International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-7S2, May 2019

# Spread Spectrum Based Digital Audio Watermarking

**Bodhvi Gaur, Chandrakala.C.B**

***Abstract*: This paper presents a comprehensive survey of the very recent advancements in the field of spread spectrum watermarking of audio data. An exhaustive literature overview of each of these in terms of the methodology, advantages, challenges, applications and future prospects has been discussed. These techniques range from aiming at improving robustness, making watermarks more imperceptible, prevention against the wide range of attacks on audio data. This paper also reveals the challenges that are faced by the researchers working in this domain, that are yet to be tackled and thus have a research potential. To the best of our knowledge there is no recent study discussing the various spread spectrum based audio watermarking techniques and thus our work will serve as the foundation for the new comer's keen on further research in this area.**

***Index Terms*: Digital Audio Watermarking, Spread Spectrum, Watermarking.**

## I. INTRODUCTION

The significant amount of transmission of data, be it images, audios or videos via open communication channels, makes their protection, against unwanted intentional attacks, a dire need. Watermarking, which refers to the technique used for inserting imperceptible and inseparable information into data in order to ensure its authenticity and confidentiality, is used for the same. A strong watermarking technique ensures robustness, imperceptibility, security, high capacity of embedded data and tolerable computational cost.

Digital Audio Watermarking (DAW) [01] refers to the inaudible reformation of audio data to embed some information in it which can later be extracted for its claim. In simple terms, audio watermarking refers to the embedding of digital data into a host (cover) audio signal. Watermarking of Audio signals is more difficult as compared to image or video watermarking because of the vast dynamic range that Human Auditory System (HAS) has as compared to the Human Visual System (HVS). Even though HAS has a large dynamic range it also carries a narrow differential range which is the reason why loud sounds generally overpower softer sounds. Also, the capacity of data that one can embed into a host signal is less in comparison to the other forms of digital data which is due to the fact that audio is a single dimension data.

Major applications of digital audio watermarking include copyright protection (refers to protecting ownership of a certain audio data), tamper detection(refers to the detection of any kind of fiddling of the host audio data), copy protection(refers to the prevention of illegal creation and circulation of copies of the actual audio data), finger printing(refers to a compressed digital representation i.e. a fingerprint created deterministically from an audio data which is used to recognize its identity and locate other such items),broadcast monitoring(refers to periodic monitoring of the embedded broadcast data to authenticate the data sent),etc.

Representing audio data in digital format be it that of music, radio shows, telephone banking, emergency calls or that of recordings for heritage preservation, offers various advantages over their representation in analog format. Its editing isn't that tough because access to the exact discrete location is possible with almost no loss of precision of data. Digital audio data gets easily transmitted over networks at low costs due to unrestrained access that is given by technology. This leads to easy copy, download, and distribution of audio data. This gives rise to the significance of authenticating, verifying and protecting digital data and this is why prevention of illegal reproduction of audio data is of importance. The reason why audio watermarking can act against these loopholes are the advantages of imperceptibility, the inseparability of the watermark from the data and the feature to go through the same transformation as that of the cover data that is carried with itself.

The major requirements [02] that need to be met for development of an effective audio watermarking model are:

- Perceptual Transparency - the host signal shouldn't get distorted or changed in any manner on the addition of the watermark.
- Robustness - the watermark should be undetachable from the host signal irrespective of the attack that the signal is exposed to.
- Security - the watermark shouldn't be detectable by unauthorized parties.
- Data Rate - a high rate is desirable i.e. the number of watermark bits that are embedded successfully into the host signal per unit time.
- Verification and Reliability - the watermark should be able to help identify the rightful owner of the host signal reliably.

The remainder of this paper is organized as follows: Audio watermarking is discussed. Use of spread spectrum in audio watermarking is explained. The most recent works in the area of spread spectrum based audio watermarking are discussed emphasizing the methodology, advantages, and drawbacks. Various spread spectrum based audio watermarking techniques are compared.

