International Journal of Security and Its Applications
Vol.9, No.12 (2015), pp.317-328
http://dx.doi.org/10.14257/ijsia.2015.9.12.31

A Robust Data Embedding Method for MPEG Layer III Audio Steganography

Mohsen Bazyar and Rubita Sudirman*

Faculty of Electrical Engineering, University Teknologi Malaysia,
81310 UTM Johor Bahru, Malaysia
Mohsenbazyar114@yahoo.com, rubita@fke.utm.my

Abstract

A new method of MP3 steganography is proposed with emphasis on increasing the robustness of data embedding based on the MP3 statistical properties, which used Modified Discrete Cosine Transform (MDCT) to compress these properties. In order to achieve this purpose, modification of selected MDCT coefficients is done through MPEG audio compression procedure. In this algorithm, secret message is embedded directly after achievement of MDCT coefficient and before quantization step in process of MP3 encoder. It is feasible to embed some hidden data by substitution of the least significant bit of the MDCT coefficients during compression of MP3 files. The proposed algorithm performance results show considerable robustness against common audio processing attacks, especially MPEG compression with lower computational complexity compared to MP3Stego. The retrieved hidden message obtained from the algorithm has the highest quality against various steganographic attacks and different signal types which shows an excellent imperceptibility for the new algorithm.

Keywords: *MP3, Steganography, Modified discrete cosine transformation, Audio data hiding*

1. Introduction

Nowadays, significant discoveries in digital information revolution and internet caused considerable changes in the world, ranging from the effect on the world economy to the way people communicate. There has been considerable interest in data hiding in the last decade. With the advent of the digital age, the proliferation of computers and the widespread usage of the Internet have created digital multimedia as a host for reliable and secure data transfer. Data security is one of the most important issues for computer forensic experts. Steganography is an art of hiding confidential information in some cover medium file. Steganography has become relevant in various forms of digital media such as images and audio. By taking advantage of lossy compression algorithms we now see steganography existing in JPEG (Joint Photographic Experts Group) images as well as MP3 (MPEG Layer3) audio. The nature of these lossy compression techniques readily allows data to be embedded such that the medium is not altered in a perceptible way. It can be used to hide secret messages in digital multimedia files which can be retrieved later by intended users. This technique can be used for secret communication and confidential information sharing via internet medium. This method of secret message passing is seen to be misused in offices by leaking important company documents though some cover medium files. Encryption techniques along with steganography can provide a high level of data security [1].

ISSN: 1738-9976 IJSIA

As audio techniques have been developed for audio streaming on the internet for the radio station, but then incorporated into social networking and communication applications such as Skype. [1]

Online gaming is also a big user of audio channels on the internet. Using the TCP/IP protocol, audio files can be uploaded, downloaded, and transmitted through the internet. This benefit of transmission makes the interest in using audio as a cover object in steganography become much stronger. There are many types of audio formats differentiated by encoding algorithms, proprietary design, compression algorithms, and standardized formats. Audio file format is commonly classified into three parts: audio formats without any compression such as AIFF and WAV. Lossless compression audio formats such as TTA, ATRAC and finally, lossy compression audio formats, such as AAC and MP3 [2-3].

MP3 digital audio files are very popular and well-known audio compression format on the Internet which uses the destructive compression methods to attain to the high compression rate. The main audio file converts to a small size. MP3 music files format are very popular on the internet and most music Producers select internet as a space to publish their new works because of low cost of sales and popularity. But steganography software using MP3 audios as hosts is rare. Hiding capacity, perceptual transparency, and robustness are three important parameters which are known as 'the magic triangle' and use in designing steganography techniques [4]. Some methods attempt to be robust against various attacks like MPEG-1 layer III compression, whereas, achieving high hiding capacity is the goal of some other steganography techniques [5]. In recent years, most of watermarking techniques which have been proposed for MP3 audio formats can be improved for the steganography purposes without any problem [6]. The goal of this study is to develop an application which can hide data in MP3 files and detect the presence of hidden messages in MP3 files using an appropriate analysis method.

