

ENHANCING SECURITY OF QUALITY SONG WITH EMBEDDING ENCRYPTED HIDDEN CODE IN TOLERANCE LEVEL (SQHTL)

Uttam Kr. Mondal¹ and J.K.Mandal²

¹Dept. of CSE & IT, College of Engg. & Management, Kolaghat, W.B, India

uttam_ku_82@yahoo.co.in

²Dept. of CSE, University of Kalyani, Nadia (W.B.), India

jkm.cse@gmail.com

ABSTRACT

Embedding secret information with song signal may hamper its audible quality as well as originality also would be infringed. Therefore, passing secret code with song signal precisely needs to measure the embedded data volume with song signal within embedding tolerance level. In this paper, embedded secret data with linear coding principles has been applied, enhanced security criteria has been maintained with the help of elliptic curve cryptographic application. Hiding information into audio signals and audible quality is maintained only by calculating the acceptance ratio of embedding data using human perception, modulation of channel capacity in modified form is used to determine the embedded data and song signal trade-off ratio for getting future guideline embedding information in song signal with correlation among embedded data is fabricated. A comparative study has been made with similar existing techniques for performance analysis and experimental results are also supported with mathematical formula based on Microsoft WAVE (".wav") stereo sound file.

KEYWORDS

Elliptic Curve Cryptographic Technique, Embedding Secret Message in Quality Song, Linear Coding, Song Authentication, Tolerance Level of Embedding Message, Encoding Lower Frequencies.

1. INTRODUCTION

Applying techniques to protect quality songs is one of the primary necessities for creative industries. Increasing/updating the processing technology is enhancing the pirated versions of original songs. Some people use the original one and make their own versions just tactically changing some parts or modifying some portions to spread business in contemporary market [1, 5]. Therefore, original investors are losing their revenues, i.e., to protect the original version of audio songs – a set of measures need to be taken [i.e., restoring the copyright property of original song (IPR)], such techniques will not hamper the originality of the audio songs but carry an authenticated secret information for shielding its own properties. Even, the modified song can be easily detected by verifying the secret information. If any part of the song is being modified, would directly change the original secret code that hidden with specified position of the audio song signal.

In this paper, a systematic approach has been applied over song signal for embedding secret information, corresponding embedding impurities acceptance ratio has been measured with the help of linear coding principles. Two layers of security has been incorporated – one based upon

linear coding principles integrated with lower frequency region and another over some selected region of song signal with the help of elliptic curve cryptographic approach. Total embedded data volume also has been computed for getting balance of embedding data and audible quality of song signal. Embedded signals with secret code can easily distinguish the original from similar available songs. Finally, incorporating channel coding theorem in modified form for calculating acceptance ratio of hiding message into song signal is done without affecting its quality. It is experimentally observed that proposed technique will not affect the song quality but provide a level of security to protect the piracy of song signal without violating the acceptance level.

Organization of the paper is as follows. Encoding and embedding secret message are presented in section 2.1 to 2.3. The authentication procedure has been depicted in section 2.4, estimation of embedded message is given in section 2.5 that of extraction in section 2.6. Experimental results are given in section 3. Conclusions are drawn in section 4. References are given at end.

2. THE TECHNIQUE

The scheme fabricates the secret key with help of linear coding technique in lower frequency region (1-200 Hz) [which is not used by audio systems] as well as encoding the lower frequencies (1-200 Hz) followed by embedding secret key in the encoded frequency region. The operation is done with an approximate estimation of limit of impurities impacted with song signal to retain the audible quality of signal. Algorithms termed as SQHTL-ELF, SQHTL-MSE and SQHTL-EED are proposed as double security measure, the details of which are given in section 2.1 and section 2.2 respectively.

2.1. Encoding Lower Frequencies (SQHTL - ELF)

Encoding lower frequencies in the specific positions of signal (1-200 Hz) is performed with help of linear coding technique over GF (8) ["GF" stand for "Galois field"]. The procedure of Encoding lower frequencies is depicted in the following algorithm.

