Secured Genetic Algorithm Based Image Hiding Technique with Boolean Functions

Krishna Bhowal, Debasree Sarkar, S. Biswas, and Partha Pratim Sarkar
(Corresponding author: Krishna Bhowal)

Department of Engineering and Technological Studies, University of Kalyani
Pin - 741235, west Bengal, West Bengal, India (Email: ykbhowal@yahoo.co.in)

(Received Sept. 29, 2014; revised and accepted Jan. 16 & June 21, 2015)

Abstract

Steganography and Watermarking are main parts of the fast developing area of information hiding. Steganography involves hiding of information in a cover media to obtain the stego media, in such a way that the cover media is supposed not to have any embedded image for its unintended recipients. This paper is based on Steganography, Watermarking and Cryptography system where image bits are embedded into higher random LSB layers of audio signals, resulting in increased robustness against noise addition. On the other hand, multi-objective Genetic Algorithm is used to minimize the deviation between original media and embedded media. The basic idea of this paper is to improve security so that probability of detecting the presence of hidden information into cover media is about to zero. For this improvement, image embedding random position numbers are converted to functions and these functions are sent using Symmetric-key encryption algorithm to the receiver end. Key distribution problem is solved by RSA algorithm. We evaluated performance based on imperceptibility, security, robustness, and hiding capacity.

Keywords: Artificial intelligence, genetic algorithm, steganography, watermarking

1 Introduction

Steganography, watermarking and fingerprinting are branches of information hiding. In a computer-based data hiding techniques in audio, secret image is hidden in digital audio signal so that they can be extracted at the receiving end with the help of a secret key and not merely to obscure its presence. The secret image is embedded by slightly altering binary sequence of audio signals.

Multimedia data hiding techniques have developed a strong basis of growing number of applications like copyright protection, authentication, tamper detection, covert communications etc. Following requirements must be satisfied in several applications [5, 6, 31, 32].

Perceptual Transparency: The main focus of this paper is on perceptually undetectable or transparent data-embedding and watermarking techniques. In many applications, such as covert communication, copyright and usage tracking, embedding metadata or additional information, the algorithms must embed data without affecting the perceptual quality of the underlying host signal.

Recovery of Data without Access to Original Signal: In most of the applications such as covert communication, data-embedding algorithms do not have access to the original audio signal while extracting the embedded signal. This inability to access the original signal limits the amount of data that can be embedded in a given host signal.

Bit Rate of Data-Embedding Algorithm: Some applications of data embedding require small amounts of information to be incorporated. On the other hand, many applications of data embedding, e.g., covert communication, require a lot of bandwidth. The ability to embed large quantities of data in a host signal will depend on how the embedding algorithm has been designed. Our algorithm can adapt large amount of information to the underlying host signal.

Robustness: Digital data are modifiable and manipulate-able using computers and widely available software. Operations that damage the host signal also damage the embedded data. Again, third parties may attempt to modify the host signal to detect of the embedded data. Basic requirement of steganography imposes that the presence of hidden information within the stego-cover media should be undetectable. There should be no perceptible difference between the watermarked and original signal, and the watermark should be difficult to remove or alter without damaging the host signal.

Security: A secure data-hiding procedure can only be broken by the authorized user has access to a secret key that controls the insertion of the data in the host signal. This requirement is very important in covert communication scenarios. Hence, a data-hiding scheme is secure if knowing the exact algorithm for embedding the data does not help an unauthorized party to detect the presence of embedded data. An unauthorized user should also be unable to extract the data in a reasonable amount of time. Basic Data hiding process has been shown in Figure 1.

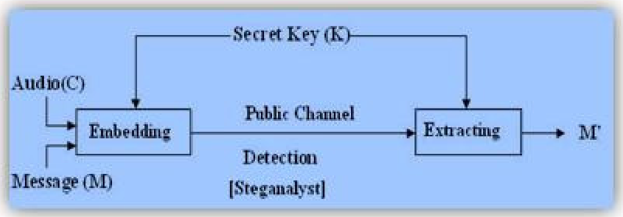

Figure 1: Basic data hiding process

LSB coding is one of the earliest techniques studied in the information hiding area of digital audio. The main advantage of the LSB coding method is a very high channel bit rate and a low computational complexity of the algorithm, while the main disadvantage is considerably low robustness against signal processing modifications. Since substitution techniques usually modify the bits of lower LSB-layers in the samples, it is easy to reveal the hidden image if the low transparency causes suspicious. In order to conceal secret image successfully, a variety of methods for embedding information in digital audio have been introduced [1, 4, 6, 7, 8, 12, 16, 21, 28, 31].

It is well known that LSB-layers bits in samples are more suspicious, so embedding the image bits other than LSB-layers could be helpful to decrease the perceptibility and to increase the robustness. The basic idea of this research work is to provide a novel method to hide the secret data from intruders at high random LSB layers. Then the secret data will be sent to the destination in safer and secure manner. The quality of sounds depends on the length of the image and size of the audio which are selected by the users. Even though it shows changes in bit level deviations in the frequency chart, as a whole we cannot determine the change in the audio. Here the technical challenge is to provide transparency and robustness which are conflicting requirements. The perceptibility and extraction of hidden information of the proposed algorithm is more challenging as well, because there is a significant number of bits flipped in a number in bit layers. On the other hand, image retrieval from random higher LSB layers is still one of the major drawbacks of the modified LSB methods. In this paper, we have been used Boolean functions to extract the hidden information location numbers.

The remainder of the paper is organized as follows: Section 2 discusses related works done by different researchers. Section 3 explains proposed work. Section 4 discusses experimental results. Section 5 highlights advantages of our approach. Section 6 concludes the paper presented here.

2 Related Works

Being a simple method, a very high level of security is not achieved in LSB insertion method. To improve security, different modified-LSB methods are proposed by different researchers. Apart from security, certain other parameters like complexity, computational load, SNR, Bit Error Rate, efficiency, etc are also considered for information hiding techniques.