Moghadam *et al*. proposed a new watermarking algorithm to MPEG-1 Layer III format during the compression of digital audio. At the proposed method, the authors modified some of the MDCT coefficients by using genetic algorithm during MP3 audio compression. With considering transparency and robustness factors, there are several approaches for modifying some of MDCT coefficient to embed the hidden data into audio file. A genetic algorithm was used to choose the best MDCT coefficients to hide the secret message in this proposed algorithm [7]. A new technique was proposed which is able to embed some secret data by replacing the MDCT coefficients using the east significant bit method (LSB) after the conversion of MP3. This method works according to the MP3 statistical properties, which is used Modified Discrete Cosine Transform to compress these properties [8-9].

Wang *et al*. proposed a new watermarking method that embed hidden message directly in the MDCT coefficient block during performing the compression process. In order to increase robustness, they embed secret data into low frequencies. The beginning of the secret data is the position of the first nonzero coefficient. The MDCT coefficients were adjusted and calculated to carry the embedded data during the MDCT phase. Embedded data is then extracted during the IMDCT process and after dequantization. To acquire the watermark coefficients, Inverse of the MDCT (IMDCT) values are analyzed. At this method the watermark is embedded only at low frequency region and it is not a self-adaptive approach even though it works directly on MP3 process. Hiding watermark at this part of the signal has significant impact on watermarked audio quality [10].

An adaptive digital audio watermarking method was proposed by Chen *et al.* [11] for MP3 compression. In this method, secret data is hidden directly in MP3 domain before quantization and after MDCT in MP3 encoder procedure. This algorithm works based on the human auditory system. Energy of subbands and Gaussian distribution are analyzed to

* Corresponding author

control of adaptability. Although, they improved their security of the watermark algorithm by adding a security key for watermark embedding in compared with the one that they proposed in [12], the robustness of this new algorithm is much better than previous one.

Here we implement a technique for hiding data in MP3 files, known as MP3 file steganography. However, steganography alone is not enough to provide a high level of security. It is also combined with encryption of the data to be embedded. We proposed a robust data embedding algorithm for MPEG layer III audio steganography based on human auditory system. In this algorithm, secret data is embedded during the Mp3 compression before quantization and after MDCT. The experimental results indicate that the proposed steganographic algorithm is robust against most common attacks especially in MP3 compression. Remaining of this paper is organized as follows. In the next section we describe the structure of MP3 file format with encoder/decoder of MPEG audio layer III. Then, we explain our proposed method, followed by discussing the experimental results.

2. MP3 Encoder/Decoder

One of the most significant developments in the digital media processing area is the invention of the MPEG-1 Layer-3 (MP3) audio compression algorithm because of its high compression ratio and quality. MPEG files because of using very sophisticated compression techniques are much smaller for the same quality compared to other audio and video coding formats and this is the main advantage of MPEG. MPEG-1 audio consist of three layers. The most well-known is MPEG-1 audio layer 3 known as MP3. MP3 creates a highly compressed signal with high perceptual quality by eliminating or reducing unnecessary frequencies. A complete MP3 system has an MP3 encoder and an MP3 decoder. The operations of the encoder and decoder are governed by a set of undetermined control signals. These are combination of both external hardware controls and software functions.

2.1. MP3 Encoder

In an MP3 system, a regular PCM audio is input to the MP3 encoder. This analog signal will be converted to an MP3 digital bitstream according to the standard MPEG-1 layer 3 encoding algorithm. An MP3 file will be generated after MP3 encoder. The control signals mean external software functions and parameter controls during MP3 encoding and decoding.

