Robust Audio Data Hiding based Probabilistic Global Search in MDCT Domain

Ei Thin Su

University Of Computer Studies (Yangon)

eithinsu.ucsy@gmail.com

Abstract

This paper managed to propose robust audio data hiding through optimization-based problem transformation method, Probabilistic Global Search Lausanne Algorithm (PGSL) in modified discrete cosine transform (MDCT) domain. We manipulate the MDCT coefficients in the selected frames with the stego key and, then implement the optimal best points searching process through PGSL Algorithm to obtain the best secret audio message embedding positions. The Lowest Bit Error Rate value are defined the optimal best coefficients. The inverse modified discrete cosine transform (IMDCT) is applied to form stego audio. In our audio data hiding scheme, the results of imperceptible (audible difference) between the stego and cover audio will be evaluated using Spectrogram and Informal Listening Tests. The robustness will be tested with standard digital signal manipulation attacks as Noise addition, MPEG compression and Resampling. The goal of proposed data hiding is to protect the privacy of secret audio message from unauthorized accesses, and to satisfy audio steganography issues as imperceptible, embedding capacity and robustness.

Keywords: *Audio Data Hiding, Audio Steganography, Modified Discrete Cosine Transform, PGSL Algorithm, optimal best points searching process.*

1. Introduction

Today's, multimedia and internet technology have been rapidly development, but these technologies have not fully strong until now, so malicious users can easily attack the security and privacy contents. Thus, it is becoming important to prevent privacy and security break. Data Hiding is one of the techniques for copyright protection and authentication control. Imperceptibility and robustness are the most important requirements for any data hiding scheme. Many applications uses of data hiding are: In watermarking applications, information is hidden for digital right management and copyright protection. In music commercial applications, hidden data is inserted within music for monitoring of the songs over broadcast radio. In audio forensics applications, inserting hidden data into audio files for the authentication of spoken words. In Steganography applications, secret message is embedded in the cover object that the existence of message can't be detectable from unauthorized users between sender and receiver.

Steganography is the art of writing hidden message in the cover object to get stego object. The purpose of Steganography is the secret communication by hiding the message to prevent from third parties.

Nowadays, to transmit secret message, a variety of transport medium have been used as a carrier such as audio, video, and image media files, text, program files, telephone signals, radio signals, etc.

Robust Audio Data Hiding is to identify the sender or owner, to hide commercial information, to protect the privacy of our personnel contents and to improve the security of our communication.

Audio Steganography Properties*:*

Audio Steganography is one of the most effective ways to prevent our privacy. The goal of audio steganography is a covert communication by hiding secret message into the cover audio file from a third party using imperceptible techniques and keeping the presence of information secret. By using these techniques any unauthorized users or attackers can't be noticed the presence of secret message within the cover object. Therefore Audio Steganography is hoped to be build in secure applications which allows users to be enjoy more privacy and security at the same time. During the past decades, many novel researches have been done in audio steganography separately.

The stego audio should be robust against standard digital signal manipulations attacks. Moreover, the differences between the original and the stego audio to be imperceptible to the human ear. Audio Steganography can be characterized by three important properties as robustness, imperceptibility and capacity. Robustness: Evaluates the audible distortion due to signal manipulation attacks such as MPEG compression, noise addition, resampling, can be performed on the stego audio before it reaches destination. These manipulations may destroy the hidden message. Imperceptibility: Measures perceptual similarity between cover audio and stego audio. The strength of audio steganography lies imperceptibility to be unnoticeable by the human ear. Payload Capacity: Indicates the total number of hidden bits and successfully recovered by the Stego media. Steganography intended at hidden communication, which needs sufficient embedding capacity.

According to the implementation process of audio steganography, secret audio embedding domain can be divided into the time domain and transform domain methods. In time domain schemes, the hidden bits are embedded directly into the time signal samples. From the view of the performance of audio steganography against attacks, the transform-domain methods are considered better than that of the time-domain methods.