In [31] a solution of supporting different audio formats and reducing the time for encoding and decoding are discussed. Data is embedded in such a way that each character requires eight 254/255 bytes.

In [6] a two steps method is proposed where data are embedded from the fourth to sixth LSB layers with minimum distortions. First, secret bit is embedded in any higher LSB layer. Second, changes white noise properties by shaping the impulse noise which is caused by the embedding bit. In [7, 8] the hearing threshold in the temporal domain is calculated which is exploited as the embedding threshold, yielding more capacity compared to uniform embedding pattern. Proposed method uses compression of information using lossless compressor, thus increasing total bit rate. In [28] an algorithm is proposed for both image and audio steganography. Regarding audio steganography, it states the technique of echo hiding. Data is embedded in the echo signal varying its parameters like decay rate, offset and amplitude. The original signal is segmented into blocks and each block is given the value 0 or 1 depending upon the secret message. The original signal is echoed and the message is embedded into it. At the receiver end, auto correlation and decoding is done to separate the secret signal and the original signal.

In [21] a robust steganographic method is proposed where data are embedded in the multiple, vague and higher LSB layers. Generally there are two types of attacks namely unintentional attacks and intentional attacks, solutions are suggested for both these type of attacks in this work. The data bits are embedded in the bit other than 1st LSB bit to stop the intentional attack. The bits other than the selected bits for embedding are altered to reduce the distortion. In [12] two LSB methods are proposed. First method is parity coding and the other is XORing of LSB. Initially embedding capability is measured by ensuring that the size of the message to be embedded is less than the cover audio signal. In parity method, the parity bit is considered before directly replacing the LSB. Depending upon the message bit to be embedded, the LSB is either flipped or retained. If the message bit is 0, LSB has to be modified in such a way that parity of the sample is even. If the message bit is 1, LSB is modified in such a way that parity of the sample is odd. In second method, XORed operation between the LSB and the next bit has to be equivalent to the message bit to be embedded. If equal, the LSB is retained otherwise LSB is flipped. Also they reduce the computational load and the capacity of the cover audio is increased. From experimental results it is found that the encryption with steganography provides better security. Embedding data in the higher LSB layers is prone to less attack than those embedded in the lower layers. But embedding in higher LSB will result in distortion. Therefore further steps have to be included to reduce these distortions.

The idea proposed by [16] is based on psycho acoustic theory of persistence and phase shifting. Persistence of hearing is based on the fact that two sounds successively with a difference of less than one-tenth of a second hit our ears, then the difference between the sounds is imperceptible. It is called the phase shift, the change of which is same as the shift in time. Author used uncompressed audio format (WAV format). In [1] the author has shifted the LSB embedding to the eight bit resulting in slight increase of robustness. However, the hiding capacity will be decreased since some of the samples are to be left unchanged to preserve the audio perceptual quality of the audio signal. In [2, 11, 15, 17, 22, 29, 33, 34], different embedding procedures are followed to hide data in image or audio file. Most of the cases they explained how to hide the information in to medium but how to extract the hidden data from medium at receiver end is not clearly mentioned. It does not raise suspicions that an important message can be possibly carried inside a harmless medium in steganography describes in [14]. Hiding a secret message in order to protect the copyright of a product is the main aims in watermarking are discussed in [3, 18, 25, 26, 30]. To demonstrate its authenticity, namely, its content originality also referred as content verification, or tamper proofing in [2]. An adversary tries to reveal the information carried by a stego-medium. In the case of watermarking, an opponent either tries to remove the watermark in order to violate copyright or to reproduce it after product tampering in order to achieve a false positive content verification.

The easiness of image retrieval is still one of the major drawbacks of the LSB and its variant, knowing by fact that embedded bits are at sixth or eighth position from the stego audio signal. To solve this problem, Boolean functions have been introduced in this paper by which we can easily extract hidden bits embedded in different random LSB positions at the receiving end.

3 Proposed Work

3.1 Best Sample selection using Genetic Algorithm

Genetic Algorithms are adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. Unlike AI systems, they do not break easily even if the inputs changed slightly, or in the presence of reasonable noise. Also, in searching a large state-space, a genetic algorithm may offer significant benefits over more typical search of optimization techniques.

A population of individuals is maintained within search space for a GA, each representing a possible solution to a given problem. Each individual is coded as a finite length vector of components, or variables, in terms of some alphabet, usually the binary alphabet 0, 1. To continue the genetic analogy these individuals are likened to chromosomes and the variables are analogous to genes. Thus a chromosome (solution) is composed of several genes (variables). A fitness score is assigned to each solution representing the abilities of an individual to 'compete'. The individual with the optimal (or generally near optimal) fitness score is sought. The GA aims to use selective reproduction of the solutions to produce 'offspring' better than the parents by combining information from the chromosomes.

The following points convinced us to use genetic algorithm in this work

1) To maintain the randomness in selection of bit level of audio sample for hiding secret bit.

2) Through GA based crossover and mutation operations on audio sample (chromosome) we may get better set (population) of audio samples than previous generation population.

3) Using the concept of fitness value in GA we may select better or best audio sample from the population of audio samples generated in the previous step.

4) We may consider fitness value is a position number of audio sample where secret bit may be embedded and for which deviation between original audio sample and stego-audio sample is minimized.

5) Using the concept of Multi-objective GA, position number of audio sample where secret bit is embedded may be used

   a. As a Fitness value to select the best chromosome (audio sample).

   b. To extract the hidden secret bit from the stego-audio at the receiving end.

3.2 Steps to Embed Image into Audio File Using Proposed Modified LSB Scheme

1) Read image file and generate byte streams.

2) Read audio file and generate byte streams, convert byte streams to 16 bit audio samples.

3) Obtain n number of chromosomes of 16 genes by inserting two image bits into 16 bits audio sample at n (2 to16) random positions.