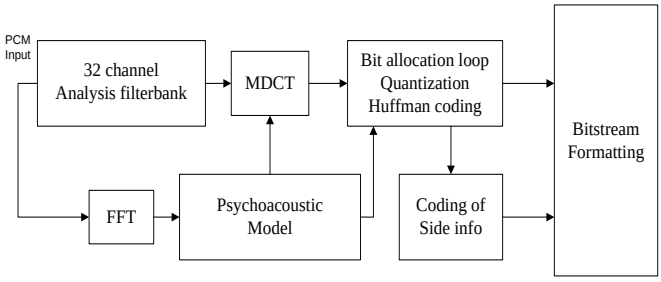


Figure 1. Block Diagram of an MP3 Encoder

Figure 1 shows the block diagram of a typical MP3 encoder [13]. The purpose of the filterbank is to transform audio signals to frequency domain signals in 32 subbands. Every subband is splited into 18 frequency lines. Therefore, the 32 subbands are then mapped into 576 (32×18) frequency lines after Modified Discrete Cosine Transform (MDCT) [14].

Empirical evidence has shown that the human ear has a limited resolution (that can be expressed in terms of critical band widths) less than 20 kHz and more than 20 Hz, which

is the range of Human Auditory System (HAS). Within a critical bandwidth the human ear blurs frequencies. Thus the filterbank creates equal-width frequency subbands that correlate to the critical bandwidths. The human ear can nominally hear sounds in the range of 20Hz to 20 kHz. Psychoacoustic model analysis is the technology on HAS to make critical banks, which is not the same as 32-equal subbands after filterbank. When in low frequency, a subband might include several critical bands, however, when in high frequencies, several subbands are included into one critical band. Unlike 32 subbands, critical bands are non-linear algorithms. Psychoacoustic model uses the weakness of the human ear to eliminate the inaudible parts of signals to make compression by using masking technology on critical bands [15-16].

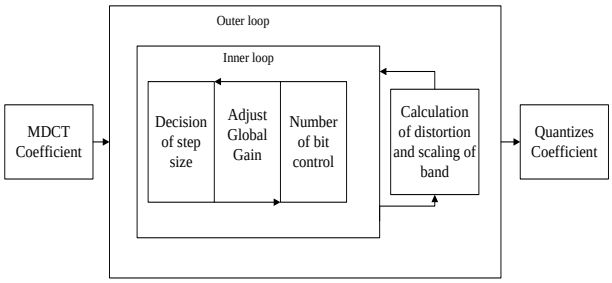


Figure 2. The Relationships between Inner Loop and Outer Loop

MDCT is lossless and is carried out on the sequential blocks of a larger dataset, where the last half of one block coincides with the first half of the next block that it means subsequent blocks are overlapped together. Since this overlapping helps to avoid artifacts stemming from the block boundaries, it makes the MDCT especially attractive for signal compression applications to the energy compaction qualities of the DCT [17].

The MDCT is performed on blocks that are windowed and overlapped 50 percentages off. The MDCT to achieve good frequency resolution is normally performed on 18 samples at a time (long blocks). It can also to achieve better time resolution and to minimize pre-echoes be performed on 6 samples at a time (short blocks). After MDCT, a non-uniform quantization maps amplitude values into finite number of bits. Non-uniform is to change sample size according to amplitude values. In this way, larger values are automatically coded with less accuracy and some noise shaping is already built into the quantization procedure. Quantization is the second time compression in MP3 file generation. Two nested loop is used during the quantization step. One of them is applied to control quantization step size, which is called inner iteration loop and the other one is used for each scale factor band to control the noise shaping factors (outer iteration loop) [18-19].

The quantized values are stored by Huffman coding. As a specific method for entropy coding, Huffman coding is lossless and keeps data in the bitstream along with the scale factors and side information. Quantized MDCT coefficients are arranged in the order from lowest to highest frequency. The whole range is divided into three sections, each coded with a different set of Huffman tables. Bit generation and CRC check make the signal frames finally to be MP3 bitstream [20-21].

