An Efficient Steganographic Approach to HideInformation in Digital Audio using Modulus Operation

Krishna Bhowal, Debasree Chanda, Susanta Biswas, and Partha Sarkar
Department of Engineering and Technological Studies, University of Kalyani, India

Abstract: *This paper presents an efficient data hiding technique where the encrypted secret message has been hidden into digital audio based on modified Exploiting Modification Direction (mEMD) technique. We put an effort to minimize the bit alterations introduced in the host audio signal during data hiding process. The proposed scheme confirms that the maximum change is less than 6.25% of the related audio sample and the average sample level error is less than 3%. The experimental results ensure that the method has a higher embedding capacity (88.2 kbps), maintaining imperceptibility (Object Difference Grades are between -0.10 and -0.31) and offer robustness against detection of intentional or unintentional audio signal attack detection. Based on imperceptibility, security, robustness, and embedding capacity - performance has been evaluated.*

Keywords: *Information security, audio steganography, watermarking, secret communication.*

Received June 5, 2016; accepted August 21, 2017

1. Introduction

Demand for Information security for the prevention of unauthorized access, is increasing day by day as using digital media in internet communications is increasing in our daily life rapidly. Steganography and watermarking are main branches of information hiding. A basic requirement of steganography should be such that the presence of hidden information within the stage-cover media will be perceptually and statistically undetectable. There should be no perceptible difference between the embedded and original signals, and it is difficult to remove or alter the watermark without damaging the host signal. Multimediadata hidingtechniqueshave developeda strongbasis fora growing number of applications like covert communications, authentication, copyright protection, tamper detection, etc.

Basic required specifications in data hiding technique may vary from application to application. The first requirement is perceptual transparency. Data hiding algorithm should be designed such a way that the difference between host audio signal and embedded audio signal are indistinguishable. Imperceptibility is the most important basic requirement in any type of data hiding technique.

Most of the data embedding algorithms which are widely used in many applications do not have to access the original host media while extracting the hidden data. Keeping in view the robustness of the system, sometimes original host audio may be used to extract hidden data from modified embedded audio signal.

Some applications of data embedding require small amounts of information to be incorporated. On the other hand, many applications of data embedding, like covert communication, require a lot of data to be incorporated. The ability to embed large quantities of data in a host medium will depend on how the embedding algorithm has been designed and also types of cover media used.

Data hiding algorithm should be designed in such a way that modification made due to signal processing operations or any other reason, should be detectable or would not affect the extraction of hidden data from embedded audio signal. The deviation between original host audio signal and embedded audio signal should be so negligible that the human auditory system would not be able to differentiate between them. Again, if some modification happened in the embedded audio signal due to active or passive attack, it should be easily detectable using or without using original audio signal.

Security is the main challenge in designing data embedding algorithm. Security requirements normally vary with the application of data embedding technique.

1. The presence of hidden information must not be able to identify by an unauthorized user, only sanctioned user be able to detect the presence of hidden information.
2. A data hiding algorithm is truly protected if the presence of hidden data into host audio is not perceived by unauthorized person even after knowing the exact algorithm applied during data hiding process.

To fulfill this requirement, sometimes data may be encrypted before embedding in a host audio signal. Therefore, to improve the security of a data hiding technique, the main challenge is to reduce the amount

of alterations required to be introduced into the digital cover media during hiding data.

In this paper, encrypted secret digits are converted to a 5-ary notational system and a digit has been embedded in a pair of audio samples based on an algorithm designed to minimize the number of alterations made into the digital audio samples during the embedding process. So a good quality of audio signal has been generated after the embedding process. The basic data hiding process has been shown in Figure 1.

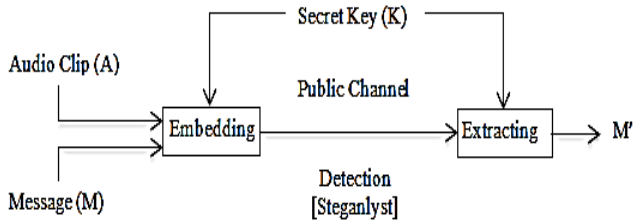


Figure 1. Basic data hiding process.

2. The Related Works

In this section some of the existing recent data hiding techniques in digital audio have been discussed. In order to hide secret information in digital audio effectively, a variety of embedding techniques have been introduced and discussed by different authors in [2, 3, 5, 7, 13]. Most of the schemes exploit sophisticated signal processing techniques for hiding secret data. To fulfill the basic requirements of data hiding technique mentioned in [25], Exploiting Modification Directions (EMD) is one of the very motivating tools. Data embedding by EMD requires that each secret digit in a (2n+1)-ary notational system is embedded on n cover media samples, where n is a system parameter [7]. To improve the embedding rate, the two-stage and the 8-ary EMD method were proposed later in [6, 15, 16]. Although, above mentioned schemes are applied when the data are embedded into the image, but a discussion has been made here as a simple reference. In this section, the EMD method, two-stage EMD and 8-ary EMD are described and later, different techniques used to hide information in audio are discussed.

