

An Improved Non-Repudiate Scheme - Feature Marking Voice Signal Communication

Remya A R, A Sreekumar, Supriya M H, Tibin Thomas
Artificial Intelligence Research Lab, Department of Computer Applications,
Department of Electronics, Cochin University of Science and Technology, Kochi, Kerala, India – 682 022
remyacusat@gmail.com, askcusat@gmail.com, supriyadoe@gmail.com, tbnmaths@gmail.com

Abstract—Guaranteeing the ownership or copyright of digital communication is of extreme importance in this digital era. Watermarking is the technique which confirms the authenticity or integrity of communication by hiding relevant information in specified areas of the original signal such that it might render it difficult to distinguish one from the other. Thus, the digital watermark can be defined as a type of indicator secretly embedded in a noise tolerant signal such as image, audio or video data.

The paper presents a voice signal authentication scheme by employing signal features towards the preparation of the watermark and by embedding it in the transform domain with the Walsh transforms. Watermark used in this technique is unique for each member participating in this communication system and makes it is very imperative in the context of signal authentication.

Index Terms—Digital Watermarking, FeatureMark, Walsh Transforms, Non-repudiation.

I. INTRODUCTION

Advances in digital technology have led to widespread use of digital communication in various areas including government, legal, banking, military etc. This in turn has increased the reproduction and re-transmission of multimedia data through both legal and illegal channels. The illegal usage causes a serious threat to the content owner's authority or proprietary right. Thus today's information based society places extreme importance on authenticating the information that is sent across the communication channels.

Digital watermarking is the technique which ensures authorized and legal use of digital communication, copyright protection, copy protection etc. that helps to prevent such misuse. Watermarking demonstrates the process of embedding copyright or ownership information to the original signal in an imperceptible form. The embedded watermark should be robust to any signal manipulations and can be unambiguously retrieved at the other end. Audio watermarking is the term coined to represent the insertion of a signal, image or text of known information in an audio signal in an imperceptible form[1].

Studies on existing watermarking schemes motivate us to develop a non-repudiated scheme which ensures the authenticity of each member participating in the communication. This paper is an enhancement on our previous work with the addition of an encryption scheme in the preparation of watermark to further improve its security. Any member participating in the communication might not be able to deny their presence of participation and thus the system guarantees an authentic, non-repudiated communication scheme.

Section 2 of this paper goes on to showcase some of the existing works in the field of audio watermarking. Section 3 describes some of the definitions and methodologies used in this work. Sections 4 and 5 demonstrate the synchronization code and analysis of Walsh transforms respectively. Section 6 explains the proposed system framework and section 7 constitutes the experimental results and discussion.

II. RELATED WORKS

Recent years have reported a rapid growth in the area of audio watermarking schemes. An audio watermarking scheme presented in the paper [2] utilizes pseudo-Zernike moment and synchronization code of the audio signal in the embedding as well as in the extraction scheme. This method is resistant towards the de-synchronization attacks. Another scheme [3] which employs statistics characteristics of an audio signal is also resistant towards de-synchronization attacks. A robust audio watermarking scheme suggested in the paper [4] is based on singular value decomposition and dither modulation quantization which comes under the blind watermarking scheme. A blind audio watermarking algorithm [5] splits up the original audio signal and hides the watermark in the Discrete Cosine Transforms (DCT) coefficients. Digital audio copyright scheme demonstrated in [6] embodies a psycho-acoustic model of MPEG audio coding to ensure the quality of the original sound. The algorithm presented in paper [7] proposes the use of adaptive quantization to overcome the synchronization attacks on digital audio watermarking. Such technique falls under the blind watermarking scheme.

An audio watermarking scheme in the time domain presented in [8] does not require the original signal for

watermark detection. An audio watermarking methodology [9] exploits the temporal and frequency perceptual masking criteria, constructs its watermark by breaking each audio clip into smaller segments and adding a perceptually shaped pseudo random noise. A digital audio watermarking in the cepstral domain is presented in [10]. Within such cepstral coefficients of the signal, the watermark is embedded using the spread spectrum techniques. Another technique [11] used is the frequency hopping watermark embedding in the spectral domain wherein the watermark is embedded in the power spectrum coefficients of the audio signal. Modified patch work algorithm in transform domain is described in [12] is robust to withstand some attacks defined by Secure Digital Music Initiative (SDMI).

Work on [13] demonstrates a method to protect the copyright of digital media over the Internet by introducing a support vector regression based audio watermarking scheme in the wavelet domain; and the watermark information is embedded into randomly selected sub-audios.

