High Embedding Capacity and Robust Audio Watermarking for Secure Transmission Using Tamper Detection

Arashdeep Kaur and Malay Kishore Dutta

Robustness, payload, and imperceptibility of audio watermarking algorithms are contradictory design issues with high-level security of the watermark. In this study, the major issue in achieving high payload along with adequate robustness against challenging signal-processing attacks is addressed. Moreover, a security code has been strategically used for secure transmission of data, providing tamper detection at the receiver end. The high watermark payload in this work has been achieved by using the complementary features of third-level detailed coefficients of discrete wavelet transform where the human auditory system is not sensitive to alterations in the audio signal. To counter the watermark loss under challenging attacks at high payload, Daubechies wavelets that have an orthogonal property and provide smoother frequencies have been used, which can protect the data from loss under signal-processing attacks. Experimental results indicate that the proposed algorithm has demonstrated adequate robustness against signal processing attacks at 4,884.1 bps. Among the evaluators, 87% have rated the proposed algorithm to be remarkable in terms of transparency.

Keywords: Digital audio watermarking, Payload, Tamper detection, Wavelet decomposition.

I. Introduction

The rapid advancement in the field of information technology in the last few years permits convenient access to digital data such as images, audio, and video. It is highly straightforward to distribute digital data over a wide network in a few seconds, without degrading the quality of the digital data. Whereas this rapid transmission has provided the user with convenient access to digital data, it has also resulted in numerous severe problems such as copyright infringement and piracy. In order to ensure secure transmission of digital data, the concept of digital watermarking has been introduced [1]. Digital watermarking can be defined as a method of embedding digital data in a host signal at the source, and efficiently extracting the concealed data at the destination. The digital signal to be transferred is known as the host signal, the digital data to be embedded in the host signal is known as the watermark, and the signal containing the watermark is known as the watermarked signal. Digital watermarking can be classified into various categories. According to the host signal used for embedding, digital watermarking can be classified into image watermarking, audio watermarking, and video watermarking. According to the watermark extraction strategy, digital watermarking can be classified into three categories: blind watermarking, semi-blind watermarking, and non-blind watermarking. In this paper, a blind audio watermarking algorithm has been presented, thus balancing the three contradictory requirements of digital watermarking.

Manuscript received Jan. 31, 2017; revised Aug. 17, 2017; accepted Oct. 30, 2017.

Arashdeep Kaur (corresponding author, arashgulati@gmail.com) is with the Department of Computer Science & Engineering and Malay Kishore Dutta (mkdutta@amity.edu) is with the Department of Electronics and Computer Engineering, Amity School of Engineering and Technology, Uttar Pradesh, India.

This is an Open Access article distributed under the term of Korea Open Government License (KOGL) Type 4: Source Indication + Commercial Use Prohibition + Change Prohibition (http://www.kogl.or.kr/info/licenseTypeEn.do).

An efficient digital audio watermarking algorithm should satisfy the three conflicting requirements: imperceptibility, robustness, and payload capacity. Imperceptibility implies that the watermark should be embedded in the host signal without affecting the quality of the host signal. Robustness refers to the behavior of the watermarking algorithm against malicious intentional or unintentional attacks over the network. The number of watermarking bits that can be reliably embedded in the host signal per unit of time is called the payload of the watermarking algorithm. In addition to these, the watermarking algorithm should be adequately secure such that it should be capable of assessing whether the signal has been tampered with or not.

There are various audio watermarking algorithms in existing literature that exhibit adequate performance in terms of robustness, capacity, and imperceptibility. An audio watermarking algorithm with a payload capacity of 848 bps, using the DCT domain, has been proposed by Hu and others [2]. Their algorithm has exhibited adequate robustness against common signal processing attacks; however, low pass filtering is still a challenge for this algorithm [3]. Hu and others [2] have proposed another audio watermarking algorithm with a payload of 473 bps and adequate robustness against low pass filtering, in addition to common signal processing attacks; however, this algorithm has not exhibited adequate performance against MP3 compression. Fallahpour and others have provided an audio watermarking algorithm with payload rate ranging from 2,000 bps to 6,000 bps. The embedding is achieved by using FFT coefficients. The experimental results indicate that this algorithm exhibits adequate transparency. The robustness against MP3 compression of their algorithm is observed to be highly adequate; however, this algorithm was not evaluated against other attacks [4]. Fallahpour and Megías [5] have presented a FFT-spectrum-based audio watermarking with a payload of 700 bps to 3,000 bps. Their algorithm is demonstrated to be robust against echo and noise addition, filtering, and compression, without significant perceptual distortion. Chen and others [6] have presented an optimization-based audio watermarking algorithm where embedding is achieved using the seventh level of discrete wavelet transform. The perceptual transparency and robustness of their algorithm have been evaluated at 1,000 bps and 2,000 bps. This algorithm is observed to be robust against common signal-processing attacks including compression and time scale modification; however, amplitude scaling, cropping, and jittering continue to be major challenges. Mosleh and others [7] proposed a robust intelligent audio watermarking solution wherein embedding is achieved using SVD, and an intelligent detector is used for watermark extraction. Their method has provided adequate imperceptibility and high robustness, albeit low payload rates. Similarly, the time-spread-echo-hiding-based scheme proposed by Hu and others [8] is not capable of achieving high payload. Another spread-spectrum audio watermarking algorithm, with 43 bps payload and adequate robustness, is provided by Li and others [9]. Moreover, Erfani and others have provided an audio watermarking algorithm, which exhibits adequate robustness against challenging signal processing attacks; however, the average payload achieved ranges from 5 bps to 15 bps [10]. Another audio watermarking scheme, provided by Kaur and others [11], is adaptive in nature, with adequate robustness and transparency; however, the maximum payload achieved as of yet is 768 bps.

Hence, it can be observed that numerous audio watermarking algorithms exist that address the issue of robustness, imperceptibility, or payload and achieve an optimal balance between these. However, the problem that continues to exist is that it is highly challenging to simultaneously achieve high payload and adequate robustness against challenging signal processing attacks such as compression, jittering, and cropping within perceptual constraints.