2.2. MP3 Decoder

In MP3 decoder, most of the steps are the inverse of the encoder, but there is no psychoacoustic modeling in decoder. Figure 3, shows the block diagram of an MP3 decoder. The decoder reconstructs the analog audio signal by decoding the MP3 bitstream, and the analog audio signal is finally turned into audible sounds. The Huffman decoder translates the variable length codes in the MP3 bitstream to spectral lines. The sample re-quantization block uses the scale factors to convert the Huffman decoded

values back to their spectral values. The re-quantized samples must be reordered for the scale factor bands that use short windows. The IMDCT transforms the frequency lines to polyphase filter subband samples. The output from the IMDCT operation is 18 time-domain samples for each of the 32 subband blocks. The synthesis polyphase filterbank transforms the 32 subband blocks of 18 time domain samples in each granule to 18 blocks of 32 PCM samples.

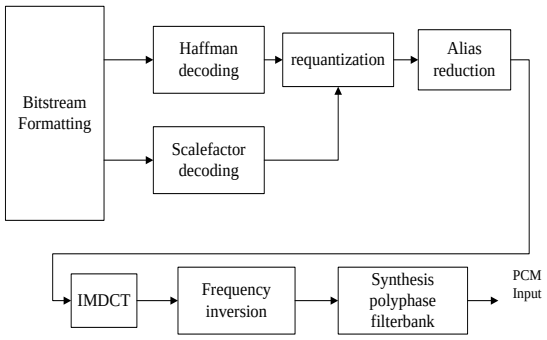


Figure 3. Block Diagram of an MP3 Decoder

3. The Proposed Audio Steganographic Algorithm

The aim of proposed algorithm is embedding secret data within MDCT coefficient during MP3 compression. We found that these coefficients don't have any significant changes after encoder. Therefore, we try to hide secret message before quantization and after MDCT.

3.1. Embedding Process

A new method to embedding secret data during the MP3 compression will be proposed at this section. Figure 4 indicates the block diagram of this algorithm.

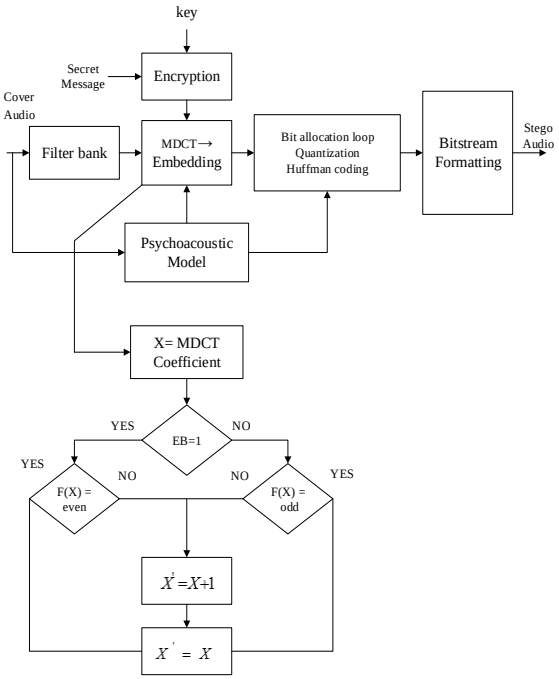


Figure 4. Block Diagram of the Embedding Process

Embedding procedure is done directly after achievement of MDCT coefficient and before quantization step in process of MP3. The keys, which are Adjustable by users, are used to improve the security of embedded algorithm. Each bit of the secret data is embedded in one bit of MDCT coefficient. Let F(X) is quotient remainder of the decimal MDCT coefficient X divided by 2. In summary, after hiding the embedded bit (EB), the new coefficients of X, (X ′) are computed as follows:

$$X' = \begin{cases} X & if\,(EB=1) \quad and \quad F(X)=even \\ X+1 & if\,(EB=1) \quad and \quad F(X)=odd \\ X & if\,(EB=0) \quad and \quad F(X)=odd \\ X+1 & if\,(EB=0) \quad and \quad F(X)=even \end{cases} \tag{1}$$

