Volume XXVI 2023
ISSUE no.2
MBNA Publishing House Constanta 2023

Scientific Bulletin of Naval Academy

SBNA PAPER • OPEN ACCESS

An analysis of steganographic and steganalytic schemes for AAC

To cite this article: M. Stoica, M. Rusu and C.Răcuciu, Scientific Bulletin of Naval Academy, Vol. XXVI 2023, pg. 86-93.

Submitted: 23.04.2023
Revised: 05.08.2023
Accepted: 29.08.2023

Available online at www.anmb.ro

ISSN: 2392-8956; ISSN-L: 1454-864X

doi: 10.21279/1454-864X-23-I2-010

An analysis of steganographic and steganalytic schemes for AAC

M Stoica, M Rusu, C Răcuciu

Bucuresti, Romania
mirela.preda@mta.ro

Abstract. AAC (Advanced Audio Coding) audio files are increasingly used, gradually replacing MP3 format files in many audio/video services, due to their low compression rate and good audio signal quality. Due to these properties, AAC audio files are frequently used for embedding information hidden by steganographic methods. This paper presents an analysis of the steganographic schemes used for audio files in AAC format, but also of the detecting and extracting methods of embedded information (steganalysis). The aim is to evaluate these steganographic and steganalytical systems, which will lead to new research directions.

1. Introduction

Steganography is gaining more and more interest, being an efficient and subtle way to transmit hidden information through multimedia files. Thus, through different steganographic methods and algorithms, the data of interest are inserted into a container and transmitted to the recipient through a public communication channel (eg Internet).

Steganography, together with its inverse procedure, steganalysis, are part of an interdisciplinary field that combines knowledge from computer science, mathematics, cryptography, etc. and is constantly evolving, along with the advance of technology and the development of new methods and algorithms.

The steganographic process must meet several requirements, such as:

- **Imperceptibility** – is an important property of steganographic algorithms, because the purpose of the process is not to draw attention to the existence of a hidden content, so that the container file presents as few as possible or no distortions at all.
- **Security** – the protection offered against unauthorized access to information.
- **Capacity** – the amount of information that can be inserted into a container and is influenced by the steganographic technique and the type of the container file.
- **Complexity** – the difficulty of implementation and the complexity of the realization of steganography.
- **Robustness** – resistance to steganalytical attacks and preservation of the integrity of the hidden content.

These properties represent the most important criteria and are used in the analysis and evaluation of steganographic techniques.

2. Advanced Audio Coding files

Due to the advantages offered by AAC files they are used as default audio format for many communication systems, such as DVB (Digital Video Broadcasting), DAB (Digital Audio Broadcasting), iTunes, the iPod, the iPhone, the PlayStation, YouTube, Twitter, Facebook).Consequenttly, they are starting to be used frequently in steganographic techniques.

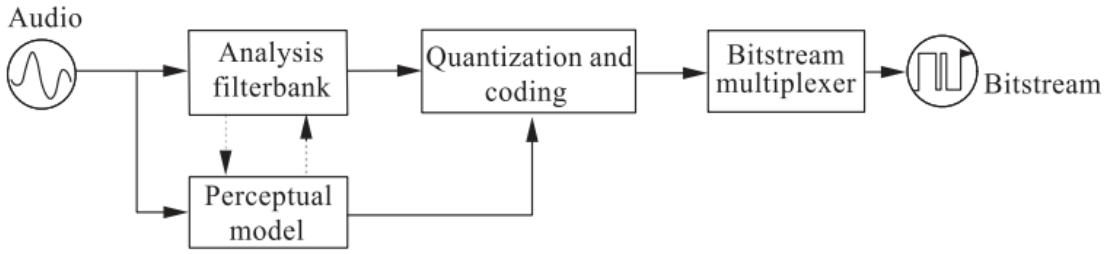

Figure 1. Block diagram of the MPEG AAC encoder[1].

In order to understand the steganographic techniques applied to these audio files, it is necessary to analyze and understand the structure of the AAC audio files (Figure 1) composed of the following blocks:

a) **Filterbank** : In this stage, a series of filters (eg MDCT, TNS) are applied to the input audio signal to help its spectral representation, forming an analysis system.

b) **Perceptual Model** : This model is based on the psychology of human hearing and can identify the masking threshold representing how much energy can be added to the audio signal without being perceptible.

c) **Quantization** : After the perceptual model stage, the audio signal is quantized, which involves assigning discrete values representative of the frequency spectrum and amplitude levels of the signal. Quantization is an important step in audio compression because it allows reducing the amount of data needed to represent the audio signal.

d) **Multiplexing** : After quantization, the audio signal is assembled to be embedded in a container audio file and involves the addition of control information, metadata and other information necessary to format and organize the audio signal within the container file (spectral coefficients quantified and coded, the scale factor)[2].