2.1. The EMD Method

The gray pixel value of the cover images is grouped in to $i_1, i_2, \ldots, i_n$ and each group is segmented in to L bits,and the decimal value of eachsecretpieceisrepresentedby Kdigitsin a($2n$+1)-ary notationalsystem as defined in Equation (1)

$$L = \lfloor K \times log_2(2n+1) \rfloor \quad (1)$$

In this system, let a secret digit be d and assume an extraction function $f(i_1, i_2, \ldots, i_n)$ as a weighted sum modulo (2n+1) as Equation (2). The different values are calculated as $s=d-f$.

$$f = f(i_1, i_2, \ldots . i_n) = \left[\sum_{j=1}^{n} (i_j \times j)\right] \bmod(2n+1) \quad (2)$$

If $f = d$, no modification is required. In case of f ≠ d and $s \leq n$, increase the value of i_s by 1. If $f \neq d$ and $s > n$, decrease the value of i_{2n+1-s} by 1.

In the extraction process, let the stego data be $i'_1, i'_2, \ldots . i'_n$ for each sub-group. A secret digit d is calculated by Equation (3).

$$\boldsymbol{f = f(i'_1, i'_2, \ldots . i'_n) = \left[\sum_{j=1}^{n} i'_j \times j\right] mod(2n+1)} \quad (3)$$

2.2. Two-stage EMD Method

In the two-stage method, the gray values of pixels p_1, p_2, …, p_n are divided into buckets with an equal interval size, m like as $\lfloor p_1/m \rfloor, \lfloor p_2/m \rfloor, \ldots . \lfloor p_n/m \rfloor$.

f and d values are calculated by Equations (1) and (2), where the new gray value of the pixel is used instead of p_i. After first stage embedding, the original EMD embedding algorithm is applied to n pixels on the second stage. If the bucket number of the pixel is changed, the second stage cannot be allowed. This method provides capacity about twice larger than the EMD method and can hide the largest secret data when m=2.

2.3. The 8-Ary EMD Method

A method proposed in [15] for high quality and high capacity embedding. Here secret messages are converted into the 8-ary notational system, and the cover image is grouped into two sequence pixels, g_1 and g_2 for all pixels and Equation (4) is applied.

$$f_e = (g_1 \times 1 + g_1 \times 3) \, mod \, 8 \quad (4)$$

In this method, one 8-ary secret digit has been embedded into two cover pixels, in which only one pixel is decreased or increased by 1 (one) if the value of extraction algorithm is not equal. The embedding capacity is 1.5 times compared to the EMD method. To improve the Embedding rate and Embedding Efficiency several schemes have been proposed in [21, 30, 33].

Along with the EMD method, some related schemes also discussed here. In [8], the selected frequency band is separated into small frames and a single secret bit is embedded into each frame. First, consider the largest Fibonacci number that is lower than each single FFT magnitude in each frame has been computed and, based on the matching secret bit to be embedded; all samples in each frame are altered. All FFT samples in a frame have been changed to the closest Fibonacci number with even index, if the secret bit is "0". If the secret bit is "1", all FFT samples in a frame have been altered to closest Fibonacci number with odd index. A new adaptive audio watermarking algorithm based on EMD is introduced in [12]. The audio signal is divided into frames and by EMB, each one is decomposed

adaptively, into intrinsic oscillatory components called Intrinsic Mode Functions (IMFs). The conventional method maintains great sound quality and is highly robust to pirate attacks, including MP3 compression have been proposed in [31] with payload 2 bps and robustness to MP3 64 kbps. The intervals are quantized and the data is embedded in the quantization indices in [20]. Wavelet extrema of the signal envelope has been used as the salient points. The method is robust to common signal processing operations, e.g., mp3 lossy compression, low pass filtering, sampling rate conversion, and Time-Scale Modification (TSM).

In general, algorithm for audio data hiding technique must be more sensible compared to imagedata hiding technique. The EMD idea may be applied in designing data hiding algorithm in digital audio without compromising the quality of the digital audio. This work proposes modified EMD method in digital audio, which yields good quality audio and a high embedding rate compared to other existing recent data hiding methods.

3. The Proposed Method

In this work, an efficient Steganographic scheme is proposed to maximize the embedding capacity when preserving the statistical property of original audio signal and limiting the Steganographic distortion at a desired level. The cover audio signal are decomposed into a series of binary sequences, called audio samples, and an optimal one among a number of candidate audio sample for representing the secret data is chosen to control the distortion level while preserve the statistical property of each cover sample.

The most important challenge in data hiding technique is security. Here security is measured in terms of non-detectability and robustness incorporated into the data hiding technique. To make it non-detectable, it is required to ensure a minimum number of alterations made in digital media during data hiding process. A scheme has been designed here to achieve this requirement and applied in these subsequent sections.

The following subsections describe the details of the information embedding and extracting procedures of the proposed scheme.

5-ary notational system has 0, 1, 2, 3, and 4 digits and the conversion is reported in Table 1.

Table 1. 10-ary digit to 5-ary digit conversion table.

Secret digits	0	1	2	3	4	5	6	7	8	9
5-ary secret digits	00	01	02	03	04	10	11	12	13	20

3.1. Information Embedding Procedure

Followings are the initial steps of the embedding procedure:

a) Converting secret message to secret digits.
b) Converting secret digits to 5-ary secret digits.
c) Encrypting 5-ary secret digits using AES.