- Step 1: Take last 4 digits of 5 consecutive rows of magnitude values of frequency set, i.e., from 1st to 4th frequency positions and put into a two dimensional array of size 4 by 4. If any of these digits greater than 7, put 7 in this position and extra value (original value - 7) with all other extra value or 0 save in somewhere for generating secure code in high frequency region [above audible range]. if the value is .0910, then take 0710 for 1st row of the taken array and save 0.0200 with row number in some place for generating secure code [which will use in the process of section 2.3].
- Step 2: Take another array of similar size and populate values with values of 2nd to 5th rows of frequency set, i.e., one position ahead of previous window (of step 1) and convert the data set as step 1 as well as save extra values for similar purpose as described in step 1.
- Step 3: Add elements of above two arrays and put the result in another array [of GF(2³), Primitive polynomial = D³+D+1 (11 decimal)] say, C and replace the value of C in the place of B of the specified frequency positions of original song signal.
- Step 4: Continue step 1 to 3 until B window (of step 2) reaches to the 200 Hz position.

The value of 1st C matrix needs to put in specified positions of higher frequency region for getting original frequency value set in the time of decoding process [which is described in section 2.6]. If the song is stereo type, then the above method may be repeated for the second channel also. Therefore, if any value changes in processing, the above relationship in lower frequency region will break and can easily detect the error.

2.2. Moulding Secret encrypted Information (SQHTL - MSE)

As the song signal has been represented by linear combination of data over Galois Field $[GF(8)]$. Therefore, added information has to be converted into similar pattern as well as security and hidden criteria need to be associated with it. After accepting a secret code, elliptic curve cryptography technique over $GF(8)$ may be applied to represent the secret code into specified hidden format. Following method may be incorporated as described below.

Step 1: Select two values of 'a' and 'b' for the equation of the elliptic curve on a binary field $F2^m$ is $y^2 + xy = x^3 + ax^2 + b$, where $b \neq 0$. Here the elements of the finite field are integers of length at most m ($=3$) bits.

Step 2: Select the domain parameters for elliptic curve over $F2^m$ are $m, f(x), a, b, G, n$ and h . G is the generator point (x_G, y_G) , a point on the elliptic curve chosen for cryptographic operations. Whereas, n is the order of the elliptic curve. The scalar for point multiplication is chosen as a number between 0 and $n - 1$, h is the cofactor where $h = \#E(F2^m)/n$. $\#E(F2^m)$ is the number of points on an elliptic curve. The public key is a point in the curve and the private key is a random number. The public key is obtained by multiplying the private key with the generator point G in the curve.

Step 3: Select some positions from song signal and make amplitude values of those positions equal for both channels [if monotype song, make two consecutive position's values equal], apply any standard public key cryptographic algorithm with generated public key through step 1 to 3 above over any of two channels amplitude value and put resultant value in the place of original value.

Step 4: As amplitude values of song are not as a binary polynomial of degree $m - 1$, therefore, step 1 of SQHTL - ELF may be used and put the extra values in similar pattern in higher frequency locations.

Therefore, another additional authenticated code will be added with song signal. But total number of amplitude values will be considered based upon the maximum tolerance level of embedding extra data over song signal without hampering its audible quality as described in section 2.5.

2.3. Embedding Extra data (SQHTL - EED)

Embedding the storage extra values [in step 1 to 2 of section 2.1 and step 4 of 2.2] and values of 1st C window in the higher frequency region will create another security criteria over original song signal without hampering its audible quality. The embedding process is as follows.

- i. Make equal the magnitude values of two channel of stereo type song as it will do not affect over audible quality of song signal [2, 3].
- ii. Separate each digit position of extra value and convert into equivalent lower magnitude value. Let, if the extra value is .0200, then, separated lower magnitude values are 0, 0.0002, 0 and 0 respectively.
- iii. Add the magnitude values in the higher frequency region of song signal [above 20,000 Hz] as follows.
Let, the magnitude value is V , V value will add to i^{th} position, then the same will add with alternate channel of $(i+1)^{\text{th}}$ position, i.e.,

$$\begin{aligned}x(20000+i, s) &= x(20000+i, s) + V \\x(20001+i, s1) &= x(20001+i, s1) + V, \\ \text{where } s &= 1 \text{ or } 2. \text{ and } s1 = [(s+1) \bmod 2+1].\end{aligned}$$

- iii. Continue the step i to iii until all extra values are added in the higher frequency region of consecutive locations. The storage 1st C matrix value also can be add in same region by similar way where each matrix element should be converted as above step ii.