Histogram based audio watermarking scheme is tested to verify its characteristics by applying different attacks including low-pass filtering, amplification etc. [14]. Multiple scrambling in the embedding process guarantees unauthorized detection of the embedded watermark and enhances the robustness. This criterion of an adaptive synchronization scheme is utilized in the watermark detection process [15] Employing DCT coefficients in the embedding and extraction procedures of audio watermarking method presented in [16] and works with neural network scheme. Many works have been conducted and are being conducted in this area of research.

III. FUNDAMENTAL THEORY

Voice signal identification can be achieved by extracting the characteristic features employed towards the development of watermark employed in this system.

A. Mel-Frequency Cepstral Coefficients [MFCC]

Mel-frequency cepstral coefficients introduced by Davis and Mermelstein in 1980's are treated as the best parametric representation of the acoustic signals employed in the recognition of speakers and have been the state-of-the-art ever since. As described in [20 - 21], Mel-scale relates perceived frequency or pitch of a pure tone to its actual measured frequency. Humans are much better at discerning small changes in pitch at low frequencies than they are at high frequencies. Incorporating this scale makes our features match more closely what is audible to humans. MFCCs are based on a linear cosine transform of a log power spectrum on a non-linear Mel-scale of frequency. MFCCs are treated as the best for speech or speaker recognition systems because it takes human sensitivity with respect to frequencies into consideration. [17-19] demonstrates audio watermarking as well as audio steganographic techniques in the cepstral coefficients.

The formula for converting from frequency to Mel-scale is:

$$M(f) = 1125 \ln(1 + f/700) \quad (\text{Eqn.1})$$

To go from Mel's back to frequency:

$$(M - 1)(m) = 700 \left(\exp \left(\frac{m}{1125} \right) - 1 \right) \quad (\text{Eqn.2})$$

B. Spectral Flux

The spectral flux also termed as spectral variation is a measure of how quickly the power spectrum varies corresponding to each frame in a short-time window. It can be defined as the squared difference between the normalized magnitudes of successive spectral distributions corresponding to successive signal frames. Timbre of an audio signal can also be derived from it [21].

$$F_r = \sum_{k=1}^N (|X_r[k]| - |X_{r-1}[k]|)^2 \quad (\text{Eqn.3})$$

A high value of spectral flux stands for a sudden change in the spectral magnitudes and therefore a possible spectral boundary at the the frame.

C. Spectral Centroid

The spectral shape of a frequency spectrum is measured with the spectral centroid. Higher the centroid values, the brighter will be the textures and the higher the frequencies. This measure characterize a spectrum and can be calculated as weighted mean of the frequencies presented in the signal, determined using a Fourier transform with their magnitudes as weights:

$$C = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \quad (\text{Eqn.4})$$

where $x(n)$ represents the weighted frequency value or magnitude of binary number n and $f(n)$ represents the center frequency of that binary number.

Centroid represents sharpness of the sound which is related to the high-frequency content of the spectrum. Higher centroid values resemble the spectra in the range of higher frequencies. The effectiveness of centroid measures to describe spectral shape makes it usable in voice signal classification activities [21].

D. Spectral Roll-Off

Spectral roll off point is defined as the N^{th} percentile of the power spectral distribution where N is usually 85% or 95%. The roll off point is the frequency below which the $N\%$ of the magnitude distribution is concentrated. In other words, spectral roll-off demonstrates the frequency below which 85% of the magnitude distribution of the spectrum is concentrated. It is the measure of spectral shape and yields higher values for higher frequencies. Both the centroid and spectral roll-off are measures of spectral shape and the spectral roll-off yields higher values for high frequencies or right skewed spectra [21-23].

The roll-off is given by $R_r = f[k]$, where K is the largest bin that satisfies

$$\sum_{k=1}^K |X_r[k]| \leq 0.85 \sum_{k=1}^N |X_r[k]| \quad (\text{Eqn.5})$$

E. QR Code

Nowadays, the use of Quick Response Codes (QR Codes) is increasing tremendously. A QR Code is represented as a 2-D code placed on a white background and consists of black modules arranged in a square pattern. Any kind of data including binary or alphanumeric can be given as input to an online QR Code generator for encoding the given information. Users with an online Bar code scanner application on their phone or PC can scan the image of the QR Code to decode and display the details hidden in it.

F. Walsh Transforms

The Walsh-Hadamard transform (WHT) is defined as a suboptimal, non-sinusoidal, orthogonal transformation that decomposes a signal into a set of orthogonal, rectangular waveforms called Walsh functions. These transformations work in the transform domain of the signal that has no multipliers and is real because the amplitude of Walsh (or Hadamard) functions have only two values, +1 or -1.