3. Steganographic methods

3.1 Traditional audio steganography

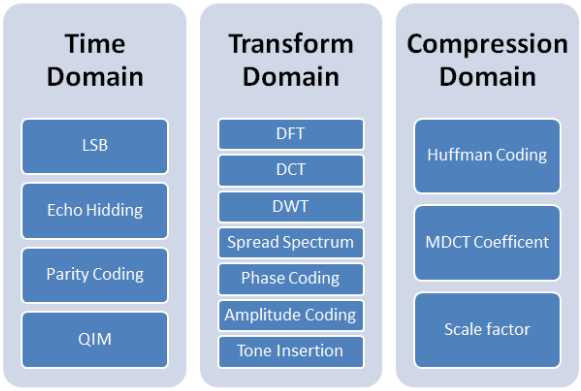

Figure 2. Traditional audio steganographic techniques

Different traditional steganographic techniques have been identified in the literature (Figure 2), divided into three categories depending on the domain where the steganography is performed: time domain, transform domain and compression domain. Time domain (spatial) techniques include the most accessible steganographic method, named LSB (Least Significant Bit), who replaces the least significant bit of the audio container values with a bit of the information to be inserted [3] , along with Echo Hidding [4.], Parity Coding [5] and QIM [6] .

In the Transform Domain, several methods have been identified based on Fourier Transform, Wavelet Transform [7] and Discrete Cosine, but also Spread Spectrum, Phase Coding [8], Amplitude Coding and Tone Insertion [9].

The Compression Domain divides the techniques according to the coding parameters that are modified in the steganographic process: Huffman coding, MDCT coefficients or scale factor.

In audio steganography that uses compression domain techniques, there are three procedures depending on the time when the process is carried out:

a) Creating the steganographic audio file and compressing it after inserting the hidden content.

The disadvantage of this method is that AAC coding is a lossy compression, so it is possible to lose some inserted information and extract it incompletely.

b) Creating a steganographic audio file on a compressed audio file.

In this case, the complexity is low, because the algorithms are incorporated in the quantized coefficients.

c) Decompressing the container, followed by inserting the steganographic content and recompressing the container.

Due to the high degree of complexity, involving decoding and inverse quantization, quantization and encoding of the stego file, this technique is rarely recommended [10].

Relatively few articles presenting steganographic schemes for AAC audio files have been identified in the literature, despite the fact that these files are widely used.

3.1.1 Huffman coding

In Huffman coding, for each symbol in the audio data stream there is a set of Huffman codes, which have the role of compressing and decompressing the audio data. But, in the coding process, escape sequence reserved for certain special situations may appear, such as rare symbols or symbols that

cannot be effectively represented using standard Huffman codes. These escape sequences can be used to insert secret information into the AAC audio data stream, making it to seem a normal variation in the data stream.

A steganographic algorithm based on this method which modifies the least significant bit (LSB) of the escape sequences to insert information with matrix encoding is described in [10]. The algorithm has a greater capacity than a spread spectrum method, it is largely undetectable, with a small difference between the histograms of the quantified coefficients before and after steganography.

The resistance to steganalytic attacks was tested with the MP3Stego application, with no significant results and low complexity, due to the fact that the length of the AAC encoding is not modified.

A simpler approach to Huffman coding is described in [11]. It does not affect the coding process nor the quality of the audio file. The algorithm extracts the Huffman coding sections and modifies them according to the inserted information. Sections are the basic units of Huffman coding, and the challenge comes from identifying the most appropriate sections to modify.

Another modality to hide data in AAC audio files is to change the sign bit of the Huffman code words[12]. This paper used an interesting technique for minimizing distortion at the expense of reduction. If the XOR of the sign bits is equal to the secret data bit, no operation will be performed, otherwise one of the sign bits will be changed to match the secret data bit. This method does not offer a high capacity, preferring a good imperceptibility of the scheme, regulated by a control factor with the role of deciding whether the absolute value of the analyzed spectral coefficient affects the audio quality, thus minimizing distortions.

3.1.2 MDCT Coefficient

The MDCT coefficients have a percentage of almost 70% in the coding bit stream of AAC files, being their main parameter, offering an embedding space for inserting hidden information.

The technique of using MDCT coefficients in inserting information into audio files is first used in a related technology named watermarking. This technology has the role of protecting the transmitted content and giving it authenticity (especially in cases of piracy, copyright, etc.) [13].