**Revised Manuscript Received on May 28, 2019.**

**Bodhvi Gaur**, Department of I&CT, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, INDIA.

**Chandrakala.C.B**, Department of I&CT, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, INDIA.

## II. AUDIO WATERMARKING

Various properties of the human audio system are exploited in order to watermark various forms of audio data. Audio Watermarking [03] can be classified into the following categories:

### *A. Time Domain Watermarking Techniques*

Also known as spatial domain watermarking techniques, these are the ones in which the watermark is embedded into the host data directly without any transformations. It carries the advantage of having the least overhead of processing data although it may not provide much robustness in comparison to its counterparts. Its various types include:

- LSB Coding - replacement of the least significant bit of the host signal is done with the watermark's bit pattern. More suited for image watermarking than audio because quality degradation could occur in the case of audio.
- Phase Coding - since the human auditory system isn't able to clearly detect the relative phase difference of its constituent spectral components, the secret data is replaced with parts of the actual data. This carries the advantage of high tolerance of noise compared to others.
- Echo Hiding - here the embedding of the watermark in the host data occurs with the introduction of an echo which leads to a richer sounding and resonance rich audio which solves the issue of Human Auditory System being sensitive to noise.
- Spread Spectrum - refers to the transmission of narrowband signal distributed over a wider bandwidth which leads to making the data much less detectable as the energy of the signal overlaps. The watermark is spread in a similar way. We will discuss this technique more closely in our paper ahead.
- Modified Audio Signal Keying - modifications are made to the time envelope of the data in order to embed a watermark which is imperceptible.

### *B. Transform Domain Watermarking Techniques*

Embedding of the watermark into the more perceptually significant portions of the data makes this technique more robust and imperceptible. The different type of quantization techniques used in audio watermarking:

- QIM-known as Quantization Index Modulation
- Mean Quantization
- Dither Modulation Quantization
- Vector Quantization

### *C. Compressed Domain Audio Watermarking Technique*

This is based on the fact that audio data can be compressed and the watermark can be embedded into the data in the compressed form itself and need of processing data in the uncompressed form is highly reduced. This is a result of a property of HAS which specify that the decrease in frequency causes decrease in the dynamic range of the auditory system too and so removal of irrelevant parts w.r.t. the perceptual aspect leads to increase in imperceptibility of the distortions.

### *D. Compressed Domain Audio Watermarking Technique*

Combination of various techniques leads to effective utilization of the benefits provided by each and therefore lead to better robustness or imperceptibility of the watermark in the audio data. Additional advantages include a negligible effect on the quality of the original signal.

## III. AUDIO WATERMARKING AND SPREAD SPECTRUM

Spread Spectrum refers to a form of electromagnetic communication in which a signal frequency or bandwidth is spread out over a broad range of frequency band making it more difficult to intercept or jam the signal or to send multiple signals over the same band. It uses much less power as compared to the narrowband transmission where the signal is transmitted at a single frequency. This spreading of the frequencies is done using predefined methods where the transmitter and receiver are informed of the pattern for ease of detecting and despreading of the data sent.

Spread spectrum watermarking, in which the spreading occurs using a pseudo noise (PN) sequence, is good at withstanding interference due to jamming, signal hiding by transmission at low power and achieving secrecy. Spread spectrum watermarking ensures effectiveness, robustness, imperceptible (because the watermark's energy is decently low in the coefficient of a single frequency) and cryptographical security along with transmission at low signal levels. An added advantage is that doesn't require the original audio data in order to extract the watermark. These properties of spread spectrum make it a popular choice for watermarking audio data and its secure transmission. This technique takes the host audio signal as the channel for communication and the watermark as the data to be transmitted. Although it is a popular choice there exist some shortcomings of this technique too. For the watermark to be successfully detected the watermark and the audio data have to be synchronized well and most importantly if the data is successfully attacked somehow then it can lead to a leak of confidential information which can be used to create a copy which can deceive the watermark detector to recognize the attacked data as unmarked. There are many ways in which spread spectrum watermarking can be achieved. Generally, spread spectrum techniques are, thus, divided into three major types:

- Direct Sequence Spread Spectrum - locally defined pseudo-noise sequence is used and the signal is spread using this predefined sequence by multiplying the sequence with the original message. This code is later used by the receiver to extract the original message. It has an upper hand in performance against jamming of signals as it makes the resultant channel more noisy thereby preventing interference.
- Frequency Hopping Spread Spectrum – the shifting of the carrier frequency occurs using a PN sequence (randomly generated or predefined).
- Time Hopping Spread Spectrum - Here code sequence (created using a pseudorandom generator) is used to vary the time period of the Radio Frequency carrier.


Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Spread spectrum watermarking has a wide use in the field of watermarking today be it image watermarking or audio watermarking. The challenges faced using spread spectrum for watermarking would include the increased complexity, increased bandwidth for transmission and the need for synchronization between the transmitter and receiver. In spread spectrum watermarking the actual signal is considered as the communication medium whereas the watermark is treated as the signal transmitted. Various techniques have been covered here which encompass all the advancements that have recently occurred in this field of spread spectrum watermarking with respect to audio data.