The forward and inverse Walsh transform pair for a signal $x(t)$ of length N are

$$y_n = \frac{1}{N} \sum_{i=0}^{N-1} x_i \text{WAL}(n, i), n = 1, 2, \dots, N - 1 \quad (\text{Eqn.6})$$

$$x_i = \frac{1}{N} \sum_{n=0}^{N-1} y_n \text{WAL}(n, i), n = 1, 2, \dots, N - 1 \quad (\text{Eqn.7})$$

It decomposes an arbitrary input vector into a superposition of Walsh functions. This is an example of generalized class of Fourier transforms. Fast Walsh transform is an efficient algorithm to compute the Walsh transform which reduces the complexity from $O(N^2)$ to $O(N \log N)$ [24].

G. Non-Repudiation

Non-repudiation allows exchange of digital communication between different groups of individuals in such a way that the people in this communicating group cannot subsequently deny their participation in the communication. It refers to a state of play where the supposed orator of a statement will not be able to effectively challenge the legitimacy of statement or contract.

Regarding digital security, the cryptologic meaning and application of non-repudiation shifts to mean:

- a) A service that provides proof of the integrity & origin of data
- b) An authentication that can be asserted to be genuine with high assurance.

Proof of data integrity is typically the easiest of these requirements to accomplish [25-26].

IV. SYNCHRONIZATION CODE

An important task associated with audio watermarking is synchronization. Any deviation or loss in synchronization results in false detection of the

embedded watermark. Synchronization is achieved by inserting synchronization codes at the beginning of the embedded watermark. Any modification in the system may result in false detection of the embedded watermark. So we need an efficient synchronization mechanism to detect and extract the exact replica of the embedded watermark from the signal. The synchronization code used in our system is orthogonal codes which are employed to improve the bandwidth efficiency of spread spectrum systems. Walsh functions are used in the watermark embedding process which has zero correlation between each other. Orthogonal spreading codes can be used in synchronization because the cross-correlation between different shifts of Walsh functions is not zero [27-29].

$$\sum_{k=0}^{M-1} \phi_i(k\tau)\phi_j(k\tau) = 0, i \neq j \quad (\text{Eqn.8})$$

where $\phi_i(k\tau)$, $\phi_j(k\tau)$: i^{th} and j^{th} members of an orthogonal set and τ is the symbol duration.

V. ANALYSIS OF WALSH TRANSFORMS

Walsh transform is the discrete analog of the Fourier transforms. Due to the fact that Walsh functions and its transforms are naturally more suited for digital computation an effort is made to gradually replace the Fourier transform by Walsh-type transforms. In this work, the digital watermark is embedded into the original voice signal by applying fast Walsh transforms on it. During the transmission process, the Feature Marked voice signal may suffer various signal manipulations including noise addition, silence addition, echo addition, re-sampling, re-quantization, low-pass filtering, band-pass filtering etc. and other de-synchronization attacks that include amplitude variation, pitch shifting, random cropping, time-scale modification etc. This may affect the selected Walsh coefficients in a non-negligent way. We have adopted the Walsh transforms for embedding the watermark under the assumption that for the case of random functions the Walsh power spectra are slowly convergent and many Walsh transform components contain approximately equal signal power. Replacing such coefficients might not degrade the signal quality.

After embedding the watermark into the selected Walsh coefficients, its inverse function is employed to cancel the changes. During the transmission process, voice signal may suffer some signal manipulations and de-synchronization attacks as mentioned above.

The effect of these attacks against the robustness criteria is demonstrated in the tables presented in the section 7. Walsh transforms yields an order of the basis vectors with respect to their spectral properties. Robustness feature of the watermark can be achieved by selecting the Walsh coefficients in such a way that detection and reconstruction of the embedded watermark does not end in any degradation of the data [30-32]

VI. PROPOSED SYSTEM FRAMEWORK

Overall system architecture at the sending window is presented in Figure 1.