In case of mono type song, the lower magnitude values can be add by separating specified positions with same channel.

2.4. Authentication

Embedding extra values in higher frequency region in specified manner creates a secure code that will use to identify the original song. The encoding data set in lower frequency set creates a unique relationship among magnitude values of song signal. The addition operation between window k^{th} and $(k+1)^{\text{th}}$ will create result window of same size which will equivalent of the original values of window that originally constitutes song signal. On the other side, the secure embedded code of specified positions of song signal as described in the section 2.2 will provide another level of authentication of original song signal.

Therefore, if any changes during processing, it will create a difference with the authenticating codes that present in the higher frequencies region of the song signal and changing a position will create difference with the hidden code in that region as well as linear coding relationship will break in lower frequency region.

2.5. Estimating Limit

Estimating limit of embedding data over song signal is one of the major issues when quality is a factor. For this purpose, an approach has been made for estimating the boundary of adding impurities without compromising its audible quality is done with the help of channel coding theorem in modified form as follows.

Find the channel capacity with N_0W , where N_0 message embedding rate (here extra data) and W is the highest magnitude value of the song signal, here highest frequency value [i.e. 20,000 Hz]. Shannon limit may be considered for generating limit with help of associated formula.
 $r = k/N$, where r is the channel transmission rate, k is number of added values (magnitude values), may also use for calculating embedding data limit with song signal.

Channel capacity can be expressed by following formula

$$C = W \log_2 \left(1 + \frac{P}{N_0 W} \right) \quad (1)$$

bits per second (here, magnitude values per sampled set of song signal)

Assigning the values of above variables with considering Shannon's limit are given as follows

$W = 20,000$ (approx) [considering audible range 20-20,000 Hz]

$C = 44,100$ [16-bit stereo audio signals sampled at 44.1 kHz]

$$P = E[X_k^2]$$

Where, X_k = frequency values of original song, $K = 1, 2, 3, \dots, L$ [L = length of song signal]

Putting above values in the equation (1), we can easily find the noise value N_0 .

Spectrum density = $N_0/2$ [maximum noise] and $(N-k)$ number of added impurities (noise) in sampled values of song signal.

Where $r \leq C$, according to channel coding theorem.

Therefore, we can conclude, if the above limit exceeds then song will lost its audible quality.

2.6. Extraction

The decoding is performed using similar calculations as encoding technique. The algorithm of the same is given below.

Algorithm:

Input: Modified song signal with embedded authenticating code in higher frequency range.

Output: Original song signal.

Method: The details of extraction of original song signal are given below.

Step 1: Apply FFT over x to get magnitude values in frequency domain of song signal, says $Y(n)$, n is the total range of frequencies of song signal.

Step 2: Find the 1st C matrix [as described in SQHTL – ELF section] from higher frequencies region and using this matrix element find original sequence of magnitude values in lower frequency region. Then remove all the secret codes from higher frequencies region (above 20,000 Hz).

Step 3: Find decrypted channel values of selected region by applying decryption process of same standard encryption technique [step 3 of 2.2] with private key [generated by step 2 of 2.2] and add the deducted value [values from above step 2] to generate the original values of that region of the selected channels.

Step 4: Apply inverse FFT to get back the sampled values of original song signal.

3. EXPERIMENTAL RESULTS

Encoding and decoding technique have been applied over 10 seconds recorded song, the song is represented by complete procedure along with results in each intermediate step has been outlined in subsections 3.1.1 to 3.1.4. The results are discussed in two sections out of which 3.1 deals with result associated with SQHTL and that of 3.2 gives a comparison with existing techniques.

3.1. Results

For experimental observation, strip of 10 seconds classical song ('100 Miles From Memphis', sang by Sheryl Crow) has been taken. The sampled value of the song is given in table 1 as a two dimensional matrix. Figure 1 shows amplitude-time graph of the original signal. SQHTL is applied on this signal and as a first step of the procedure which is performed SQHTL over input song signal. The output generated in the process is shown in figure 2. Figure 3 shows the difference of frequency ratio of original and modified song after embedding secret code. From figure 3, it is seen that the deviation is very less and there will not affect the quality of the song at all.