In [14], a steganography algorithm is proposed based on the adjustment of the MDCT coefficient, identifying the section with the lowest value of the MDCT coefficient, which uses the genetic algorithm to optimize the modification of this coefficient. The complexity of the method is relatively low and the choosing of the coefficient to be modified is based on the XOR operation of the four quantified MDCT coefficients. Thus, depending on the parity, if it is not the same, the fourth coefficient will be modified to insert the information. Otherwise, it will not be modified and the process will be repeated until the coding is completed. The method proves imperceptibility and good insertion capacity.

3.1.3 Scale Factor

In the AAC coding process, an important part is represented by the scale factor, which is responsible for adjusting the amplitude levels of the spectrum before quantization and ensuring the efficient and faithful representation of the signal.

A steganographic scheme that uses these scale factors to insert information is described in [1]. The algorithm modifies the quantization and coding process by changing the scale factors, without degrading the audio quality.

In order to have an increased security of the inserted data, in the algorithm described by [15] cryptography is introduced in the steganographic process, so that the data to be inserted in the scale factor are first encrypted with the 3DES algorithm.

3.2. Adaptive audio steganography

The need to implement some adaptive steganography algorithms is due to the reduced capacity of incorporation and low resistance to modern steganananalysis systems of the traditional methods [16].

Thus, adaptive steganography techniques were developed, given the fact that adaptive steganography obtained very good results for image steganography. In the case of AAC audio files, adaptive steganography techniques are based on STC (Trellis Syndrome) codes, used for an enhanced masking effect, or combined with other methods (QMDCT+STC).

Adaptive steganography schemes aim to use equal-length entropy code substitution (EECS) by designing a content-aware distortion function based on the psychoacoustic model. A generalized adaptive framework Huffman coding mapping (AHCM) algorithm is proposed in [17], but it is intended for MP3 encoded audio files.

In [18] a QMDCT scheme is presented, modifying the coefficients that use the syndrome-trellis code (STC) technique and a distortion function, obtaining superior results to traditional techniques in terms of insertion capacity, security and imperceptibility .

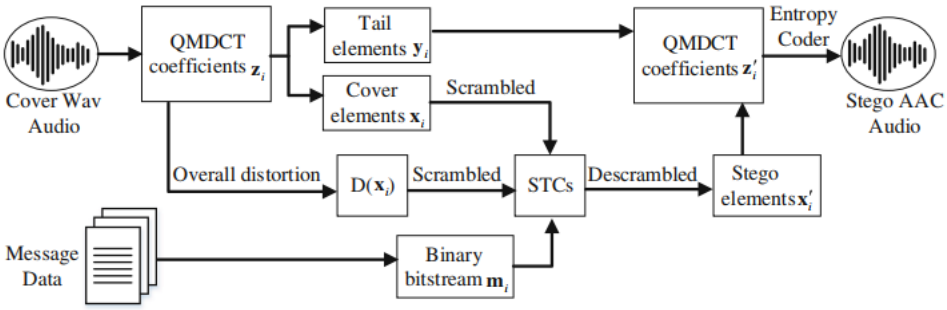

Figure 3. The embedding process of [18]

The difference between QMDCT and MDCT steganography is that in the case of QMDCT, the steganographic information is hidden in the quantized coefficients of the transform, while in the case of MDCT, the steganographic information is hidden in the coefficients of the MDCT transform before quantization.

All previously mentioned adaptive steganographic schemes change the AAC compression parameters, but in the paper [19] the insertion method is performed by changing the Time Domain parameters and the trellis-syndrome code (STC) technique.

4. Steganalysis methods

For the detection of steganographic schemes in AAC files, some steganalytic systems made for other compressed audio files (MP3) or made for different steganographic methods (eg Huffman Coding), do not return the expected results.

In the case of steganographic files made with Huffman coding methods, which do not modify the MDCT coefficients, they are not detected by the instruments that perform the steganalysis of MDCT schemes. In the case of files created by modifying Huffman coding, Markov Models gave results, which were initially used successfully for steganalysis of image files. Such an algorithm is described in [20], which aims to extract the Markov transition probabilities of adjacent SFB (scale factor band) codes and identify stego files using the C-MAC (Calibrated matrix of adjacent codebook) feature, for an accuracy system elevation.

[21] represents a steganalytic scheme for detecting stego files that have incorporated information by modifying the MDCT coefficient. The consequence of the MDCT coefficient change is also reflected in the static characteristic of the inter-intra frame difference, and this change is used in a classifier to perform steganalysis. The paper proposes 16 sets of features to identify the effects of MDCT modification.