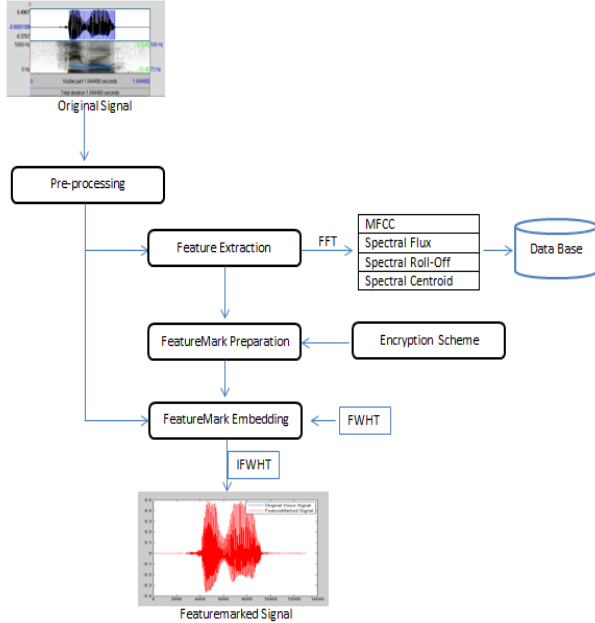


Figure1: System Architecture – Sending Module

A. Non-Repudiate Audio Feature Marking Scheme

Algorithm 1:

Input : Original Audio Signal

Output: Feature Marked Audio Signal

Step 1: Perform the pre-processing
- framing and windowing

Step 2: Extract feature vectors using the fast Fourier transformations and save it into a database
- Mel-frequency cepstral coefficients
- spectral flux
- spectral roll off
- spectral centroid

Step 3: Feature values extracted are given to an encryption scheme

$$y = \bar{x} \cup_k$$

y is obtained by inserting $\bar{x} \bmod 7$ at the k^{th} position of \bar{x} , where

$$\bar{x} = x_1 x_2 \dots x_n \cdot x_{n+1} x_{n+2} \dots x_{n+k}$$

and $0 \leq k \leq m$ and $m = n + k$

Step 4: Results obtained are given as input to an online QR Code generator, and the generated QR Code will now on be termed as the Feature-Mark in our system

Step 5: Synchronization code embedding

This system involves insertion of orthogonal codes towards the beginning of each watermark in the signal.

Step 5: Embed the Feature-Mark to the original signal by performing the fast Walsh-transforms.

Step 6: Perform the inverse fast Walsh transforms and send the Feature-Marked signal to the desired recipient

Step 7: Detect the presence of the feature mark to confirm the authenticity of the received signal

Step 8: Obtain the feature vectors from the extracted FeatureMark using an online QR Code scanner

Step 9: Perform the decryption scheme to obtain the exact feature values

Step 10: Feature comparison to confirm the authenticity of the voice signal

B. System Framework

The proposed scheme is tested with signal duration of 1s, 2s... 60s, 240s, 300s with a collection of around 50 samples. Input signals are pre-processed with framing and windowing methodology. The number of frames depends on the signal duration. The frame rate is taken as 100. The feature extraction module considers each of these frames and extracts the feature values that are directly related to the computable characteristics of time-domain signals. A copy of these extracted values is also stored in the database.

Extracted feature vectors are encrypted using the following method:

$$y = \bar{x} \cup_k, \quad (\text{Eqn.9})$$

y is obtained by inserting $\bar{x} \bmod 7$ at the k^{th} position of \bar{x} , where $\bar{x} = x_1 x_2 \dots x_n \cdot x_{n+1} x_{n+2} \dots x_{n+k}$ and $0 \leq k \leq m$ and $m = n + k$

The encryption method is implemented as an improvement to our earlier work where the extracted features are employed directly in the watermark preparation module. Encryption scheme helps in preserving the feature values from creating fake Feature-Marks in case of imitated voice. The obtained encrypted-feature values aid in the preparation of the signal dependent Feature Mark. Embedding module function is a two-step process: (i) the first phase involves the insertion of synchronization code to detect the presence of watermark in the signal; and (ii) the second phase involves the actual watermark bit insertion. Orthogonal codes or Walsh codes are generated and used as the synchronization code in this system. The actual watermark bit insertion is done by transforming the host audio using the Walsh transforms and inserts each watermark bit by replacing the appropriate values of the Walsh coefficients.

According to the length of the audio signal, upon embedding more than one module function and on completion of the embedding task, the system performs the inverse Walsh transform to hide the changes occurred to the original host signal. This ensures that

both the host audio signal as well the new FeatureMarked signal exhibits similar statistical properties. The FeatureMarked voice signal can now be send across the communication channel. The extraction module starts its functioning upon receipt of the FeatureMarked signal. On receipt of the FeatureMarked signal the system searches for the synchronization code and then starts fetching the FeatureMark bits between each synchronization codes. To accomplish this, the system performs the reverse operations of the embedding scheme. The extracted FeatureMark is then processed further to confirm the authenticity in case of any suspicion/dispute or non-repudiate behavior in future. The purpose of authentication or non-repudiation module is to compare the values present in the database as well with the extracted feature values in the receiving end to confirm its ownership. Thus, the entire system acts as a strong non-repudiated scheme to guarantee legal time-domain signal communications, in the real-world applications.