4) Apply following GA operator based insertion algorithm to generate better next generation population:

   a. Let pos be the image bit insertion position into the audio sample;

   b. Let $fm(pos)$ be mutation operation on pos position;

   c. Let $fc1(start, end)$ be crossover operation from start to end by 1 and $fc0(start, end)$ be crossover operation from start to end by 0.

   d. If pos = 1, then take no action;

   e. If pos = 2 to 16 then do the following

      i. If image bit is 0 and audio bit is 1 for pos = i;

      ii. If audio bits on 1 to (i-1) positions are holding 0s, then perform $fc1(1, i-1)$ operation;

      iii. If audio bits on 1 to (i-1) positions are holding 1s and on (i+1) position holding 0, then perform $fc0(1, i-1)$ and $fm(i+1)$ operations.

      iv. If image bit is 1 and audio bit is 0 for pos = i;

      v. If audio bits on 1 to (i-1) positions are holding 1s, then perform $fc0(1, i-1)$ operation;

      vi. If audio bits on 1 to (i-1) positions are holding 0s and on (i+1) position holding 1, then perform $fc1(1, i-1)$ and $fm(i+1)$ operations;

      vii. If audio bit on (i+1) and (i-1) positions are holding 0/1 and 1/0 respectively, take no action;

      viii. If audio bit and image bit is same, then no action is to be taken as there will be no deviation between two samples.

5) Now, the best chromosome has been selected, where best one is the chromosome (audio sample) which has the minimum deviation compare to the original 16 bit audio sample.

6) Here fitness value represents the position number for which we get the best chromosome. Again, the position number, best chromosome and distortion are closely related as selection of the best chromosome will reduce the distortion.

7) Fitness value is representing two things here:

   a. Position number which is very important at the receiving end to extract the image.

   b. Distortion which again very important regarding security (distortion can convinced hacker to hack the image).

   So, multi-objective GA is used here.

8) Secret-Bit-Insertion Positions have been stored in Position Arrays during this embedding process.

9) Stego-audio byte streams have been written into audio file.

10) Boolean functions have been generated from the Position Arrays are described in the next section.

Boolean Functions Generation from Position Arrays:

Let S be size of the Position Array.

S_{bin} be binary representation of $S - 1$.

N be number of bits required for S_{bin}.

Suppose A_{ij} be a Position Array where i = 0 to S - 1, j = 0 to 1. Here $A_{i,0}$ represents Array index and $A_{i,1}$ represents Position numbers.

$IPA_N(A_{i,0})$ be N bits binary representation of Index i of Position Array, where i = 0 to S - 1;

$IPA_3(A_{i,1})$ be 3 bits binary representation of Index i of Position Array, where i = 0 to S - 1;

Now we get a Matrix $M_{i,N+3}$ by combining $IPA_N(A_{i,0})$ and $IPA_3(A_{i,1})$. Elements of this matrix are 0 or 1.

$$\begin{aligned}
MSB(i) &= MSB \quad \text{bit set of Position Array} \\
&= MSB(IPA_3(A_{i,1})) \\
&= M_{i,N+1} \\
MdSB(i) &= \text{Middle bit set of Position Array} \\
&= MdSB(IPA_3(A_{i,1})) \\
&= M_{i,N+2} \\
LSB(i) &= \text{LSB bit set of Position Array} \\
&= LSB(IPA_3(A_{i,1})) \\
&= M_{i,N+3}.
\end{aligned}$$

where $i = 0$ to $S - 1$. Again,

$$\begin{aligned}
MSB(i)_{minterm} &= \text{i's where } M_{i,N+1} = 1 \\
MdSB(i)_{minterm} &= \text{i's where } M_{i,N+2} = 1 \\
LSB(i)_{minterm} &= \text{i's where } M_{i,N+3} = 1 \\
SOP(MSB(i)_{minterm}) &\longrightarrow f_x(a, b, \cdots, N \text{ terms}) \\
SOP(MdSB(i)_{minterm}) &\longrightarrow f_y(a, b, \cdots, N \text{ terms}) \\
SOP(LSB(i)_{minterm}) &\longrightarrow f_z(a, b, \cdots, N \text{ terms})
\end{aligned}$$

Here Size of the Position Array need to be sent to the receiver.

The functions and size of the Position Array has been encrypted using Shared Key AES encryption algorithm.

The Shared Key has been encrypted using Public Key RSA algorithm.

3.3 Steps to Extract Hidden Image from Stego-Audio File

Deviation between audio samples and stego-audio samples has been minimized during the proposed embedding method. So stego-audio is almost equal to the original audio. By getting the image hidden position numbers i.e., Position Array, we can easily extract the hidden image from audio file as follows:

1) Read stego-audio file and generate byte streams, convert byte streams to 16 bit audio samples.

2) Decrypt the Shared Key using the receiver Private Key of RSA.

3) Decrypt the functions and size of the Position Array using Shared Key of AES algorithm.

4) The position numbers of the secret (image) bits have been extracted using the Boolean functions and size of the Position Arrays is described below.

   We have f_x $(a, b, c, \cdots, N$ terms), $f_y(a, b, c, \cdots, N$ terms) & $f_z(a, b, c, \cdots,$ upto N terms) and $IPA_N(A_{i,0})$ for i = 0 to S - 1.

   For index i = x, x has been converted to binary number of N bits like $b_1, b_2, ..., b_N$.

   Assigning $a = b_1, b = b_2, c = b_1, ...$;

   Using f_x(a, b, c,.... upto N terms) we get bit b11;

   Using f_y(a, b, c,.... upto N terms) we get bit b12;

   Using f_z(a, b, c,.... upto N terms) we get bit b13;

   Finally the bit pattern b11b12b13 is generated;

   The secret bit position numbers has been generated by converting this bit pattern to decimal number for index i=x and hidden image bits has been extracted from audio file.

5) The image bits have been converted to bytes and original image has been generated from image bytes.

3.4 Transferring Embedding Position Numbers to the Receiver end in Terms of Boolean Functions: An Example

Following Boolean algebra terms are used to generate Boolean functions: - midterm, Sum of Product, Minimization etc. Let the size of the Position Array is 8 and the corresponding positions of image bits are 2,4,6,1,7,4,6,5.