Thus, we focus on transform-domain technique and transformation is done by applying MDCT as secret message embedding domain. The MDCT is a type-IV DCT; the basic DCT is type II. MDCT is a significant purpose in perceptual audio coding and particularly appealing for signal compression applications. MDCT has an advantage that it can be done Perfect Reconstruction. Thus, high imperceptible can get in audio data hiding by applying it. MDCT belongs to the compaction qualities of the DCT. So, it can give large payload capacity in audio data hiding system. Moreover, MDCT can avoid signal compression errors. Thus, stego audio withstands one of signal modification attack as MPEG compression thus can satisfy the robustness issues by using MDCT in audio Steganography applications. As a result of these advantages, the MDCT is utilized in our audio data hiding system.

In our proposed system, robust audio data hiding scheme based on PGSL (Probabilistic Global Search Lausanne) Algorithm is presented. In this scheme, firstly input cover audio is transformed by using MDCT (Modified Discrete Cosine Transform) and manipulated the MDCT coefficients in the selected frames depending on the stego key which is generated by the pseudo random number generator. To take the efficient audio Steganography, the data hiding without perceptible is important. Thus, in our propose scheme, searching the lowest bit error rate coefficients as the optimal best points to embed the secret audio message through Probabilistic Global Search. Afterwards, the inverse modified discrete cosine transform (IMDCT) is employed to get the stego audio.

The rest of our presented paper is organized as follows: Section 2 introduces the related work of our proposed system. Section 3 discusses Background Theory and Section 4 described Proposed Robust Audio data hiding based PGSL. Section 5 predict Expected outcome. Section 6 Concludes the paper and Section 7 mention References of our proposed system.

2. Related Work

Audio data hiding is and active research area that has been strongly motivated to solve the privacy protection problem of secret communication. Many excellence audio steganography techniques have been proposed to be effective, nevertheless, and due to the challenging nature of audio signal processing, there remains much to do.

The previous methods in the audio steganography are Low-bit Encoding, Phase coding, Echo data hiding and Spread Spectrum coding. General methods used (lower sensitivity of Human Auditory System) by a little changing of frequency, phase and amplitude of audio sound files have also been proposed. In *Low-bit encoding*, the secret message is substituted with the least significant bit (LSB) of each sample of the audio cover file. This method is simple and can be used to embed larger messages, but cannot protect the hidden message from small modifications. Thus, it is rarely used in the real commercial applications.

Phase coding is based on the phase components of sound as imperceptible which the differences between cover and stego audio to the human ear. Message bits are encoded as phase shifts in the phase spectrum of a digital signal. This leads to inaudible encoding. A characteristic feature of Phase coding is the low data transmission rate that the secret message is encoded only in the first segment of the audio signal. An increase in the length of the segment would have a ripple effect by altering the phase relations between the frequency components of the segment, thus making detection easier. Hence, the Phase Coding method is used only when a small amount of data embedded. In *Echo hiding*, secret message is embedded by introducing an echo into the discrete audio signal. Echo hiding allows for a higher data transmission rate than least significant bit insertion method. *Spread Spectrum coding* method spreads the secret message across the frequency spectrum of the audio signal. It is robust than Phase coding and Low-bit coding but use complex algorithms.

In [1], the author proposed Genetic-Algorithm Based-Approach for audio Steganography that can withstand attacks like noise addition than simple LSB method. In [2], the authors presented audio Steganography by Cepstrum Modification which combines psychoacoustical masking property of HAS with the decorrelation property of speech cepstrum, that achieved high imperceptible embedding and correct data recovery but low payload. The method of [3] has proposed increasing the capacity of LSB-based audio Steganography using a Novel Embedding Method. This can be achieved against noise addition or MPEG compression better than standard LSB method. In [4], the author proposed Speech Information Hiding Using Fractional Cosine Transform. They find in their analysis that the 3rd parameter of FRCT presents property of high frequency, then they use low bit embedding method and over zero rate detection method for embedding secret information into speech signal in FRCT domain. They discoveries in their study provide a new thought for application of FRCT in speech information hiding.

