consists of following two steps.

i. Identification of redundant bits in a cover-file. Redundant bits are those bit that can he modified without corrupting the quality or destroying the integrity of the cover-file.

ii. To embed the secret information in the cover file, the redundant bits in the cover file is replaced by the bits of the secret information.

2.1.2 Encoding Secret Messages in Audio

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. To put this in perspective, the (HAS) perceives over a range of power greater than one million to one and a range of frequencies greater than one thousand to one making it extremely hard to add or remove data from the original data structure. The only weakness in the (HAS) comes at trying to differentiate sounds (loud sounds drown out quiet sounds) and this is what must be exploited to encode secret messages in audio without being detected.

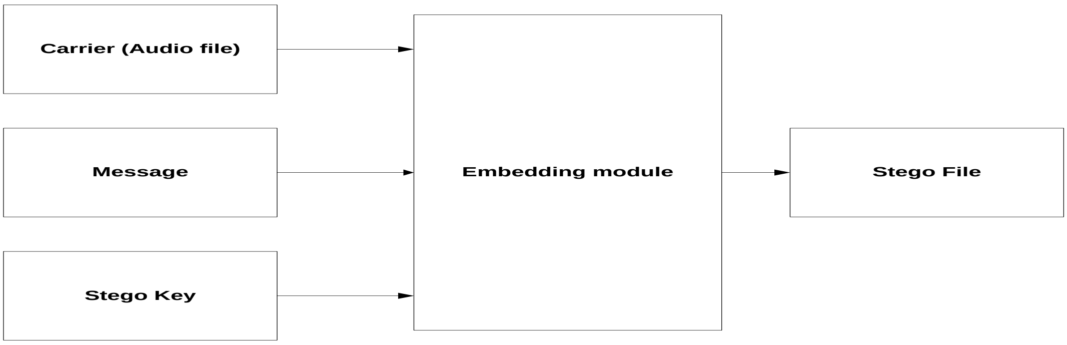


Figure 4 Basic Audio Steganography model

III. Advantages Of Audio Steganography

1. 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
2. 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.
3. 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.
4. Many sources and types makes statistical analysis more difficult :
- Greater amounts of information can be embedded without audible degradation
5. 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.

IV. Disadvantages Of Audio Steganography

4.1. 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.

4.2. 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.

4.3. Commercialized audio Steganography have disadvantages that the existence of hidden messages can be easily recognized visually and only certain sized data can be hidden.

4.4.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.

V. Applications

There are many applications for information hiding in today's world. However, information hiding can be used in ethical ways; there are some ways that digital data hiding could be misused. Sometimes a method cannot be easily categorized in either of Steganography or watermarking categories. Occasionally, the boundaries between these two disciplines have been blurred. However category can be specified by identifying an application which the method is used for. So, without any classification of the application, most important and common application of data hiding is presented.

5.1 Secret communication

Some people may use information hiding in order to hide data and be in confidential communication. For example, Steganography is a trusted way for those who want to have an undisclosed communication. This is actually the area of Steganography rather than watermarking.

5.2 Secure storage

Another use of information hiding is in the area of security storage. Obviously many types of sensitive information such as medical records of patients or prescription drug information need security during storage and transmission because in case those are accessible for unauthorized person could lead to illegal activities as identity theft as well as insurance fraud (Boneh et al., 1998).

5.3 Covert communication

Some people or organizations need a covert communication for their business operations. For example, military can use information hiding for some technical activities as sending battle plans that if they fall into the wrong hands, then entire tactic would be compromised (Kurak et al., 1992; Kirovski et al., 2002).

5.4 Fingerprinting

The recipients or originator of a specific copy of media file could be traced by watermarking. The applied technique has to comprise a high robustness against intentional and unintentional attacks (Wang et al., 2003; Yacobi et al., 2001; Dittmann et al., 2000). For example, before distributing numerous copies of multimedia products to recipients, they can be watermarked by different serial or identity numbers (Wu et al., 2004; Trappe et al., 2003; Chenyu et al., 2003; Hong et al., 2003).

5.5 Copy right protection

While a considerable portion of economic resources is dedicated to the creation of intellectual property, particularly in industrial societies, the cost of reproducing such intellectual creations typically constitutes only a small fraction of the creation (Cox et al., 2003; Bloom et al., 1999; Pan et al., 1995). In the copy right protection, a watermark which contains the information of the owner is embedded into the host media (Noll et al., 1993; Dittmann et al., 2000). The watermark is supposed to be robust and enables the owner to prove his ownership in case is needed. Also with fragile watermarking, watermarks are used to verify if the host signals are tampered (Termont et al., 2000; Termont et al., 1999). Furthermore, in the copy control application, watermarks control access policy or limit to a certain copy (Depovere et al., 1999; Kalker et al., 2000; Craver et al., 2001).