3.1.1. Original recorded song signal (10 seconds)

The values for sampled data array $x(n,2)$ from the original song is given in table 1. Whereas the graphical representation of the original song, considering all rows (441000) of $x(n,2)$ is given in the figure 1, table 2 is showing properties of song signals.

Sl no	$x(k,1)$	$x(k,2)$
...
	0	0.0001
	0.0000	0.0000
	-0.0009	-0.0009
	-0.0006	-0.0007
	-0.0012	-0.0012
	-0.0014	-0.0014
	-0.0016	-0.0017
	-0.0023	-0.0022
	-0.0027	-0.0027
	-0.0022	-0.0021
...

Table 1. Sampled data array $x(n,2)$.

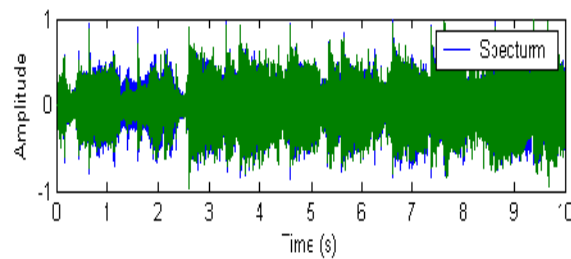


Table 2. Properties of input song signals

Figure 1. Original song ('100 Miles From Memphis', sang by Sheryl Crow)

Properties/ [attack type/bit rate (bps)]	Value
1. Bit Rate (kbps)	1411
2. Audio sample size(bit)	16
3. Channels (2- stereo/ 1-mono)	2
4.Audio sample rate(kHz)	44.1
4. Amplitude compression	0
5. Echo addition	0
6. All-pass filtering	0
7. Equalization	0
8. Noise addition	0
9. Audio format	PCM

3.1.2. Modified song after encoding lower frequencies and adding secure code (10 seconds)

The graphical representation of the modified song signal is shown in the figure 2.

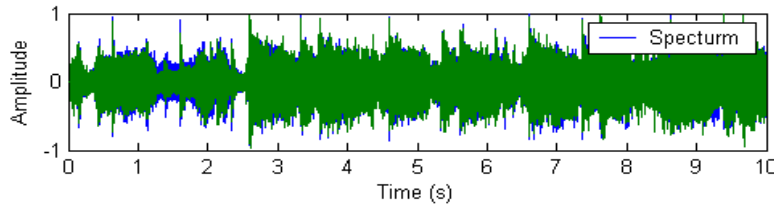


Figure 2. Modified song with secure code

3.1.3 The difference of magnitude values between original and modified signals

The graphical representation of the magnitude differences of original and modified songs is shown in the figure 3.

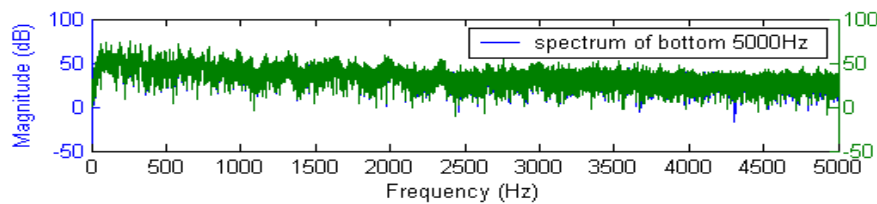


Figure 3. The difference of magnitude values between signals shown in figure 1 and 2.

3.1.4 Estimating limit of embedded code

Applying equation (1) to above song [16-bit stereo type sampled at 44.1 kHz] , hidden extra data (noise) that added for authenticating original song is very less than the maximum noise (0.0128), will not affect over all song audible quality. Because, only about 900 positions have been altered from 441000 sampled values of taken song signal.