VII. EXPERIMENTAL RESULTS AND DISCUSSIONS

The proposed scheme is tested with signal duration ranging from 1second to 5minutes, the same was found to be acceptable with around 50 sample signals. Recorded Signals are pre-processed with framing and windowing methods. Rather than processing the signal with the sampling theorem, a frame-based processing divides continuous sequences input and output samples into frames of N samples. Number of frames depends on the length of the signal with a frame rate of 100 and the suggested scheme employs the Hamming windowing method to reduce or avoid the signal discontinuities towards the beginning and ending of the segmented frames which can be represented as the equation given below:

$$W_H = 0.54 + 0.46 \cos\left(\frac{n\pi}{Q}\right), \text{(Eqn.10)}$$

for $|n| \leq Q$ and 0 for other values

Then, the individual data chunks obtained for the time-domain signal are given as input to the feature extraction module.

The feature extraction module extracts the Mel-frequency cepstral coefficients, spectral flux, spectral roll-off and spectral centroid values for each frame. The Mel-frequency cepstral coefficients are restricted to 20 which results in a data matrix of 20 X 256 coefficients. Embedding such a huge data volume in the audio will expose the furtiveness of the system hence-forth vector quantization is implemented resulting in a data matrix of 20 X 1 coefficients. The process of feature extraction employed in extracting MFCC values is depicted using the following figure 2:

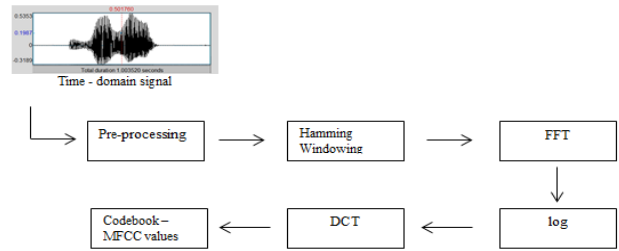


Figure2: MFCC evaluation

A copy of the original feature values as well as the encrypted values is saved into the database. The FeatureMark prepared with the help of an online QR Code generator is streamed to fit in the host audio signal. Streaming involves pre-processing, scrambling, scaling and reshaping; scrambling is performed using Arnold transforms as specified below:

$$FM_1 = fm_1(p, q), 0 \leq p < M, 0 \leq q < N \text{ (Eqn.11)}$$

The transformation results in a one-dimensional sequence of ones and zeroes as shown below:

$$FM_2 = fm_2(r) = fm_1(p, q), \text{ (Eqn.12)}$$

where $0 \leq p < M, 0 \leq q < N, r = p \times N + q$ and

$$fm_2(r) \in 0,1$$

Then scaling and reshaping is done to fit the FeatureMark into the original time-domain signal. Examples for generated FeatureMark are represented in the figure 3.



Figure3: Sample FeatureMark Employed in this System

Orthogonal codes or Walsh codes as well as the streamed FeatureMark is embedded into the host audio by replacing the peak coefficient values which are obtained by transforming the signal using fast-Walsh-Hadamard transform. Number of Walsh coefficients selected is directly proportional to the size of the synchronization code as well as the FeatureMark. An example of Walsh code generation is represented in the following figure 4:

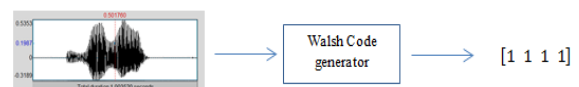


Figure4: Sample Walsh Code Generation

Embedding module ends by performing the inverse fast-Walsh-Hadamard transform to neutralize the changes that had occurred in the original signal properties as part of the transformation. Now, the FeatureMarked voice signal can be send across the communication channel, upon receipt of the signal the FeatureMark extraction module, detects and extracts the

authentication content of the signal by performing the following operations. Authentication of the voice signal can be confirmed by comparing feature values that are extracted and decrypted as mentioned in the algorithm with the database values as well as by performing the feature extraction module at the receiving end, in case of further clarification.

Reliability test is conducted in this study by calculating the Bit Error Rate (BER) of the extracted FeatureMark and is measured by

$$BER = \frac{B}{M \times N} \times 100\% \quad (\text{Eqn.13})$$

where B corresponds to the number of erroneously detected bits. BER values for the FeatureMarked signals are calculated and depicted in tables 1 – 6.