Thus, the correlations between the adjacent intra-frame MDCT coefficients and the coefficients that have the same frequency in the adjacent frame (inter-frame) are evaluated. The method consists in

dividing the frames according to the type of block, long frames and short frames, and for each type the correlation of the intra- and inter-frame MDCT coefficients is analyzed, including calculating the first-order and second-order difference between these coefficients. Additionally, for each adjacent relationship, the Markov transition probability and the cumulative density (Joint Density) are calculated (Figure 4).

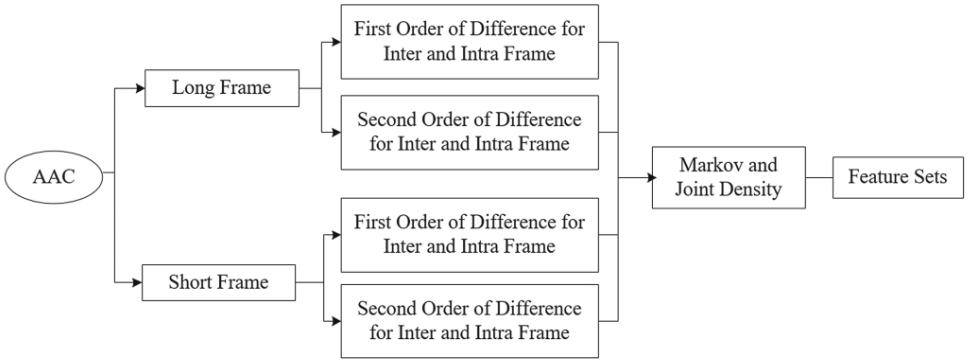

Figure 4. Diagram of feature extraction presented in [21]

Following the analyses, it was highlighted that the intra-frame correlations are more relevant compared to the inter-frame ones, and the steganalytic system obtained an accuracy of 85.34%, a superior result compared to the existing systems.

[22] uses a deep residual network, because several convolutional layers can be integrated without affecting the performance of the network, as is the case with CNN networks. This steganalytic system obtained a very good accuracy for all three domains of AAC steganography (Huffman coding, MDCT and scale factor), but also for encoded MP3 files. The Spec-ResNet network has 30 layers and was trained with spectrograms with different window sizes, in order to extract the most varied features.

A similar analytical scheme is also described in [23] and is dedicated to steganography which involves the modification of MDCT coefficients and Huffman coding, reaching an accuracy of 94% and respectively, 85.5%. The process starts with filtering the data using high pass filters with the role of preprocessing the audio data and removing the residues, following that the matrix of QMDCT coefficients is given as input to the neural network, as any steganographic processing modifies the QMDCT coefficients. To test the steganalytic network, AAC stego audio files were created using three different algorithms, MIN [24], SIGN [12] and HCM [25].

5. Conclusion

The article presents an extensive analysis of the existing steganographic schemes for AAC audio files, along with steganalysis schemes. An effective AAC audio steganography scheme should be able to hide a large amount of secret information in a subtle, undetectable way without introducing significant distortion to the underlying audio file and without affecting the audio quality of the file. It should also be able to withstand various types of statistical analysis and be relatively fast and easy to implement. Adaptive steganographic methods are a major improvement over traditional methods, just as steganalytic methods based on neural networks have superior results to classical methods.

References

[1] S. Xu, P. Zhang, P. Wang and H. Yang. (2009). Performance analysis of data hiding in MPEG-4 AAC audio, in Tsinghua Science and Technology, vol. 14, no. 1, pp. 55-61, https://doi.org./10.1016/S1007-0214(09)70007-0 .

[2] Neubauer, C., & Herre, J. (2000). Audio watermarking of MPEG-2 AAC bit streams. Audio Engineering Society Convention 108.

[3] Salem Atoum, M., M Alnabhan, M., & Habboush, A. (2017). Advanced LSB Technique for Audio Stenography. https://doi.org/10.5121/csit.2017.70409 .

[4] Ko, BS, Nishimura, R., & Suzuki, Y. (2005). Time-spread echo method for digital audio watermarking. IEEE Transactions on Multimedia, 7(2). https://doi.org/10.1109/TMM . 2005.843366.

[5] Kaur, Ramandeep & Thakur, Abhishek. (2014). Enhanced steganographic method using LSB, Parity and spread spectrum technique for Audio Signals.10.13140/RG.2.2.16950.91203.

[6] Chen, B., & Wornell, GW (2001). Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Transactions on Information Theory, 47(4). https://doi.org/10.1109/18.923725 .