VI. Audio Steganography Techniques

This provides the features of the human auditory system, which are important while dealing with the audio Steganography technique. Further, this considers the requirement of an efficient Steganography strategy and different audio Steganography techniques involving both time and frequency domain. Then we discuss the background about discrete cosine transform (DCT) and discrete wavelet transform (DWT).

6.1 Features of Human Auditory System (HAS)

Note that audio Steganography is more challenging than an image Steganography technique due to wider dynamic range of the HAS in comparison with human visual system (HVS) . Human ear can perceive the power range greater than 109: 1 and range frequencies of 103:1 . In addition, human ear can hear the low ambient Gaussian noise in the order of 70dB . However, there are some useful features such as the louder sounds mask the corresponding slow sounds. This feature can be used to embed additional information like a watermark. Further, HAS is insensitive to a constant relative phase shift in a stationary audio signal, and, some spectral distortions are interpreted as natural, perceptually non-annoying ones. Two properties of the HAS dominantly used in Steganography algorithms are frequency (simultaneous) masking and temporal masking :

6.1.1 Frequency masking:

Frequency (simultaneous) masking is a frequency domain phenomenon where low levels signal (the maskee) can be made inaudible (masked) by a simultaneously appearing stronger signal (the masker), if the masker and maskee are close enough to each other in frequency. A masking threshold can be found and is the level below which the audio signal is not audible. Thus, frequency domain is a good region to check for the possible areas that have imperceptibility.

6.1.2 Temporal masking:

In addition to frequency masking, two phenomena of the HAS in the time domain also play an important role in human auditory perception. Those are pre- masking and post-masking in time. However, considering the scope of analysis in frequency masking over temporal masking, prior is chosen for this thesis. Temporal masking is used in application where the robustness is not of primary concentration.

VII. Requirements Of The Efficient Steganography Technique

According to IFPI (International Federation of the Phonographic Industry), audio Steganography algorithms should meet certain requirements. The most significant requirements are perceptibility, reliability, capacity, and speed performance.

7.1. Perceptibility:

One of the important features of the Steganography technique is that the watermarked signal should not lose the quality of the original signal. The Signal to Noise Ratio (SNR) of the watermarked signal to the original signal should be maintained greater than 20dB . In addition, the technique should make the modified signal not perceivable by human ear.

7.2. Reliability:

Reliability covers the features like the robustness of the signal against the malicious attacks and signal processing techniques. The watermark should be made in a way that they provide high robustness against attacks. In addition, the watermark detection rate should be high under any types of attacks in the situations of proving ownership. Some of the other attacks summarized by Secure Digital Music Initiative (SDMI), an online forum for digital music copyright protection, are digital-to-analog and analog-to-digital conversions, noise addition, band-pass filtering, time-scale modification, echo addition, and sample rate conversion .

7.3 Capacity:

The efficient Steganography technique should be able to carry more information but should not degrade the quality of the audio signal. It is also important to know if the watermark is completely distributed over the host signal because, it is possible that near the extraction process a part of the signal is only available. Hence, capacity is also a primary concern in the real time situations .

7.4 Speed:

Speed of embedding is one of the criteria for efficient Steganography technique. The speed of embedding of watermark is important in real time applications where the embedding is done on continuous signals such as, speech of an official or conversation between airplane pilot and ground control staff. Some of the possible applications where speed is a constraint are audio streaming and airline traffic monitoring. Both embedding and extraction process need to be made as fast as possible with greater efficiency .

7.5 Asymmetry:

If for the entire set of cover objects the watermark remains same; then, extracting for one file will cause damage watermark of all the files. Thus, asymmetry is also a noticeable concern. It is recommended to have unique watermarks to different files to help make the technique more useful.

VIII. Problems And Attacks On Audio Signals

As discussed in Section 7 the important requirements of an efficient Steganography technique are the robustness and inaudibility. There is a tradeoff between these two requirements; however, by testing the algorithm with the signal processing attacks that gap can be made minimal. Every application has its specific requirements, and provides an option to choose high robustness compensating with the quality of the signal and vice-versa. Without any transformations and attacks every Steganography technique performs efficiently. Some of the most common types of processes an audio signal undergoes when transmitted through a medium are as follows-

8.1 Dynamics:

The amplitude modification and attenuation provide the dynamics of the attacks. Limiting, expansion and compressions are some sort of more complicated applications which are the non-linear modifications. Some of these types of attacks are re-quantization .

8.2 Filtering:

Filtering is common practice, which is used to amplify or attenuate some part of the signal. The basic low pass and high pass filters can be used to achieve these types of attacks.

8.3Ambience:

In some situations the audio signal gets delayed or there are situations where in people record signal from a source and claim that the track is theirs. Those situations can be simulated in a room, which is of great importance to check the performance of an audio signal.