In [5], the author presented Telephony Speech Enhancement by Data Hiding to improve intelligibility and perceived quality of telephone speech. Data is hiding based on the perceptual masking principle; the inaudible spectrum components within the telephone bandwidth can be removed without degrading the speech quality, providing a hidden channel to transmit extra information. While this hidden signal is imperceptible to the human ear, it can be extracted at the receiver end. They showed that robust to quantization errors and channel noises. In [6], the authors proposed “Robust Information Hiding in Speech Signal Based on Pitch Period Prediction”. In this system, first estimates the pitch period of short time speech signal, and then predicts the next period data by the previous one with one parameter according to the least-mean-square-error criterion. Based on the basic idea of Scalar Costar Scheme (SCS), the secret information is embedded into the predict parameter and thus generates the stego speech. It can resist to attacks such as additive noise, resampling, especially to speech codecs employed in GSM and CDMA.

In our proposed system, audio message is hidden within the best coefficients of selected frames in the MDCT Domain through Probabilistic Global Search for audio steganography applications with robustly, imperceptibility as well as large payload capacity.

3. Background Theory

This section describes PGSL Algorithm and MDCT domain for our proposed robust audio data hiding in detail.

3.1. Modified Discrete Cosine Transform

The modified discrete cosine transform (MDCT) is type-IV discrete cosine transform (DCT-IV). The basic DCT is type II. MDCT provides better energy compaction than DCT.

Properties of MDCT

MDCT can be directly estimated the capacity of a transformation scheme by its ability to pack input data into as few coefficients as potential. This allow without audible distortion in the reconstructed audio. It avoids errors between block boundaries. It is a significant purpose in perceptual audio coding and particularly appealing for signal compression applications. Hence, it can meet the capacity issue which is one of the important issues of our audio data hiding scheme. It has the principle of time-domain aliasing cancellation (TDAC) that is to cancel the errors by adding subsequent overlapped IMDCT blocks (inverse modified discrete cosine transform). So it can be done perfect reconstruction. Because of this advantage, we can fulfill imperceptible issues that the differences between the original and stego audio in our audio data hiding scheme. Moreover, it deflects signal compression errors. This fact can perfect robustness issues in audio steganography that the stego audio withstands even MPEG compression which is one of the signal modification attacks. Therefore, the most important audio data hiding issues of imperceptible, robustness as well as capacity can be accomplished in our audio steganography by applying the MDCT transform-domain in general. Thus, we choose MDCT domain as the secret audio message embedding domain in our proposed robust audio data hiding technique.

3.2. PGSL Algorithm

PGSL (Probabilistic Global Search Lausanne Algorithm) is the global optimization algorithm**.** It has been developed by Raphael and Smith in 2002. It is applied in the field of structural engineering to solve the optimization problems for bridge diagnosis. It is a direct search algorithm, to find the minimum of user defined objective function by using global sampling [7]. In PGSL, optimal solutions can be defined through focusing search around sets of good solutions. Tests on benchmark problems having multi-parameter non-linear objective functions exposed that PGSL is advanced algorithms and performs better than genetic algorithms. Moreover, PGSL performs better than other approaches even increased the problem size.

In the fields of diagnosis, design and control, PGSL has already established to be valuable for engineering tasks. PGSL helps to solve complicated problems in a high convergence speed. There is no need to a large number of interrelated parameters. It has the advantage of fewer search parameters.PGSL algorithm has been first used to design the interdigital transducer (IDT) structure of the surface acoustic waves (SAW transducer) [8]. To aim at a great advance in the transducer characteristics and decreasing in the lead time of the product design. In PGSL Algorithm, a uniform probability density function is accepted for entire search space in the beginning of search. When good solutions are found, increased the probabilities in these regions. Better sets of points are found in the neighborhood of good sets of points. So, search space is more focus on an area of best points then the convergence is achieved. PGSL algorithm allows four nested cycles: Sampling Cycle, Probability Updating Cycle, Focusing Cycle and Subdomain Cycle

Sampling Cycle- Number of samples are randomly generated according to the current probability density function. Each point is evaluated by user-defined objective function and selected the best point. Probability Updating Cycle- The sampling cycle is brought up number of probability updating cycle times. After each iteration, the probability density function of each variable is modified using the probability-updating algorithm. This ensures that the sampling frequencies in regions containing good solutions are increased and regions containing bad points are decreased. Focusing Cycle-Search is focused on the interval containing the best solutions after a number of Probability Updating Cycle, by further subdivision of interval. Subdomain Cycle-Search space is increasingly narrowed by selecting a subdomain of smaller size center on the best point after each Focusing Cycle.