[7] Sheikhan, M., Asadollahi, K., & Shahnazi, R. (2011). Improvement of embedding capacity and quality of DWT-based audio steganography systems. World Applied Sciences Journal, 13(3).

[8] Dong, X., Bocko, MF, & Ignjatovic, Z. (2004). Data hiding via phase manipulation of audio signals. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 5. https://doi.org/10.1109/icassp.2004.1327126.

[9] Yousif, SA, M. wahbi, T., & Sayed, MH (2017). Audio Steganography Using Tone Insertion Technique. International Journal of Computer Applications Technology and Research, 6(6). https://doi.org/10.7753/ijcatr0606.1004.

[10] Wang , Y., Guo, L., Wei, Y., & Wang, C. (2010). A steganography method for AAC audio based on escape sequences. Proceedings - 2010 2nd International Conference on Multimedia Information Networking and Security, MINES 2010. https://doi.org/10.1109/ MINES.2010.178.

[11] Jie Zhu, Rang-Ding Wang, Juan Li and Di-Qun Yan, 2011. A Huffman Coding Section-based Steganography for AAC Audio. Information Technology Journal, 10: 1983-1988.

[12] Zhu, J., Wang, R., & Yan, D. (2010). The sign bits of Huffman codeword-based steganography for AAC audio. 2010 International Conference on Multimedia Technology, ICMT 2010. https://doi.org/10.1109/ICMULT.2010.5629745

[13] Pinel, J., Girin, L., Baras, C., & Parvaix, M. (2010). A high-capacity watermarking technique for audio signals based on MDCT-domain quantization. 20th International Congress on Acoustics 2010, ICA 2010 - Incorporating Proceedings of the 2010 Annual Conference of the Australian Acoustical Society, 5.

[14] Li, C., Zhang, X., Luo, T., & Tian, L. (2020). Audio Steganography Algorithm Based on Genetic Algorithm for MDCT Coefficient Adjustment for AAC. Proceedings - 2020 IEEE International Symposium on Multimedia, ISM 2020. https://doi.org/10.1109/ISM.2020.00026

[15] Wei, Y., Guo, L., & Wang, Y. (2010). Controlling bitrate steganography on AAC audio. Proceedings - 2010 3rd International Congress on Image and Signal Processing, CISP 2010, 9. https://doi.org/10.1109/CISP.2010.5647484

[16] Zhang, Z., Yi, X., & Zhao, X. (2020). An AAC steganography scheme for adaptive embedding with distortion minimization model. Multimedia Tools and Applications, 79(37–38). https://doi.org/10.1007/s11042-020-09344-0

[17] Yi, X., Yang, K., Zhao, X., Wang, Y., & Yu, H. (2019). Ahcm: Adaptive huffman code mapping for audio steganography based on psychoacoustic model. IEEE Transactions on Information Forensics and Security, 14(8). https://doi.org/10.1109/TIFS.2019.2895200

[18] Zhang, Z., Yi, X. & Zhao, X. An AAC steganography scheme for adaptive embedding with distortion minimization model. Multimed Tools Appl 79, 27777–27790 (2020). https://doi.org/10.1007/s11042-020-09344-0 .

[19] Luo, Weiqi & Zhang, Yue & Li, Haodong. (2017). Adaptive Audio Steganography Based on Advanced Audio Coding and Syndrome-Trellis Coding. 177-186. 10.1007/978-3-319-64185-0_14.
[20] Ren, Yanzhen & Xiong, Qiaochu & Wang, Lina. (2016). Steganalysis of AAC using calibrated Markov model of adjacent codebook. 2139-2143. 10.1109/ICASSP.2016.7472055.
[21] Ren, Yanzhen & Xiong, Qiaochu & Wang, Lina. (2017). A Steganalysis Scheme for AAC Audio Based on MDCT Difference Between Intra and Inter Frame. 217-231. 10.1007/978-3-319-64185-0_17.
[22] Ren, Y., Liu, D., Xiong, Q., Fu, J. and Wang, L. (2019). Spec-resnet: a general audio steganalysis scheme based on deep residual network of spectrogram. arXiv preprint arXiv:1901.06838.
[23] Wei, Z. and Wang, K. (2022). Lightweight AAC Audio Steganalysis Model Based on ResNeXt. Wireless Communications and Mobile Computing.
[24] WANG Yu-jie1 ,2 , GUO Li1, WANG Cui-ping1.(2011). Steganography Method for Advanced Audio Coding. Journal of Chinese Computer Systems., 32(7): 1465-1468
[25] Z. Jie. (2012). The research on information hiding in MPEG-2/4 advanced audio coding, [Ph.D. Thesis], Ningbo University, Ningbo.