8.4 Conversion and lossy compression:

Audio generation is done at a particular sampling frequency and bit rate; however, the created audio track will undergo so many different types of compression and conversion techniques. Some of the most common compression techniques are audio compression techniques based on psychoacoustic effect (MPEG and Advanced Audio Codec (AAC)). In addition to that, it is common process that the original audio signal will change its sampling frequencies like from 128Kbps to 64Kpbs or 48 Kbps. There are some programs that can achieve these conversions and perform compression operation. However, for testing purposes we have used MATLAB to implement these applications. Attacks like re-sampling and mp3 compression provide some typical examples.

8.5 Noise:

It is common practice to notice the presence of noise in a signal when transmitted. Hence, Steganography algorithm should make the technique robust against the noise attacks. It is recommended to check the algorithm for this type of noise by adding the host signal by an additive white Gaussian noise (AWGN) to check its robustness.

8.6 Time stretch and pitch shift:

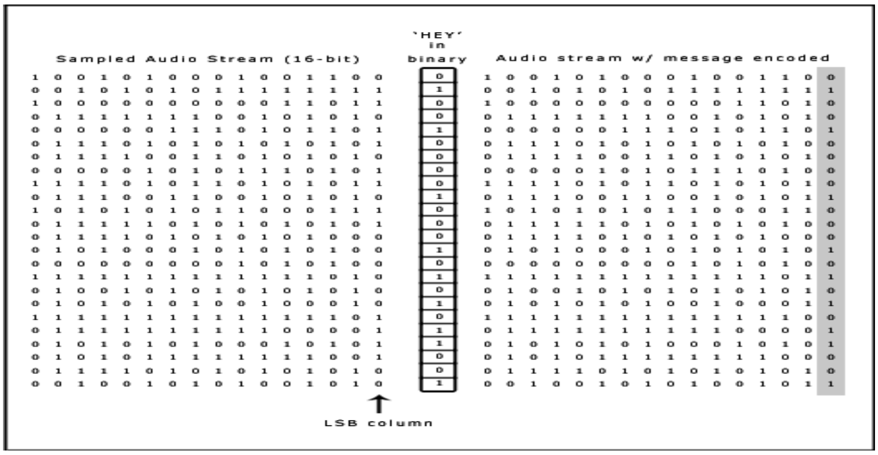
These attacks change either the length of the signal without changing its pitch and vice versa. These are some de-synchronization attacks which are quite common in the data transmission. Jittering is one type of such attack.

IX. Audio Steganography Techniques – An Overview

An audio Steganography technique can be classified into two groups based on the domain of operation. One type is time domain technique and the other is transformation based method. The time domain techniques include methods where the embedding is performed without any transformation. Steganography is employed on the original samples of the audio signal. One of the examples of time domain Steganography technique is the least significant bit (LSB) method. In LSB method the watermark is embedded into the least significant bits of the host signal. As against these techniques, the transformation based Steganography methods perform Steganography in the transformation domain. Few transformation techniques that can be used are discrete cosine transform and discrete wavelet transform. In transformation based approaches the embedding is done on the samples of the host signal after they are transformed. Using of transformation based techniques provides additional information about the signal. In general, the time domain techniques provide least robustness as a simple low pass filtering can remove the watermark. Hence time domain techniques are not advisable for the applications such as copyright protection and airline traffic monitoring; however, it can be used in applications like proving ownership and medical applications. Steganography techniques can be distinguished as visible or non-blind Steganography and blind Steganography, here we present typical Steganography strategies such as LSB coding, spread spectrum technique, patchwork technique, and quantization index modulation (QIM).