Also, the FeatureMarked voice signal quality is measured by calculating the PSNR values as follows:

$$PSNR = 10 \log_{10} \frac{R^2}{MSE} \quad (\text{Eqn.14})$$

where R is the maximum fluctuation in the input image data type and MSE stands for the Mean Square Error, a matrix used to compare the image quality.

A. Performance Tests

The plots of a short portion of the original voice signals and the corresponding FeatureMarked signals are shown in Figures (5) and (6) which does not reveal any difference between these two signals. The extracted feature marked image without being attacked is obtained with BER = 0.0 and PSNR = 0.0.

B. Robustness Tests

Robustness nature of the proposed watermarking scheme against common signal processing and desynchronization attacks is determined by splitting the tests in two stages.

The first stage constitutes tests for the common signal processing functions:

Noise Addition: Added white Gaussian noise with 50

Silence Addition: A silence of 0.5 seconds duration is inserted at the beginning of the FeatureMarked signal

Echo Addition: Added an echo signal with a delay of 0.5 seconds

Re-sampling: FeatureMarked voice signals are down sampled to frequencies 22.05, 11.025, and 8 kHz and then up sampled to its original 44.1 kHz

Re-quantization: FeatureMarked voice signals with 16-bit are quantized to 8-bit and then back to its original 16-bit

Low-pass filtering: It is done with cut off frequencies 1 kHz and 200 Hz

Band-pass filtering: It is done with cut off frequencies 1 kHz and 200 Hz

Later stage constitutes tests against de-synchronization attacks such as:

Amplitude Variation: FeatureMarked voice signal is amplified to its double as well as to half for conducting this test

Pitch Shifting: This attack results in frequency fluctuation to the signal and is conducted in our

study with a deviation of 1 degree higher and 1 degree lower

Cropping: It is performed randomly at different locations of the signal

Time-scale modification: Watermarked voice signal was lengthened (slow-down) and shortened (double speed) for evaluating the robustness nature.

C. Experiment # 1

Experiment # 1 is performed on single channel sounds with around 27 samples and was found successful on all the signals.

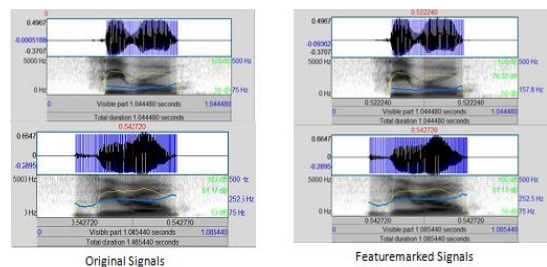


Figure5: Comparison between Original and Watermarked Files

TABLE 1. Robustness test for common signal manipulations (in BER)

Original Signal	Attacks Free	Noise Addition	Silence Addition	Echo Addition
Host Audio 1	0	0.0001	0.0001	0.0002
Host Audio 2	0	0.0000	0.0000	0.0001
Host Audio 3	0	0.0001	0.0003	0.0001
Host Audio 4	0	0.0002	0.0000	0.0001

TABLE 2. Robustness test for common signal manipulations (in BER)

Original Signal	Re-Sampling	Re-Quantization	Low-pass filtering	Band-pass Filtering
Host Audio 1	0.0001	0.0001	0.0002	0.0002
Host Audio 2	0.0002	0.0001	0.0004	0.0005
Host Audio 3	0.0002	0.0002	0.0002	0.0003
Host Audio 4	0.0001	0.0002	0.0032	0.0040

TABLE 3. Robustness test for de-synchronization attacks (in BER)

Original Signal	Amplitude Variation	Pitch Shifting	Random Cropping	Time-Scale modification
Host Audio 1	0.0005	0.0039	0.0500	0.0042
Host Audio 2	0.0003	0.0033	0.0060	0.0007
Host Audio 3	0.0002	0.0041	0.0040	0.0009
Host Audio 4	0.0062	0.0042	0.0039	0.0005

D. Experiment # 2

Experiment # 2 is performed on multi-channel sounds with around 23 samples and was found successful on all the signals.