The main contribution of this study is a secure and high-payload audio watermarking algorithm that exhibits adequate robustness against challenging signal processing attacks. The strategic selection of the wavelet sub-band permits the determination of the highest feasible payload within perceptual constraints. The robustness of the proposed algorithm is achieved by using the Daubechies wavelets in the selected sub-band. The security of the proposed method has been ensured using the tamper detection at the receiver end. The payload of 4,884.1 bps with adequate robustness has been strategically achieved using the detailed coefficients of the wavelet decomposition. The proposed mathematical model for the embedding of watermark introduces negligible changes in the host signal in order to render it imperceptible to the human ear. The embedding quantization function is adequate to achieve signal-to-noise ratio (SNR) values higher than 20 dB for all the audio samples. The proposed algorithm is observed to exhibit highly adequate robustness against cropping, jittering, re-quantization, noise, resampling, compression, and low pass filtering at such high payloads.

Further, the extraction process has been designed and implemented such that it is feasible to extract the exact watermark notwithstanding the application of severe and challenging signal processing attacks during transmission. The comparative analysis of the proposed algorithm with the extant high-payload audio watermarking algorithms

https://doi.org/10.4218/etrij.2017-0092

indicate the higher efficiency and performance of the proposed algorithm in terms of the maintenance of an optimal equilibrium between imperceptibility, robustness, and payload.

The rest of the paper is organized as follows: Watermark embedding and extraction is discussed in Section II. Section III presents the experimental results and comparative study. The conclusions are drawn in Section IV.

II. Proposed Methodology

This section discusses the proposed embedding and extraction watermarking algorithm in detail. In this study, the detailed coefficients of wavelet transform are used for embedding the watermark. These low frequency components exhibiting high energy values of wavelet transform are observed to be robust against malicious attacks. The quantization values of delta have been selected such that it inserts minimal distortions in the watermarked signal, resulting in adequate perceptual transparency.

1. Discrete Wavelet Transformation

The discrete wavelet transform is the implementation of wavelet transform adhering to previously defined rules, using a fixed set of wavelet scales. To evaluate the DWT of an audio signal H it is passed through a series of low pass and high pass filters. Initially, the convolution of the audio samples and the low pass filter with impulse response G is determined as expressed in (1):

$$y[n] = (H * G)[n] = \sum_{k=-\infty}^{\infty} H[k]G[n-k]. \quad (1)$$

The high-pass filter L is used simultaneously to decompose the signal and obtain detailed coefficients. The approximation coefficients are obtained as the output of the low-pass filter. The filter outputs are then further sub-sampled as follows:

$$y_{\text{low}}[n] = (H \times G)[n] = \sum_{k=-\infty}^{\infty} x[k]G[2n-k], \quad (2)$$

$$y_{\text{high}}[n] = (H \times G)[n] = \sum_{k=-\infty}^{\infty} x[k]L[2n-k]. \quad (3)$$

Using the sub-sampling operator, $\downarrow(1)$ can be expressed as:

$$(y \downarrow k)[n] = y[kn]. \quad (4)$$

Equations (2) and (3) can also be expressed as

$$y_{\text{low}} = (x \times G) \downarrow 2,$$

$$y_{\text{high}} = (x \times H) \downarrow 2.$$

This decomposition process is repeated until a certain defined level in order to generate wavelet scales [12], [13].

2. Watermark Embedding

In the proposed algorithm, the watermark is embedded using discrete wavelet transform. The detailed procedure of watermark embedding is provided in Algorithm 1.

Algorithm 1.

Step 1: Read host audio signal H and the watermark w.

Step 2: Evaluate *size*(H) and *size*(w).

Step 3: *Encrypted watermark* = *Arnold*(w). Convert the encrypted watermark to 1-D vector

Step 4: Divide H into non-overlapping frames with number of frames equal to *size*(w).

Step 5: Read individual frame from H of size f.

Step 6: DWT of each frame
[C A] = *DWT* (*frame*,'*db1*')

Step 7: *index* = A(1); *Det_coef* = C((*index* + 1):*index* + A(2));
M_Det_coef = *max*(*Det_coef*):

Step 8: Update the maximum value in the third level detailed coefficient of DWT.
If *encrypted watermark*(i) == 1
M_Det_coef = *M_Det_coef-mod*(*M_Det_coef*,μ) + *del*1
else
M_Det_coef = *M_Det_coef-mod*(*M_Det_coef*,μ) + *del*2
end

Step 9: Take inverse DWT with the modified detailed coefficient.

Step 10: Merge all the frames to obtain the watermarked signal.

The host audio signal, H $\{h_k | 1 \le k \le m\}$, of length m is segmented into various frames of equal length, say, f. The number of frames of an audio signal is equal to the number of bits to be watermarked. The image watermark, w $\{w_{i,j} | 1 \le i, j \le n\}$, is a two-dimensional binary sequence of length n. The image watermark is scrambled using the Arnold cat map encryption method, as expressed in (5), in the proposed algorithm to ensure the security of the watermark.

$$\begin{bmatrix} a_{n+1} \\ b_{n+1} \end{bmatrix} = \text{mod}\left(\begin{bmatrix} u & v \\ w & x \end{bmatrix} \begin{bmatrix} a_n \\ b_n \end{bmatrix}, K \right), \quad (5)$$

where u, v, w, and x are positive integers such that $ux - vw = +1$, and a_n, b_n, a_{n+1}, and b_{n+1} are integers in $\{0, 1, 2, \ldots, K-1\}$. Then, the watermarked signal is generated by merging all the updated frames and subsequently transferred over the network. Figure 1 illustrates the watermark embedding procedure in detail.