## IV. ATTACKS ON AUDIO WATERMARKING

A number of attacks exist in the field of audio watermarking [04] which can, in turn, affect the watermarking technique's robustness. The attacks can majorly be divided into these types:

- Dynamics - refers to the attacks that have an impact on the loudness characteristic of the audio data. These can range from primary attacks like increasing or decreasing to compressing or expanding to modifying frequency.
- Filter - filters are used to allow only the required portions of the signal to pass and block the rest by modifying the spectrum. High pass filter blocks all values below a preset value and Low pass blocks above a value. Equalizers are another category which is similar to filters which are used to vary particular portions of the spectrum.
- Ambience - the broadcast of an audio within a room is dealt with here. Reverb and delay are responsible for any changes made to this.
- Conversion – converting a form of audio to another. It can occur in the form of resampling (sampling at different sampling frequencies) or inversion (the sign of the sample is changed).
- Loss compression - the size of the data is compressed to 10 or even lesser times the actual data and is based on the psychoacoustic effects of audio.
- Noise - this is the effect of a majority of attacks that occur. Noise can be generated due to various sources the major ones being hardware, etc. Generation of noise for the destruction of the watermark is a common attack.
- Modulation - most of the audio handling software deal with modulation effects such as vibrato, amodulation, amplitude modulation, remodulation, etc. and are therefore used for attacking watermarks.
- Time stretch (also known as pitch shift) - they do one of the following: vary the audio's duration without affecting the pitch or vice versa which is done to fine tune or fit the audio into time windows.
- Sample Permutation - attacks performed to make the embedded watermark useless. Eg. permutating sample or dropping them.
- Resampling - downsample to half the original rate of sampling and then unsample.

The purpose of the proof is to check for typesetting or conversion errors and the completeness and accuracy of the text, tables, and figures. Substantial changes in content, e.g., new results, corrected values, title, and authorship, are not possible and cannot be processed.

## V. DISCUSSION ON SPREAD SPECTRUM AUDIO WATERMARKING TECHNIQUE

### *A. An Algorithmic Digital Audio Watermarking in Perceptual Domain Using Direct Sequence Spread Spectrum [05]*

**Methodology.** This paper deals with a model based on psychoacoustic auditory components (mimics the human hearing system) using direct sequence spread spectrum. It focuses on the continuous frequency masking that occurs during a hearing process which is responsible for the production of the final masking threshold which leads to the formation and shaping of the audio watermark which is hard to catch for the human ear. Also, segmentation of the signal occurs in order to overcome the issue of large lengths of the audio signal which are hard to process. Fast Fourier transform is used to convert original signal to frequency domain signal. The masking threshold is decided using spread energy per critical band. The result of the frequency domain is further converted to the time domain. This is the initial processing of the frames of the original data. Embedding happens using both the psychoacoustic model and the DSSS. After deciding the masking threshold for the particular frame of audio with the help of psychoacoustic model with frame size 2048 and FFT of 2048 points, PN sequence is generated which used to modulate the FFT watermark bits. Using this the shaping of the watermark signal is done for imperceptibility. After this watermark is embedded into the actual signal in time domain. The extraction of the watermark and its detection occurs as follows: Linear prediction filtering and linear sequence analysis are both done simultaneously after which the left audio signal is forwarded to the matched filtering procedure where the PN sequence is applied and output of this is the required extracted watermark.

**Advantages.** Attacks withstanding capacity: can successfully withstand echo suppression attack. Gives minimal error in case of Noise addition (White noise 50 dB), amplitude compression and equalization attack. Gives a decent stand against bandpass filtering attack, Compression, resampling (44.1 kHz to 22.05 kHz and back to 41.1kHz), re-quantization (16 bit to 8 bit) and MP3 attack (MPEG 1audio layer 3 compression-96kbps).