Embedding encrypted 5-ary secret digits are used in a subsequent pair of audio samples by applying proposed scheme.For each 5-ary secret digit d_i(i=0 to 4) and for a pair of audio samples S_j, S_k for (j=1, 3, 5…, k=2, 4, 6…) following steps are followed:

- *Step* 1: Calculate S_{jk}=(S_j+S_k).
- *Step* 2: Calculate f_i by Equation (5), where the value of x_i is selected to satisfy the condition fi=di, abs(x_i)<n=5 and– $\left\lfloor \frac{n}{2} \right\rfloor \leq x_i \leq \left\lfloor \frac{n}{2} \right\rfloor$, where abs($x_i$) means absolute value of x_i.

$$f_i = (S_{jk} + x_i) mod\ n \tag{5}$$

- *Step* 3: Using x_i selected in the previous step, calculate S_j' and S_k' from S_j and S_k by Equations (6) and (7) respectively as per the following rules:

As – $\left\lfloor \frac{n}{2} \right\rfloor \leq x_i \leq \left\lfloor \frac{n}{2} \right\rfloor$, values of x_i are in the range -2, -1, 0, +1, +2 for n = 5 and consider x_i' andx_i'' so that $x_i = x_i' + x_i''$

Based on the values of x_i following cases have been considered:

- *Case* 1: For x_i = 0, no changes will be happened in both of the samples, i.e., $S_j' = S_j$ and $S_k' = S_k$
- *Case* 2: For $x_i = +1$, $S_j' = S_j + x_i$ or $S_k' = S_k + x_i$, select sample based on logic defined in step iv and step v.
- *Case* 3: For $x_i = -1$, $S_j' = S_j + x_i$ or $S_k' = S_k + x_i$, select sample based on logic defined in step iv and step v.
- *Case* 4: For $x_i = +2$, $S_j' = S_j + x_i'$ and $S_k' = S_k + x_i''$, where $x_i' = x_i'' = 1$.
- *Case* 5: For $x_i = -2$, $S_j' = S_j + x_i'$ and $S_k' = S_k + x_i''$, where $x_i' = x_i'' = -1$.

$$S_j' = (S_j + x_i) \tag{6}$$

$$S_k' = (S_k + x_i) \tag{7}$$

- *Step* 4: In case of x_i=+1 and x_i=-1 in the previous steps, calculate number of bits altered or flipped (NBF) in S_j' and S_k' compare to S_j and S_k respectively and represents them by NBF(S_j') andNBF(S_k') respectively.
- *Step* 5: Now ifNBF(S_j') <=NBF(S_k') then choose the sample $S_j$for embedding secret digitd_i else choose the sample $S_k$for embedding secret digit d_i.

Here, for a particular secret digit d_i,value of x_i is fixed, that means deviation between S_j' andS_j or $S_k'$and S_k will remain same after embedding process. But to minimize the number of bits flipped during embedding process, above procedure has been adopted.

- *Step* 6: So, some of the samples having the lesser possibility to alter values or bits during embedding process and other samples some time having no possibility to alter bit at all i.e., it remains unaltered. Here, maximum alteration may happen with decimal value +1 or -1.

Here, we can consider $S'_{jk} = (S'_j + S_k)$ or $(S_j + S'_k)$or $S'_{jk} = (S'_j + S'_k)$, but for simplicity we consider $S'_{jk} = (S'_j + S'_k)$.

For validating the correctness of information embedding scheme, we give the Theorem 1:

- *Theorem* 1: For abs(x_i)<n,f_i=(S_{jk}+x_i) mod n=d_i for $-\left\lfloor \frac{n}{2} \right\rfloor \leq x_i \leq \left\lfloor \frac{n}{2} \right\rfloor$, where abs($x_i$) means absolute value of x_i
- Proof: Least Significant Digit (LSD) in S_j are in the range of 0 to 9 and 5-ary secret digits d_i are in the range of 0 to 4. Now if we calculate x_i according to the Equation (5), where 0≤ LSD(S_j)≤ 9 and 0≤ d_i≤4, five cases have to be considered for each of S_j to make f_i=d_i.

For each sample value ofS_j,f_i is calculated where $f_i = (S_j + x_i) \bmod n = d_i$and x_i is chosen based on min(abs(x_i)) but $S'_j = (S_j + x_i)$ is calculated based on the original value of x_i butabs(x_i), where abs(x_i) means absolute value of x_i .
For example, for LSD(S_j)= 0
For d_i= 0, f_i=(S_j+x_i)mod n=0 for x_i=min (abs(0), abs(5))=0
For d_i=1, f_i=(S_j+x_i)mod n=1 for x_i=min (abs(1), abs(-4))=1
For d_i=2, f_i=(S_j+x_i)mod n=2 for x_i=min (abs(2), abs(-3))=2
For d_i=3, f_i=(S_j+x_i)mod n=3 for x_i=min (abs(3), abs(-2))=-2
For d_i=4, f_i=(S_j+x_i)mod n=4 for x_i=min (abs(4), abs(-1))=-1

The probable cases have been reported in Table 2 to justify the above range of values for x_i.

Table 2. For different LSD (Sj) values, corresponding values of d_i and x_i.