9.1 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:



```
                                                   'HEY'
                                                     in
    Sampled Audio Stream (16-bit)                  binary     Audio stream w/ message encoded
1 0 0 1 0 1 0 0 0 1 0 0 1 1 0 0                      0       1 0 0 1 0 1 0 0 0 1 0 0 1 1 0 0
0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1                      1       0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1
1 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1                      0       1 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0
0 1 1 1 1 1 1 1 0 0 1 0 1 0 1 0                      0       0 1 1 1 1 1 1 1 0 0 1 0 1 0 1 0
0 0 0 0 0 0 1 1 1 0 1 0 1 1 0 1                      1       0 0 0 0 0 0 1 1 1 0 1 0 1 1 0 1
0 1 1 1 0 1 0 1 0 1 0 1 0 1 0 1                      0       0 1 1 1 0 1 0 1 0 1 0 1 0 1 0 0
0 1 1 1 1 0 0 1 1 0 1 0 1 0 1 0                      0       0 1 1 1 1 0 0 1 1 0 1 0 1 0 1 0
0 0 0 0 0 1 0 1 0 1 1 1 0 1 0 1                      0       0 0 0 0 0 1 0 1 0 1 1 1 0 1 0 0
1 1 1 1 0 1 0 1 1 0 1 0 1 0 1 1                      0       1 1 1 1 0 1 0 1 1 0 1 0 1 0 1 0
0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 0                      1       0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 1
1 0 1 0 1 0 1 0 1 1 0 0 0 1 1 1                      0       1 0 1 0 1 0 1 0 1 1 0 0 0 1 1 0
0 1 1 1 1 1 0 1 0 1 0 1 0 1 0 1                      0       0 1 1 1 1 1 0 1 0 1 0 1 0 1 0 0
0 1 1 1 1 0 1 0 1 0 1 0 1 0 0 0                      0       0 1 1 1 1 0 1 0 1 0 1 0 1 0 0 0
0 1 0 1 0 0 0 1 0 1 0 1 0 1 0 0                      1       0 1 0 1 0 0 0 1 0 1 0 1 0 1 0 1
0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0                      0       0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0                      1       1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1
0 1 0 0 1 0 1 0 1 0 1 0 1 0 1 0                      0       0 1 0 0 1 0 1 0 1 0 1 0 1 0 1 0
0 1 0 1 0 1 0 1 0 0 1 0 0 0 1 0                      1       0 1 0 1 0 1 0 1 0 0 1 0 0 0 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1                      0       1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0
0 1 1 1 1 1 1 1 1 1 1 0 0 0 0 1                      1       0 1 1 1 1 1 1 1 1 1 1 0 0 0 0 1
0 1 0 1 0 1 0 1 0 0 0 1 0 1 0 1                      1       0 1 0 1 0 1 0 1 0 0 0 1 0 1 0 1
0 1 0 1 0 1 1 1 1 1 1 1 1 0 0 1                      0       0 1 0 1 0 1 1 1 1 1 1 1 1 0 0 0
0 1 1 1 1 0 1 0 1 0 1 0 1 0 1 0                      0       0 1 1 1 1 0 1 0 1 0 1 0 1 0 1 0
0 0 1 0 0 1 0 1 0 1 0 0 1 0 1 0                      1       0 0 1 0 0 1 0 1 0 1 0 0 1 0 1 1
                              ↑
                         LSB column
```

Figure 5 Example of LSB coding

9.2 Standard LSB ALGORITHM:

It performs bit level manipulation to encode the message. 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.

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

9.3 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.

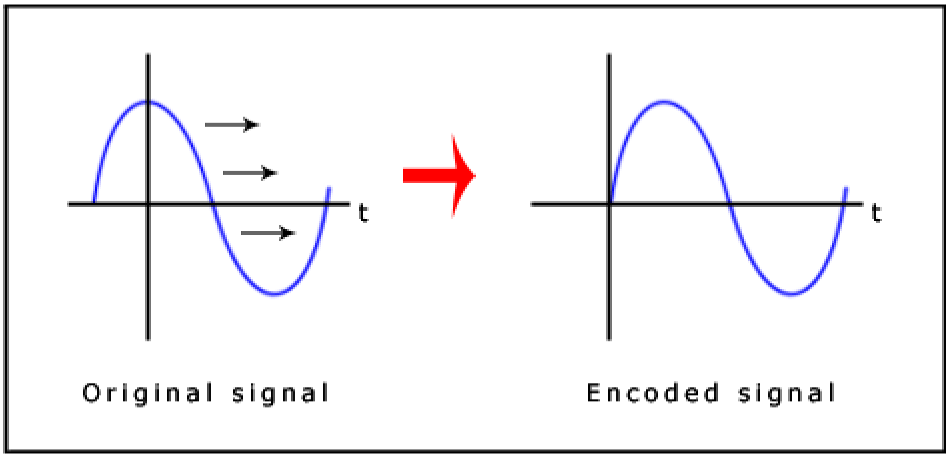


Figure 6 Phase Shift coding

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.

9.4 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.