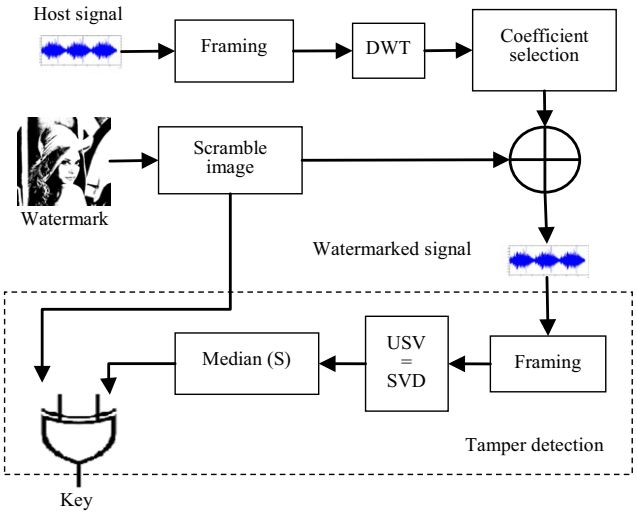


Fig. 1. Watermark embedding procedure.

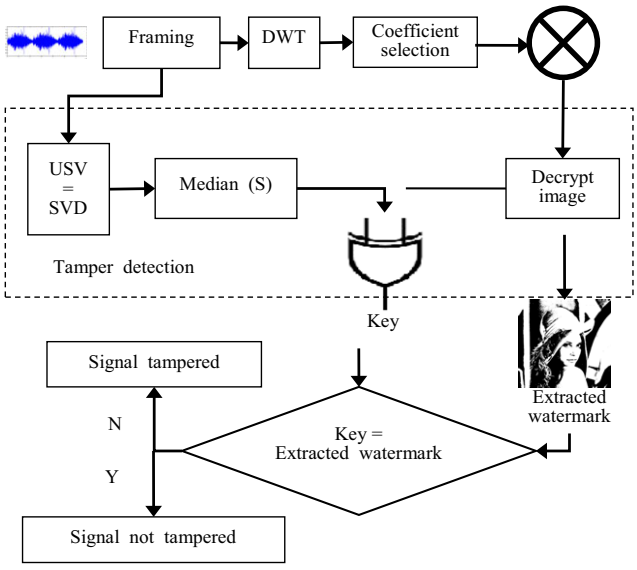


Fig. 2. Watermark extraction procedure.

3. Watermark Extraction

Watermark extraction in the proposed audio watermarking algorithm is completely blind. There is no requirement of the watermark or the original audio signal to extract the watermark. In addition to blind watermark extraction, tamper detection is also performed in this study. Figure 2 presents the basic overview of the watermark extraction procedure used in this study. The method followed to examine whether that signal has been disturbed on the network during transfer is discussed in the following section. The watermark extraction is described in detail in Algorithm 2.

Algorithm 2.

Step 1: Read the watermarked audio signal.

Step 2: Follow Step 4 to Step 7 of Algorithm 1 to obtain the maximum value of the detailed coefficient of wavelet transform in a frame.

Step 3: If $mod(M\_Det\_coef,\mu) > divide(\mu,2)$
$Extracted\ bit == 1$,
Else
$Extracted\ bit == 0$
end.

Step 4: Merge all the extracted watermarking bits and resize back to the original 2-D size to obtain M.

Step 5: $Extracted\ watermark = Decrypt(M)$ by using Arnold cat map.

4. Tamper Detection

To ensure the security of the watermarked signal, we have used tamper detection with watermarking in this study.

Algorithm 3.

Step 1: Read individual frame from H of size f and convert into the square matrix.

Step 2: SVD of each frame
$[U\ S\ V] = SVD\ (frame)$

Step 3: Singular values $S = \{s_1, s_2, s_3, \ldots, s_f\}$

Step 4: Calculate Median (M) of Singular Values
$Med = median\ \{s_1, s_2, s_3, \ldots, s_f\}$

Step 5: Round off the values of Med for each frame to generate the cover key for the host signal.

Step 6: Obtain the encrypted watermark and covert it to 1-D matrix.

Step 7: Secret key (S_k) = XOR (*cover key*, *encrypted watermark*) where XOR denotes logical XOR operation.

The detailed algorithm to generate the secret key for the tamper detection performed at the embedding stage, is presented in Algorithm 3. Each audio frame is selected individually and converted into a square matrix. Singular value decomposition (SVD) is then applied to the square matrixes of the individual frames. The median of the singular values is evaluated for the individual frames. These values are then rounded off, and logical XOR operation is performed on the encrypted watermark and the rounded median of the singular values for generating the secret key.

The secret key S_k is transferred over the network along with the watermarked signal. This secret key is not required for extracting the watermark. The watermark extraction is completely blind in nature in this study. This secret key is used at the receiver end to examine whether the signal has been tampered with or not. The detailed

https://doi.org/10.4218/etrij.2017-0092

algorithm followed for the tamper detection during extraction is presented in Algorithm 4.

Algorithm 4.

Step 1: Read the received watermarked signal.

Step 2: Create the cover key for the received audio signal by following Step 1 to Step 5 of Algorithm 1.

Step 3: If Secret Share access is authorized, then
Encrypted key = XOR (S_k, *cover key*)
where XOR denotes logical XOR operation.

Step 4: *Recovered key* = *decrypt* (*Encrypted Key*) by using Arnold cat map.

Step 5: D = *diff*(*recovered key*, *extracted watermark*)

Step 6: If $D == 0$, the signal is not tampered.

Else if $D >= 0.9$ && $D < 0$, signal is tampered albeit with minimal distortions.

Else signal is tampered (requires attention).

If there is any difference in the extracted watermark and the recovered key, it implies that the signal has been tampered with during transfer. If there is no change, it implies that the signal is not tampered or that very minimal distortions have been performed, which can be omitted.

III. Experimental Results

This section presents the results of the tests performed to evaluate the performance of the proposed algorithm. In this study, a set of 140 audio signals of various genres including country music, jazz, folk, pop, rock, metal, and classical were used for testing the efficiency of the proposed algorithm. Multiple audio signals were used for each genre, in order to produce accurate results. The audio signals used in this study are mono audio signals, which are sampled at 44.1 kHz; each sample is quantized using 16 bits.