Here Indexed Position Array and Equivalent Binary Representation are explained in Table 1 and Table 2.

Now MSB bit set X= [0,1,1,0,1,1,1,1]. So, Midterm for MSB is [1,2,4,5,6,7]. And Sum of Product $X = a'b'c + a'bc' + ab'c' + abc' + abc + ab'c$.

Table 1: Indexed position array

Index	Position
0	2
1	4
2	6
3	1
4	7
5	4
6	6
7	5

Table 2: Equivalent binary representation

a	**b**	**c**	**X**	**Y**	**Z**
0	0	0	0	1	0
0	0	1	1	0	0
0	1	0	1	1	0
0	1	1	0	0	1
1	0	0	1	1	1
1	0	1	1	0	0
1	1	0	1	1	0
1	1	1	1	0	1

Minimization of X:

Pass 0: $a'b'c + a'bc' + ab'c + abc' + abc + ab'c$;

Pass 1: $a'b'c + ab'c$ reduce to $b'c$; $a'bc' + abc'$ reduce to bc'; $ab'c + ab'c'$ reduce to ab'; $abc' + abc$ reduce to ab;

Pass 2: $ab' + ab$ reduce to a. Finally $X = a + b'c + bc'$. Middle bit set Y=[1,0,1,0,1,0,1,0]. So Midterm of Y is [0,2,4,6].

Sum of Product $Y = a'b'c' + a'bc' + ab'c' + abc'$.

Minimization of Y:

Pass 0: $a'b'c' + a'bc' + ab'c' + abc'$;

Pass 1: $a'b'c' + a'bc'$ reduce to $a'c'$; $ab'c' + abc'$ reduce to ac'; $a'c' + ac'$ reduce to c';

Pass 2: c'; Y = c';

LSB bit set Z=[0,0,0,1,1,0,0,1]. So Midterm of Z is [3,4,7]. Sum of Product $Z = a'bc + ab'c' + abc$.

Minimization of Z:

Pass 0: $a'bc + ab'c' + abc$;

Pass 1: $a'bc + abc$ reduce to bc;

Pass 2: $bc + ab'c'$.

So, the three functions are given below

$$\begin{aligned} X &= a + b'c + bc'; \\ Y &= c'; \\ Z &= bc + ab'c'. \end{aligned}$$

Continuing with the previous example, for the input of the binary of (0-7), we get three outputs from the functions X, Y and Z. Extraction of position number explained in Table 3.

Table 3: Index to position number conversion

a	b	c	X	Y	Z
0	0	0	0	1	0
0	0	1	1	0	0
0	1	0	1	1	0
0	1	1	0	0	1
1	0	0	1	1	1
1	0	1	1	0	0
1	1	0	1	1	0
1	1	1	1	0	1

Example 1. *For the Index 7, $a = 1$, $b = 1$, $c = 1$,*

$$\begin{aligned} X &= a + 0.1 + 1.0 = 1 \\ Y &= 0 \\ Z &= 1.1 + 1.0.0 = 1 \end{aligned}$$

So $XYZ = 101$, i.e., is equal to 5.

Same way will get the position numbers 2, 4, 6, 1, 7, 4, 6.

4 Experimental Results

Among the various image file format, the image which is smaller in size has been considered in this work. The JPG file is wonderfully small in size, often compressed to perhaps only 1/10 of the size of the original data. JPEG files achieve a smaller file size by compressing the image in a way that retains detail which matters most, while discarding details deemed to be less visually impactful. It supports 8-bit grayscale images and 24-bit color images (8 bits each for red, green, and blue). Here a 24-bit 64×64 color JPEG image has been hidden in the audio file. JPEG file have been read in Java as like below:

BufferedImage originalImage = ImageIO.read(new File(”rocket.jpg”));

Proposed LSB information hiding algorithm has been tested on 5 audio sequences from different music styles (classic, jazz, country, pop, rock). The audio experts were selected so that they can represent a broad range of music genres, i.e. audio clips with different dynamic and spectral characteristics. The image has been embedded in all music pieces using the proposed and standard LSB algorithm. Clips were 44.1 kHz sampled mono audio.wav files, represented by 16 bits per sample. Duration of the samples ranged from 10 to 15 seconds.

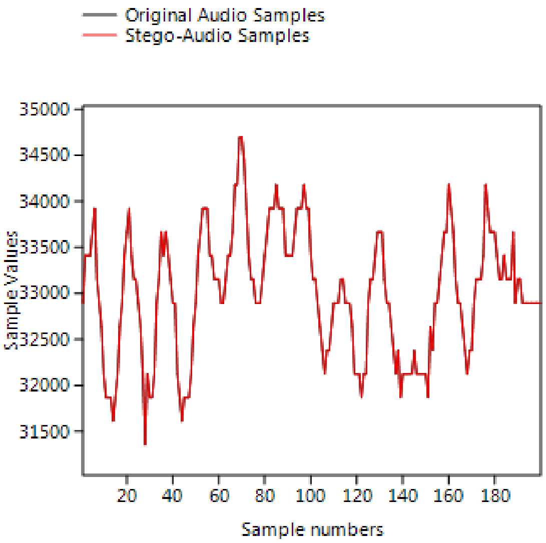

Figure 2: Negligible deviation between host audio samples & watermarked audio samples

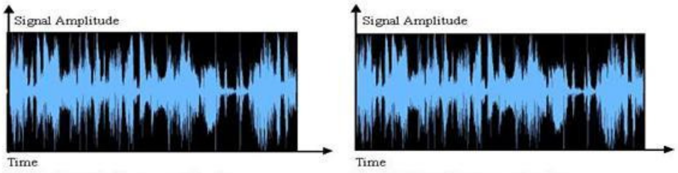

Figure 3: Negligible deviation between host audio wave & watermarked audio wave

4.1 Audio Quality Evaluation

Here 200 audio samples of both original audio and stego-audio has been considered to measure the sample level similarities between these two types of audio samples. From Figure 2, it is clear that statistical signal change (signal amplitude) due to bit embedding is very negligible compare to the original signal.