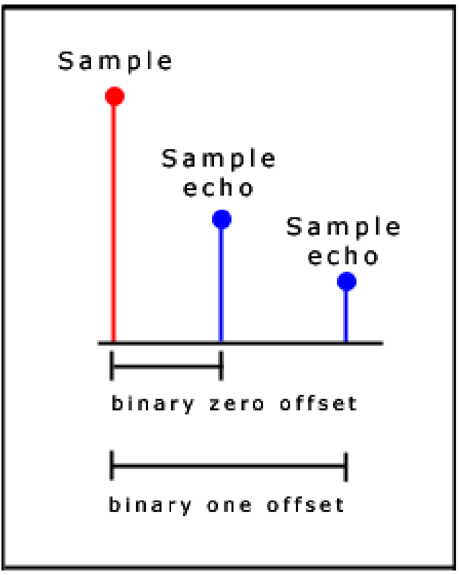


Figure 7 Echo Hiding

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

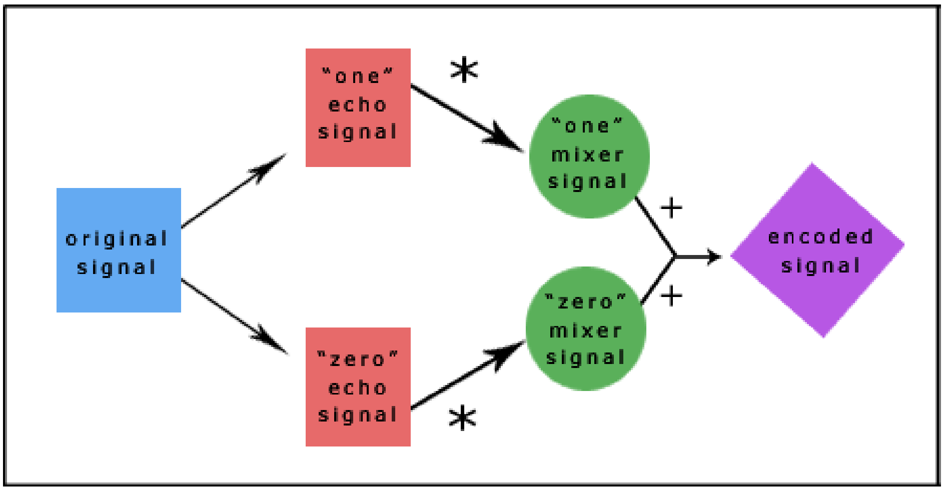


Figure 8 implementation of Echo Hiding process

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.

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.

9.5 Spread Spectrum

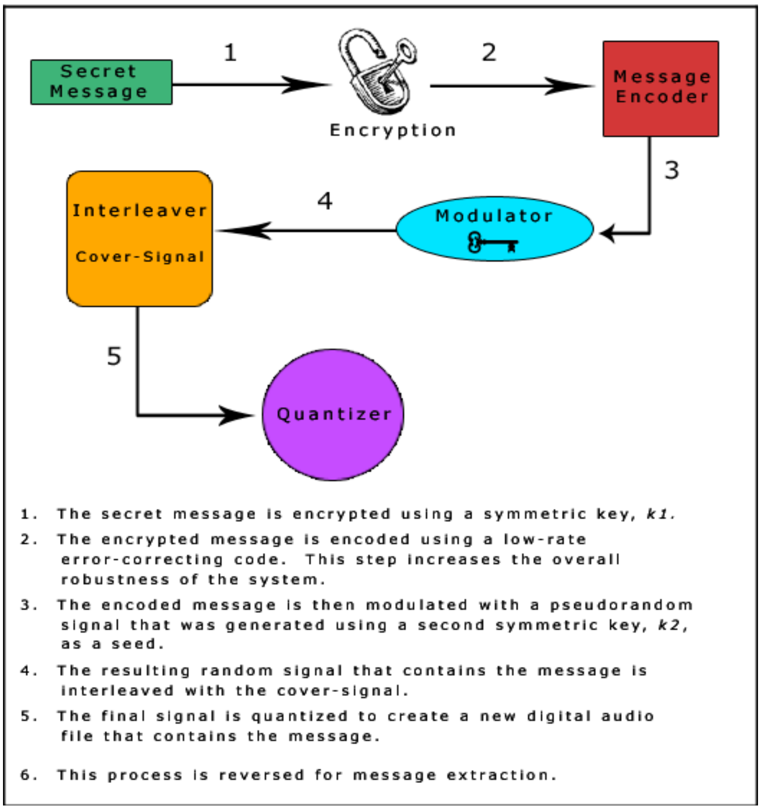
Spread spectrum systems encode data as a binary sequence which sounds like noise but which can be recognised by a receiver with the correct key. The technique has been used by the military since the 1940s because the signals are hard to jam or intercept as they are lost in the background noise. Spread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

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

The following procedural diagram illustrates the design:



1. The secret message is encrypted using a symmetric key, *k1*.
2. The encrypted message is encoded using a low-rate error-correcting code. This step increases the overall robustness of the system.
3. The encoded message is then modulated with a pseudorandom signal that was generated using a second symmetric key, *k2*, as a seed.
4. The resulting random signal that contains the message is interleaved with the cover-signal.
5. The final signal is quantized to create a new digital audio file that contains the message.
6. This process is reversed for message extraction.

Figure 9 Procedural diagram of Spread Spectrum Design

X. Discrete Cosine Transform

The discrete cosine transform is a technique for converting a signal into elementary frequency components. The DCT can be employed on both one-dimensional and two dimensional signals like audio and image, respectively. The discrete cosine transform is the spectral transformation, which has the properties of Discrete Fourier Transformation [41]. DCT uses only cosine functions of various wave numbers as basic functions and operates on realvalued signals and spectral coefficients. DCT of a 1-dimensional (1-d) sequence and the reconstruction of original signal from its DCT coefficients termed as inverse discrete cosine transform (IDCT) can be computed using equations.