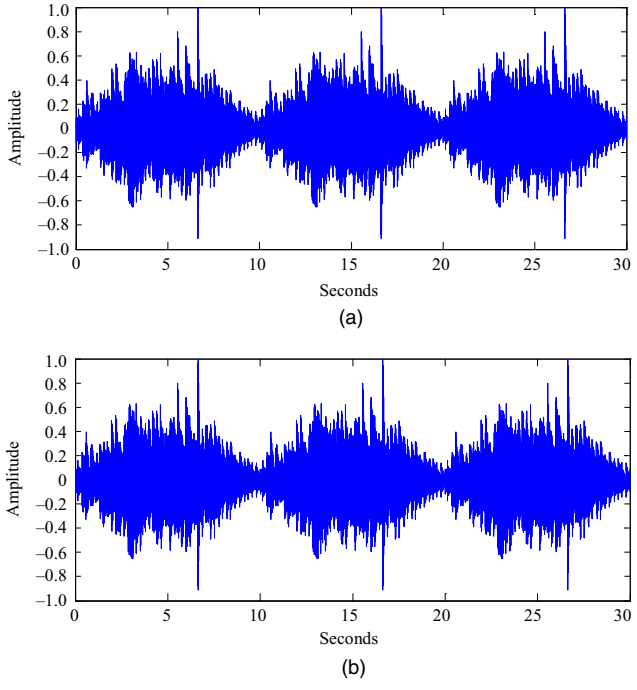


Fig. 3. (a) Host signal and (b) watermarked signal.

(a) (b)

Fig. 4. (a) Embedded watermark and (b) extracted watermark.

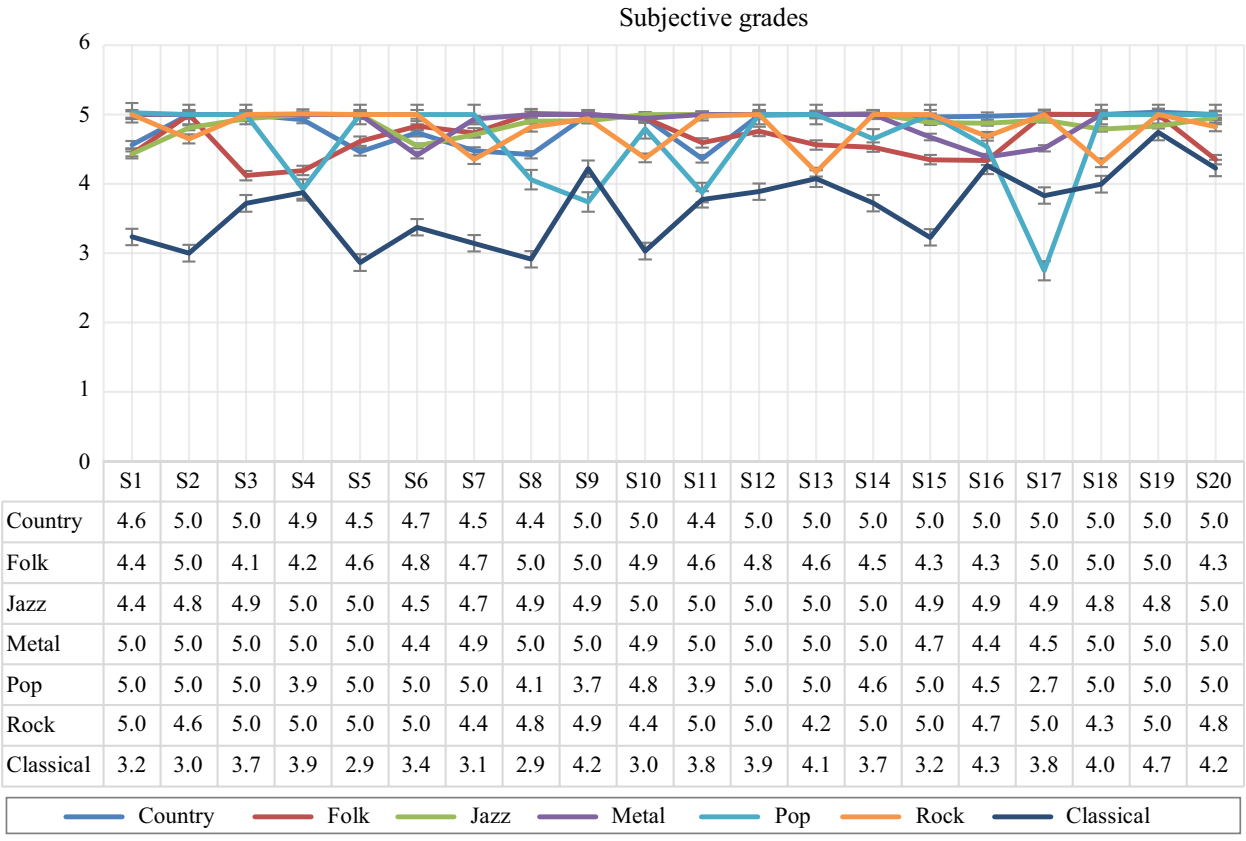
Table 1. Subjective scores.

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

Table 2. Attacks description.

Attack	Description	Labels
No attack	N/A	A1
AWGN	SNR, 30 dB	A2
Resampling	44.1 - > 22.05 - > 44.1	A3
Re-quantization	16 - > 8 - > 16	A4
LPF	2nd order Butterworth, 11 kHz	A5
MP3 compression	64 kbps	A6
Cropping1	25%	A7
Cropping2	18%	A8
Cropping3	10%	A9
Jittering1	1/2,000	A10
Jittering2	1/1,000	A11
Jittering3	1/500	A12

LPF: Low pass compression.