LSD(S_j)	Secret Digits d_i					Values of x_i				
	d_0	d_1	d_2	d_3	d_4	x_0	x_1	x_2	x_3	x_4
0	0	1	2	3	4	0	1	2	-2	-1
1	0	1	2	3	4	-1	0	1	2	-2
2	0	1	2	3	4	-2	-1	0	1	2
3	0	1	2	3	4	2	-2	-1	0	1
4	0	1	2	3	4	1	2	-2	-1	0
5	0	1	2	3	4	0	1	2	-2	-1
6	0	1	2	3	4	-1	0	1	2	-2
7	0	1	2	3	4	-2	-1	0	1	2
8	0	1	2	3	4	2	-2	-1	0	1
9	0	1	2	3	4	1	2	-2	-1	0

- **An example of embedding process**:

let n= 5, -2≤x_i≤2, secret digit d_1=2, audio sample S_1=32690 and audio sample S_2 = 32671.
First, calculate S_{12} = 32690+32671=65361.
Then f_1 is calculated by f_1=(S_{12}+x_i)mod n= (65361-2) mod 5= 4, (65361-1) mod 5=0, (65361-0) mod 5=1, (65361+ 1) mod 5=2, (65361+2) mod 5=3. Here f_1 is equal to d_1 for x_i=1, so x_i=1 is considered in the next steps of embedding process.

Then, calculate $S'_1 = (S_1 + x_i)$=(32690+1)= 32691. Now, binary representation of 32690 is 01111111 10110010 and 32691 is 01111111 10110011. HenceNBF(S'_1)=1.
Here $|S_1 - S'_1| <= 1$ and deviation is 0.003%.

Again, calculate $S'_2 = (S_2 + x_i)$ = (32671 + 1) = 32672. A binary representation of 32671 is 0111111110011111 and 32672 is 0111111110100000. Hence NBF(S'_2) = 6.
Here also $|S_2 - S'_2| <= 1$ and deviation is 0.003%.

In this example,NBF(S'_1) <NBF(S'_2), so S_1 will be chosen as a candidate sample for embedding secret digit d_1=2. That means, value of $S'_{12} = (S'_1 + S'_2) = 32691 + 32671 = 65362$.

3.2. Information Extracting Procedure

Following steps are required to extract embedded information from embedded audio file.

a) Extracting encrypted 5-ary secret digits from the embedded audio samples by applying proposed scheme.
b) Decrypting encrypted 5-ary secret digits using AES.
c) Converting 5-ary secret digits to 10-ary secret digits.
d) Converting secret digits to the message.

For each pair of embedded audio samples S'_j and S'_k Calculate $S'_{jk} = (S'_j + S'_k)$.

To extract a secret digit $d_i$from audio samples S'_j and S'_k Equation (8) is used.

$$d_i = S'_{jk} \bmod n$$

For validating the correctness of information extracting scheme, we give the Theorem 2.

- *Theorem* 2: For i>0, j=1, j=j+2, k=2, k=k+2,

$$S'_{jk} \bmod n = d_i \quad (8)$$

- Proof: We calculate $S'_{jk} = (S'_j + S_k)\ or\ (S_j + S'_k)$ but for simplicity we considered $S'_{jk} = (S'_j + S'_k)$ in the embedding section.

For a particular value of S_{jk}, we have

$$S'_{jk} = (S'_j + S_k)\ or\ (S_j + S'_k)$$
$$= (S_j + x_i + S_k)\ or\ (S_j + S_k + x_i)$$

By Equations (6) and (7)

$$= S_j + S_k + x_i$$
$$= S_{jk} + x_i$$

Again, we have

$$S'_{jk} = (S'_j + S'_k)$$

$$= S_j + x_i' + S_k + x_i'' \text{ , where } x_i = x_i' + x_i''.$$
$$= S_j + S_k + x_i$$
$$= S_{jk} + x_i$$

From Equation (5), we have $fi=(S_{jk}+x_i) \bmod n$ for $f_i=d_i$
So, $d_i=(S_{jk}+x_i) \bmod n = S'_{jk} \bmod n$

Continuation of the previous example, for the stego-samples $S_1' = 32691$ and $S_2' = 32671$, the secret digit $d_1 = 2$ is extracted as like below:

$S'_{12} = (S_1' + S_2') = 32691 + 32671 = 65362.$
$d_1 = S'_{12} \bmod n = 65362 \bmod 5 = 2.$

4. Experimental Results and Discussion

The proposed data hiding scheme is tested on10 digital audiosequencesfromdifferentmusictypes likeclassic, jazz, country, pop, rock etc.Message stored in a file has been embedded in allmusicpieces usingtheproposed algorithms. Clipswere44.1kHz sampledaudiofilesare representedby16bitspersample.

Lengthoftheclipsrangedfrom10to20seconds.Inthisse ction,performance of our data hiding process has been measured and discussedbased on audio quality, embedding capacity, embedding complexity and security through someexperimentalresults.

4.1. Measurement of Similarity Between Original and Embedded Audio

The most familiar measure of similarity between two quantities is the linear correlation coefficient. If there is a series of n original audio samples X and a series of n embedded audio samples Y written as x_i and $y_i$where i = 1, 2, 3,… n, the sample correlation coefficient can be used in correlation r between X and Y. The audio sample correlation coefficient is written in Equation (9).