In our audio data hiding scheme, the best coefficients which have the lowest bit error rate value are seeked based on the optimized PGSL (Probabilistic Global Search Lausanne) Algorithm to embed the secret audio message with robustly and imperceptibility.

4. Proposed Robust Audio data hiding based PGSL

This section describes Pre-processing; PGSL based optimal best points searching process, secret audio message embedding and extraction procedure in detail.

4.1. Pre-Processing

Clean-host speech from the TIMIT database, with 16 bit speech samples at the rate of 16,000 per second are used as the cover audio for embedding process. And then windowing and transformation are applied.

In signal representation, the Modified Discrete Cosine Transform (MDCT) is a powerful tool in high quality audio coding. MDCT (type IV DCT) can get better compaction quality of the Basic DCT (type II DCT). So, the secret audio can be transmitted with high data transmission rate to reach the destination. MDCT achieves reduction of block effect as it has time-domain aliasing cancellation (TDAC) properties. Thus, it could be achieve perfect reconstructed without error. In that case, all of the secret massage bits can correctly retrieved at the receiver site. Moreover, MDCT can avoid compression errors. Hence, in our audio data hiding scheme can satisfy MPEG compression attack which is one of the digital signal manipulation attacks. According to these effects, we applied Modified Discrete Cosine Transform as the secret audio message embedding domain in our proposed audio data hiding scheme.

The random secret key is generated by using the pseudo random number generator for embedding the secret audio message into cover audio.

4.2. PGSL based optimal best points searching process

To get the optimized position for secret message embedding process, the best coefficients of selected frames in MCDT (modified discrete cosine transform) domain are searched through PGSL (Probabilistic Global Search Lausanne) algorithm by using the objective function.

Define objective function and the parameters: To find the best embedding points, the smallest bit error rate (BER) value are defined as the objective function. The minimum, maximum values, and the number of cycle times in PGSL algorithm.

4.3. Embedding Procedure

The embedding procedure performs at the sender side, firstly define ***Input*** as: take 16 bit speech samples at the rate of 16,000 per second as the original cover audio.

Windowing and Transformation with modified discrete cosine transform (MDCT) are applied into non-overlapping frames on the cover audio. For security, the secret key is generated by pseudo random number generator.

MDCT coefficients in the selected frames of the transform domain is manipulated with the secret key. Each bit of secret message is embedded in the best coefficients of each frame. The best coefficients which have the Lowest Bit Error Rate values between cover and stego audio are searched through PGSL Algorithm.

After the embedding, the Inverse Transformation (IMDCT) is applied to form new audio sequence as stego audio.

4.4. Extraction Procedure

In the receiver site, extract the stego object to view the secret information. The embedded bits are recovered from the stego samples by reversing the embedding procedure.The receiver can extract the embedded message with the correct key and restore the stego audio to the original audio.

The above embedding and extraction procedures were carried out in the MDCT transform of the host samples. We will be expected to correct data recovery; moreover, the error-free secret communication to be established withstand digital signal manipulation attacks on the stego audio.

The audible difference between the Original cover audio and the Stego Audio (imperceptibility) will be evaluated by Spectrogram and Informal Listening Test. Whether the stego audio to be roubstness or not, test will be applied with Standard Digital Signal Manipulation attacks as MPEG Compression, Noise addition and Resampling.

In Figure 1, the flow of our proposed system based on PGSL is displayed.