	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Country	4.6	5.0	5.0	4.9	4.5	4.7	4.5	4.4	5.0	5.0	4.4	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Folk	4.4	5.0	4.1	4.2	4.6	4.8	4.7	5.0	5.0	4.9	4.6	4.8	4.6	4.5	4.3	4.3	5.0	5.0	5.0	4.3
Jazz	4.4	4.8	4.9	5.0	5.0	4.5	4.7	4.9	4.9	5.0	5.0	5.0	5.0	5.0	4.9	4.9	4.9	4.8	4.8	5.0
Metal	5.0	5.0	5.0	5.0	5.0	4.4	4.9	5.0	5.0	4.9	5.0	5.0	5.0	5.0	4.7	4.4	4.5	5.0	5.0	5.0
Pop	5.0	5.0	5.0	3.9	5.0	5.0	5.0	4.1	3.7	4.8	3.9	5.0	5.0	4.6	5.0	4.5	2.7	5.0	5.0	5.0
Rock	5.0	4.6	5.0	5.0	5.0	5.0	4.4	4.8	4.9	4.4	5.0	5.0	4.2	5.0	5.0	4.7	5.0	4.3	5.0	4.8
Classical	3.2	3.0	3.7	3.9	2.9	3.4	3.1	2.9	4.2	3.0	3.8	3.9	4.1	3.7	3.2	4.3	3.8	4.0	4.7	4.2

Fig. 5. Imperceptibility test in terms of subjective grades.

Table 3. Imperceptibility results for audio samples of different genres (SNR in dB).

Sample	Country	Folk	Jazz	Metal	Pop	Rock	Classical
S1	31.953	31.032	31.011	36.768	35.167	38.418	22.646
S2	39.030	35.706	33.661	37.694	38.455	32.539	21.001
S3	35.549	28.836	34.571	39.560	37.701	40.702	26.032
S4	34.480	29.346	37.207	40.089	27.468	35.083	27.135
S5	31.238	32.299	36.598	37.763	38.373	39.404	20.057
S6	33.155	33.849	31.833	30.895	37.614	35.007	23.614
S7	31.324	33.150	32.968	34.523	35.970	30.456	21.997
S8	30.947	35.090	34.335	35.526	28.421	33.726	20.389
S9	38.194	37.424	34.379	39.116	26.172	34.561	29.535
S10	34.662	34.642	35.741	34.610	33.567	30.606	21.215
S11	30.527	32.143	35.421	39.547	27.129	34.859	26.430
S12	34.942	33.297	35.876	37.494	37.542	39.322	27.222
S13	38.499	31.928	37.219	38.990	39.074	29.188	28.529
S14	38.302	31.705	35.100	37.797	32.533	38.717	26.050
S15	34.740	30.434	34.199	32.721	36.909	37.486	22.599
S16	34.838	30.368	34.113	30.693	31.711	32.793	29.817
S17	36.650	35.017	34.419	31.584	19.230	35.562	26.821
S18	36.339	35.776	33.514	38.870	36.590	30.118	27.965
S19	35.233	35.444	33.824	39.620	36.762	38.940	33.217
S20	35.667	30.446	34.681	37.337	39.873	33.752	29.599
Average	**34.814**	**32.897**	**34.533**	**36.560**	**33.813**	**35.062**	**25.593**

https://doi.org/10.4218/etrij.2017-0092

Table 4. Robustness against common signal processing attacks in terms of NCC.

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Classical	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.88	0.89	0.89	0.89	0.89	0.88	0.89	0.88	0.89	0.88	0.89	0.88	0.88	0.88	0.88	0.88	0.89	0.89
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.97	0.97	0.97	0.98	0.96	0.91	0.97	0.91	0.98	0.98	0.98	0.98	0.96	0.98	0.96	0.98	0.98	0.96	0.98	0.98
	A5	1.00	1.00	0.95	1.00	1.00	1.00	0.97	1.00	0.93	1.00	0.98	0.98	1.00	1.00	1.00	1.00	0.94	1.00	0.96	0.92
	A6	0.97	0.96	0.79	0.62	0.99	0.98	0.84	0.98	0.53	0.93	0.70	0.72	0.95	0.71	0.98	0.93	0.41	0.95	0.38	0.28
Country	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.88	0.88	0.88	0.89	0.89	0.88	0.88	0.89	0.88	0.88	0.89	0.88	0.88	0.88	0.88	0.88	0.89	0.88
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.97	0.98	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
	A5	0.88	0.13	0.86	0.60	0.93	0.77	0.88	0.81	0.73	0.53	0.52	0.68	0.63	0.30	0.65	0.61	0.57	0.47	0.89	0.35
	A6	0.53	0.01	0.26	0.18	0.48	0.20	0.48	0.41	0.22	0.11	0.18	0.24	0.10	0.04	0.30	0.21	0.22	0.07	0.41	0.02
Folk	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.88
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.98	0.98	0.98	0.98	0.97	0.98	0.98	0.98	0.97	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
	A5	0.86	0.90	0.99	0.96	0.95	0.97	0.97	0.92	0.63	0.86	0.97	0.70	0.96	0.76	0.95	0.97	0.57	0.75	0.90	0.91
	A6	0.54	0.39	0.75	0.53	0.71	0.36	0.59	0.56	0.39	0.31	0.61	0.17	0.64	0.25	0.55	0.66	0.19	0.29	0.44	0.42
Jazz	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.89	0.88	0.89	0.89	0.89	0.88	0.89	0.88	0.89	0.88	0.89	0.88	0.89	0.88	0.88	0.88	0.88	0.88
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.98	0.98	0.98	0.98	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
	A5	0.99	0.78	0.75	0.86	0.51	0.97	0.92	0.67	0.97	0.45	0.87	0.88	0.89	0.53	0.71	0.88	0.66	0.92	0.74	0.38
	A6	0.74	0.43	0.41	0.45	0.12	0.69	0.48	0.31	0.50	0.14	0.36	0.44	0.38	0.14	0.29	0.57	0.24	0.50	0.33	0.15
Metal	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.88	0.89	0.89	0.88	0.88	0.89	0.88	0.88	0.88	0.89	0.89	0.88	0.88	0.88	0.89	0.88	0.89	0.88
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.98	0.98	0.97	0.97	0.98	0.98	0.98	0.98	0.97	0.97	0.98	0.98	0.98	0.98	0.97	0.98	0.98	0.97	0.98	0.98
	A5	0.15	0.23	0.08	0.06	0.08	0.51	0.19	0.33	0.10	0.31	0.05	0.27	0.08	0.05	0.32	0.45	0.72	0.13	0.12	0.29
	A6	0.01	0.08	0.01	0.00	0.00	0.04	0.06	0.11	0.01	0.05	0.02	0.03	0.00	0.00	0.04	0.09	0.16	0.02	0.05	0.05