### *B. A New Informed Spread Spectrum Embedding for Robust Audio Watermarking [06]*

**Methodology.** A new technique has been discussed in this paper which proposes to introduce distortions in the host in the form of mean square errors in order to fully utilize the potential of the host signal in order to improve the robustness of the embedded watermark by a large extent. The association between the watermark and the host is analyzed to classify three categories which are host-ignored (the host is ignored while embedding), host-rejected(host is cancelled while embedding ) and host-utilized(host is utilized to extract the watermark later on in the process or refrain


Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

from rejecting the host unless required) spread spectrum techniques. Compensation signal modulation is used for achieving these. The usual approach for measuring the extent of distortions is Mean Square Error but in case of the audio signal, this may not suffice because perceptual distortions need to be considered too and therefore the proposed strategy handles this gap. The proposed technique called the WOHI-SS: watermark oriented host inversion spread spectrum exploits the fact that the Human auditory system cannot differentiate between signals that are inverted in terms of absolute phase. In this method, inversion of the host signal occurs based on the relation between the host and the watermark. Random and sudden switches in the phases of the adjacent audio frames will lead to the creation of periodic beat like sounds which can be evaded by soothing out the distortions by making changes to the phase spectrum of every block of the various blocks in the spectrum calculated using L-point short time Fourier Transform. Inverse short time Fourier Transform is used to get the required time domain frames. This is the processing of the host. Experimental results are in agreement with the claims of the authors.

**Advantages.** Robustness improvement. In terms of the perceptual quality, although the proposed technique is responsible for additional noise generation due to the processing of the host, it is taken care of by the soothing operations which reduce the perceptual distortions. In terms of robustness, the proposed technique is effective against attacks like resampling(44.1 kHz to 22.05kHz , then back to 44.1kHz), requantization (16 bit to 8), compressions, low pass filters(cutoff frequency 12 kHz),MP3(compression at bit rate 128 kbps), Advanced Audio Coding(AAC- bit rate 128kbps) and Additive white Gaussian noise(with relative power of -30 dB).

**Shortcomings.** The proposed system may fail to effectively create inaudible perceptual distortions for certain kinds of audios with a large amount of harmonics like instrument pieces.

**Future prospects.** Improvement in the perceptual quality and extension of the technique for other data forms like images or videos.

### *C. Channel Capacity Analysis of the Generalized Spread Spectrum Watermarking in Audio Signals [07]*

**Methodology.** This paper discusses the need for considering channel capacity while presenting a generalized spread spectrum watermarking technique. Channel capacity refers to the highest rate of data that can be transmitted via the channel. In the traditional spread spectrum techniques for audio signals information about the underlying signal, on which the watermark is embedded, isn't used unlike the technique presented here where it's proven that the GSSW outperforms the traditional techniques since it possesses a higher channel capacity than its counterpart. The expression for binary watermark's capacity is also presented. Using the information about the host acts as a compensation for the interference caused by the host signal to the overall watermarked signal thereby aiding the increase in capacity. Host-vector is made up of elements that are independent and belong to a Gaussian distribution. It's mathematically proven that the traditional technique fails to attain a high channel capacity. It is also shown that the highest channel capacity can be obtained when information about the interference is missing. Using the optimal parameter for the expression obtained it's finally proven that GSSW can outperform its counterpart for higher WNR. Larger channel capacity is an indication of a larger rate of transmission.

**Advantages.** Focus on maximizing channel capacity by considering the information of the host too while embedding the watermark, unlike the traditional techniques which don't do so.

**Future prospects.** Building real-time experiments in order to showcase the above-proven results on actual audio data.

### *D. Channel Capacity Analysis of the Multiple Orthogonal Sequence Spread Spectrum Watermarking in Audio Signals [08]*

**Methodology.** This paper deals with maximizing the channel capacity by embedding several orthogonal sequences in parallel watermarks. The most appropriate number of sequences is then decided based on the capacity results obtained for the combination. The traditional Spread spectrum watermarking scheme comes with interference from the host audio which can lead to a very high loss in the robustness of the watermark. To improve this shortcoming an improved spread spectrum watermarking technique was created where projections of the host were used to compensate for the loss but it uses only a single SS sequence. In another technique, a pair of orthogonal sequences were used in order to help in decreasing the interference. Taking inspiration from the two, the authors present a technique where embedding several orthogonal sequences in the host signal for parallel transmission of watermarks can be achieved via various channels. Overall this newly proposed technique proves to provide a higher channel capacity in comparison to the existing improved spread spectrum technique i.e. a better rate of transmission. The host is assumed to have elements that are independent and following the Gaussian distribution. Channel capacities are of two kinds: Continuous and discrete. Discrete channel capacity is maximized by considering the bits/sample as the measure for the transmission rate. Due to the introduction of data loss by the decoder for the watermark bit, capacity wise continuous channels outperform discrete channels. The channel capacity in the proposed method depends on the number of watermark bits (K) that are transmitted in per N audio signal sample. The conclusion is that the channel capacity of the technique proposed above is K times the ISS technique's capacity.

