Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9794160, 1 page
https://doi.org/10.1155/2023/9794160

Retraction

Retracted: A New Blind IoT-Based MP3 Audio Watermarking Scheme for Content Integrity Checking and Copyright Protection

Wireless Communications and Mobile Computing

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Masmoudi, M. Charfeddine, S. Alsharif, and C. Ben Amar, "A New Blind IoT-Based MP3 Audio Watermarking Scheme for Content Integrity Checking and Copyright Protection," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 5075827, 12 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 5075827, 12 pages
https://doi.org/10.1155/2022/5075827

Research Article

A New Blind IoT-Based MP3 Audio Watermarking Scheme for Content Integrity Checking and Copyright Protection

Salma Masmoudi,[1] **Maha Charfeddine**,[1] **Sameer Alsharif**,[2] **and Chokri Ben Amar**[2]

[1]*REsearch Groups in Intelligent Machines (REGIM Lab), University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, 3038, Sfax, Tunisia*
[2]*Department of Computer Engineering, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Maha Charfeddine; maha.charfeddine.tn@ieee.org

Received 21 January 2022; Revised 12 March 2022; Accepted 30 April 2022; Published 25 May 2022

Academic Editor: Alessandro Bazzi

Rapid development in the area of information and embedded technology, mobile communication networks, IoT multimedia applications, compression, and distribution over Internet has led to a significant need to answer the question: "How to protect, secure, and authenticate multimedia documents?" One of the proposed responses to this challenge is digital watermarking, which hides an inaudible watermark in digital multimedia content. Despite the large number of proposed competent watermarking algorithms in the literature, few are suitable for the compression domain, mainly the authentication application. In this context, this paper depicts a new MP3 audio watermarking scheme for copyright protection and content integrity checking operating directly in the compression domain using Huffman data and side information features. This scheme overcomes the problem of computational time as it operates directly on the compressed bitstream. In addition, it provides enough embedding space by using Huffman data features. A strong advantage of our scheme is that it has been used successfully for both authentication and copyright protection applications. Experimental results have revealed that the proposed watermarking schema has achieved very competitive results compared to others from the literature in terms of inaudibility, robustness, and especially capacity ratio. Our approach offers also good values of ODG and NC even after double recompression-StirMark attacks.

1. Introduction

With the speedy progress of the multimedia and Internet technologies, the combination of multimedia devices and services in the IoT (Internet of Things) become a central task [1] especially to design new IoT multimedia protocols and ensure copyright protection and authentication of digital media content. Digital watermarking [2, 3] has been suggested to solve several multimedia security difficulties facing. This technology, which presents a vital research branch of multimedia data hiding, embeds additional information as a watermark in the host files and then extracts it when necessary. This watermark data can meet the requirements of certain applications such as authentication [4–6], copyright protection [2, 7], indexation, and watermark tracing [8, 9]. Audio content is an important part of the media streaming as it can be used to improve multimedia applications for many purposes [10]. Audio watermarking schemes should respect some fundamental properties [2]. The most important ones are inaudibility, robustness, security, capacity or data rate (data payload), and computational complexity. It is important to maintain a tradeoff as a result of these conflicting characteristics.

The extensive use of compressed audio and video data on the Internet makes compressed audio contents sensitive and produces financial losses for musical artists that are caused by illegally copying and distributing the audio content. It explains the necessity of designing copyright protection and integrity control techniques suitable for compression files. However, changing the content of compressed files without

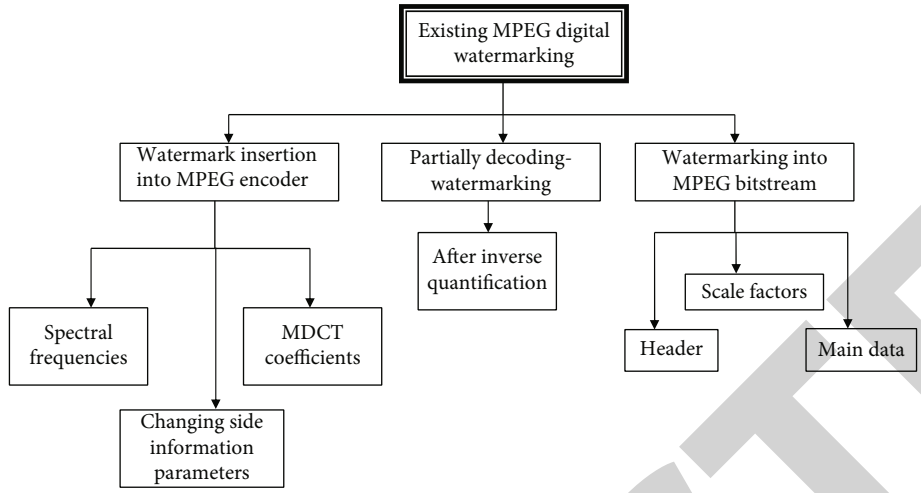


Figure 1: Categorization of existing digital audio watermarking works for MPEG compressed files.

touching the quality raises the question of how to ensure the watermark embedding without inaudible modifications. The compressed signal is too reduced to find positions to hide the watermark bits.

Most existing watermarking schemes in the literature hide signatures in uncompressed signals [2, 7, 11]. These schemes could be classified into time domain, frequency domain, or wavelet domain approaches. Some of these techniques would persist to the decompression and recompression attacks [7]. One possible method to watermark a compressed domain bitstream is to decode the input, apply the signature embedding process, and finally re-encode the watermarked carriers. This process can guarantee the watermark robustness. However, as a significant disadvantage, it complicates the computational time since it uses the compression process which is not satisfactory for online applications. For this reason, additional watermarking schemes working on the compressed host should be considered. However, based on our state of the art, few of prior watermarking methods are working on the compressed bitstream.

This paper suggests a new blind IoT based MP3 audio watermarking approach using the compressed MP3 bitstream directly. The suggested method uses Huffman-data, MP3 recompression calibration, and side information features. In this paper, the following sections are discussed: Section 2 gives an overview of the MP3 encoders and decoders. Section 3 describes the most relevant works in audio watermarking for MP3 encoded audio, revealing their strengths and weaknesses. Finally, Section 4 describes the proposed MP3 audio watermarking approach, working directly in the compressed space for copyright protection and data integrity check.

2. Literature Survey

Different IoT protocols and standards are used nowadays with a high variety of data type and a wide range of services and generated files. One of the important exchanged data is audio information using MP3 standard format. IoT ecosystems should cover the confidentiality, privacy and integrity

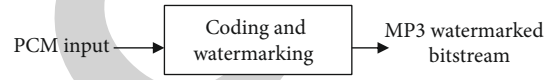


Figure 2: General scheme of Watermark insertion process in the compression Encoder

of such sensitive information where different approaches are used as cryptography, steganography, and watermarking [12]. The watermarking scheme excels in terms of payload capacity and speed mainly in the detection (extraction) process. Therefore, we take advantage of using it for our proposed based audio approach.