http://onlinelibrary.wiley.com/journal/10.4218/(ISSN)2233-7326

Table 4. Continued

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Pop	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.88	0.89	0.88	0.88	0.89	0.88	0.89	0.88	0.88	0.88	0.89	0.88	0.89	0.88	0.88	0.89	0.88	0.88
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.97	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.98	0.98	0.98
	A5	0.43	0.40	0.46	0.86	0.30	0.28	0.47	0.93	0.96	0.95	0.78	0.58	0.37	0.87	0.82	0.75	0.99	0.50	0.74	0.03
	A6	0.06	0.12	0.13	0.49	0.05	0.05	0.07	0.56	0.72	0.43	0.37	0.21	0.08	0.44	0.41	0.32	0.87	0.12	0.34	0.01
Rock	A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A2	0.88	0.88	0.88	0.88	0.89	0.89	0.89	0.89	0.89	0.88	0.89	0.88	0.89	0.89	0.88	0.88	0.88	0.88	0.88	0.88
	A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A4	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.98	0.97	0.98	0.98	0.98	0.97	0.98
	A5	0.08	0.51	0.67	0.58	0.03	0.56	0.49	0.46	0.54	0.64	0.55	0.09	0.95	0.18	0.13	0.97	0.73	0.80	0.13	0.38
	A6	0.02	0.10	0.41	0.13	0.01	0.13	0.09	0.07	0.08	0.19	0.13	0.01	0.53	0.02	0.00	0.63	0.26	0.27	0.02	0.05

All these samples were subjected to perceptual transparency and robustness analysis at a payload of 4,884.1 bps, in the experimental setup. Moreover, all these samples were used for examining the robustness of the proposed algorithm under various attacks listed in Table 2.

A sample host-audio-signal and the corresponding watermarked audio signal are presented in Figs. 3(a) and (b), respectively, representing the visible similarity between the host and watermarked audio signals. The binary image watermark of dimensions 221 × 221 is used for embedding in this study. A sample pair of embedded and extracted watermarks is presented in Figs. 4(a) and (b), respectively.

Different tests were performed using the watermarked signal and the extracted watermark to examine the effectiveness of the proposed algorithm. The SNR and subjective grades (SGs) are calculated for the watermarked signal to evaluate the audio quality of the watermarked signal. To examine the quality of the extracted watermark at the receiver end, the normalized correlation coefficient (NCC) and bit error rate (BER) are evaluated in this study. The mathematical formula to calculate SNR, NCC, and BER at 4,884.1 bps payload are expressed in (6), (7), and (8), respectively.

$$\mathrm{SNR}(S_o, S_w) = 10 \log_{10} \frac{\sum_{i=1}^{L} S_o^2(i)}{\sum_{i=1}^{L} [S_o(i) - S_w(i)]^2} \mathrm{d}B, \quad (6)$$

where S_o and S_w are the original and watermarked audio signals, respectively, and L is the length of the audio signal.

$$\mathrm{NCC} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} I_o(i,j) I_E(i,j)}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} 3(i,j)} \times \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} I_E^2(i,j)}}, \quad (7)$$

where I_o and I_E denote the original and extracted binary watermark images, respectively.

$$\mathrm{BER} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} I_o(i,j) \oplus I_E(i,j)}{M \times M}, \quad (8)$$

where I_o and I_E denote the original and extracted binary watermark images, respectively, and $\oplus$ is the exclusive OR operator (XOR).

In addition to SNR, SGs have also been evaluated in this study, which are a subjective measure to examine the

https://doi.org/10.4218/etrij.2017-0092

Table 5. Robustness against cropping and jittering attacks in terms of NCC.

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Classical	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.94	0.94	0.93	0.92	0.96	0.96	0.95	0.97	0.92	0.93	0.93	0.92	0.94	0.93	0.97	0.93	0.92	0.94	0.92	0.92
	A11	0.99	0.99	0.99	0.99	1.00	1.00	0.99	1.00	0.99	0.99	0.99	0.99	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Country	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.93	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
	A11	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Folk	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
	A11	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Jazz	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
	A11	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Metal	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.93	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
	A11	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 5. Continued

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Pop	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.92	0.92	0.92	0.93	0.92	0.92	0.92	0.92	0.93	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.93	0.92	0.92	0.92
	A11	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Rock	A7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	A10	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.93	0.92	0.92	0.92	0.92	0.92	0.92	0.92
	A11	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	A12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

imperceptibility of the proposed algorithm. The SGs have been calculated by calculating the average of all the scores given by the listeners. The watermarked audio signals and the host audio signals were made available to the ten listeners, and five of which exhibit a reasonable knowledge of music. The watermarked signals were scored by the 10 listeners based on their quality, as presented in Table 1.

The attacks listed in Table 2 have been applied to the watermarked signals to examine the robustness of the proposed algorithm by evaluating NCC and BER. An algorithm should have an NCC value closer to 1 and a BER closer to 0 to be adequate in terms of robustness [14].

In this study, the values used for experimentation during embedding and extraction of del1 and del2 are 0.16 and 0.8, respectively. The value of μ is 0.24. The frame size of nine samples per frame is used for experimentation, resulting in a payload of 4,884.1 bps within perceptual constraints and robustness.

The imperceptibility results of the proposed algorithm in terms of the SGs and SNR are presented in Fig. 5 and Table 3, respectively. It is observed form the experimental results that the proposed algorithm exhibits adequate perceptual quality as the subjective scores lie closer to or are equal to five and the SNR values are above 20 dB. It is observed from the experimental results that the SNR values are above 20 dB. According to established audio watermarking standards, an audio watermarking algorithm exhibiting SNR equal to or above 20 dB indicates adequate perceptual transparency of the designed algorithm.

Tables 4 and 5 present the robustness results of the proposed algorithm against common and challenging signal processing attacks in terms of the correlation coefficient. The experimental results provided in the tables indicate adequate robustness of the proposed algorithm as the values of NCC for all the signals under the different attacks approach one.