**Advantages.** Maximization of channel capacity by embedding several orthogonal sequences in parallel watermarks.

**Shortcomings.** The probability of bit error can increase but increase in the overall transmission rates of the bits is seen.

### *E. Spread Spectrum Audio Watermarking Using Multiple Orthogonal PN Sequences and Variable Embedding Strengths and Polarities [09]*

**Methodology.** This paper focuses on providing a high capacity of embedding along with high robustness against attacks especially the ones like noise and compression attacks. High computational efficiency is another of its highlights. It ensures that host interference is reduced



Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

to its maximum extent. In the process of watermark extraction, the host data undergoes discrete cosine transform. Then, generation of orthonormal PN sequences occurs where each of these is representative of a sequence of watermark consisting of several bits. This is followed by the embedding of watermark sequence on the host audio along with insertion of orthonormal PN sequences generated above. DCT coefficient modification also occurs. In the process of watermark extraction comparisons of the correlation between the watermark section and orthonormal PN sequence are used to extract watermark bits that were embedded. The previously mentioned spread spectrum audio watermarking techniques aren't able to achieve high capacity of embedding without compromising the robustness. The increase in embedding capacity in this technique is achieved by using several orthogonal PN sequences where each sequence considers various watermark bits. The orthogonal property helps in easy extraction while decoding which further leads to better robustness. The bound on the strength of the watermark, embedding of stronger watermarks on stronger hosts in DCT domain and consistency of the embedding strength's decay with the strength of the host is ensured by this technique. Exploitation of DCT coefficients nature, controlling the strength of the PN sequence is possible which helps the proposed technique prove its laurels. Simulation results, based on around 50 audio clips, clearly testify the above-mentioned claims of higher robustness, higher embedding capacity and computational efficiency. The audio signals are attacked under the following categories: Re-quantization attack, noise attack (15 dB, 10dB), amplitude attack (1.5 times, 2.2times), MP3attack (MPEG1 based compression - 96kbps,64kbps), AAC (Advanced Audio Coding) attack (MPEG4 based compression - 96kbps,64kbps), high pass (100kHz) and low pass (8kHz) filtering attack. The proposed technique can withstand all of these in a much better manner in comparison to its counterparts. It also performs better in comparison to its non-polarized version in terms of robustness. The rate of embedding can be decided by changing the number of PN sequences or the DCT segment's length.

**Advantages**. High robustness along with high capacity of embedding. Specially robust against noise(15 dB, 10dB), compression, re-quantization, amplitude(1.5 times,2.2 times),closed loop , MP3 ( 96kbps,64kbps), AAC(96kbps,64kbps), low pass filtering attack(8kHz) and high pass filtering attack(100kHz).This technique provides high computational efficiency. Decent balance between robustness and imperceptibility is ensured. Unlike code division multiplexing based spread spectrum watermarking schemes where several PN sequences overlap the host, in this technique only one PN sequence is added which leads to much fewer distortions as far as embedding is concerned.

**Shortcomings.** This technique isn't robust against desynchronization attack.

**Future prospects.** Tackling the problem of de-synchronization attack.

### *F. Spread spectrum audio watermarking based on perceptual characteristic aware extraction [10]*

**Methodology.** This technique involves exploitation of the perceptual characteristics of the audio signal that is watermarked before the extraction process uses correlation. Perceptual analysis methods are used in order to know about the embedded watermark's spectral power structure. An extraction technique based on estimation equalization and correlation is used. Perceptual analysis methods long with shaping are used for better estimation of the watermark that was embedded. High capacity of embedding is achieved along with low perceptual distortions. The structure of the embedded watermark is estimated at the receiver's end and the watermark is embedded keeping in mind the auditory distortions. Both of the above-mentioned tasks are done using the perceptual analysis. Perceptual characteristic aware (PCA) is used in the extractor before the watermarked signal proceeds to correlation detection. This proposed technique utilizes each spectral band's advancements towards getting the maximum performance. The contributions as stated by the authors include the following: firstly, a unique extraction technique based on estimation, equalization and correlation which promotes perceptual analysis and equalization for the extractor. Secondly, Perceptual shaping by making use of subband analysis. Thirdly, continuous calculation for obtaining the most optimum value of coefficients of equalization. Lastly, a refined method for perceptual analysis which leads to the choice of the most appropriate spectral components calculation of the masking threshold. The two major parts discussed in the given technique are the embedder possessing high quality and the blind extractor. First, the watermark is estimated at the extractor's end and then the analysis, equalization and correlation extraction is covered.