Considering the importance of audio watermarking in the compression domain, this section presents the most relevant works describing audio watermarking schemes for MP3 compressed files [13]. We can categorize them according to the tree structure presented in Figure 1. The insertion of watermark information is achieved by one of the three proposed approaches: after partial decompression, throughout the compression stage, or directly into the compressed MP3 audio bitstream.

The watermark can be embedded after a subsequent decompression phase and recompression afterward. Otherwise, the watermark embedding process can be applied inside the MP3 compressor. In this situation, the embedding time and the required time for audio signal compression are the same. An additional approach consists of hiding the watermark in the compressed MP3 Bitstream [13]. For this case, the hiding space and the robustness performance are much reduced. Therefore, multiple watermarks state a problem for the robustness of the system face to signal attacks.

2.1. Watermark Insertion Process in Compression Encoder. This watermarking concept is founded on inserting a watermark during encoding step. It performs encoding and watermarking simultaneously (see Figure 2). This approach provides low consuming time and low computational complexity, high robustness, and maximum inaudibility performance. In this approach, the embedded information can be hidden in parallel with one of the different levels of MP3

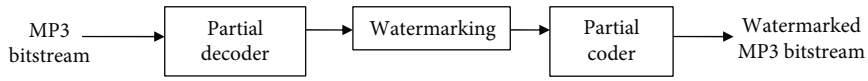


Figure 3: General scheme of partially decoding watermarking re-encoding.

encoding scheme which are quantification, space transform, and entropy coding. One of the best ways to insert a watermark during MP3 encoding is after MDCT transformation of frequency sub-band samples obtained from the analysis of the polyphase filter bank [13–18]. However, other existing solutions embed watermarks by changing some side information parameters [18]. Watermark embedding methods using MDCT coefficients have various typical features. Watermark is in general a binary sequence generated randomly (often created by a secret key). It may be a gray-scale binary image, or short audio signal. The generated watermark is then embedded into the spectral coefficients after MDCT transformation using common algorithms of digital watermarking such as spread spectrum (SS) modulation [15, 19, 20].

2.1.1. Watermark Insertion Process into MDCT Coefficients. In 2008, Chen et al. [13, 17] suggested an adaptive digital audio watermarking approach during the MP3 compression process. In the first scheme, the signature is inserted in the process of MP3 encoder after MDCT and before quantization, thus exploiting the human auditory system. In addition, it applies the Gaussian distribution analysis on frames and original audio energy of sub-bands so as to ensure adaptive control. Regarding the watermark retrieving process, the scheme has succeeded to be blind. As evaluation performance, this algorithm warrants the robustness against MP3 compression and survives to most common attacks. To boost the signature security, authors in [13, 17] used an enhanced algorithm to hide watermark into various frequency sub-regions and not only in low or middle frequency ones. Blind detection is then performed by calculating correlation coefficient to retrieve watermark without using original audio signal. This algorithm improved the watermarking ratio and the robustness to MP3 compression attack.

In 2018, LI Chen et al. [19] proposed a watermarking scheme, in compressed domain, based on calculating the low frequency energy value of the MP3 frame channels. The process of embedding and detection are operating, respectively, during MP3 encoding and decoding processes. Using the MDCT coefficients generated during the MP3 encoding stage, the low frequency energy of channels was calculated. Then, the watermarking process is operated by modifying some MDCT coefficient chosen during the quantization process with a fixed step and according to a best value of the ratio between the energy of the left and right channels. Experiments show a good inaudibility and robustness results against several attacks mainly MP3 recompression attacks with an average value of NC equal to 0.95.

2.1.2. Watermark Insertion Process by Changing Side Information Parameters. In 2017, Su et al. have announced in [18] a new MP3 audio watermarking scheme using window switching strategy. This semi-fragile watermarking algorithm uses the feature of window switching during encoding stage to be able to localize tamper. This technique achieves hiding process by developing a mapping relation between the MD5 (Message Digest 5) of chosen watermark and the type of window. In addition, the authors of this paper describe the tamper detection and identification processes by analyzing the hidden authentication information. The experiments show the efficiency of this scheme in terms of time consuming, imperceptibility, robustness against some attacks, and accuracy for tamper detection. The main limit of this scheme resides in the fact that it cannot survives to the attacks of MP3 recompression.

2.2. Partial Decoding/Re-encoding Watermarking. This technique can be appropriate essentially for on fly inserting. It is based on MP3 decoder principle (see Figure 3). Embedding watermark is done after subsequent audio decompression followed by recompression which generates a degradation in transparency and robustness. For instance, the MP3 Bitstream is subject to bitstream demultiplexing followed by side information and scale factors to extract the quantized and coded spectral values. To obtain the spectral representation, a Huffman decoding and an inverse quantization process are applied using decoded side information and scale factors [21, 22].

In 2012, a novel algorithm was proposed by Subramanyam and Emmanuel [21] based on a dual encryption and compression process. The embedding process is started by a simultaneous process of compression and encryption. Then, the resulting signal is partially decoded to embed watermark in the quantized frequency coefficients. The choice of candidate coefficient is performed similarly to encryption process. Then, the new coefficients (modified coefficients) are changed and recompressed to build the watermarked audio signal. This watermarking approach has shown a good robustness results against general transformations and attacks such as resampling, lowpass and highpass filtering, and recompression.

In 2017, Wenhui et al. have proposed in [22] a watermarking algorithm taking as input an MP3 audio signal and using unipolar quantization and wavelet transform. This algorithm starts by decompressing the MP3 audio file and then applies the unipolar quantization to make changes in the low frequency coefficients selected from the third-order discrete wavelet transform. Finally, watermarked MP3 audio signal is generated. Experiments prove good auditory transparency, good robustness face to lowpass filtering, whitening, resampling, and cropping attacks and also a rapidity in the extraction process of the watermark.

2.3. Watermarking in the Compressed Bitstream. The term watermarking in a compressed bitstream denotes that the

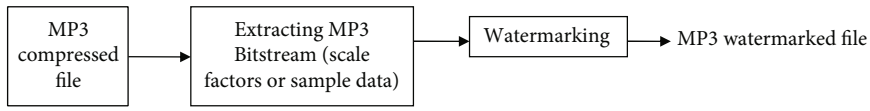


Figure 4: General scheme of the concept of watermark insertion in compressed bitstream.

insertion of watermarks is carried out directly in the bitstream domain without decoding and re-encoding the audio signal.

This offers many benefits in many applications:

(i) Robustness: The result of the bitstream watermarking is already in compressed form. Therefore, the degradation of the confidential data by partial recompression is avoided

(ii) Improvement of sound quality by avoiding cascaded encoding/decoding stages: Using this type of bitstream watermark embedder results in an improvement of sound quality

(iii) Low computational complexity

In the literature, many papers proposed to embed the watermark into the compressed bitstream directly. Such works choose to insert some sample data [22, 23], scale factors [3, 24, 25] and header parameters [26] (see Figure 4).