At this algorithm, the binary values of secret data bits (0 and 1) are embedded by calculating the quotient remainder of the decimal MDCT coefficient divided by 2. For example, suppose one of the MDCT coefficients is 51 and embedding bit is 1. Remainder of the 51 divided by 2 is 1. But according to the proposed algorithm, in order to embedding "1", remainder must be zero, therefore, this coefficient plus by 1 and it convert to 52. But if the embedding bit is zero, no need to change value of coefficient because based on algorithm, to replace the zero, remainder value should be 1. Consequently, embedding additional bits can be continued base on even or odd state of the decimal values of the MDCT coefficients. With considering the new value of MDCT coefficients, the probability of losing the hidden data bits and being changed at quantization step are reduced. Results indicate that such replacement do not effect on quality of audio.

3.2. Extracting Process

Figure 5, shows the extracting procedure of proposed algorithm. Reconstructed MDCT coefficients are analyzed to extract the hidden message after dequantization procedure. If the quotient remainder of the IMDCT coefficient after divided by 2 is zero, the corresponding embedded bit is '1', otherwise, the corresponding embedded bit is equal to '0'.

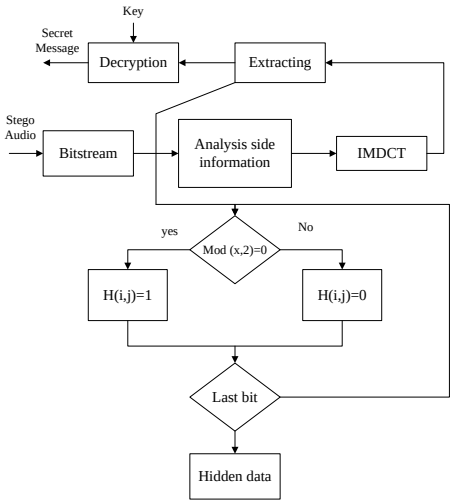


Figure 5. Block Diagram of the Extracting Process

4. Experimental Results

The following experiments were considered to evaluate the performance of the proposed steganographic algorithm.

An MP3 encoder and decoder (The LAME 3.96.2) according to the ISO/IEC 11172 standard was used as an open source. We added the proposed steganographic algorithm in the original LAME source code. Six music clips have been used to evaluate the experimental results. Music 1, 2 and 3 are classic, jazz and pop respectively, and music 4, 5 and 6 are three different types of speech audios. The secret data are selected randomly with binary value 0 and 1.

Six types of 16-bit mono PCM audio signals with sampling frequency of 44.1 kHz are chosen to hide secret data as different carrier audio files. The length of the selected audio sample to evaluate the steganographic algorithm is between 5 to 20 seconds. The experimental results are similar to 16-bit stereo and mono audio signals with different sampling frequency (32-kHz, 44.1-kHz, and 48-kHz) for the proposed algorithm.

The Bit Error Rate (BER) values are all zero and PSNR's values are changed between 64 and 73dB. Experimental results show that the BER values for different compression and embedding bit rates on same 44-kHz 16-bit stereo audio are all zero.

Table 1. Performance on Different Types of Audio

Audio	SNR	PSNR	BER
Classic	89.04	97.17	0
Jazz	75.87	86.22	0
Pop	68.23	81.42	0
Speech1	73.35	84.97	0
Speech2	72.11	83.14	0
Speech3	66.18	79.73	0

Four various types of one wav audio (32k-16bit-stereo, 32k-16bit-mono, 48k-16bit-stereo, 48k-16bit-mono) with different features and the length of 20 seconds are used to prove the performance of proposed algorithm against different MP3 compression rates. The steganographic algorithm has good performance in all audio wave formats.