If the extractor has knowledge about the embedded watermark, its extraction would be an easier task. The nontonal masker, which proves to be of better worth than the tonal masker due to its property of not being influenced by the channel noise or the process of watermarking itself, are used for perceptual analysis and shaping. The experimental results back the author's claim of their model outperforming the ordinary correlation detector.

**Advantages.** This technique is able to withstand the following attacks: AWGN (addition of random white noise-30 dB), Re-quantization (16 bit to 8), Resampling (22.05 kHz), Low pass filtering (10kHz), lossy compression (MP3 and AAC) (64 kbps/channel). It is pretty robust and shows extremely good performance.

**Shortcomings:** If the host and watermark are white Gaussian in nature, this technique has no contribution.
Also, poor performance in case of de-synchronization attacks like cropping and replication is evident. To be able to tackle it manual preprocessing is required where blind extraction may not make much sense.

**Future prospects:** Development of a technique to tackle the issue of desynchronization attacks while retaining the other advantages.

### *G. Spread Spectrum-Based High Embedding Capacity Watermarking Method for Audio Signals [11]*

**Methodology.** In the technique proposed in this paper high embedding capacity is ensured along with decent imperceptibility and robustness. It majorly focuses on increasing the capacity by ensuring the following: firstly, each audio segment has multiple watermarks


Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

embedded in it, secondly, reduction of interference by the host during the embedding process and lastly adaptive adjustment of the amplitude of the PN sequence during embedding of the watermark. It provides an effective way to ensure robustness and imperceptibility along with high embedding capacity without compromising the perceptual quality much unlike the other proposed techniques. The host signal undergoes discrete cosine transformation in order to get the required DCT coefficients. Then those at risk of filtering or compression attacks are removed and the left DCT coefficients are then further used for watermark embedding. Before embedding, the coefficients obtained above undergo segmentation and every segment is then fragmented into a pair. A seed PN sequence is then shifted rotationally to give almost orthogonal PN sequences which also play the role of secret keys which adds an additional aspect of security. Insertion of the PN sequence in the corresponding pair of fragments is then carried out. During this, the amplitude adjustment of the PN sequence ensures the highest perceptual quality. Property of audio segments is an important factor in the adaptive determination of the scaling factor using "analysis by synthesis" technique. This factor if kept too small will sacrifice robustness on one hand and result in increased perceptual quality on the other and so its proper selection is of extreme importance. During the process of extraction, the host interference is brought down to minimal by making use of the similarity between the different fragments of the original data. The simulation results capture the ability of the model against the below-mentioned attacks.

**Advantages.** High imperceptibility and robustness ensured along with increased embedding capacity without compromising with the perceptual quality. Attacks that this technique is able to withstand: closed loop attack, re-quantization attack, noise attack, amplitude attack(from 16 bits to 8), MP3 (MPEG1-layerIII-128 kbps and 96 kbps compressions) attack, AAC attack(MPEG4-128 kbps and 96 kbps compressions), high pass filtering(50 Hz and 100 Hz cut-off) and low pass filtering (12 kHz and 8 kHz cut-off) attack. This technique shows almost 100% rate of detection in the above attacks.

### *H. Host cancelation-based spread spectrum watermarking for audio anti-piracy over Internet [12]*

**Methodology.** This paper deals with the aspect of data redundancy w.r.t. the host signal which possesses a short time stationary property which can be exploited to majorly extract the power of the host using linear prediction filter. The two major aspects, first, Removal of redundancy of the host signal using Levinson Durbin Algorithm and second, improving the modulation of spread spectrum which overall leads to decreasing the distortions, while embedding the watermark, by folds are focused. At the extractor's end also redundancy removal is done in order to obtain matching as far as filtering is concerned and refining the performance. ISS(improved spread spectrum) is considered as a base for comparison of the current technique where ISS uses an informed method neglecting the redundancy of data which leads to the introduction of localized distortions which can further lead to low perceptual quality of the host, the current technique exploits redundancy to provide high transparency and robustness. The host audio is assumed to consist of a Deterministic(DC) and a Stochastic component(SC).DC consists of the vital power of the audio and is made of data that can be determined priory whereas SC has data that doesn't depend on the prior samples and can be thought of as having a near Gaussian property. Embedding distortions decrease the localized distortions while embedding and host suppression contribute to a higher performance. The core idea revolves around combining cancellation of host interference with redundancy removal at the embedder and extractor end. A two-step correlation technique is used in order to ensure robustness-traditional detector used for correlation and pre-processing with the help of linear prediction filtering. The computational complexity is taken care of by the FFT (Fast Fourier transform) used in the perceptual analysis. Parameters used for evaluation of the given technique is signal to watermark ratio and objective difference grade.