Figure 3.a shows the waveform of host audio and Figure 3.b shows the waveform of watermarked-audio. From these two Figures it is clear that after increasing the bit level of embedding, the audio signals are not differentiable by the general people.

4.2 Audio Quality Measurements

Here brief descriptions of the quality measures used have been introduced. The original signal (the cover audio) is denoted x(i), i = 1 to N while the distorted signal (the stego-audio) as y(i), i = 1 to N.

Signal-to-Noise Ratio (SNR): The SNR is very sensitive to the time alignment of the original and distorted audio signal [27]. The SNR is measured as equation no. (1), Table 4 and Table 5 showing the experimental result for 5 categories of audio file.

$$SNR = 10log_{10}\frac{\sum_{i=1}^{N} x^2(i)}{\sum_{i=1}^{N}(x(i)-y(i))^2} \quad (1)$$

Table 4: SNR values and capacities comparison with similar works (1 bit per sample)

	Embedding	1 bit per 16 bits sample			
	Music Genre	SNR (dB)		Capacity (%)	
		PM	SW	PM	SW
1	Classic	83.42	33 to 76	6.25	2 -12.5
2	Jazz	82.67	32 to 80	6.25	2 -12
3	Country	82.94	31 to 80	6.25	2 -12.5
4	Pop	83.15	38 to 82	6.25	2 -12.5
5	Rock	83.27	39 to 83	6.25	2 -12.5

Table 5: SNR values and capacities comparison with similar works (2 bits per sample)

	Embedding	2 bits per 16 bits sample			
	Music Genre	SNR (dB)		Capacity (%)	
		PM	SW	PM	SW
1	Classic	71.32	33 to 76	12.5	2 -31
2	Jazz	70.62	32 to 80	12.5	2 -32
3	Country	70.35	31 to 80	12.5	2 -33
4	Pop	71.04	38 to 82	12.5	2 -34
5	Rock	71.14	39 to 83	12.5	2 -34

PM means Proposed Method; SW means Similar Works [1, 4, 6, 7, 8, 12, 16, 21, 28, 31].

4.3 Correlation Based Measures

The similarity between two digital audio samples can also be quantified in terms of the correlation function [9, 27]. These ensure measurement of the similarity between two audios, hence in this sense they are complementary to the difference-based measures: Some correlation based measures are given in Equations (2), (3) and (4). Structural content:

$$C1 = \frac{1}{K}\sum_{k=1}^{K}\frac{\sum_{i=0}^{N-1} x(i)^2}{\sum_{i=0}^{N-1}(y(i))^2} \quad (2)$$

Normalized cross-correlation measure:

$$C1 = \frac{1}{K}\sum_{k=1}^{K}\frac{\sum_{i=0}^{N-1} x(i) * y(i)}{\sum_{i=0}^{N-1} x(i)^2} \quad (3)$$

Czenakowski distance (CZD): A metric that is useful for comparing vectors with strictly non-negative components, like in the case of audio samples, is given by the Czenakowski distance:

$$C = \frac{1}{N}\sum_{i=0}^{N-1}(1 - \frac{2 * min(x(i), y(i))}{x(i) + y(i)}) \quad (4)$$

The Czenakowski coefficient (also called the percentage of similarity) measures the similarity among different samples, communities, and quadrates.

Obviously as the difference between two audio samples tends towards zero $\epsilon = x(n) - y(n)$ tends to 0, all the correlation-based measures tend towards 1, while as ϵ^2 tends to G^2 they tend towards 0.

Recall also that distance measures and correlation measure are complementary, so that under certain conditions, minimizing distance measures is tantamount to maximizing the correlation measure. Table 5 is explaining the experimental result for CZD.

Table 6: Correlation based measure of the proposed algorithm

	Music Genre	Sample Size	CZD
1	Classic	16 bits	0.00001888249326
2	Jazz	16 bits	0.00002231742249
3	Country	16 bits	0.00002079240629
4	Pop	16 bits	0.00001666038440
5	Rock	16 bits	0.00001739731132

Experimental results show that the two audio clips (original audio sequence and embedded-audio signal) cannot be discriminated by people. Results of subjective tests showed that perceptual quality of watermarked-audio, if embedding is done using the proposed algorithm, is higher in comparison to standard LSB embedding method. This confirms that described algorithm succeeds in increasing the depth of the embedding layer and also randomizing the bit layer without affecting the perceptual transparency of the watermarked-audio signal.

Therefore, significant improvement in robustness against signal processing manipulation can be obtained, as the hidden bits can be embedded higher LSB layers deeper than in the standard LSB method. The proposed algorithm flips bits in more than one bit layers of the watermarked-audio during the embedding procedure. This property may increase the resistance against Steganalysis that identifies the used LSB layer by analyzing the noise properties of each bit layer.

4.4 Capacity and Detection Probability

The capacity depends on the embedding function, and may also depend on properties of the cover. For example, least-significant-bit (LSB) replacement with one bit per sample in an eight-bit audio achieves a net capacity of 12.5%, or slightly less if one takes into account that each audio is stored with header information which is not available for embedding. If the sample size is 16-bit then net capacity will be 6.25% or slightly less. It is intuitively clear, often demonstrated and theoretically studied that longer secret images require more embedding changes and thus are statistically better detectable than smaller ones. Hence, capacity and embedding rate are related to security.

The purpose of information hiding is to hide the existence of a secret image and also increasing robustness. Therefore, the security of a data hiding technique is judged by the impossibility of detecting the image content and extracting the hidden image after detection. However, sometimes, Cryptography also is used to increase the level of security. In this paper one image bit and two image bits have been embedded in a 16-bit sample separately and have been compared the result.

4.4.1 Detection Probability (Embedding Location Number-wise)

Here eight (8) 16-bit samples has been used to embed 8 image-bits. The opponent has to detect 8 bits to get 1 byte of information.