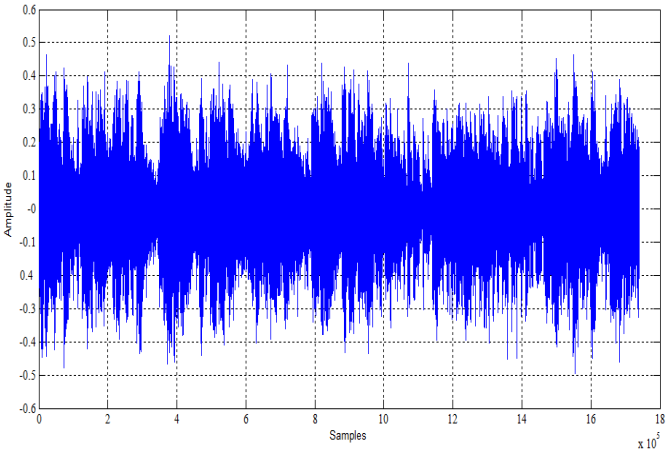
Where $\bar{x}$ is the mean of original audio samples X and $\bar{y}$ is the mean of embedded audio samples Y, S_x **and** S_y are the sample standard deviations of X and Y. Correlation coefficients are calculated for ten different categories of audio clips (denoted as A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9 and A_{10}) in MATLAB and value of r is almost 1 in all the cases. The experimental result is presented in Table 3.The data presented in the table determines their absolute similarity.

Table 3. Measurement of correlation coefficient.

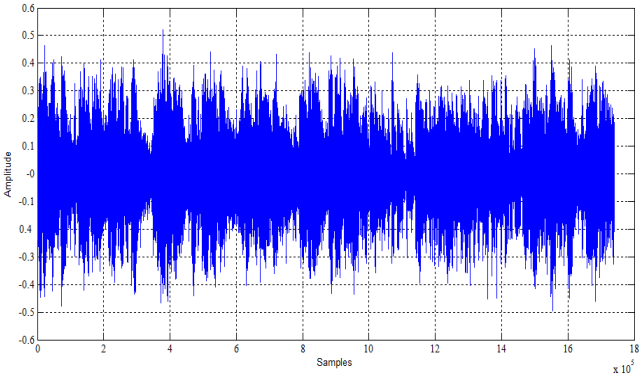
Audio types	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}
Value of	0.999	1.000	0.999	0.999	0.999	0.999	1.000	0.999	0.999	0.999

The similarity between wave form of original audio signal and embedded audio signal are reported in Figure 2-a and 2-b.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{(n-1)S_xS_y} \quad (9)$$



a) Original pop song wave form.



b) Embeddedpop song wave form.

Figure 2. Similarity of wave forms of original and embedded pop songs.

4.2. Objective Quality Measurements

Here imperceptibility quality assessment has been performed using both Signal-to-Noise Ratio (SNR) and Object Difference Grade (ODG) measurements. ODG is an appropriate measurement of audio distortions, since it is assumed to provide an precise model of the Subjective Difference Grade (SDG) results which may be obtained by a group of expert listeners. In this section, we performed ODG measurement, where ODG=0 means no degradation happened in digital host audio and ODG= -4 means a very annoying distortion happened in host digital audio.

The SNR values are calculated using the original digital audio and embedded digital audio files described in later section, whereas the ODG measurements are provided using the advanced ITU-R BS.1387 standard defined in [26] and is implemented by the software tool EAQUAL in [17]. ODG values of the ten embedded signal are reported in Table 4. All ODG values of the embedded audio signal are between -0.10 and -0.31 which determines their good qualities.

4.3. Subjective Quality Measurements

Subjective quality measurements have been performed to evaluate the inaudibility of our proposed data hiding scheme defined in [27, 28]. SDG listening tests are necessary to evaluate the perceptual quality of the

embedded digital audio, since the final decision is made by human acoustic perception. The ten participants were nominated for these subjective listening tests, five of them were experts in music and the rest of the five was general listeners. All of the participants are presented with the original and the embedded digital audio signal and were asked to report any difference between the two signals, using five-points SDG: (5: imperceptible, 4: perceptible but not annoying,3: slightly annoying, 2: annoying 1: very annoying). The output of the subjective tests is an average of the quality ratings called a Mean Opinion Score (MOS). The SDG experimental results reported in Table 4 and shows that the perceived quality of the embedded signal is imperceptible (about 5.0 in all cases). We can confirm convenient imperceptibility of the secret message in the digital audio signal.

4.4. Audio Quality Measurement by SNR

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

SNR: The SNR is very effective tool to make the difference between the original and embedded audio signal [23]. The SNR is used to judge the quality of the embedded audio. In general, if the SNR value is higher than the standard measurement of 50 dB, then the secret data which are embedded in the cover media are imperceptible to the human auditory system. The SNR is measured using Equation (10) and reported in Table 4 for 10 categories of audio clips.

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

In our work, during data hiding in digital audio, sometimes need to alter the 1st or 2nd LSB bits and hence distortion of the audio signal will occur. This distortion generates random errors in the original audio signal, which we referred to as a BER.

Table 4. ODG, SDG, BER AND SNR values comparison between different Audio types.