In the following, $f_{dct}(x)$ is original sequence while $C_{dct}(u)$ denotes the DCT coefficients of the sequence.

$$C_{dct}(u) = \alpha(u) \sum_{x=1}^{N_{1t}-1} f_{dct}(x) \cos\left[\frac{\pi(2x+1)u}{2N_{1t}}\right], for\ u\ =\ 0,1,2,\ldots,N_{1t}\text{ -1}$$

$$f_{dct}(x) = \sum_{u=1}^{N_{1t}-1} \alpha(u) C_{dct}(u) \cos\left[\frac{\pi(2x+1)u}{2N_{1t}}\right], for\ x\ =\ 0,1,2,\ldots,N_{1t}\text{ -1}$$

$$where\ \alpha(u) = \begin{cases} \sqrt{\frac{1}{N_{1t}}} for\ u = 0 \\ \sqrt{\frac{2}{N_{1t}}} for\ u \neq 0 \end{cases}$$

From the equation for $C_{dct}(u)$ it can be inferred that for u = 0, the component is the average of the signal also termed as dc coefficient in literature. And all the other transformation coefficients are called as ac coefficients. Some of the important applications of DCT are image compression and signal compression.

The most useful applications of two-dimensional (2-d) DCT are the image compression and encryption. The 1-d DCT equations, discussed above, can be used to find the 2-d DCT by considering every row as an individual 1-d signal. Thus, DCT coefficients of an M×N two dimensional signals $C_{dct2}(u,v)$ and their reconstruction $f_{dct2}(x,y)$ can be calculated by the equations below.

$$C_{dct2}(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{M_{2t}-1} \sum_{y=0}^{N_{2t}-1} f_{dct2}(x,y) \cos\left[\frac{\pi(2x+1)u}{2M_{2t}}\right] \cos\left[\frac{\pi(2y+1)v}{2N_{2t}}\right]$$

$$f_{dct2}(x,y) = \sum_{u=0}^{M_{2t}-1} \sum_{v=0}^{N_{2t}-1} \alpha(u)\alpha(v) C_{dct2}(u,v) \cos\left[\frac{\pi(2x+1)u}{2M_{2t}}\right] \cos\left[\frac{\pi(2y+1)v}{2N_{2t}}\right]$$

$$\text{where } u\ \&\ x\ \in\ 0,1,2,\ldots.,M_{2t}\text{ -1 and } v\ \&\ y\ \in\ 0,1,2,\ldots..,N_{2t}\text{ -1}$$

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N_{2t}}} for\ u = 0 \\ \sqrt{\frac{2}{N_{2t}}} for\ u \neq 0 \end{cases} \quad \& \quad \alpha(v) = \begin{cases} \sqrt{\frac{1}{N_{2t}}} for\ v = 0 \\ \sqrt{\frac{2}{N_{2t}}} for\ v \neq 0 \end{cases}$$

Some of the properties of DCT are de-correlation, energy compaction, separability, symmetry and orthogonality. DCT provides interpixel redundancy for most of natural images and coding efficiency is maintained while encoding the uncorrelated transformation coefficients. DCT packs the energy of the signal into the low frequency regions which provides an option of reducing the size of the signal without degrading the quality of the signal.

10.1. Discrete Wavelet Transform (DWT)

Majority of the signals in practice are represented in time domain. Time-amplitude representation is obtained by plotting the time domain signal. However, the analysis of the signal in time domain cannot give complete information of the signal since it cannot provide the different frequencies available in the signal.

Frequency domain provides the details of the frequency components in the signal which are importance in some applications like electrocardiography (ECG), graphical recording of heart's electrical activity or electroencephalography (EEG), an analysis of electrical activity of human brain. The frequency spectrum of a signal is basically the frequency components (spectral components) of that signal. The main drawback of frequency domain is it does not provide when in time these frequencies exist.

There are considerable drawbacks in either time domain or frequency domains, which are rectified in wavelet transform. Wavelet Transform provides the time-frequency representation of the signal. Some of the other types of time-frequency representation are short time Fourier transformation, Wigner distributions, etc. There are different types of wavelet transforms such as **continuous wavelet transform (CWT)** and discrete wavelet transform (DWT). CWT provides great redundancy of reconstruction of the signal whereas DWT provides the sufficient information for both analysis and synthesis signal and is easier to implement as compared to CWT.

A complete structure of wavelet contains domain processing analysis block and a synthesis block. Analysis or decomposition block decomposes the signal into wavelet coefficients. The reconstruction process is the inverse of decomposition process. Here, the block takes the decomposed signal and synthesizes (near) original signal. A view of the wavelet process is shown in Figure 3.6. From the figure the original signal is decomposed in the analysis block and the signal is reconstructed using the synthesis block. Filters used in the analysis and synthesis block