2.3.1. Watermarking in Compressed Bitstream Using Scale Factors. In 2006, Koukopoulos and Stamatiou proposed in [3] an algorithm ensuring an efficient and a blind digital watermarking scheme operating for MP3 files directly in the compressed data domain. This algorithm outperforms by using the semantic information to construct the watermark and offers high performances, copyright protection, and authentication applications. Experiments show that this algorithm can survive to the conventional attacks applied to audio data, but is inappropriate to survive to decompression/compression attacks.

In 2008, Takagi et al. [25] proposed an MP3 watermarking method operating directly in compressed bitstream for mobile terminals. The embedding process is achieved by changing the scale factors. In this algorithm, the authors analyze the modification of scale factors' LSB to guarantee a high embedding speed with minimum distortion. The evaluation of this algorithm indicates that the embedding payload can reach 3 bits per frame to retrieve rapidly the hidden bits without destroying the transparency of the signal. The insertion ratio of the proposed scheme is sufficient to hide both digital watermark and its digital certificate.

In 2011, Ting-ting et al. proposed in [23] a new algorithm that uses scale factors and embeds watermark directly in the MP3 Bitstream. The watermark is embedded by slightly modifying some random scale factors selected using a linear congruential generator. To secure their bitstream, the authors used an Arnold transform to scramble the watermark. Experimental results show a good result of inaudibility.

2.3.2. Watermarking in Compressed Bitstream Using Sample Data. In 1998, Nahrstedt proposed in [24] to choose to study the sensitivity of the human hearing system face to the sample data modification when embedding the watermark. To reduce the distortion rate, the uthors select some samples using spacing parameters

Masmoudi et al. introduced in [26] a novel blind audio watermarking scheme for MP3 bitstreams. The suggested solution exploits encoded MP3 data and decompression requirements to select the positions of embedding. The watermark retrieving process is based on a secret key. This scheme ensures good results in terms of imperceptibility, robustness, and payload.

2.3.3. Watermarking in Compressed Bitstream Using Header Parameters. In 2014, Bailong et al. [6] suggest a system that exposes a MP3 digital audio watermarking scheme without changing the host audio data. Based on the characteristics of encoding process and the MP3 frame structure, this scheme uses a part of the header of the host MP3 frame (especially the private bit) to illustrate the consistency between Main data and watermark. The most important advantage is that the embedding information process offers a very high transparency as the Main data are not affected. To improve the security performance, this scheme is enhanced by adding an encryption level using Arnold transform for the secret information. To avoid the problem of packet loss during transmission and to increase the rebuilding quality of the watermark, a synchronization information is added to the signature. Experiments confirm low processing time and good results in terms of transparency and robustness against disturb attacks

3. Proposed MP3 Watermarking Scheme for Copyright Protection and Integrity Verification Applications

Based on the literature survey, we can notice a lack of proposed solutions to the watermarking problem of MP3 files operating directly in the compressed bitstream devoted for authentication application. The existing systems offer good inaudibility and low complexity results but gives a low ratio of insertion and low robustness performances. Consequently, we intend to improve our previous work in [26] and propose a new approach achieving the control integrity and the authentication requirements. We use a fragile content watermarking approach [27] combining robust watermarking and fragile content features.

The proposed watermarking scheme is blind since we do not need the original compressed file for both watermark detection and control integrity steps. In addition, no specific operations are needed to perform those processes. In the watermark detection process, we use only the secret key to search the embedding positions from the watermarked audio signal. For the control integrity process, only the

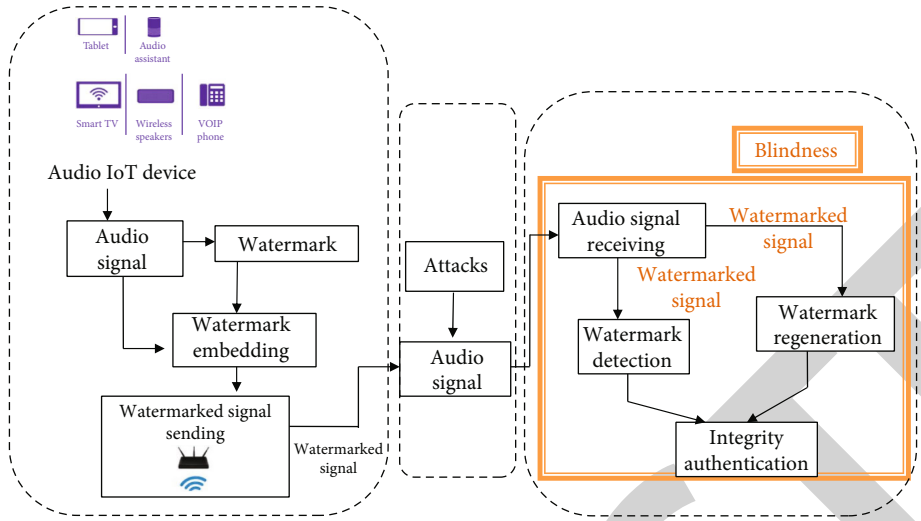


Figure 5: Integration of audio watermarking scheme in an IoT architecture.

watermarked audio signal is necessary to regenerate the features needed to check the content alterations if existing. The idea of using blind scheme enhances not only the security aspect but also the rapidity propriety making our system suitable for real time-based multimedia IoT applications.

For the concept of Internet of Things (IoT), the variety of data type transferred throughout the networks has been emerged for a wide range of services with a high volume of generated files. One of the important parts of data diversity is audio type of files using MP3 format. In addition, the IoT ecosystem should cover the confidentiality, privacy, and integrity of sensitive information where different approaches are used such as cryptography, steganography, and watermarking. The watermarking scheme excels in term of payload capacity and speed mainly in the detection process. Therefore, we take advantage of using it in our proposed system.

Figure 5 highlight the blindness need and the use of IoT multimedia context.

The section below details the process of watermark hiding, retrieving, and also the proposed integrity control scheme.

3.1. Watermark Insertion Process. As shown in Figure 6, the proposed watermarking hiding process uses an MP3 bitstream as input. The watermark is hidden in the Huffman data of the compressed bitstream without needs of decoding. This watermarking technique is described in the paper [26]. It confirms the integrity propriety of the MP3 audio files. This contribution is founded on a preliminary study of the MP3 side information features, MDCT distribution, and recompression effects. This study helps us to construct the watermark and to select the embedding positions (to construct the secret key). The proposed watermark embedding process is preceded by a step of silence deletion [28], and it is mainly composed of four parts: feature extraction, watermark construction, recompression calibration, and watermark embedding.