Probability to detect an embedded bit position = $\frac{1}{16}$;

Probability to detect 8 embedded bit positions = $\frac{1}{16} \times \frac{1}{16}$ upto 8 terms = $\frac{1}{16^8}$;

If the length of a image is N bytes, then the probability to extract whole image = ($\frac{1}{2} \times \frac{1}{2}$ upto 8 terms)* N terms = $\frac{1}{16^{N*8}}$.

4.4.2 Decoding Probability (bit (0/1)-wise)

Probability to decode an embedded bit = $\frac{1}{2}$;

Probability to decode 8 embedded bits = $\frac{1}{2} \times \frac{1}{2}$ upto 8 terms = $\frac{1}{2^8}$.

If the length of an image is N bytes, then the probability to extract whole image = ($\frac{1}{2} \times \frac{1}{2}$ upto 8 terms) * N terms = $\frac{1}{2^{N*8}}$.

Figure 4 show histogram of the number of modified bit layers in a 10 sec audio sample (116892x16 bits in total) for the proposed LSB algorithm. It is clear that number of flipped bits per bit layers is distributed over all bit layers in the proposed algorithm. In the case of standard LSB algorithm, LSB data hiding techniques can easily detect the bit layer where the data hiding was performed. It is a much more challenging task in the case of the proposed algorithm, because there are a significant number of bits flipped in 16 bit layers and the adversary cannot identify exactly which bit layer is used for the data embedding.

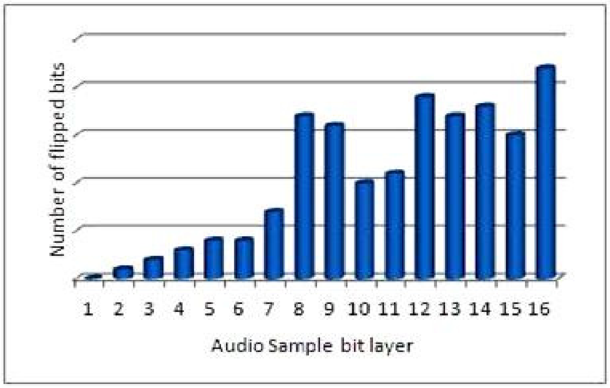

Figure 4: Number of flipped bits per bit layer for the proposed algorithm

4.5 Security Analysis

Detection, extraction, destruction and manipulation of the hidden data in a watermarked or stego object are the common attacks in Steganography and Watermarking techniques. While there has been quite some effort in the steganalysis of digital images, steganalysis of digital audio is relatively unexplored. Many attacks that are malicious against image Steganography algorithms cannot be implemented against audio Steganography schemes. Consequently, embedding information into audio seems more secure due to the nature of audio signals to be high-capacity data streams necessitates the scientifically challenging statistical analysis.

The attacks to a data hiding technique mainly include passive attack, active attack, and extracting attack. A passive attacker only wants to detect the existence of the embedded image, while an active attacker wants to destroy the embedded image. The purpose of an extracting attacker is to obtain the image hidden in the stego-object. So there are three kinds of security measures for different attackers respectively, i.e., detectability, robustness and difficulty of extraction. Usually the problem of steganography only concerns the detectability so in many literatures delectability is referred to the security of a stegosystem [10]. The problem of Watermarking concerns the detectability and robustness both. In this section the security of our data hiding process is discussed.

Westfeld [35] addressed the steganalysis of the MP3Stego algorithm. Ozer et al. [24] proposed a universal audio steganalysis technique that is effective on both watermarking and steganographic data-embedding methods. The basic idea in [24] rests on the statistical evidence that the distortion measures computed between signals and their de-noised versions have statistically distinguishable distributions for cover-signals and stego-signals. These statistically distinguishable features are used in steganalyzer design to classify cover-signals from stego-signals.

The audio steganalysis algorithm proposed by Liu et. al. [20] uses the Hausdorff distance measure to measure the distortion between a cover audio signal and a stego audio signal. I.Avcibas [13] proposed the use of stream data mining for steganalysis of audio signals of high complexity. Their approach extracts the second order derivative based Markov transition probabilities and high frequency spectrum statistics as the features of the audio streams. The variations in the second order derivative based features are explored to distinguish between the cover and stego audio signals. This approach also uses the Mel-frequency cepstral coefficients [19], widely used in speech recognition, for audio steganalysis.

The above mentioned steganalysis schemes are designed mainly based on the Analysis of Variance (ANOVA), Sequential Floating Search Method (SFS), Regression Analysis Classifier, Support Vector Machine Classifier etc.

By randomizing the embedding approach and choosing the higher LSB layer, the algorithm to estimate the cover statistics can be effectively disabled. The steganalyst cannot make any consistent assumptions about the hiding process even if the embedding algorithm is known to everyone as per the Kerckhoffs principle. Hiding in a randomized manner is quite attractive, and we explore its simplest realization in this paper.

So, most of the technique will not work in our proposed hiding scheme.

According to the experimental result of Signal-to-noise ratio (SNR) and Czenakowski distance (CZD), it is clear that human auditory system will not able to distinguish between original audio and stego-audio.

Here statistical analysis has been performed using [23] on original audio samples and stego audio samples and data have been introduced below and in Tables 7 and 8.

Table 7: Test for equal means (ANOVA)

	Sum of sqrs	df	Mean square
Within groups	1.86199E08	398	467836
Total	1.86199E08	399	
$omega^2$:	-0.002506		
HOV(Levene):	0.9937		
Medians p:	0.9939		

Between groups:

$$\begin{aligned} Sum\_of\_sqrs &= 58.5225; \\ df &= 1; \\ Meansquare &= 58.5225; \\ F &= 0.0001251; \\ p(same) &= 0.9911. \end{aligned}$$

Welch F test in the case of unequal variances: $F = 0.0001251$, $df = 398$, $p = 0.9911$.

Intra-class Correlation statistics: ANOVA

Between raters: Sum of sqrs = 58.5225; *df* = 1; Mean square=58.5225; F=23.96; p(same)=0.9911.