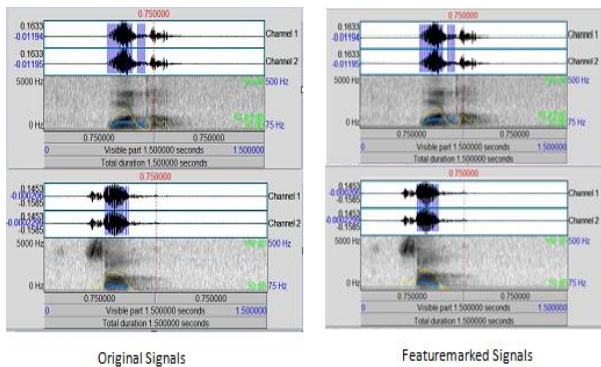


Figure6: Comparison between Original and Watermarked Files

TABLE 4. Robustness test for common signal manipulations (in BER)

Original Signal	Attacks Free	Noise Addition	Silence Addition	Echo Addition
Host Audio 5	0.0000	0.0002	0.0001	0.0002
Host Audio 6	0.0000	0.0000	0.0000	0.0001
Host Audio 7	0.0000	0.0003	0.0001	0.0002
Host Audio 8	0.0000	0.0001	0.0002	0.0003

TABLE 5. Robustness test for common signal manipulations (in BER)

Original Signal	Re-Sampling	Re-Quantization	Low-pass filtering	Band-pass Filtering
Host Audio 5	0.0004	0.0001	0.0040	0.0009
Host Audio 6	0.0004	0.0002	0.0039	0.0080
Host Audio 7	0.0003	0.0002	0.0050	0.0031
Host Audio 8	0.0002	0.0001	0.0700	0.0003

TABLE 6. Robustness test for De-synchronization attacks (in BER)

Original Signal	Amplitude Variation	Pitch Shifting	Random Cropping	Time-Scale Modification
Host Audio 5	0.0020	0.0004	0.0400	0.0040
Host Audio 6	0.0040	0.0034	0.0050	0.0050
Host Audio 7	0.0004	0.0040	0.0020	0.0010
Host Audio 8	0.0037	0.0040	0.0039	0.0040

VIII. ONCLUSIONS

The proposed method functions as a secure, robust voice authentic system that guarantees anon-repudiated scheme for voice signal communication. The use of voice signal features; its classification and feature marking, offers an improved scheme for authentic voice communication. The use of an online QR Code generator to create the Feature Mark from the encrypted feature values and an online QR Code scanner to extract the feature values is vital to this work. The QR Code brings in the ease of comparison in the authentication module. Repudiation can be avoided by performing the feature extraction module at the receiver end and cross check with the Feature Mark values in case of denial of participation by any member.

IX. UTURE WORK

This system can be extended for implementation in real time communication systems for the purpose of guaranteeing the government, military, banking etc. transactions or telephonic communications.

ACKNOWLEDGEMENTS

This work was funded by the Department of Science and Technology, Government of India under the INSPIRE Fellowship (IF110085).

REFERENCES

- [1] Stefan K, Fabien A P, Information hiding techniques for steganography and digital watermarking. Artech House, London, UK, 2000.
- [2] Wang XY, Ma TX, Niu PP (2011) A pseudo-zernike moment based audio watermarking scheme robust against desynchronization attacks. Computers & Electrical Engineering 37(4):425–443.
- [3] Wang XY, Niu PP, Yang HY (2009) A robust digital audio watermarking based on statistics characteristics. Pattern Recognition 42(11):3057–3064.
- [4] Bhat V, Sengupta I, Das A (2011) An audio watermarking scheme using singular value decomposition and dither-modulation quantization. Multimedia Tools and Applications 52(2-3):369–383.
- [5] Mierswa I, Morik K (2005) Automatic feature extraction for classifying audio data. Machine learning 58(2-3):127–149.
- [6] JongwonSeok, Jinwoo Hong, and Jinwoong Kim, A Novel Audio Watermarking Algorithm for Copyright Protection of Digital Audio, ETRI Journal, Volume 24, Number 3, June 2002.
- [7] Lin Y, Abdulla WH (2008) Multiple scrambling and adaptive synchronization for audio watermarking. In: Digital watermarking, Springer, pp 440–453.