3.2 Comparison with existing systems

Various algorithms [6] are available for embedding information with audio signals. They usually do not care about the quality of audio but we are enforcing our authentication technique without changing the quality of song. A comparison study of properties of our proposed method with Data hiding via phase manipulation of audio signals(DHPMA)[4] and Secret Data Hiding within Tolerance Level of Embedding in Quality Songs (DHTL)[8] before and after embedding secret message/modifying parts of signal (16-bit stereo audio signals sampled at 44.1 kHz.) is given in table 2, table4 and table5. Average absolute difference (AD) is used as the dissimilarity measurement between original song and modified song to justify the modified song. Whereas a lower value of AD signifies lesser error in the modified song. Normalized average absolute difference (NAD) is quantization error is to measure normalized distance to a range between 0 and 1. Mean square error (MSE) is the cumulative squared error between the embedded song and the original song. A lower value of MSE signifies lesser error in the embedded song. The SNR is used to measure how much a signal has been tainted by noise. It represents embedding errors between original song and modified song and calculated as the ratio of signal power (original

song) to the noise power corrupting the signal. A ratio higher than 1:1 indicates more signal than noise. The PSNR is often used to assess the quality measurement between the original and a modified song. The higher the PSNR represents the better the quality of the modified song. Thus from our experimental results of benchmarking parameters (NAD, MSE, NMSE, SNR and PSNR) in proposed method obtain better performances without affecting the audio quality of song.

Table 4 gives the experimental results in terms of SNR (Signal to Noise Ratio) and PSNR (Peak signal to Noise Ratio). Table 5 represents comparative values of Normalized Cross-Correlation (NC) and Correlation Quality (QC) of proposed algorithm with DHPMA and DHTL. The Table 6 shows PSNR, SNR, BER (Bit Error Rate) and MOS (Mean opinion score) values for the proposed algorithm. Here all the BER values are 0. The figure 4 summarizes the results of this experimental test. It shows this algorithm's performance is stable for different types of audio signals.

Table 3. Metric for different distortions

SI No	Statistical parameters for differential distortion	Value using SQHTL	Value using DHTL	Value using DHPMA
1	MD	0.0126	0.0132	3.6621e-004
2	AD	0.0016	0.0017	2.0886e-005
3	NAD	0.0065	0.0063	0.0063
4	MSE	5.7526e-006	5.7434e-006	1.4671e-009
5	NMSE	2.2053e+004	2.2066e+004	8.4137e-005

Table 4. SNR and PSNR

SI No	Statistical parameters for Differential distortion	Value using SQHTL	Value using DHTL	Value using DHPMA
1	Signal to Noise Ratio (SNR)	40.0534	39.0012	40.7501
2	Peak Signal to Noise Ratio (PSNR)	53.2064	52.408	45.4226

Table 5. Representing NC and QC

Audio (10s)	SNR	PSNR	BER	MOS
Song1	41.5560	63.1024	0	5
Song2	38.1030	51.442	0	5
Song3	36.0123	53.83	0	5
Song4	40.8701	63.5501	0	5

Table 6. Showing SNR, PSNR BER, MOS

SI No	Statistical parameters for Corelation distorstion	Value using SQHTL	Value using DHTL	Value using DHPMA
1	Normalised Cross-Correlation(NC)	1	1	1
2	Correlation Quality (QC)	-0.0124	-0.0128	-0.5038

The quality rating (Mean opinion score) is computed by using equation (2).

$$Quality = \frac{5}{1 + N * SNR} \quad (2)$$

Where N is a normalization constant and SNR is the measured signal to noise ratio. The ITU-R Rec. 500 quality rating is perfectly suited for this task, as it gives a quality rating on a scale of 1 to 5 [7]. Table7 shows the rating scale, along with the quality level being represented.

Table 7. Quality rating scale

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

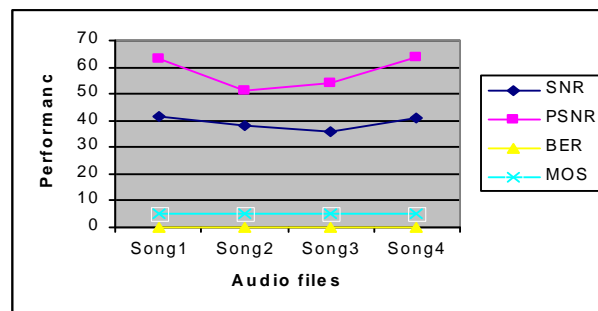


Figure 4. Performance for different audio signals

4. CONCLUSION AND FUTURE WORK

In this paper, an algorithm for encoding the lower frequency region with linear coding technique as well as embedding some secret code with the help of elliptic curve cryptographic technique in some specified region of song signal has been proposed. Another second layer of security also has been incorporated considering the extra deducted values of lower frequency region by sequentially adding in higher frequency region which will not affect the song quality but it will

ensure to detect the distortion of song signal characteristics. Additionally, the proposed algorithm is also very easy to implement.