Between cases: Sum of sqrs = 1.86198E08; *df* = 199; Mean square=935670; F=3.831E05.

Within cases: Sum of sqrs = 544.5; *df* = 200; Mean square=2.7225;

Residual: Sum of sqrs = 485.977199; *df* = 2.4421;

Total: Sum of sqrs = 1.86199E08; *df* = 399; 95% confidence has been explain in Table 8.

Table 8: Intra-class correlation statistics(95% confidence)

Model 1	Individual	ICC(1,1)1	[1, 1]
	Mean	ICC(1,k)1	[1, 1]
Model 2	Individual	ICC(2,1)1	[1, 1]
	Mean	ICC(2,k)1	[1, 1]
Model 3	Individual	ICC(3,1)1	[1, 1]
	Mean	ICC(3,k)1	[1, 1]

From the above statistical analysis it is very much clear that statistical attack most of the time will fail to detect the hidden image from the stego-audio.

In our scheme, it is very difficult to detect the embedding information from stego audio. Again, if opposition add any noise with the stego audio randomly, there is a possibility of destroying or modifying information embedded into the audio file. To avoid this type of situation, we may use parity bit error checking like familiar technique or we may use hamming code for error detection and correction in our future work.

5 Advantages of Our Approach

- Embedding position numbers of image bits into audio file are sent to the receiver by converting them to Boolean functions which is more secured.

- Boolean functions are transmitted to receiver using digital signature concept which is very secure and reliable.

- Described algorithm succeeds in not only increasing the depth of the embedding layer but also layers are chosen randomly without affecting the perceptual transparency of the audio signal.

- Two-way robustness (to know the actual position of the image bit) are there, First, insertion positions are randomly chosen, Second, LSB layers are most of the time are high LSB layers.

- Embedding image into audio file causes minimal embedding distortion to the host audio, since optimization is done using GA operators.

- The hidden information detection of the proposed algorithm is more challenging as well, because there is a significant number of bits flipped in random higher LSB bit layers.

- In addition, listening tests showed that perceptual quality of stego-audio is higher in the case of the proposed method than in the standard LSB method.

6 Conclusions

This paper presents a novel bit-modification algorithm for modified LSB data hiding technique where image bit positions are transmitted to the receiver using digital signature concept which is very secure and reliable. The key idea of the algorithm is to embed the image bit which will cause negligible embedding distortion of the host audio. Listening test shows that described algorithm succeeds in increasing the depth of the embedding layer from lower to higher random LSB layers without affecting the perceptual transparency of the audio signal. The detection and extraction of hidden information of the proposed algorithm is more challenging as well, because there is a significant number of bits flipped in random higher LSB bit layers. On the other hand, position numbers are converted to Boolean functions and functions are transmitted using AES and RSA algorithms to the receiver end to make it more secure.

References

[1] M. A. Ahmed, L. M. Kiah, B. B. Zaidan, and A. A. Zaidan, "A novel embedding method to increase capacity and robustness of low-bit encoding audio steganography technique using noise gate software logic algorithm," *Journal of Applied Sciences*, vol. 10, no. 4, pp. 59–64, 2010.

[2] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Systems Journal*, vol. 45, no. 3, pp. 313–336, 1996.

[3] H. Berghel and L. B. Gorman, "Protecting ownership rights through digital watermarking," *IEEE Computers*, vol. 29, no. 7, p. 101–103, 1996.

[4] K. Bhowal, D. Bhattacharyya, A. J. Pal, and T. H. Kim, "A GA-based audio steganography with enhanced security," *Telecommunication Systems Journal, Springer*, vol. 52, no. 4, pp. 2197–2204, 2013.

[5] L. Boney, A. Tewfik, and K. Hamdy, "Digital watermarks for audio signals," in *Proceedings of the Third IEEE International Conference on Multimedia Computing and Systems*, pp. 473–480, 1996.

[6] N. Cvejic and T. Seppnen, "Reduced distortion bit-modification for LSB audio steganography," *Journal of Universal Computer Science*, vol. 11, no. 1, pp. 56–65, 2005.

[7] D. Ahmad and P. Mohammad, "Adaptive and efficient audio data hiding method in temporal domain," in *7th International Conference on Information, Communications and Signal Processing (ICICS'09)*, pp. 1–4, 2009.

[8] A. Delforouzi and M. Pooyan, "Adaptive digital audio steganography based on integer wavelet transform," *Circuits Systems Signal Processing*, vol. 27, no. 2, pp. 247–259, 2008.

[9] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Transactions on Communications*, vol. 43, no. 12, pp. 2959–2965, 1995.

[10] J. Fridrich, M. Long, "Steganalysis of LSB encoding in color images," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME'00)*, pp. 1279–1282, 2000.

[11] K. Gopalan, "Audio steganography using bit modification," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 421–424, 2003.

[12] H. B. Kekre, A. Archana, R. Swarnalata, and A. Uttara, "Information hiding in audio signals," *International Journal of Computer Applications*, vol. 7, no. 9, pp. 14–19, 2010.

[13] I. Avcibas, "Audio steganalysis with content-independent distortion measures," *IEEE Signal Processing Letters*, vol. 13, no. 2, p. 92–95, 2006.

[14] N. F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," *IEEE Computers*, vol. 31, no. 2, pp. 26–34, 1998.

[15] H. Jouhari and E. M. Souidi, "A novel embedding scheme based on walsh hadamard transform," *Journal of Theoretical and Applied Information Technology*, vol. 32, no. 1, pp. 55–60, 2011.

[16] K. B. Samir, U. P. Tuhin, and Ra. Avishek, "A robust audio steganographic technique based on phase shifting and psycho-acoustic persistence of human hearing ability," *International Journal of Computing and Corporate Research*, vol. 1, no. 1, 2011.

[17] Md. S. Khan, V. V. Bhasker, and V. S. Nagaraju, "An optimized method for concealing data using audio steganography," *International Journal of Computer Applications*, vol. 33, no. 4, pp. 25–30, 2011.