The NCC values have also been represented graphically, as depicted in Fig. 6. Figure 7 presents the graphical representation of the BER values of proposed algorithm; it depicts that the extracted watermark exhibits few or no erroneous bits.

Table 6 presents the comparative analysis of the proposed algorithm with a few of the existing important audio watermarking algorithms. In this paper, a comparison with the state-of-art is presented on the basis of all the three critical design parameters. It is observed from the table that watermark embedding proposed algorithm can achieve a payload of 4,884.1 bps which is

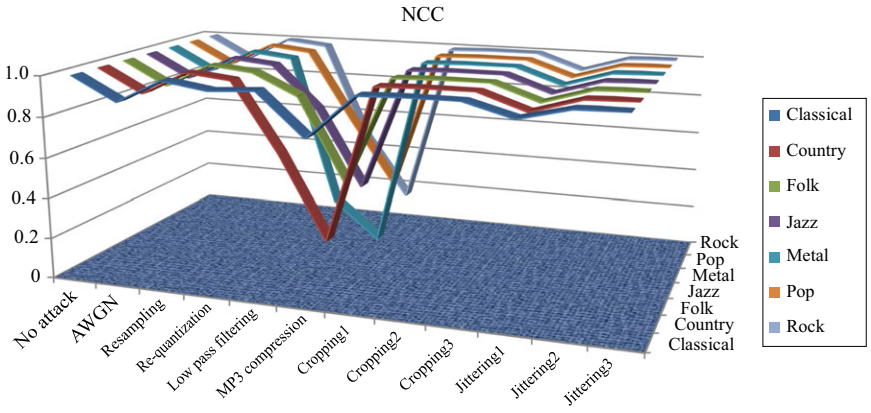


Fig. 6. Plot depicting NCC values of various audio signals under signal processing attacks.

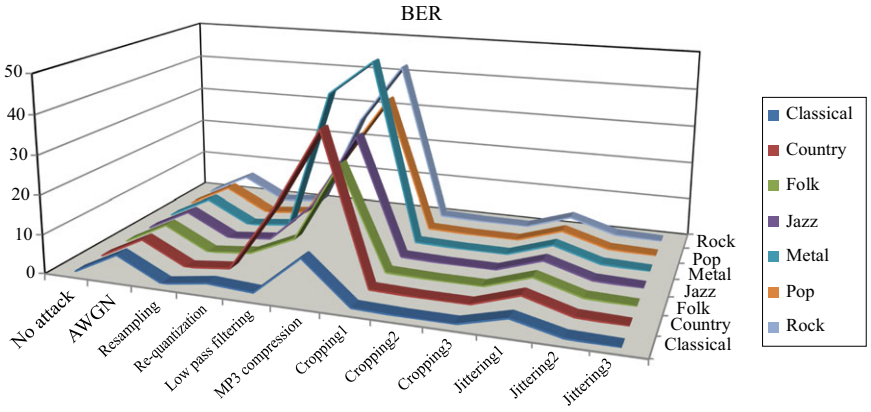


Fig. 7. Plot depicting BER (%) of various audio signals under signal processing attacks.

Table 6. Comparative study.

Reference	SNR (dB)	Payload (bps)	BER (%)		
			MP3 compression	Cropping	Jittering
[8]	>14.58	NR	3.18 (128 kbps)	NR	NR
[12]	>25	172	24.18 (32 kbps)	NR	NR
[13]	29.5	4.26	1.56 (32 kbps)	0 (18%)	6.25 (1/500)
[15]	>20	10.72	5.71 (128 kbps)	NR	NR
[16]	>40	3	15 (128 kbps)	0 (20%)	0 (1/300)
[17]	NR	86	0 (56 kbps)	0.48	NR
[18]	48.33	172.3	3.0273 (128 kbps)	0.9984	NR
[19]	NR	196	2 (32 kbps)	0 (0.11%)	NR
[20]	>35	102.4	0.48 (128 kbps)	0 (0.11%)	NR
[21]	>37	320	0 (32 kbps)	NR	NR
Proposed	**>25**	**4,884.1**	**0.02 (64, 96, 128, 192, 256 kbps)**	**0 (10%, 18%)**	**0 (1/500)**

adequately high in the field of audio watermarking. Moreover, the SNR of the proposed algorithm at such (high) payloads is over 25 dB; this is comparable with all the other methods, indicating the adequate perceptual transparency of the proposed algorithm. The low or zero BER% of the proposed algorithm indicates highly adequate robustness, as it is highly challenging to achieve an adequate robustness against challenging signal processing attacks in the field of audio watermarking at high payload.

IV. Conclusion

This paper presented a high-payload audio watermarking algorithm that is robust against challenging signal-processing attacks such as cropping, compression, jittering, re-quantization, noise, resampling, and low-pass filtering, with NCC equal to or closer to 1 for all the audio samples. A payload of 4,884.1 bps has been achieved with

SNR ranging from 21 dB to 40 dB. The selection of high-frequency detailed coefficients of wavelet-decomposition for embedding the watermark has rendered the proposed algorithm robust against intentional or unintentional attacks at high payloads. In addition to this, the use of a security code has rendered the proposed algorithm suitable for tamper detection and secure transmission. The comparative study of the proposed algorithm with existing algorithms indicates that the proposed algorithm exhibits adequate robustness at such high payloads. In the future, the emphasis may be to address other severe attacks such as time scale and pitch scale modifications.

References

[1] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for Data Hiding," *IBM Syst. J.*, vol. 35, no. 3/4, 1996, pp. 131–336.

[2] H.T. Hu, L.Y. Hsu, and H.H. Chou, "Perceptual Based DWPT-DCT Framework for Selective Blind Audio Watermarking," *Signal Process.*, vol. 105, Dec. 2014, pp. 316–327.

[3] H.T. Hu and L.Y. Hsu, "Robust, Transparent and High Capacity Audio Watermarking in DCT Domain," *Signal Process.*, vol. 109, no. C, Apr. 2015, pp. 226–235.

[4] M. Fallahpour and D. Megías, "High Capacity Method for Real-Time Audio Data Hiding Using the FFT Transform," *Adv. Inform. Sec. Appl.*, vol. 36, 2009, pp. 91–97.