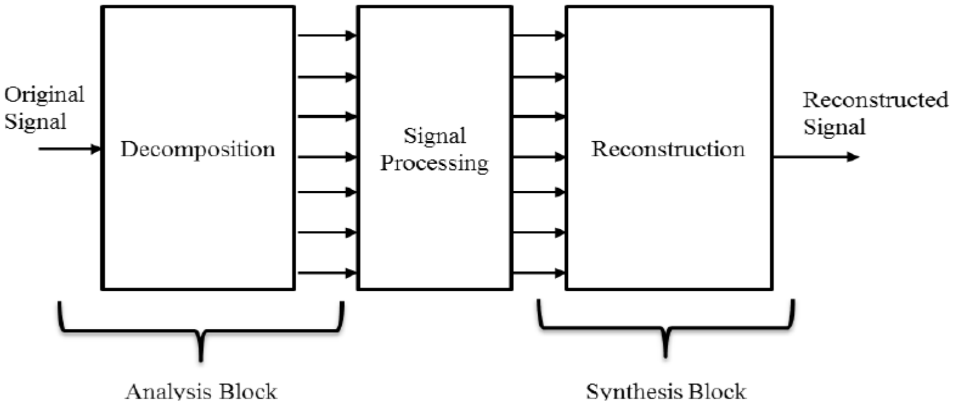


Figure 10 Basic block view of wavelet functionality

The operation of 1-level discrete wavelet transform decomposition is to separate high pass and low pass components. Thus, process involves passing the time-domain signal x[n] through a high pass filter $g_0[n]$ and down sampling the signal obtained yields detailed coefficients (D). And, passing x[n] through low pass filters $h_0[n]$and down sampling generated approximate coefficients (A). The working principle is shown in Figure 11.

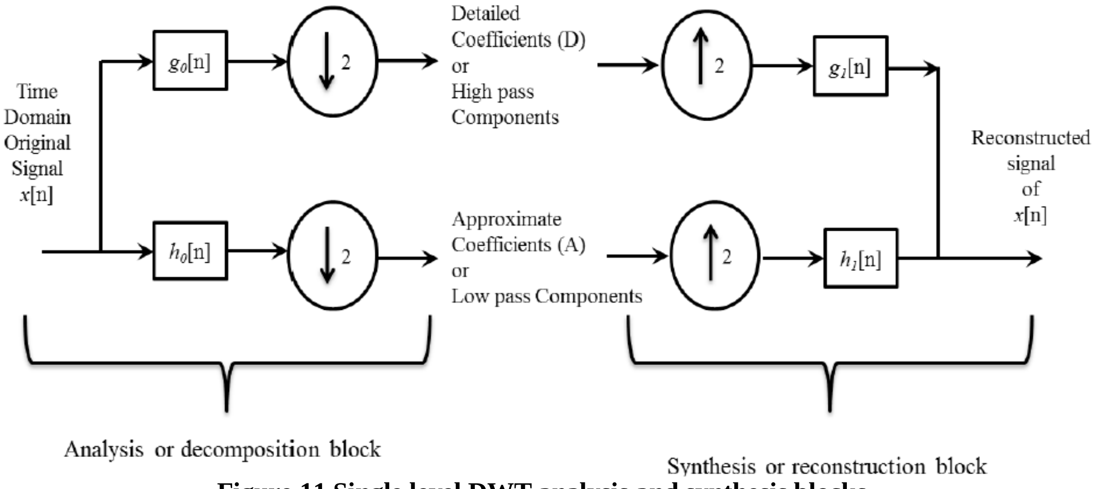

Time Domain Original Signal x[n]
g0[n]
h0[n]
2
Detailed Coefficients (D) or High pass Components
Approximate Coefficients (A) or Low pass Components
g1[n]
h1[n]
Reconstructed signal of x[n]
Analysis or decomposition block
Synthesis or reconstruction block

Figure 11 Single level DWT analysis and synthesis blocks

For the multi-level operation the 1-level DWT procedure is repeated by taking either the low frequency components or the high frequency components or both as in wavelet packets as the input to the one level analysis block. It can be observed that every time some portion of the signal corresponding to some frequencies being removed from the signal. The most common decomposition components chosen are low frequency coefficients. The 3-level DWT decomposition is shown in Figure 12. A1 and D1 are the first level decomposition coefficients of signal x[n]. At the second level A1 is further decomposed into A2 and D2; and A2 is further decomposed into A3 and D3 as explained earlier.

For the reconstruction of the decomposed signal, A3 and D3 are used to find low pass coefficients at level-2 as explained in the single level reconstruction process. The obtained level-2 low- pass signal with D2 is used to obtain low pass coefficients at level-1. The level-1 low frequency components with D1 are used to find the reconstructed original signal.

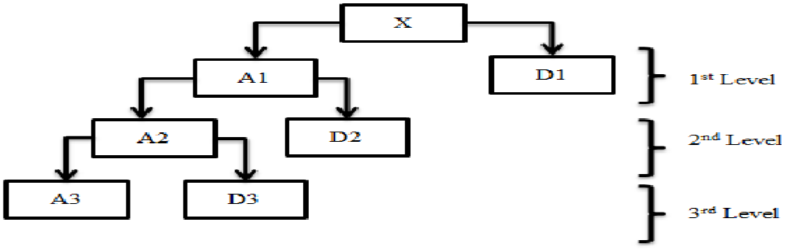


Figure 12 Level DWT decomposition of signal x[n]

From Figure 12, the reconstruction processes can be interpreted and is the inverse of the decomposition process. The approximate coefficients are up-sampled and passed through a low pass filter h_1[n], similarly, detailed coefficients are up-sampled and passed through high pass filter g_1[n]. The obtained samples from these filters are convoluted to obtain the reconstructed signal of x[n].