Table 2. Comparison of SNR Values

Audio	ER (%)	128 kbps		256 kbps	
		MP3stego	proposed	MP3stego	proposed
Classic	0	68.56	78.68	72.35	72.38
Jazz	0	45.03	52.32	51.54	58.08
Pop	0	65.09	65.86	79.04	84.31
Speech1	0	50.12	54.77	55.79	55.80
Speech2	0	58.46	60.19	63.13	64.11
Speech3	0	53.34	58.63	59.98	62.24

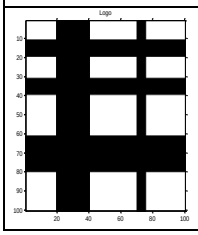
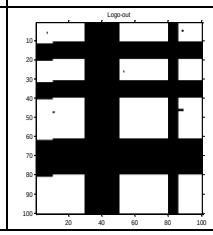
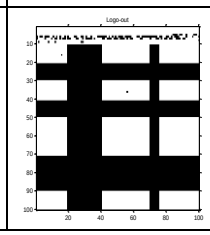
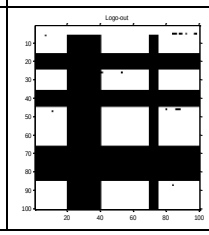
Different types of signal processing methods are applied to the carrier audio file to test the robustness of the steganographic algorithm. These attacks are including noise, quantization, low-pass filtering attack, mp3 compression which are introduced as below.

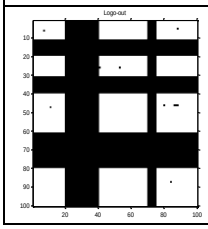
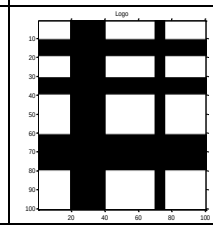
a) Additive white Gaussian noise (AWGN): noise 10 dB is added to the carrier audio file.
b) Re-quantization: 16 bit carrier audio file sampled at 44.1 kHz re-quantized to 8bit and again back to original one.

c) Low-pass filtering: Chebyshev filters with 0.5% ripple and cutoff frequency near 4,025 Hz is used
d) MP3 compression 64 kbps: The MP3 compression is applied to the carrier audio file with the bit rate of 64 kbps and then back to the original wave format.
e) MP3 compression 128 kbps: The MP3 compression is applied to the carrier audio file with the bit rate of 128 kbps and then back to the original wave format.

The extracted hidden data for the previous mention attacks are summarized in Table 3.

Table 3. The Retrieved Hidden Data after Different Attacks

Attack	AWGN	Re-quantization	Low-pass
			

Mp3-64kps	Mp3-128kps
	

The embedding algorithm is proposed to hide secret message within MP3 audio format during MP3 encoding. A high percentage of those who heard the MP3 file couldn't distinguish between the audio file with hidden message and original one. In both embedding and extracting secret message, the equal audio compression rate was used. The new steganographic algorithm has good performance by testing most popular format of audio signal. One MP3 audio is selected for analyzing the waveform in time domain before and after steganography, as shown in Figure 6. First photo shows the audio carrier signal waveform, second photo shows the audio stego signal waveform and the last one shows the difference between the audio carrier signal waveform and the audio stego signal waveform. It can be seen from Figure 6 that the difference is so small that human beings cannot perceive the difference. This leads to the imperceptibility of steganography.