This technique is developed based on the observation of characteristics of different songs with human audible characteristics and an approach is also made for estimating the embedded extra data limit with the help of Shannon's limit in the channel encoding scheme. It also can be extended to embed an image into an audio signal instead of text and audio. The perfect estimation of percentage of threshold numbers of sample data of song that can be allow to change for a normal conditions will be done in future with all possibilities of errors in song signal processing.

REFERENCES

- [1] Mondal, Uttam Kr., Mandal, J.K.: A Practical Approach of Embedding Secret Key to Authenticate Tagore Songs(ESKATS), Wireless Information Networks & Business Information System Proceedings (WINBIS'10), ISSN 2091-0266, organized by Rural Nepal Technical Academy (Pvt.) Ltd , Nepal, Volume 6, Number 1, pp 67-74(2010).
- [2] Mondal, Uttam Kr., Mandal, J.K. : A Novel Technique to Protect Piracy of Quality Songs through Amplitude Manipulation (PPAM), International Symposium on Electronic System Design (ISED 2010), ISBN 978-0-7695-4294-2 ,pp 246-250(2010).
- [3] Uttam Kr. Mondal and J.K. Mandal, "A Fourier Transform Based Authentication of Audio Signals through Alternation of Coefficients of Harmonics (FTAT)", First International Conference on Parallel, Distributed Computing Technology and Applications(PDCTA 2011), Lecture Notes in Computer Science Series (CCIS), Vol 203, ISSN 1865-0929, ISBN 978-3-642-24036-2, pp 76-85.
- [4] Xiaoxiao, Dong, Mark, F., Bocko, Zeljko Ignjatovic: Data Hiding Via Phase Manipulation of Audio Signals , IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004) , ISBN 0-7803-8484-9 ,Vol 5, pp 377-380(2004).
- [5] Erten, G., Salam, F.:Voice Output Extraction by Signal Separation ", ISCAS '98 ,ISBN 07803-4455-3,Vol 3, pp 5 - 8(1998).
- [6] Katzenbeisser, S., Petitcolas, F.A.P.: Information Hiding Techniques for Steganography and Digital Watermarking, Artech House, Norwood, MA, ISBN 978-1-58053-035-4(2000).
- [7] Arnold, M.: Audio watermarking: Features, applications and algorithms, IEEE International Conference on Multimedia and Expo ,New York, NY, Vol 2, pp. 1013-1016, (2000).
- [8] Mondal Uttam Kr., Mandal J K.: Secret Data Hiding within Tolerance Level of Embedding in Quality Songs (DHTL), , Second International Conference on Computer Science, Engineering and Application (ICCSEA 2012), Vol 2, ISSN 1867-5662, ISBN 978-3-642-30110-0, pp 753-761.

Authors

Uttam Kr. Mondal, has received his Bachelor of Engineering (B.E) degree in Information Technology in 2004 and Master of Technology (M.Tech) in Information Technology in 2006 from University of Calcutta, India. He has now working as an Asst. Professor in department of Computer Science & Engineering and Information Technology in College of Engg. & Management, Kalaghat, West Bengal, India. His research areas include cryptography & Network Security, Audio signal authentication. He has 21 publications in National and International conference proceedings and journal.



Jyotsna Kumar Mandal, M. Tech. (Computer Science, University of Calcutta), Ph.D. (Engg., Jadavpur University) in the field of Data Compression and Error Correction Techniques, Professor in Computer Science and Engineering, University of Kalyani, India. Life Member of Computer Society of India since 1992 and life member of cryptology Research Society of India. Dean Faculty of Engineering, Technology & Management, working in the field of Network Security, Steganography, Remote Sensing & GIS Application, Image Processing. 25 years of teaching and research experiences. Eight Scholars awarded Ph.D. and 8 are pursuing. Total number of publications 189.