From Figure 11 it is clear that the original signal can be reconstructed by combining the highest level available decomposed coefficients. In other words x[n] can be reconstructed using high and low pass filters g1[n] and h1[n], respectively.

Future Work

Disadvantages associated with this proposed system are a low data transmission rate due to the fact that the each bit of secret message is embedded into one segment and size of segment is approx 4 to 16 samples. It means that utilization of samples is very poor . . As a result, this method can be used when only a small amount of data needs to be concealed. Otherwise this can be proved as a good method for audio Steganography.

In future we will modify and improve this technique so that more data can be embedded into cover signal.

XI. Conclusion

Steganography is an information hiding technique where secret message is embedded into unsuspicious cover signal. An effective audio steganographic scheme should possess the following three characteristics: Inaudibility of distortion (Perceptual Transparency), Data Rate (Capacity) and Robustness. These characteristics (requirements) are called the magic triangle for data hiding.

We have presented a high capacity and high stego-signal quality audio Steganography scheme based on samples comparison in DWT domain where selected coefficent of a segment are compared with pre determined threshold value T and based on comparison bits are embedded. The strength of our algorithm is depend on the segment size and their strength are enabled the algorithm to achieve very high embedding capacity for different data type that can reach up to 25% from the input audio file size with lest of 35 dB SNR for the output-stego signal.

The proposed scheme was tested for different hiding capacity and the results showed that it has excellent output quality. From the tests we find the proposed algorithm support high capacity rate reach up to 4 kb/sec and that is form above 25% from the size of the input audio cover file at SNR above 35 dB for the output signal.

Acknowledgement

I would like to thank my guide Dr. Namrata Dhanda, (Associate Professor, Department of Computer Science, GITM, Lucknow) who provided special assistance, support and encourage me by their thoughts, my special thanks to Goel Institute of Technology and Management, Lucknow.

References

[1] "Information Hiding: A survey" (pdf). Proceedings of the IEEE (special issue) 87 (7): 1062–78. doi:10.1109/5.771065. Retrieved 2008-09-02.

[2] A New Text Steganography Method By Using Non-Printing Unicode Characters, Akbas E. Ali, Eng. & Tech. Journal, Vol.28, No.1, 2010

[3] B.r., Roshan Shetty; J., Rohith; V., Mukund; Honwade, Rohan; Rangaswamy, Shanta (2009). Steganography Using Sudoku Puzzle. pp. 623–626. doi:10.1109/ARTCom.2009.116.

[4] Chvarkova, Iryna; Tsikhanenka, Siarhei; Sadau, Vasili (15 February 2008). "Steganographic Data Embedding Security Schemes Classification". Steganography: Digital Data Embedding Techniques. Intelligent Systems Scientific Community, Belarus. Retrieved 25 March 2011.

[5] Johnson, Neil; Duric, Zoran; Jajodia, Sushil (2001). Information hiding: steganography and watermarking: attacks and countermeasures. Springer. ISBN 978-0-7923-7204-2.

[6] Krzysztof Szczypiorski (4 November 2003). "Steganography in TCP/IP Networks. State of the Art and a Proposal of a New System - HICCUPS". Institute of Telecommunications Seminar. Retrieved 17 June 2014.

[7] Kundur D. and Ahsan K. (April 2003). "Practical Internet Steganography: Data Hiding in IP". Texas Wksp. Security of Information Systems. Retrieved 16 June 2014.

[8] Steganography, Technique of Sending Random Passwords on Receiver's Mobile (A New Technique to Hide Information File with an Image) by Shubhendu S. Shukla, Vijay Jaiswal, Anurag Singh, Sumeet Gupta, IOSR Journal of Computer Engineering (IOSR-JCE) Volume 15, Issue 3 (Nov. - Dec. 2013), PP 17-25

Author Profile:

Author (Sumeet Gupta[1]), is a research scholar in the field of Information Technology and Computer Science Engineering. He writes the 5 international research papers in his respected field in international reputed journals. He attend 5 national and 3 international seminars along with his studies.

Dr. Namrata Dhanda[2] is working as Associate Professor and Head of Department in the Department of Computer Science and Engineering/Information Technology at Goel Institute of Technology and Management, Lucknow. She has teaching experience of 14 years. Prior to her current assignment she has taught at Amity University, Lucknow and Babu Banarasi Das National Institute of Technology and Management, Lucknow at different positions. Her areas of interest include Automata theory, Database Management Systems and Algorithm design and Analysis.