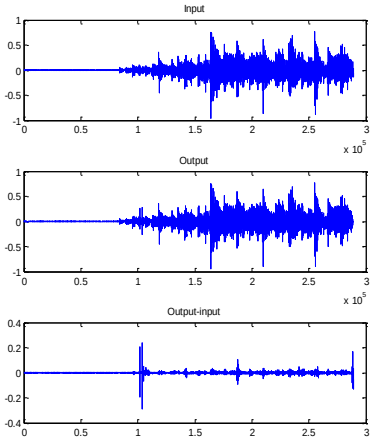


Figure 6. Comparison of the Waveforms Before and After Embedding

ROC curve [22], which is a plot of false positive rate (false alarm) against true positive rate (missed detection), shows the efficiency of the steganographic methods. A random set of MP3 files is used to analyze the performance of proposed algorithm. The steganography algorithm is efficient and the embedding method is very strong, if the curve is close to the quadrant bisector. As can be seen in Figure 7, the ROC curve is moderately coincident with bisector line. In relation with it, closeness of lines illustrates the robustness of steganographic method once the proposed algorithm is applied. Robustness of algorithm is being fully affected by the gap between two lines as the distance increases, robustness would significantly decrease. Moreover, high level of security in embedding methods is the direct consequence of coincidence between ROC curve and bisector.

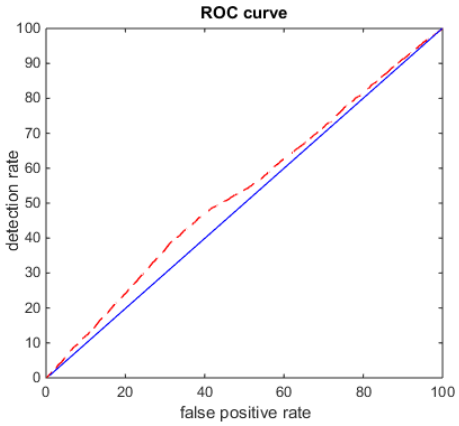


Figure 7. ROC Curve for Proposed Method

5. Conclusion

Having low robustness against common attacks as noise addition or MPEG compression that try to reveal the secret data and minimize the embedding distortion of the cover audio without affecting the perceptual transparency of the MP3 audio signal are two problems of embedding techniques of MP3 steganography. Therefore, a new algorithm is proposed to resolve these two problems.

The retrieved hidden message obtained from the proposed algorithm has the highest quality against various steganographic attacks and different signal types which shows an excellent imperceptibility for the new algorithm. A high percentage of those who heard the MP3 file couldn't distinguish between the audio file with hidden message and original one. The algorithm performance showed the qualified robustness against common attacks, especially MPEG compression compared to MP3Stego.

Acknowledgments

The authors would like to thank Universiti Teknologi Malysia for funding the research under vote 4F558. We also would like to thank our research group in supporting and giving positive comment to improve our paper.

References

[1] A. Abraham, M. Paprzycki, "Significance of steganography on data security", Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on, vol. 2, **(2004)** April, pp. 347-351.

[2] X. Dong, M. F. Bocko, Z. Ignjatovic, "Data hiding via phase manipulation of audio signals", In Acoustics, Speech, and Signal Processing, Proceedings. (ICASSP'04). IEEE International Conference on, vol. 5, **(2004)** May, pp. V-377.