(i) Features extraction: Extracted audio features represent the fragile content watermark. This watermark is robust against many signal attacks, and it can also detect content manipulation. When we use MP3 files, we focus on the features of MPEG audio. To avoid time consuming problem, we use features extracted directly from the MP3 bitstream without decoding. Two kinds of data are efficient as features: the encoded sub-band values and the encoded data in the header-like fields (scale factors, header, and side information). This proposed scheme uses side information and MDCT distribution. The feature "main_data_begin" will be used to select the embedding frame by calculating the frame offset. Moreover, the MDCT distribution and recompression calibration step are employed to pick the inserting positions. However, the other side information features [29] (scfi, part2_3_length, big_values, global gain, block type, table_select, and scalefac scale) are calibrated with MP3 recompression and embedded as a mark to control the carrier signal integrity. This step is summarized in Figure 6(b)

(ii) Watermark construction: The embedded watermark is a set of side calibrated information features robust to the content manipulation attacks. To avoid the embedding capacity problem, we use a checksum function. Instead of inserting the feature vector, we insert only the checksum vector. The checksums can be compared to the recalculated, attacked, and watermarked bitstream feature checksums to detect content modifications. The hash function MD5 [30] is computed as checksum of the feature vector. The used watermark in this work is a 128-bit binary sequence

(iii) Recompression calibration: We calibrate recompression to preserve the embedding positions. The original MP3 bitstream is denoted by Xc, which

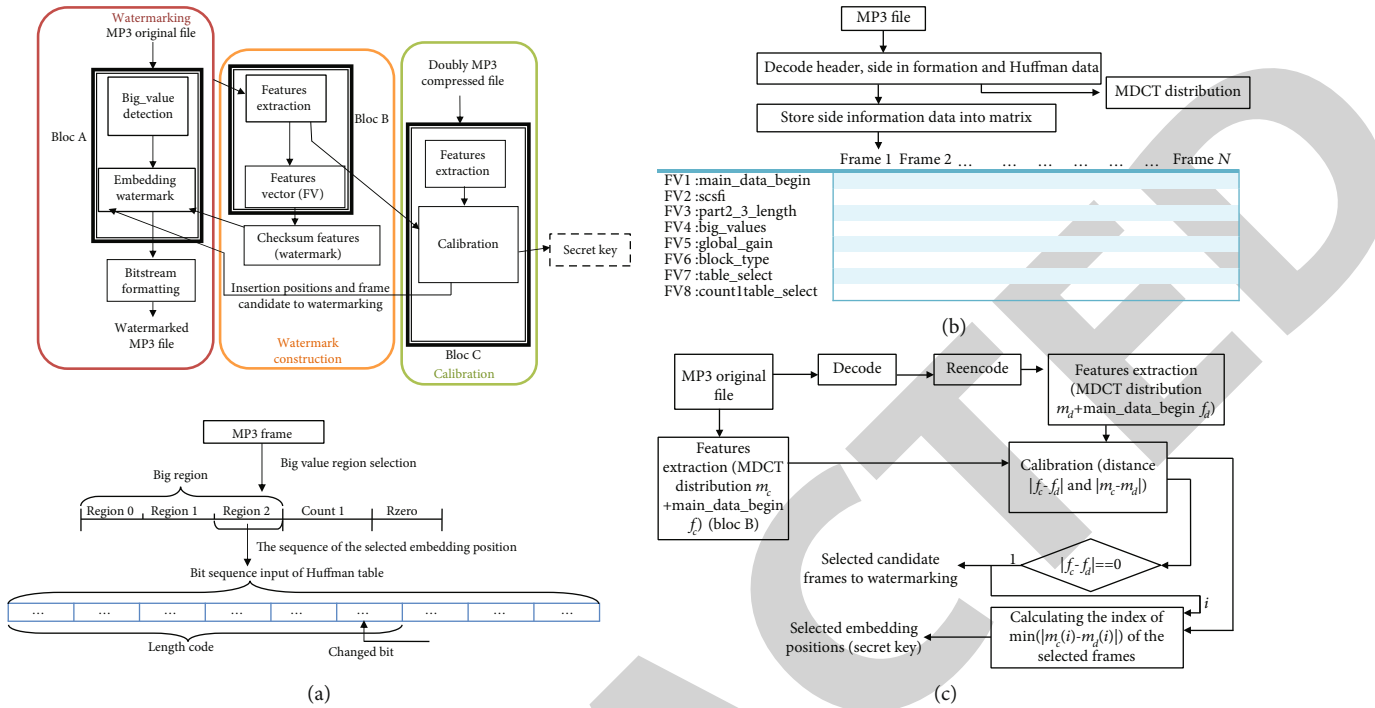


FIGURE 6: Proposed watermarking embedding schemes. (a) watermark embedding (Bloc A). (b) feature extraction (Bloc B). (c) recompression calibration (Bloc C).

contains N frames, and Xd is the doubly MP3 compressed bitstream achieved by decoding and coding again. The process of calibration (represented in Figure 6(c)) is summarized as follows:

(1) For the i^{th} frame in Xc and Xd, get the MDCT distribution and the main_data_begin feature, respectively

(2) Repeat Step 1 until reaching the end. Eventually, we get two MDCT distributions m_c and m_d and two vectors of main_data_begin features f_c and f_d of the two bitstreams Xc and Xd (original and double compressed bitstreams)

(3) Calculate the frame offsets using the term $|f_c-f_d|$, and select frames with zero offset

(4) Select the insertion positions by calculating the index of $\min(|m_c-m_d|)$ for each frame with zero offset

(5) Take the insertion position as input to the embedding process, and save it like a secret key used inside the watermark detection mechanism

(iv) Watermark embedding: First, the host MP3 audio file undergoes a step of silence trimming [28]. The second step is using a partial MP3 decoder to extract the header, scale factors, side information, and then the Huffman data of each frame. In the third step, we use the Huffman decoder to detect the significant value region. This watermarking algorithm uses the Huffman data codes to boost the embedding capacity. More details can be retrieved in [26]. As illustrated in Figure 6(a), we use the bits of Huffman data codes as candidate bits to be in the significant values region (region2) of the mp3 frame selected in the calibration step. These bits are picked out using the calibration of MDCT distribution. Region2 holds spectral coefficients in the range 5 to 14 KHz at 44.1 KHz sampling rate [29]. Most of the spectral energy coefficients are concentrated in region0 and region1 of the signal due to the energy compaction properties of MDCT [29, 31]. Therefore, any modification in this region introduces lower noise in the host signal. The candidate bit should also verify that after embedding, the index of Huffman table does not change. The embedding strategy is substitutive (we substitute the located bit by the current watermark bit).

3.2. Watermark Detection Stage. The lefthand side of Figure 7 summarizes the watermark detection stage. The extraction mechanism is blind. It consists to retrieve all the hidden watermarks, which does not necessitate the original audio. In this process, we require the embedding insertion positions. Such positions compose the secret key of our scheme. This procedure can be done easily as we have no needs to the partial decoding step. Experimental results demonstrate a high capacity of the proposed system in term of inaudibility and a best robustness against several attacks.