Audio types	ODG	SDG	BER	SNR(dB)
A_1	-0.31	4.9	0.01	92.95
A_2	-0.25	5.0	0.01	93.26
A_3	-0.30	5.0	0.01	92.48
A_4	-0.11	5.0	0.01	92.65
A_5	-0.12	5.0	0.01	92.31
A_6	-0.29	4.9	0.01	93.01
A_7	-0.23	5.0	0.01	92.97
A_8	-0.28	5.0	0.01	92.75
A_9	-0.09	5.0	0.01	92.91
A_{10}	-0.10	5.0	0.01	92.72

Minimization of bit alteration during embedding process is the main idea of the suggested algorithm and this preserve the quality of the host audio signals. To measure the quality of the embedded audio, the Bit Error Rate (BER) metric is one of the best tools. The ratio of altered bits and the total amount of embedded bits is defined as BER, as expressed by Equation (11).

$$BER = \frac{100}{l} \sum_{i=0}^{l-1} \begin{cases} 1, & S_i' = S_i \\ 0, & S_i' \neq S_i \end{cases} \quad (11)$$

where l is the bit length of audio file, S_i is the i^{th} bit of the originalaudio and S_i' is the i^{th} bit of the embedded audio.The Objective Difference Grade (ODG), Subjective Difference Grade (SDG), Bit Error Rate (BER) and SNR values for different audio clips are reported in Table 4. For simplicity, 10 audio clips are denoted by A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9 and A_{10}.

4.5. Embedding Time Complexity Measurement

We calculate x_i using $f_i=(S_{jk}+x_i)$ mod $n=d_i$ where $-2 \leq xi \leq 2$ to embed a secret digitd_i. That means, to select x_i a fixed number, i.e., 5 numbers of iterations are required for each secret digit. Now if the length of the audio sample is 'm' and for each secret bit we are considering a pair of audio sample, so m/2 number of iterations are required to consider whole audio clip. Total time complexity is 5*(m/2), i.e., (5/2)*m. So the data embedding complexity is O(m), i.e., linear.

4.6. Detection Analysis

Embedding information into digital audio seems to be more secure and safe due to the nature of audio signal which are high-capacity data streams. This high-capacity data stream necessitates the scientifically challenging statistical analysis. The Steganalysis schemes in [1, 10, 11, 18, 19, 22, 29] are designed mainly based on different statistical tools.In thiswork, detection probability is measured in terms of linearity regression, classification tree, vector support, and projection which are selected as the classification algorithms in the classifier.The cover audio signal and the stego audio signal are objectives to be classified with the condition that the original audio signal and the stego audio signal are both blind to the classifier. A system with a judgment result very close to 50% indicates absolutely secure system.

Normally, change of characteristics of embedded audio signal depends on the algorithm used for secret data embedding. For this reason, to assure the generality of the analysis, the system analyses as many characteristics as possible. Basic classifier structure is shown in Figure 3.

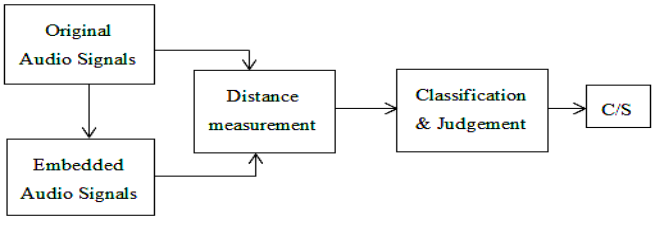


Figure 3. Basic classifier structure.

To test the detection probability, there are three kinds of characteristics which are measured in the classifier [4, 22].

1. The SNR, segment SNR (SNRseg), Minkowsky Measurement (MM) and Czenakowski distance (CAD) are considered in time domain measurements.
2. The logarithm likelihood ratio, COSH distance, Itakura-Saito distance, multitude spectrum distortion, inverse spectrum distance, and phase spectrum distortion are considered in frequency domain measurements.
3. The BARK distortion is considered in perceptual domain measurements.

Generally, the two classifier results are true detection and false detection. Here, the testing sets include 10 audio clips of 44.1 kHz, 16 bits, mono and stereo PCM bit streams. And the test results are reported in Table 5.

According to the data presented in Table 5, the measurement probability of our proposed data hiding algorithm is about 50% while the probability for LSB, ECHO, Hide4PGP, Steghide [24] are much higher. The test result determines that our data hiding algorithm is comparatively more secure.

Table 5. Detection performance comparison.

Probability of detection / probability of false detection				
Algorithm	**Linearity regression**	**Vector support**	**Classification tree**	**Projection**
LSB	0.72 / 0.38	0.79 / 0.42	0.70 / 0.60	0.82 / 0.44
ECHO	0.78 / 0.26	0.81 / 0.36	0.84 / 0.38	0.79 / 0.22
Hide4PGP	0.75 / 0.40	0.86 / 0.38	0.60 / 0.43	0.82 / 0.32
Steghide	0.67 / 0.44	0.71 / 0.37	0.69 / 0.35	0.70 / 0.39
[32]	0.52 / 0.41	0.66 / 0.42	0.56 / 0.42	0.48 / 0.44
Proposed	0.51 / 0.44	0.64 / 0.43	0.55 / 0.42	0.49 / 0.45

4.7. Robustness

Robustness of data hiding technique are defined as the modification made due to conventional digital signal processing operations or any other intentional attacks on embedded audio signal, should be detectable or would not affect the extraction of hidden data from embedded audio signal. The common attacks are AddNoise, BassBoost, Echo addition, LSB Zero etc. In our scheme, using original digital audio signal we can easily detect above mentioned attacks as like below. Let S_{jk} and AS'_{jk} are original and embedded digital audio samples. The difference between S_{jk} and S'_{jk} is limited and is maximum value is 2 as per the algorithm proposed here. We can identify whether any modification is happening between 3rd bit to 16th bit of embedded digital audio due to common attacks by $|S_{jk} - S'_{jk}|>2$. So the common attack detection probability is about 87.5%.