- [8] Bassia P, Pitas I, Nikolaidis N (2001) Robust audio watermarking in the time domain. *Multimedia, IEEE Transactions on* 3(2):232–241.
- [9] Swanson MD, Zhu B, Tewfik AH, Boney L (1998) Robust audio watermarking using perceptual masking. *Signal Processing* 66(3):337–355.
- [10] Lee SK, Ho YS (2000) Digital audio watermarking in the cepstrum domain. *Consumer Electronics, IEEE Transactions on* 46(3):744–750.
- [11] Cvejic N, Seppänen T (2004) Spread spectrum audio watermarking using frequency hopping and attack characterization. *Signal processing* 84(1):207–213.
- [12] Yeo IK, Kim HJ (2003) Modified patchwork algorithm: A novel audio watermarking scheme. *Speech and Audio Processing, IEEE Transactions on* 11(4):381–386.
- [13] Xu X, Peng H, He C (2007) Dwt-based audio watermarking using support vector regression and subsampling. In: *Applications of Fuzzy Sets Theory*, Springer, pp 136–144.
- [14] Mali MD, Khot S (2012) Robustness test analysis of histogram based audio watermarking. In: *Wireless Networks and Computational Intelligence*, Springer, pp 611–620.
- [15] Lin Y, Abdulla WH (2008) Multiple scrambling and adaptive synchronization for audio watermarking. In: *Digital watermarking*, Springer, pp 440–453.
- [16] Wang C, Ma X, Cong X, Yin F (2005) An audio watermarking scheme with neural network. In: *Advances in Neural Networks–ISNN 2005*, Springer, pp 795–800.
- [17] Gopalan K, Audio steganography by cepstrum modification. In: *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP'05). IEEE International Conference on, IEEE*, vol 5, pp v–481.
- [18] Gopalan K, Robust watermarking of music signals by cepstrum modification. In: *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, IEEE*, pp 4413–4416.
- [19] Kraetzer C, Dittmann J (2007) Mel-cepstrum based steganalysis for voip-steganography. *Proceedings of SPIE, Security, Steganography, and Watermarking of Multimedia Contents IX* 6505:650,505–1.
- [20] <http://practicalcryptography.com/>.
- [21] <http://en.wikipedia.org/wiki/>.
- [22] http://sovarr.c4dm.eecs.qmul.ac.uk/wiki/Spectral_Rolloff.
- [23] http://www.paradisedata.com/collateral/articles/AN_035OptimisedSpectralRollOffApplicationNote.pdf /.
- [24] <http://www.mathworks.com/>.
- [25] Coffey T, Saidha P (1996) Non-repudiation with mandatory proof of receipt. *ACM SIGCOMM Computer Communication Review* 26(1):6–17.
- [26] Kremer S, Markowitch O, Zhou J (2002) An intensive survey of fair non-repudiation protocols. *Computer communications* 25(17):1606–1621.
- [27] He X, Scordilis MS (2008) Efficiently synchronized spread spectrum audio watermarking with improved psychoacoustic model. *Journal of Electrical and Computer Engineering* 2008.
- [28] http://paginas.fe.up.pt/~hmiranda/cm/Pseudo_Noise_Sequences.pdf/.
- [29] http://www.math.wpi.edu/MPI2008/TSC/TSCeindli_jk.pdf/.
- [30] Tzafestas S (1983) Walsh transform theory and its application to systems analysis and control: an overview. *Mathematics and Computers in Simulation* 25(3):214–225.
- [31] Abbasi, Shuja A and Alamoud, ARM and others, Design of Real Time Walsh Transform for Processing of Multiple Digital Signals, *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 3, No. 2, April 2013, pp. 197-206.
- [32] Goresky, Mark, and Andrew Klapper, Arithmetic Correlations and Walsh Transforms, *IEEE Transactions On Information Theory*, Vol. 58, No. 1, January 2012.

Remya A R received the post-graduation degree in software engineering under the Faculty of Technology, Cochin University of Science & Technology (CUSAT), Kochi, in 2007. She was working as the assistant systems engineer in Tata Consultancy Services during the year 2007-2011. Presently, she is a PhD Candidate at the Department of Computer applications, CUSAT, under the supervision of Dr. A Sreekumar.

Dr. A Sreekumar received the PhD degree under the Faculty of Technology, Cochin University of Science & Technology (CUSAT). He is an associate professor and a research supervisor in the Department of Computer Applications, CUSAT.

Dr. Supriya M H received the PhD degree under the Faculty of Technology, Cochin University of Science & Technology (CUSAT). She is an associate professor and a research supervisor in the Department of Electronics, CUSAT.

Tibin Thomas received the Post-graduation degree from the Department of Mathematics, Cochin University of Science and Technology (CUSAT). He is an assistant professor in the Department of Mathematics, St. Berchmans College, Changanassery, Kerala, and also a PhD candidate in the Department of Computer Applications, CUSAT.

How to cite this paper: Remya A R, A Sreekumar, Supriya M H, Tibin Thomas, "An Improved Non-Repudiate Scheme-Feature Marking Voice Signal Communication", *IJCNIS*, vol.6, no.2, pp.1-8, 2014. DOI: 10.5815/ijcnis.2014.02.01