3.3. Integrity Verification. During integrity verification, the hidden features are compared with the recalculated ones (as for hash functions in cryptography). If any modification is perceived, the current contents and hidden watermark will

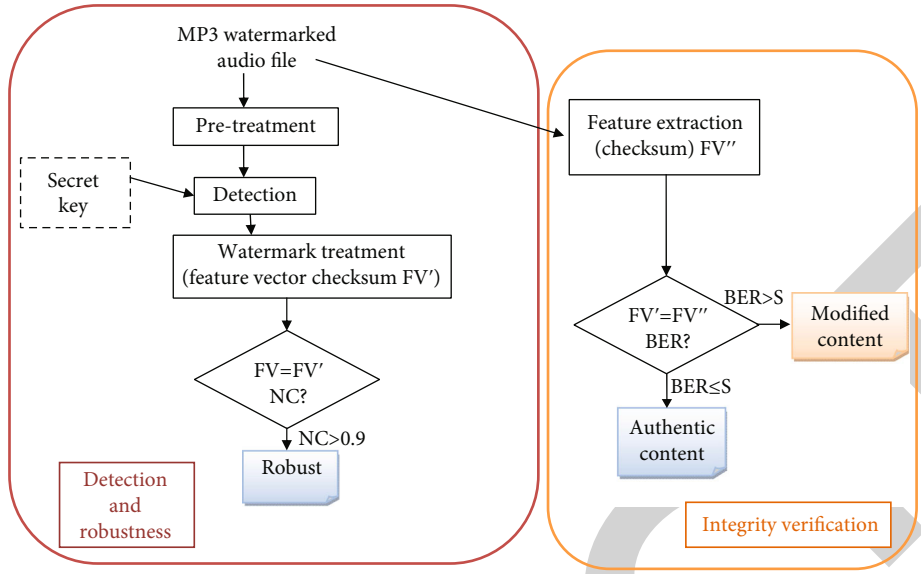


Figure 7: Proposed watermarking scheme detection and Integrity verification process.

Table 1: Audio description.

Audio file	Time (s)	Rate (Kps)	Description
Classic.mp3	140	128	Classic music
Blue1.mp3	140	128	Blue music
Blue2.mp3	120	128	Blue music
Country.mp3	140	128	Country music
Pop1.mp3	140	128	Pop music
Pop2.mp3	120	128	Pop music
Folk.mp3	120	128	Folk music
Quran1.mp3	80	128	Quran
Quran2.mp3	127	128	Quran
Recordings.mp3	145	128	Women speech

be different, and the system throws an alert message. Here, we speak about a watermarking scheme fragile to modification operations, but able to handle content preserving operations (such manipulations that do not modify the content).

To control the integrity for this scheme, we compare the checksum of the vector of features extracted from the watermarked files with the extracted watermark (original embedded features), as described in the righthand side of Figure 7.

3.4. Experimental Results. In this part, the evaluation of our suggested technique is presented. The experiments use various stereo audio MP3 files with a compression rate of 128 kbps. Such audio segments (see Table 1) contain multiple styles, such as blues, pop, classical country, folk, Quran, and some recorded audio (with content vulnerability). The watermark used in this paper has a size of 128 bits due to MD5 algorithm checksum.

The tests are carried out on a machine with a Core i3 Intel processor with 2 GHz frequency and 4 GB RAM using MATLAB 17-b. The average time of feature extraction, significant region detection, watermark inserting, and retrieving are established (see Table 2).

The computation time of each process is competitive, and it demonstrates the effectiveness of our watermarking algorithm to fulfill the requirements of MP3 audio authentication across wireless networks.

3.4.1. Watermarking Method Performance

(i) Inaudibility tests: Transparency performance ensures that the watermarking scheme does not degrade the host Bitstream significantly. Otherwise, the watermark embedding process did not introduce a distinguishable noise in the host carrier. The objective difference grade (ODG) measure is used [32]. ODG can take a value between −4 and 0. The closer the value of ODG to 0, the more degradation is imperceptible. The results for some MP3 digital audio are presented in Figure 8. The achieved ODG values show that the watermark transparency is confirmed by ODG values around −1

(ii) Robustness: Robustness measurement determines the persistence of the hidden signature. The normalized correlation (NC) is used as an evaluation metric. NC is used to calculate the correlation between the hidden mark and the retrieved bits as expressed in

$$NC = \frac{\sum_{i,j=1}^{n} bin(i,j) * bin'(i,j)}{\sqrt{\sum_{i,j=1}^{n} bin'(i,j)^2 * \sum_{i,j=1}^{n} bin(i,j)^2}}, \quad (1)$$

$bin(i,j)$ and $bin'(i,j)$ are the hidden and the obtained watermarks, respectively. Hidden and retrieved watermarks are considered equivalent if $NC \geq 0.9$. In the case of an ideal interchange with

Table 2: Time computation.

	Feature extraction/frame	Big region detection/frame	Watermark embedding	Detection process
Average time (s)	**0.2**	**0.3**	**4**	**3**

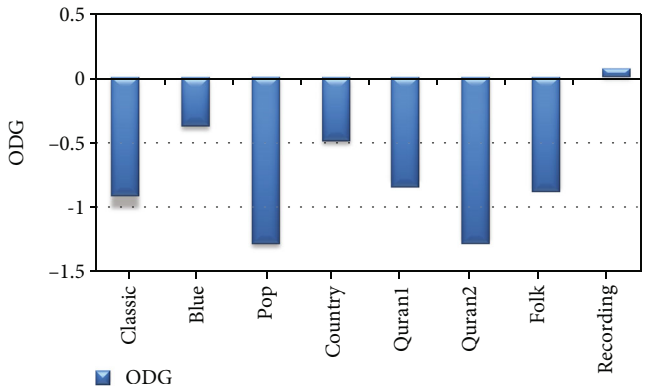


Figure 8: Transparency results (ODG) of the suggested scheme.

no attacks, our watermarking algorithm assures error-free detection (NC = 1) . The attacks of the StirMark benchmark are used to check the robustness of our algorithm. To guarantee the robustness of the watermarking algorithm, it should survive different attacks and signal distortions. To check the robustness against audio degradation and manipulation, we calculate the NC values of the hidden and retrieved watermarks.

(iii) Robustness against MP3doubly compression: The upper part of Figure 9 shows that the proposed technique gives good results of robustness against MP3 doubly compression with different rates. The most values of NC are greater than 0.8