4.8. Performance Comparison

The Proposed scheme has been compared with some recent steganography and watermarking schemes in audio. Each data hiding scheme has different embedding algorithms and also properties. For this reason, it is difficult to establish an impartial comparison of the proposed scheme with some data hiding schemes in audio signal. In this section, few latest and relevant audio data hiding techniques have been chosen for comparison. The PSNR / SNR values of data hiding in image are compared with our proposed work is reported in Table 6 for simple reference. Table 7 provides a performance comparison between the proposed data hiding algorithm and several recent data hiding techniques in audio signal. This comparison determines that our proposed blind information hiding scheme managed higher embedding rate without compromising the statistical property of the host audio signal, i.e., generate a good quality audio.

Table 6. SNR / PSNR values comparison between data hiding in image and proposed work.

Scheme	**PSNR(dB) / SNR(dB)**
Chang *et al*. [6]	41.87
Khan *et al*. [14]	47.1
Mielikainen [21]	54.76
Westfeld [30]	56.44
Zhang and Wang [32]	57.80
Zhang and Wang [33]	51.80
Proposed Method	93.26

Table 7. Performance comparisons with recent works.

Scheme	**Capacity (bps)**	**SNR(dB)**	**ODG**
Fallahpour and Megías [8]	683 to 3 k	35 to 61	- 0.3 to -1.1
Garcia-HernandezaMichelb *et al*. [9]	5472-82518	93	0.0 to -0.97
Khaldi and Boudraa [12]	46.9 to 50.3	26.38	-0.4 to -0.6
Mansour and Tewfik [20]	4.3	29.3	Not reported
Xiang *et al*. [31]	2	42.8 to 44.4	-1.66 to -1.88
Proposed	88200	93.26	-0.10 to -0.31

5. Conclusions and Future Scope

Thispaperpresents an efficient blind data hiding scheme where encrypted secret digits are embedded into digital audio samples by minimizing bit alternations introduced during the embedding process. To achieve this, two-step technique is considered. First, secret digits are converted to 5-ary notational systems. Second, an effective scheme has been introduced where minimum number of bit alteration is considered as a criterion to select an audio sample for embedding secret digit. From the experimental result, it is clear that a better embedding function has been designed to improve the imperceptibility (ODG are-0.10 to -0.31, SNR is 93.26 dB) with higher embedding rate (88.2 kbps). The detection analysis ensures that the detection of secret message from the stego-audio is much more challenging. Thekeyideaof the proposed algorithmis that, with high embedding rate, a very good quality of embedded audio is generated which is almost indistinguishable compared to the original audio. There is a scope here to enhance

the proposed algorithm for increasing the robustness against different types of attacks and we are leaving it as a future scope.

References

[1] Avcibas I., "Audio Steganalysis with Content-independent Distortion Measures," *IEEE Signal Processing Letters*, vol. 13, no. 2, pp. 92-95, 2006.

[2] Bender W., Gruhl D., Morimoto N., and Lu A., "Techniques for Data Hiding," *IBM Systems Journal*, vol. 35, no. 3-4, pp. 313-336, 1996.

[3] Bhowal K., Bhattacharyya D., Jyoti Pal A., and Kim T., "A GA based Audio Steganography With Enhanced Security," *Telecommunication Systems*, vol. 52, no. 4, pp. 2197-2204, 2013.

[4] Bo Y., "Extensive Bark Spectral Distortion Measurement for Objective Speech Quality Assessment," *Journal of University of Electronic Science and Technology of China*, vol. 35, no. 3, pp. 343-345, 2006.

[5] Bohme R., *Advanced Statistical Steganalysis, Principles of Modern Steganography and Steganalysis*, Springer, 2010.

[6] Chang C., Tai W., and Chen K., "Improvements of EMD Embedding for Large Payloads," *in Proceedings of 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Kaohsiung, pp. 473-476, 2007.

[7] Cox I., Miller M., and Kalker T., *Digital Watermarking and Steganography*, Morgan Kaufmann Publishers, 2008.

[8] Fallahpour M. and Megías D., "Audio Watermarking Based on Fibonacci Numbers," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 8, pp. 1273-1282, 2015.

[9] Garcia-HernandezaMichelb J., RamonParra-M., Feregrino-Uribec R., and Cumplidoc R., "High Payload Data-Hiding in Audio Signals Based on A Modified OFDM Approach," *Expert Systems with Applications*, vol. 40, no. 8, pp. 3055-3064, 2013.

[10] Harmsen J. and Pearlman W., "Steganalysis of Additive-Noise Modelable Information Hiding," *in Proceedings of Security and Watermarking of Multimedia Contents*, Santa Clara, pp. 131-142, 2003.

[11] Ker A., "Steganalysis of LSB Matching in Greyscale Images," *IEEE Signal Processing Letters*, vol. 12, no. 6, pp. 441-444, 2005.

[12] Khaldi K. and Boudraa A., "Audio Watermarking Via EMD," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 3, pp. 675-680, 2013.

[13] Khan M., Bhasker V., and ShivaNagaraju V., "An Optimized Method for Concealing Data using Audio Steganography," *International Journal of Computer Applications*, vol. 33, no. 4, pp. 25-30, 2011.