[18] E. Koch and J. Zhao, "Toward robust and hidden image copyright labeling," in *IEEE Workshop Nonlinear Signal and Image Processing*, pp. 452–455, 1995.

[19] Q. Liu, A. H. Sung, and M. Qiao, "Novel stream mining for audio steganalysis," in *Proceedings of the 17th ACM International Conference on Multimedia*, pp. 95–104, 2009.

[20] Y. Liu, K. Chiang, C. Corbett, R. Archibald, B. Mukherjee, and D. Ghosal, "A novel audio steganalysis based on higher-order statistics of a distortion measure with hausdorff distance," *11th International Conference on ISC*, LNCS 5222, pp. 487–501, 2008.

[21] Z. Mazdak, A. M. Azizah, B. A. Rabiah, M. Z. Akram, and A. Shahidan, "A genetic-algorithm-based approach for audio steganography," *World*

Academy of Science, Engineering and Technology, vol. 54, pp. 360–363, 2009.

[22] N. Cvejic and T. Seppanen, "Increasing the capacity of LSB based audio steganography," in *Proceedings of 5th IEEE International Workshop on Multimedia Signal Processing*, pp. 336–338, 2002.

[23] O. Hammer, *Past 3.x - the Past of the Future*, 2013. (`http://folk.uio.no/ohammer/past/`)

[24] H. Ozer, I. Avcibas, B. Sankur, and N. Memon, "Steganalysis of audio based on audio quality metrics," in *Proceedings of SPIE Security Watermarking Multimedia Contents V*, vol. 5020, p. 55–66, 2003.

[25] I. Pitas and T. H. Kaskalis, "Applying signatures on digital images," in *IEEE Workshop Nonlinear Signal and Image Processing*, pp. 460–463, 1995.

[26] J. J. Quisquater, O. Bruyndonckx, and B. Macq, "Spatial method for copyright labeling of digital images," in *IEEE Workshop Nonlinear Signal and Image Processing*, pp. 456–459, 1995.

[27] S. R. Quackenbush, T. P. Barnwell, and M. A. Clements, *Objective Measures of Speech Quality*, Prentice Hall, 1988.

[28] S. Gurvinder, S. K. Dey, S. Dubey, and S. Katiyal, "Increasing the efficiency of echo hiding digital audio steganography," in *4th National Conference on Computing for National Development (INDIACom'10)*, 2010.

[29] S. K. Pal, P. K. Saxena, and S. K. Mutto, "The future of audio steganography," in *Pacific Rim Workshop on Digital Steganography*, 2002.

[30] Van R. G. Schyndel, A. Z. Tirkel, and C. F. Osborne, "A digital watermark," in *Proceedings of IEEE International Conference on Image Processing (ICIP'94)*, vol. 2, pp. 86–90, 1994.

[31] R. Sridevi, A. Damodaram, and S. V. L. Narasimham, "Efficient method of audio steganography by modified LSB algorithm and strong encryption key with enhanced security," *Journal of Theoretical and Applied Information Technology*, vol. 5, no. 6, pp. 768–771, 2009.

[32] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust audio watermarking using perceptual masking," *Signal Processing. Special Issue on Watermarking*, vol. 66, no. 3, pp. 337–355, 1997.

[33] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," in *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1064–1087, 1998.

[34] Y. C. Tseng, Y. Y. Chen, and H. K. Pan, "A secure data hiding scheme for binary images," *IEEE Transactions on Communications*, vol. 50, no. 8, pp. 1227–1231, 2002.

[35] A. Westfeld and A. Pfitzmann, "Attacks on steganographic systems," in *Information Hiding*, LNCS 1768, pp. 61–66, Springer-Verlag, 1999.

Krishna Bhowal obtained his M. Tech from West Bengal University of Technology in 2010. He is a research scholar at the Dept. of Engineering and Technological Studies, University of Kalyani, Kolkata. Currently Mr. Bhowal is working as a Assistant Professor at Academy of Technology, a Degree Engineering College, Kolkata, India. He has about 7 years of experience in teaching. His area of interest includes Audio Steganography, Watermarking, Cryptography; He has published 3 International Journal papers and 2 research papers in International IEEE Conference. He is a member of IEEE.

Debasree (Chanda) Sarkar obtained her Ph.D in Engineering from Jadavpur University in the year 2005. She has obtained her M.E. from Bengal Engineering College (Presently known as Bengal Engineering and Science University, Shibpur) in the year 1994. She earned her B.E. degree in Electronics and Telecommunication Engineering from Bengal Engineering College (Presently known as Bengal Engineering and Science University, Shibpur) in the year 1991. She is presently working as Scientific Officer at the Dept. of Engineering and Technological Studies, University of Kalyani. Her area of research includes Microstrip Antenna, microstrip Filter, Frequency Selective Surfaces. She has contributed to numerous research articles in various journals and conferences of repute.

S. Biswas obtained his Ph.D in engineering from Jadavpur University in the year 2004. He obtained his M.E from Jadavpur University and B.E from Bengal Engineering College (Presently known as Bengal Engineering and Science University, Shibpur) in the year 1994 and 1990 respectively. He is presently working as Scientific Officer at the Dept. of Engineering & Technological Studies, University of Kalyani. He has more than 14 years of teaching experience. His area of interest includes, Artificial Neural Network, Image Processing, Frequency Selective Surfaces, Microstrip Antennas.

Partha Pratim Sarkar obtained his Ph.D in Engineering from Jadavpur University in the year 2002. He has obtained his M.E. from Jadavpur University in the year 1994. He earned his B.E. degree in Electronics and Telecommunication Engineering from Bengal Engineering College (Presently known as Bengal Engineering and Science University, Shibpur) in the year 1991. He is presently working as Senior Scientific Officer at the Dept. of Engineering and Technological Studies, University of Kalyani. His area of research includes Microstrip Antenna, microstrip Filter, Frequency Selective Surfaces and Artificial Neural Network. He has contributed to numerous (more than 110 publications) research articles in various journals and conferences of repute. He is a life fellow of IETE, and fellow of IE (India).