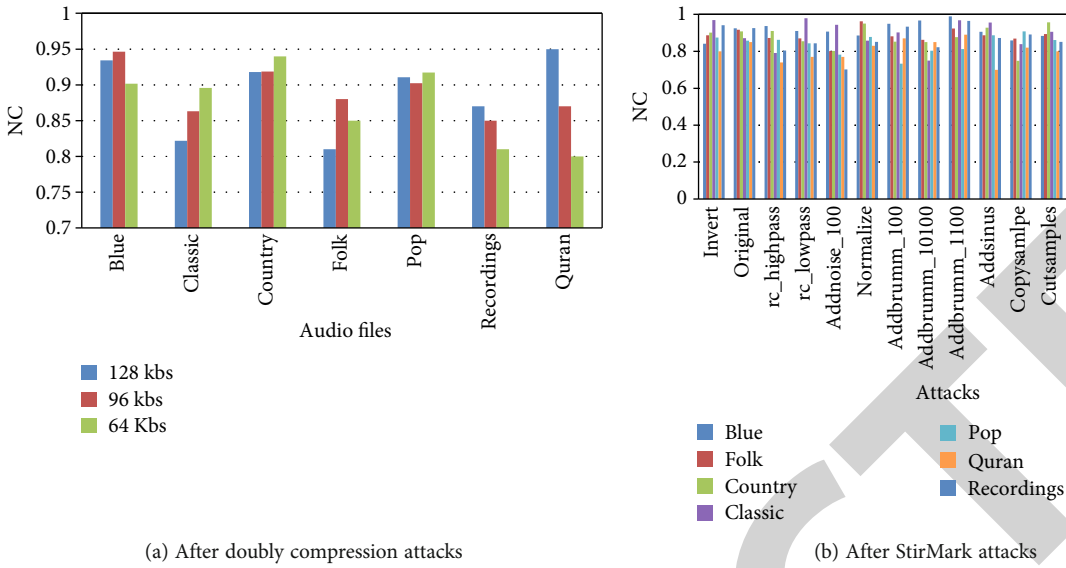
(iv) Robustness against StirMark attacks. Usually, applying attacks to the watermarked audio signal is done in the decompression domain. Therefore, the treatment of the MP3 watermarked audio signal requires the subsequent steps. First, the watermarked MP3 audio signal will be decompressed to have the possibility to load it by the audio editing software. Then, some attacks from the StirMark benchmark are applied, such as additive noise (fftnoise, dynnoise, addsinus, addbrummm, and echo), filtering (highpass and lowpass), and content transformation (copying, slicing, and flipping samples) [33]. Lastly, the audio signal attacked is recompressed and reconstituted to obtain a new MP3 bitstream. The lower part of Figure 9 displays the NC values of the hidden and retrieved marks of decompressed and attacked watermarked bitstream. Although the test signal (MP3 watermarked and attacked audio) is doubly attacked (decompression+StirMark attack NC values are close to 1 in most cases), the results confirm the robustness of the suggested scheme face to different manipulation

(v) Comparison with previous works: We made a comparative study between our proposed approach and the MP3 audio watermarking based works cited in [6, 18, 21, 26]. The paper [6] provides the evaluation values of the transparency criterion, payload, and the robustness against disturbing attack only. Therefore, we compare the performance of our proposed method with those of [18, 21, 26]. In [26], it presents our previous scheme results against inaudibility, robustness, and embedding capacity. As the works suggested in [6, 26] operate directly on MP3 bitstream, we compare their payload-based performance with our proposed work in which our current scheme inserts 128 bits due to use of MD5 algorithm checksum. In contrast, the schemes in [6, 26] can embed one bit and 0.499 bits per frame, respectively. The proposal of [18] is one of the recent watermarking works that uses embedding throughout the MP3 encoding process, and it provides high transparency with good robustness results. Moreover, the paper [21] uses a partial MP3 decoder for MPEG layer III watermarking. This method gives also reasonable robustness in terms of inaudibility performances. The metrics used in [18, 21, 26] are ODG and normalized correlation values to measure the transparency and the persistence of the hidden signature, respectively. Figure 10 illustrates the results of inaudibility of our proposed algorithm compared to [26].

We show that the new scheme has improved the inaudibility. The comparison of the persistence of the hidden signature between our current proposed method and those of [18, 26] is shown in Figure 11.

It is clear from the obtained results that our proposed approach achieves best results in terms of robustness with respect to many attacks, essentially recompression attacks. The normalized correlation value of our suggested scheme varies from [0.8 0.9], while the NC value of the work in [18] differs from [0.05 0.29].

Furthermore, compared with our previous work of [26], the new proposed scheme enhances the NC values against some important attacks, such as recompression, invert, normalize, and filters. Moreover, we notice that our new approach provides good results compared to the scheme announced in [18] when the watermarked audio undergoes specific attacks such as resampling, lowpass filter, and inserting an echo. However, both algorithms show comparative values once the watermarked audio signal is attacked by a highpass filter. Besides, our technique guarantees an average normalized correlation value equal to 0.88, better than the average NC value provided by the technique in [18] which is equal to 0.852.



(a) After doubly compression attacks

(b) After StirMark attacks

Figure 9: NC values corresponding to the hidden watermark and retrieved after MP3 doubly compression and StirMark attacks.

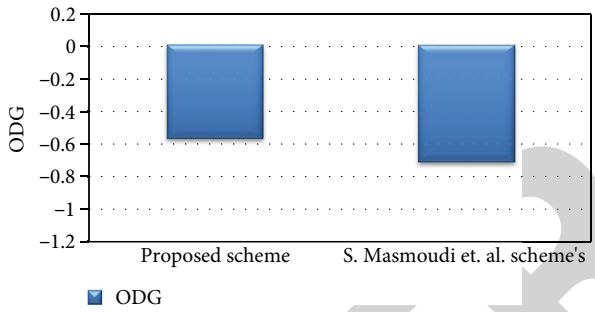


Figure 10: ODG values for the proposed scheme and the scheme described in [26].

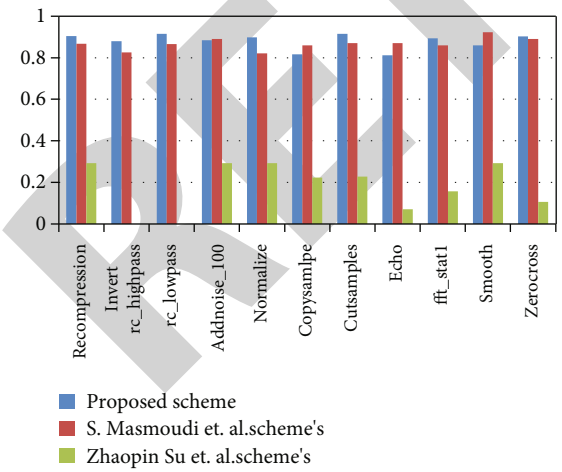


Figure 11: correlation values between the embedded and retrieved watermarks of the attacked watermarked audio for the suggested algorithm and the algorithms used in [18, 26].

3.4.2. Integrity Control. This section describes and evaluates the MP3 audio file content integrity checking. The used metric is the BER, which is the ratio exposing the number of error bits over the received total bits. It compares the value of hidden watermark bits W(i,j), and the retrieved watermark W'(i,j), as follows:

$$BER = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{w(i,j)\oplus w'(i,j)}{n*m}. \quad (2)$$

The sizes of W and W' are n and m, respectively, and the $\oplus$ denotes the xor operation. To quantify the watermarked audio content integrity changes against StirMark attack, we compare the hidden watermark (original content side information features) and the recalculated content side information features of the attacked watermarked MP3 audio file. If no attack occurs, the bit rate error (BER) equals zero. The evaluation test procedure is described as follows:

(1) Input the MP3 audio bitstream

(2) Select the important MP3 side information features characterizing the cover bitstream

(3) Extract features

(4) Create the watermark by applying MD5 to generate a feature checksum

(5) Apply watermarking algorithm

(6) Attack the MP3 watermarked bitstream (decompression + attack + recompression)

(7) Obtain the hidden watermark (W) by applying watermark detection algorithm

(8) Generate the attacked watermarked Bitstream features and calculate their checksums (W')

(9) Check the integrity by comparing W and W' to decide if the content changes or not

Table 3: Integrity verification against MP3 doubly compression using feature checksum.