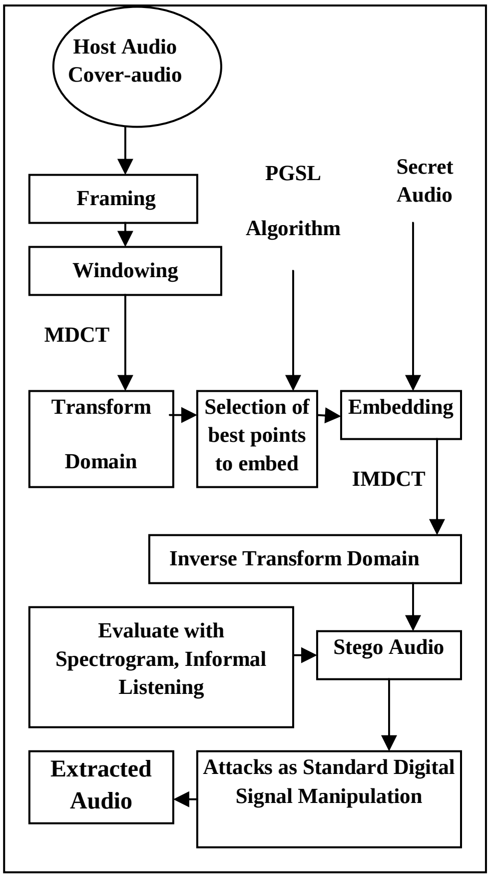

Figure. 1. Block Diagram of Proposed System

5. Expected Outcome

Proposed system will use clean-host speech from the TIMIT database, with 16 bit speech samples at the rate of 16,000 per second are used as the cover audio. At first, the host audio samples are transformed with modified discrete cosine transform (MDCT) to get MDCT coefficients values. The MDCT coefficients in the selected samples of the host speech are manipulated with the random key, and the secret message bits are embedded in the best coefficients. Optimization-based Probabilistic Global Search Algorithm is applied to get the best coefficients. For security, random key is used by generate with pseudo random number generator at the transmitter. Inverse transform (IMDCT) are applied to get the stego audio. Imperceptibility between original cover audio and stego audio will be evaluated with Informal Listening Tests and Spectrograms. The stego audio will be tested with standard digital signal manipulations attacks as Noise addition, MPEG compression and Resampling. Robustness will be measured with bit error rate (BER). Proposed method addresses the problems of having low robustness against signal manipulation. Better imperceptibility and robustness as well as large capacity can be expected upon the optimize evaluation of audio data hiding for Audio Steganography applications.

6. Conclusion and Future Work

In this paper, we have introduced a technique of robust audio data hiding for audio steganography. As an ongoing research, the embedding procedure has been build with the optimization algorithm and currently testing with the clean host speech with the TIMIT data sets. We are now building the embedding process and testing with PGSL algorithm. The perfect experimental results emerged from our testing will be our future work. We believe that our proposed system can provide in the security and an authentication of audio steganography applications with successfully.

7. References

[1]. M. Zamani, A. A. Manaf, R. B. Ahmad, A. M. Zeki, and S. Abdullah[5] "A Genetic-Algorithm-Based Approach for Audio Steganography ", World Academy of Science, Engineering and Technology 54 2009.

[2]. K. Gopalan "Audio Steganography by Cepstrum Modification"*ICASSP, IEEE*, 2005.

[3]. N. Cvejic & T. Seppanen "Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method ", *IEEE 2004*.

[4]. Y. Tao, Y. Bao, "Speech Information Hiding Using Fractional Cosine Transform," First International Conference on Information Science and Engineering, 2009.

[5]. S. Chen ; H. Leung, H. Ding, "Telephony Speech Enhancement by data hiding "Instrumentation and Measurement, *IEEE Transactions on*, Feb 2007.

[6]. Chunyan, B. Sen, "Robust Information Hiding in Speech Signal Based on Pitch Period Prediction," Computational and Information Sciences, International Conference on, pp. 533-536, 2010.

[7]. B. Raphael, I.F.C. Smith, "A direct stochastic algorithm for global search "Applied Mathematics and Computation 146 (2003) 729–758.

[8]. G. Huang, Q. Zhao, L. Zhao, S. Li, J. Lv ,F. Wang, J. Liao "Optimization of the SAW Transducer Design by Probabilistic Global Search Lausanne "Progress In Electromagnetic Research Symposium, Hangzhou, China, March 24-28, 2008