[14] Khan Z., Shah M., Naeem M., Mahmood T., Amin N., and Shehzad D., "Threshold-based Steganography: A Novel Technique for Improved Payload and SNR," *The International Arab Journal of Information Technology*, vol. 13, no. 4, pp. 381-386, 2016.

[15] Lee C., Chang C., and Wang K., "An Improvement of EMD Embedding Method for Large Payloads By Pixel Segmentation Strategy," *Image and Vision Computing*, vol. 26, no. 12, pp. 1670-1676, 2008.

[16] Lee C., Wang Y., and Chang C., "A Steganographic Method With High Embedding Capacity By Improving Exploiting Modification Direction," *in Proceedings of 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Kaohsiung, pp. 497-500, 2007.

[17] Lerch A., *Zplane Development, EAQUAL-Evaluate Audio QUALity, version:0.1.3 alpha*, 2002.

[18] Liu Q., Sung A., and Qiao M., "Novel Stream Mining for Audio Steganalysis," *in Proceedings of 17th ACM International Conference on Multimedia*, Beijing, pp. 95-104, 2009.

[19] Liu Y., Chiang K., Corbett C., Archibald R., Mukherjee B., and Ghosal D., "A Novel Audio Steganalysis based on Higher-Order Statistics of a Distortion Measure with Hausdorff Distance," *in Proceedings of International Conference on Information Security*, Berlin, pp. 487-501, 2008.

[20] Mansour M. and Tewfik A., "Data Embedding in Audio Using Time-Scale Modification," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 3, pp. 432-440, 2005.

[21] Mielikainen J., "LSB Matching Revisited," *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 285-287, 2006.

[22] Ozer H., Avcibas I., Sankur B., and Memon N., "Steganalysis of Audio Based on Audio Quality Metrics," *in Proceedings of Security and Watermarking of Multimedia Contents V*, Santa Clara, pp. 55-66, 2003.

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

[24] Shuzheng X., Peng Z., Pengjun W., and Huazhong Y., "Performance Analysis of Data Hiding in MPEG-4 AAC Audio," *Tsinghua Science and Technology*, vol. 14, no. 1, pp. 55-61, 2009.

[25] Swanson M., Zhu B., and Tewfik A., "Current State of the Art, Challenges and Future

Directions for Audio Watermarking," *in Proceedings of IEEE International Conference on Multimedia Computing and Systems*, Florence, pp. 19-24, 1999.

[26] Thiede T., Treurniet W., Bitto R., Schmidmer C., Sporer T., Beerends J., Colomes C., Keyhl M., Stoll G., Brandenburg K., and Feiten B., "PEAQ-The ITU Standard for Objective Measurement of Perceived Audio Quality," *Journal of the Audio Engineering Society*, vol. 48, no. 1-2, pp. 3-29, 2000.

[27] Unoki M., Imabeppu K., Hamada D., Haniu A., and Miyauchi R., "Embedding Limitations With Digital-Audio Watermarking Method Based on Cochlear Delay Characteristics," *Journal of Information Hiding Multimedia Signal Proces*sing, vol. 2, no. 1, pp. 1-23, 2011.

[28] Wang S. and Unoki M., "Speech Watermarking Method Based on Formant Tuning," *IEICE Transactions on Information and Systems*, vol. E98-D, pp. 29-37, 2015.

[29] Westfeld A. and Pfitzmann A., "Attacks on Steganographic Systems, in Information Hiding," *in Proceedings of International Workshop on Information Hiding*, Berlin, pp. 61-66, 1999.

[30] Westfeld A., "A Steganographic Algorithm," *in Proceedings of 4th International Workshop Information Hiding, Lecture Notes in Computer Science*, Berlin, pp. 289-302, 2001.

[31] Xiang S., Kim H., and Huang J., "Audio Watermarking Robust Against Time-Scale Modification and MP3 Compression," *Signal Processing*, vol. 88, no. 10, pp. 2372-2387, 2008.

[32] Zhang X. and Wang S., "Dynamically Running Coding in Digital Steganography," *IEEE Signal Processing Letters*, vol. 13, no. 3, pp. 165-168, 2006.

[33] Zhang X. and Wang S., "Efficient Steganographic Embedding By Exploiting Modification Direction," *IEEE Communications Letters*, vol. 10, no. 10, pp. 781-783, 2006.

Krishna Bhowal is a research scholar in University of Kalyani, India. He is presently working as an Assistant Professor at Academy of Technology, Kolkata, India. His area of interest includes Audio Steganography, Watermarking, and Cryptography. He has published 9 research articles in various journals and conferences.

Debasree Chanda obtained her Ph.D in Engineering from Jadavpur University in the year 2005. She is presently working as Associate Professor Rank at the Dept. of Engineering and Technological Studies, University of Kalyani. Her area of research includes Microstrip Antenna, microstrip Filter, Frequency Selective Surfaces.

Susanta Biswas obtained his Ph.D in engineering from Jadavpur University in the year 2004. He is presently working as Associate Professor Rank at the Dept. of Engineering & Technological Studies, University of Kalyani. His area of interest includes, Artificial Neural Network, Image Processing, Frequency Selective Surfaces, Microstrip Antennas.

Partha Sarkar obtained his Ph.D in Engineering from Jadavpur University in the year 2002. He is presently working in the rank Professor 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 270 publications) research articles in various journals and conferences of repute.