MP3 file		Classic	Country	Blue	Folk	Recordings	Quran
128-128	Ber	0.18	0.28	0.31	0.24	0.34	0.38
	Integrity	Ok	Ok	Ok	Ok	Ok	Ok
128-96	Ber	0.33	0.33	0.34	0.31	0.35	0.39
	Integrity	No	No	No	No	No	No
128-64	Ber	0.32	0.35	—	—	—	—
	Integrity	No	No	—	—	—	—
128-256	Ber	0.35	0.38	0.41	0.48	0.4	0.36
	Integrity	No	No	No	No	No	Ok
128-320	Ber	0.35	0.38	0.42	0.48	0.39	0.34
	Integrity	No	No	No	No	No	Ok

Table 4: Integrity Verification against StirMark attacks using feature checksum.

MP3 file		Classic.mp3	Country.mp3	Blue.mp3	Folk.mp3	Recordings.mp3	Quran.mp3
Nothing	Ber	0.2	0.28	0.31	0.28	0.34	0.38
	Integrity	Ok	Ok	Ok	Ok	Ok	Ok
Echo	Ber	0.19	0.22	0.31	0.27	0.34	0.37
	Integrity	Ok	Ok	Ok	Ok	Ok	Ok
Addnoise	Ber	0.27	0.29	0.32	0.31	0.34	0.39
	Integrity	No	No	No	No	Ok	No
Addbrumm 100	Ber	0.21	0.29	0.31	0.27	0.34	0.38
	Integrity	No	No	Ok	Ok	Ok	Ok
Addbrumm 10100	Ber	0.25	0.3	0.32	0.27	0.35	0.386
	Integrity	No	No	No	Ok	No	No
Cut samples	Ber	0.35	0.37	0.34	0.34	0.35	0.34
	Integrity	No	No	No	No	No	Ok
Copy sample	Ber	0.33	0.33	0.35	0.33	0.35	0.385
	Integrity	No	No	No	No	No	No
Highpass filter	Ber	0.21	0.27	0.31	0.27	0.345	0.38
	Integrity	No	Ok	Ok	Ok	No	Ok
Lowpass filter	Ber	0.25	0.3	0.31	0.28	0.34	0.39
	Integrity	No	No	Ok	Ok	Ok	No
Invert	Ber	0.2	0.28	0.3	0.27	0.34	0.38
	Integrity	Ok	Ok	Ok	Ok	Ok	Ok
Normalize	Ber	0.2	0.28	0.31	0.27	0.34	0.38
	Integrity	Ok	Ok	Ok	Ok	Ok	Ok
Voice remove	Ber	0.27	0.29	0.31	0.28	0.35	0.37
	Integrity	No	No	Ok	Ok	No	Ok

Table 3 and Table 4 show the experiments results after applying a StirMark benchmark audio attack [33], and MP3 doubly compression for the MP3 watermarked audio. The hidden feature checksum is detected and compared to the recalculated feature checksum vectors. The attacks preserving the audio content such as "normalize," "invert," and "amplify" give same error rates as in the "nothing" attack applied in StirMark benchmark or after an ideal exchange (BER = 1). An error rate equal or less than those obtained after no attacks can be considered a threshold to discriminate content-conserving attacks and content-changing attacks. Content manipulations like inserting noise (addnoise) or humming (addbrumm), voice removal, sample removal, and copying have higher error rates than the case of absence of attack. The MP3 doubly compression has an error rate equal or less than the threshold when it occurs at the same bitrate of the host, but it will be considered a content manipulation attack when the MP3 bitrate is increased or decreased. The results show that some attacks

like filters (lowpass filter and highpass filter) and voice removal may be considered to check the content preserving, in some cases as audio recording context. The results show also that the bit error rate is related to the strength degree of the attack. In other words, low values of noise lead to low error rates.

4. Conclusion

The purpose of this paper was to propose a new blind IoT based MP3 audio watermarking scheme operated in the compressed domain. We have presented a literature review discussing the audio watermarking techniques for MPEG encoded files. We classified them into three different approaches and compared them based on their robustness, inaudibility, insertion ratio, and complexity. Since the audio watermarking in the compressed domain is not well addressed in the literature, some of the existing algorithms are used for online transmission and authentication applications. The embedding capacity ratio in these cases is small and depends largely the used audio stream. Consequently, we proposed a new MP3 watermarking approach using the calibration of the recompression process based on the MDCT distribution to guarantee a maximum robustness against decompression and recompression attacks. In addition, the used watermark is constructed from a set of side information features that permits, first, to detect the content manipulation attacks and, second, to be robust for preserving content attacks. Furthermore, this scheme is blind since we do not need the original compressed file for both watermark detection and control integrity processes. Other advantage of the proposed solution lies in its speed detection process of the watermark and the integrity control of the MP3 file. This makes the scheme suitable for authentication application across wireless network.

The proposed scheme is tested for copyright protection and authentication applications. In future works, we plan to use this scheme to detect forensics in the compressed video files to look for fake videos, mainly in COVID 19 crisis.

Data Availability

No data were used to support this study.

Ethical Approval

Hereby, the authors consciously assure that for the manuscript/insert title, the following is fulfilled: (a) this material is the authors' original work, which has not been previously published elsewhere. (b) The paper is not currently being considered for publication elsewhere. (c) The paper reflects the authors' research and analysis truthfully and completely. (d) The paper properly credits the meaningful contributions of coauthors and coresearchers. (e) The results are appropriately placed in the context of prior and existing research. (f) All sources used are properly disclosed (correct citation). Copying of text must be indicated as such by using quotation marks and giving proper reference. (g) All authors have been personally and actively involved in the substantial work leading to the paper and will take public responsibility for its content.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by Taif University Researchers Supporting Project number (TURSP-2020/348), Taif University, Taif, Saudi Arabia.

References

[1] A. Karaagac, E. Dalipi, P. Crombez, E. D. Poorter, and J. Hoebeke, "Light-weight streaming protocol for the internet of multimedia things: voice streaming over NB-IoT," *Pervasive and Mobile Computing*, vol. 59, article 101044, 2019.

[2] I. Cox, M. Miller, and I. Bloom, *Digital Watermarking*, Academic Press, USA, 2002.

[3] D. Koukopoulas and Y. Stamatiou, "A watermarking scheme for MP3 audio files," *International Journal of Computer and Information Engineering*, vol. 2, no. 8, pp. 2831–2838, 2006.

[4] F. Chaabane, M. Charfeddine, and C. Ben Amar, "A QR-code based audio watermarking technique for tracing traitors," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, pp. 51–55, Nice, France, August 2015.