[5] M. Fallahpour and D. Megías, "Audio Watermarking Based on Fibonacci Number," *IEEE/ACM Trans. Audio Speech Language Proc.*, vol. 23, no. 8, Aug. 2015, pp. 1273–1282.

[6] S.-T. Chen, C.-Y. Hsu, and H.-N. Huang, "Wavelet-Domain Audio Watermarking Using Optimal Modification on Low-Frequency Amplitude," *IET Signal Proc.*, vol. 9, no. 2, 2015, pp. 166–176.

[7] M. Mosleh, H. Latifpour, M. Kheyrandish, M. Mosleh, and N. Hosseinpour, "A Robust Intelligent Audio Watermarking Scheme Using Support Vector Machine," *Front. Inform. Technol. Electron. Eng.*, vol. 17, no. 12, Dec. 2016, pp. 1320–1330.

[8] P. Hu, D. Peng, Z. Yi, and Y. Xiang "Robust Time-Spread Echo Watermarking Using Characteristics of Host Signals," *IEEE Electron. Lett.*, vol. 52, no. 1, 2016, pp. 5–6.

[9] R. Li, S. Xu, and H. Yang, "Spread Spectrum Audio Watermarking Based on Perceptual; Characteristic Aware Extraction," *IET Signal Proc. Lett.*, vol. 10, no. 3, Apr. 2016, pp. 266–273.

[10] Y. Erfani, R. Pichevar, and J. Rouat, "Audio Watermarking Using Spikegram and a Two-Dictionary Approach," *IEEE Trans. Inform. Forensics Secur.*, vol. 12, no. 4, Apr. 2017, pp. 840–852.

[11] A. Kaur, M.K. Dutta, K.M. Soni, and N. Taneja, "Localized & Self Adaptive Audio Watermarking Algorithm in Wavelet Domain," *J. Inform. Security Appl.*, vol. 33, no. 1, Apr. 2017, pp. 1–15.

[12] S. Wu, J. Huang, D. Huang, and Y.Q. Shi, "Efficiently Self-Synchronized Audio Watermarking for Assured Data Transmission," *IEEE Trans. Broadcast*, vol. 51, no. 1, Mar. 2007, pp. 69–76.

[13] W. Li, X. Xue, and P. Lu, "Localized Audio Watermarking Technique Robust Against Time Scale Modification," *IEEE Trans. Multimedia*, vol. 8, no. 1, Feb. 2006, pp. 60–69.

[14] A. Kaur, M.K. Dutta, K.M. Soni, and N. Taneja, "A Blind Audio Watermarking Algorithm Robust Against Synchronization Attacks," *IEEE Int. Conf. Signal Process., Comput. Contr.*, Solan, India, Sept. 26–28, 2013, pp. 1–6.

[15] R. Wang, D. Xu, J. Chen, and C. Du, "Digital Audio Watermarking Algorithm Based on Linear Predictive Coding in Wavelet Domain," *IEEE Int. Conf. Signal Process., (ICSP)*, Beijing, China, 2004, pp. 2393–2396.

[16] S. Xiang and J. Huang, "Histogram Based Audio Watermarking Against Time Scale Modification and Cropping Attacks," *IEEE Trans. Multimedia*, vol. 9, no. 7, Nov. 2007, pp. 1357–1372.

[17] C.Y. Chang, W.C. Shen, and H.J. Wang, "Using Counter-Propagation Neural Network for Robust Digital Audio Watermarking in DWT Domain," *IEEE Int. Conf. Syst., Man Cybernetics*, Taipei, Taiwan, Oct. 8–11, 2006, pp. 1214–1219.

[18] P.K. Dhar and T. Shimamura, "Entropy-Based Audio Watermarking Using Singular Value Decomposition and Log-Polar Transformation," *IEEE Int. Midwest Symp. Circuits Syst.*, Columbus, OH, USA, Aug. 4-7, 2013, pp. 1224–1227.

[19] V. Bhat, I. Sengupta, and A. Das, "An Audio Watermarking Scheme Using Singular Value Decomposition and dither-modulation Quantization," *Multimed. Tools Appl.*, vol. 52, no. 2–3, Apr. 2011, pp. 369–383.

[20] M. Hemis, B. Boudraa, and T. Merazi-Meksen, "Optimized Audio Watermarking Scheme with Swarm Intelligence," *IEEE First Int. Conf. New Technol. Inform. Commun. (NTIC)*, Mila, Algeria, Nov. 8–9, 2015, pp. 1–6.

[21] A. Kaur, M.K. Dutta, K.M. Soni, and N. Taneja, "Hiding Biometric Features in Audio Signals Using Gram-Schmidt Orthogonalisation," *Int. J. Electron. Security Digit. Forensics*, vol. 8, no. 1, Dec. 2016, pp. 63–81.

https://doi.org/10.4218/etrij.2017-0092

Arashdeep Kaur is associated with the Department of Computer Science and Engineering, Amity School of Engineering and Technology, Noida, Uttar Pradesh, India. She completed her B.Tech in Computer Science and Engineering with honors from Punjab Technical University, Jalandhar, Punjab, India, and her M.Tech in Computer Science and Engineering from Guru Nanak Dev Engineering College, Ludhiana, India. She completed her PhD from Amity University, Uttar Pradesh, India in the field of multimedia security and digital watermarking. Her research interests include watermarking, multimedia data security, image processing, and medical imaging. She has published numerous journal and conference papers of repute.

Malay Kishore Dutta is associated with the Department of Electronics and Communication Engineering, Amity School of Engineering and Technology, Noida, Uttar Pradesh, India. He completed his M.Tech in Electronics Engineering from Central University, Tezpur, Assam, India, securing the first position (Gold Medalist). He completed his PhD in the area of multimedia data security and signal processing from Uttar Pradesh Technical University, India. His research interests include multimedia data security, watermarking, and steganographic algorithms for multimedia. He has published numerous peer-reviewed research papers in international journals and conferences. He has filed three patents in the area of signal processing and is the principal investigator of various DST-funded projects.