[3] S. K. Moon, R. S. Kawitkar, "Data security using data hiding", In Conference on Computational Intelligence and Multimedia Applications. International Conference on, vol. 4, **(2007)** December, pp. 247-251.
[4] P. Jayaram, H. R. Ranganatha, H. S. Anupama, "Information Hiding Using Audio Steganography–A Survey", The International Journal of Multimedia & Its Applications (IJMA), vol. 3, no. 3, **(2011)** August, pp.86-96.
[5] K. Rabah, "Steganography-the art of hiding data", Information Technology Journal, vol. 3, no. 3, **(2004)**, pp. 245-269.
[6] M. Bazyar, R. Sudirman, "A Recent Review of MP3 Based Steganography Methods", International Journal of Security and Its Applications, vol. 8, no. 4, **(2014)**, pp. 405-414.
[7] N. Moghadam, H. Sadeghi. "Genetic content-based MP3 audio watermarking in MDCT domain", Watermark, vol. 1, no. 2, **(2005)**.
[8] R. Pinardi, F. Garzia, R. Cusani, "Peak-Shaped-Based Steganographic Technique for MP3 Audio", Journal of Information Security, **(2013)**, pp.12-18.
[9] M. Bazyar, R. Sudirman, "A New Method to Increase the Capacity of Audio Steganography Based on the LSB Algorithm ", Jurnal Teknologi , vol.74, no. 6, **(2015)**, pp.49-53.
[10] C. T. Wang, T. S. Chen, W. H. Chao, "A New Audio Watermarking Based on Modified Discrete Cosine Transform of MPEG/Audio Layer III", in Proc. of the IEEE International Conference on Networking, Sensing and Control, vol. 2, **(2004)**, pp. 984–989.
[11] B. Chen, J. Zhao, "An adaptive and secure audio watermarking algorithm robust to MP3 compression", In Instrumentation and Measurement Technology Conference, 2009. I2MTC'09, **(2009)** May, pp. 629-632.
[12] B. Chen, J. Zhao, D. Wang, "An Adaptive Watermarking Algorithm for MP3 Compressed Audio Signals", in Proc. of the IEEE I2MTC 2008, Victoria, BC, Canada, **(2008)** May, pp. 1057-1060.
[13] H. O. Oh, J. S. Kim, C. J. Song, Y. C. Park, D. H. Youn "Low power MPEG/audio encoders using simplified psychoacoustic model and fast bit allocation", Consumer Electronics, IEEE Transactions on. vol. 47. vo 3, **(2001)**, pp. 613-621.
[14] H. L. Dai, D. He, "An Efficient and Robust Zero-Watermarking Scheme for Audio Based on DWT and DCT", Microelectronics & Electronics, **(2009)**, pp. 233-236.
[15] ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 13818-3, "Generic Coding of Moving Pictures and Associated Audio: Audio", **(1994)**.
[16] Z. Zhou, L. Zhou, "A Novel Algorithm for Robust Audio Watermarking Based on Quantification DCT Domain", in Proc. of the Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2007), vol.1, **(2007)** November, pp. 441–444.
[17] X. Dai, M. D. Wagh, "An MDCT hardware accelerator for MP3 audio", In Application Specific Processors, 2008. SASP 2008. Symposium on IEEE. **(2008)** June, pp. 121-125.
[18] X. Y. Wang, P. P. Niu, H. Y. Yang, "A robust digital audio watermarking based on statistics characteristics", Pattern Recognition , **(2009)** pp. 3057 -3064.
[19] R. Yu, X. Lin, S. Rahardja, C. C. Ko, "A Statistic Study of the MDCT Coefficient Distribution for Audio", IEEE International Conference on Multimedia and Expo, Taipei, **(2004)** June, pp. 1483-1486.
[20] M. Qiao, A. Sung, Q. Liu, "Feature mining and intelligent computing for MP3 steganalysis", in: Proceedings of International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, **(2009)**, pp. 627–630.
[21] J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, C. Spenger, "MP3 Surround: Efficient and compatible coding of multi-channel audio", Audio Engineering Society Convention 116. Audio Engineering Society, **(2004)**.
[22] I. Avcıbas, "Audio steganalysis with content-independent distortion measures", IEEE Signal Processing Letters vol. 13, no. 2, **(2006)** February.

Authors

Mohsen Bazyar, he was born in Boshehr, Iran in 1986, and has since lived in Boshehr for the past 27 years. His primary school was Okhovat and secondary was Hedayat, Boshehr. He received his Bachelor and Master's degree from the Islamic Azad University of Boshehr in May of 2008 and 2011 respectively, and his Ph.D. in Electrical Engineering has been started from Universiti Teknologi Malaysia in July of 2013.

Rubita Sudirman, she received her Bachelor and Master degree from the University of Tulsa and her Ph.D. in Electrical Engineering from Universiti Teknologi Malaysia. Her research interests include Biomedical Electronics & Biosignal (EOG, EEG, EMG), Rehabilitation Engineering, Speech Processing.