[5] F. Chaabane, M. Charfeddine, and C. Ben Amar, "A survey on digital tracing traitors schemes," in *2013 9th International Conference on Information Assurance and Security (IAS)*, pp. 85–90, Gammarth-Tunisia, December 2013.

[6] B. Yang, P. Wu, Y. Jing, and J. Mao, "Lossless and secure watermarking scheme in MP3 audio by modifying redundant bit in the frames," in *2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering*, pp. 154–157, Xi'an, China, November 2013.

[7] M. El'Arbi, M. Charfeddine, S. Masmoudi, M. Koubaa, and C. Ben Amar, "Video watermarking algorithm with BCH error correcting codes hidden in audio channel," in *IEEE symposium series in computational intelligence*, pp. 164–170, Paris-France, 2011.

[8] S. Masmoudi, M. Charfeddine, and C. Ben Amar, "A robust audio watermarking technique based on the perceptual evaluation of audio quality algorithm in the multiresolution domain," in *The 10th IEEE International Symposium on Signal Processing and Information Technology*, pp. 326–331, Luxor-Egypt, Decmeber 2011.

[9] N. B. Puhan and A. T. Ho, "Secure authentication watermarking for localization against the Holliman Memon attack," *Multimedia Systems*, vol. 12, no. 6, pp. 521–532, 2007.

[10] J. Lee, F. De Simone, and T. Ebrahimi, "Influence of audio-visual attention on perceived quality of standard definition multimedia content," in *2009 International Workshop on Quality of Multimedia Experience*, pp. 13–18, San Diego, California, U.S.A, July 2009.

[11] N. Cvejic and T. Seppanen, *Digital Audio Watermarking Techniques and Technologies Applications and Benchmarks: Applications and Benchmarks*, IGI Global, Pennsylvania, United States, 2008.

[12] R. Wazirali, R. Ahmad, A. Al-Amayreh, M. Al-Madi, and A. Khalifeh, "Secure watermarking schemes and their approaches in the IoT technology: an overview," *Electronics*, vol. 10, no. 14, article 1744, 2021.

[13] B. Chen, J. Zhao, and D. Wang, "An adaptive watermarking algorithm for MP3 compressed audio signals," in *2008 IEEE Instrumentation and Measurement Technology Conference*, pp. 1057–1060, Victoria, BC, Canada, May 2008.

[14] W. Ching-Te, C. Tung-Shou, and C. Wen-Hung, "A new audio watermarking based on modified discrete cosine transform of MPEG/audio layer III," *Proc IEEE International Conference on Networking, Sensing and control*, vol. 2, pp. 984–989, 2004.

[15] H. Ruan and E. Yaz, "A novel chaotic watermarking scheme for MP3 audio signals," in *2004 IEEE Electro/Information Technology Conference*, pp. 62–65, Milwaukee, WI, USA, August 2004.

[16] N. Moghadam and H. Sadeghi, "Genetic content based MP3 audio watermarking in MDCT domain," *Watermark*, vol. 1, no. 2, article 3, 2005.

[17] B. Chen, J. Zhao, and D. Wang, "An adaptive and secure watermarking algorithm robust to MP3 compression," in *2009 IEEE Instrumentation and Measurement Technology Conference*, pp. 5–7, Singapore, May 2009.

[18] S. Zhaopin, C. Lejie, Z. Guofu, J. Jianguo, and Y. Feng, "Window switching strategy based semi-fragile watermarking for MP3 tamper detection," *Multimedia Tools and Applications*, vol. 76, no. 7, pp. 9363–9386, 2017.

[19] L. I. Chen, W. A. N. G. Kexin, and T. I. A. N. Lihua, "Audio watermarking algorithm in MP3 compressed domain based on low frequency energy ratio of channels," *Journal of Computer Applications*, vol. 38, no. 8, article 2301, 2018.

[20] B. Lei and I. Yann Soon, "Perception-based audio watermarking scheme in the compressed bitstream," *AEU International Journal of Electronics and Communications*, vol. 69, no. 1, pp. 188–197, 2015.

[21] A. V. Subramanyam and S. Emmanuel, "Audio watermarking in partially compressed- encrypted domain," in *2012 IEEE international conference on systems, man, and cybernetics (SMC)*, pp. 14–17, Seoul, Korea, October 2012.

[22] W. Wenhui, G. Xuan, X. Zhiting, and W. Renyi, "MP3 audio watermarking algorithm based on unipolar quantization," in *International Conference on Emerging Internetworking, Data & Web Technologies*, vol. 6, pp. 448–456, Springer, Cham, 2018.

[23] C. Ting-Ting, W. Xin-Fang, and J. Cun-Yun, "MP3 audio digital watermark algorithm based on compressed domain," *Computer Engineering*, vol. 37, no. 10, pp. 204–206, 2011.

[24] L. Qiao and K. Nahrstedt, "Non-invertible watermarking methods for MPEG encoded audio," *Security and Watermarking of Multimedia Contents*, vol. 3675, pp. 194–202, 1998.

[25] K. Takagi, S. Sakazawa, and Y. Takishima, "Light weight MP3 watermarking method for mobile terminals," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 91, no. 9, pp. 2546–2554, 2008.

[26] S. Masmoudi, M. Charfeddine, and C. Ben Amar, "A semi-fragile digital audio watermarking scheme for MP3-encoded signals using Huffman data," *Circuits Systems and Signal Processing*, vol. 36, no. 6, pp. 1–16, 2020.

[27] M. Steinebach and J. Dittmann, "Watermarking-based digital audio data authentication," *EURASIP Journal on Applied Signal Processing*, vol. 3, no. 10, pp. 1001–1015, 2003.

[28] E. Mezghani, M. Charfeddine, and C. Ben Amar, "Audio silence deletion before and after MPEG video compression," in *2013 International Conference on Computer Applications Technology (ICCAT)*, pp. 1625–1629, Sousse, Tunisia, January 2013.

[29] D. Pan, "A tutorial on Mpeg/audio compression," *IEEE Multimedia*, vol. 2, no. 2, pp. 60–74, 1995.

[30] D. Simitopoulos, N. Zissis, P. Georgiadis, V. Emmanouilidis, and M. G. Strintzis, "Encryption and watermarking for the secure distribution of copyrighted MPEG video on DVD," *Multimedia Systems*, vol. 9, no. 3, pp. 217–227, 2003.

[31] A. Servetti, C. Testa, and J. C. De Martin, "Frequency-selective partial encryption of compressed audio," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, Hong Kong, China, April 2003.

[32] International Union of Télécommunications (UIT), "Recommandation UIT-R B.S. 1387, Method of Objective Measurement of Perceived Sound Quality," 2001.

[33] A. Lang, J. Dittmann, R. Spring, and C. Vielhauer, "Audio watermark attacks: from single to profile attacks," in *Proceedings of the 7th Workshop on Multimedia and Security*, pp. 39–50, New York, NY, USA, August 2005.