**Advantages.** High robustness and accuracy along with a high performance w.r.t. the extraction process. Around 50 audio signal samples were used for testing against the following attacks: White noise attack(20 dB SNR ratio Gaussian channel used),Amplitude scaling(200%),Low pass filtering(10-kHz),Resampling(44.1 kHz to 22.05 kHz then back to 41.1 kHz), Re-quantization (16 bit to 8 bit),MP3 compression(64kbps) and AAC compression(64kbps). It was able to withstand all of these decently. For covert communications, the audio can overcome extremely poor communication channels without compromising the security of the data. Also, unauthorized access to audios will not allow removal of watermark without compromising the host signal's quality.

## VI. CONCLUSION

Spread Spectrum Audio Watermarking holds in itself abundant potential for research opportunities due to its wide range of benefits and practical applications in today's scenario. To the best of our knowledge, this paper covers the recent technological advancements in this field with respect to their methodologies, the pros, the cons and the applications. The techniques mentioned above aim at improving the main characteristics responsible for an secure and efficient transmission of the watermarked audio signals such as , channel capacity, robustness, imperceptibility, performance, ability to withstand the widest variety of attacks, maintaining perceptual quality, securing the data being transmitted, increased computational efficiency , reduced host interference and reduced distortions. The techniques discussed include exploring data redundancy of the host signal, multiple watermark embeddings in a single audio segment along with adaptive adjustment of PN sequence, perceptual characteristics of the audio signal, embedding several orthogonal sequences in parallel watermarks, improvement in the channel capacity by considering the information of the host, etc. Challenges that provide ground for future work include improvement of perceptual quality for techniques involving introduction of distortions in the host in order to improve robustness, performing real-time experiments to verify and further the results favoring generalized spread spectrum watermarking considering channel capacity over its counterparts, minimizing the increased probability of bit error on maximization of channel capacity by embedding multiple


Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

orthogonal sequences in parallel, tackling de-synchronization attack without compromising with the other advantages that the spread spectrum audio watermarking techniques offers, and extension of the scope of the techniques discussed in the paper for videos. This paper will act as a stepping stone for the researchers interested in working in this field to explore the vast potentials of this field and overcome the research gaps that still prevail.

**Table 1.** Advantages and Shortcomings of Spread Spectrum based Audio Watermarking Techniques

| Reference Number | Advantages | Shortcomings |
|---|---|---|
| [05] | **Robustness (Withstands)**<br>• Echo suppression attack<br>• Noise addition(White noise 50 dB)<br>• Amplitude compression equalization attack<br>• Pass filtering attack<br>• Compression (MP3 attack-MPEG 1audio layer 3 compression-96kbps)<br>• Resampling(44.1 kHz to 22.05kHz and back to 41.1kHz)<br>• Requantization (16 bit to 8 bit) | |
| [06] | **Robustness (Withstands)**<br>• Resampling(44.1 kHz to 22.05 kHz, then back to 44.1kHz)<br>• Requantization(16 bit to 8)<br>• Compression<br>• Low pass filters(cutoff frequency 12 kHz)<br>• MP3(compression at bit rate 128 kbps)<br>• Advanced Audio Coding(AAC- bit rate 128 kbps)<br>• Additive white Gaussian noise (with the relative power of -30 dB) | **Perceptual quality**<br>• Additional noise generation due to the processing of the host<br>Handled by the soothing operations which reduce the perceptual distortions |
| [07] | **Maximizes channel capacity** | |
| [08] | **Robustness (Withstands)**<br>• White noise attack (20 dB SNR ratio Gaussian channel used)<br>• Amplitude scaling (200%)<br>• Low pass filtering(10-kHz)<br>• Resampling(44.1 kHz to 22.05 kHz then back to 41.1 kHz)<br>• Requantization(16 bit to 8 bit)<br>• MP3 compression (64kbps)<br>• AAC compression(64kbps) | **Bit error per channel**<br>May increase when the distortions in the audio are less. |
| [09] | **Robustness (Withstands) :**<br>• Noise(15 dB, 10dB)<br>• Compression<br>• Requantization<br>• Amplitude(1.5 times,2.2 times)<br>• Closed loop<br>• MP3(96kbps,64kbps)<br>• AAC (96kbps,64kbps)<br>• Low pass filtering attack (8kHz)<br>• High pass filterng attack (100kHz).<br>**High computational efficiency**<br>**Fewer distortions**<br>Embedding(since only one PN sequence added) | Cannot withstand desynchronization attack |
| [10] | **Robustness (Withstands) :**<br>• AWGN(addition of random white noise-30 dB)<br>• Requantisation(16 bit to 8)<br>• Resampling(22.05 kHz)<br>• Low pass filtering (10kHz)<br>Lossy compression (MP3 and AAC-64 kbps/chan) | |
| [11] | **Robustness (Withstands):**<br>• Closed loop attack<br>• Re-quantization attack<br>• Noise attack<br>• Amplitude attack(from 16 bits to 8)<br>• MP3 (MPEG1-layerIII-128 kbps and 96 kbps compressions) attack<br>• AAC attack(MPEG4-128 kbps and 96 kbps compressions)<br>• High pass filtering(50 Hz and 100 Hz cut-off)<br>• Low pass filtering (12 kHz and 8 kHz cut-off) Attack<br>**High Imperceptibility**<br>**Increased embedding capacity**<br>Without compromising the perceptual quality | |
| [12] | **Robustness (Withstands)**<br>• White noise attack (20 dB SNR ratio Gaussian channel used)<br>• Amplitude scaling (200%)<br>• Low pass filtering (10-kHz) | |


Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

| | | |
|---|---|---|
| | • Resampling (44.1 to 22.05 kHz then back to 41.1 kHz)<br>• Requantization (16 bit to 8 bit)<br>• MP3 compression (64kbps)<br>• AAC compression (64kbps) | |

## REFERENCES


1. R. F. Olanrewaju and O. Khalifa, "Digital audio watermarking; techniques and applications," *International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, 2012, pp.830-835
2. S. P. Singh Chauhan and S. A. M. Rizvi, "A survey: Digital audio watermarking techniques and applications," *4th International Conference on Computer and Communication Technology (ICCCT)*,Allahabad,2013,pp.185-192.
3. Jain R., Trivedi M.C., Tiwari S. "Digital Audio Watermarking: A Survey "*Advances in Intelligent Systems and Computing*, vol 554. Springer, Singapore, 2018.
4. Nematollahi M.A., Vorakulpipat C., Rosales H.G. "Audio Watermarking. In: Digital Watermarking", *Springer Topics in Signal Processing*, vol 11. Springer, Singapore, 2017
5. ] N. K. Trivedi and M. C. Trivedi, "An Algorithmic Digital Audio Watermarking in Perceptual Domain Using Direct Sequence Spread Spectrum," *2014 Fourth International Conference on Communication Systems and Network Technologies*, Bhopal, 2014, pp. 805-809.
6. P. Zhang, Y. Li, J. Jiang, Y. Fan, Q. Hao and X. Ma, "A new informed spread spectrum embedding for robust audio watermarking," *2016 IEEE 13th International Conference on Signal Processing (ICSP)*, Chengdu,2016,pp.282-286.
7. Y. Zhang, Z. Xu and B. Huang, "Channel Capacity Analysis of the Generalized Spread Spectrum Watermarking in Audio Signals," in *IEEE Signal Processing Letters*, vol. 22, no. 5, pp. 519-523, May 2015.
8. Z. Xu, C. Ao and B. Huang, "Channel Capacity Analysis of the Multiple Orthogonal Sequence Spread Spectrum Watermarking in Audio Signals," in *IEEE Signal Processing Letters*, vol. 23, no. 1, pp. 20-24, Jan.2016.
9. Y. Xiang, I. Natgunanathan, D. Peng, G. Hua and B. Liu, "Spread Spectrum Audio Watermarking Using Multiple Orthogonal PN Sequences and Variable Embedding Strengths and Polarities," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 529-539, March 2018.
10. R. Li, S. Xu and H. Yang, "Spread spectrum audio watermarking based on perceptual characteristic aware extraction," in *IET Signal Processing*, vol. 10, no. 3, pp. 266-273, 5 2016.
11. Y. Xiang, I. Natgunanathan, Y. Rong and S. Guo, "Spread Spectrum-Based High Embedding Capacity Watermarking Method for Audio Signals," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2228-2237, Dec. 2015.
12. Li, R., Xu, S., Rong, B., and Yang, H. "Host cancelation-based spread spectrum watermarking for audio anti-piracy over Internet", in *Security Comm. Networks, 9: pp.4691–4702, 2016.*


## AUTHORS PROFILE

**Bodhvi Gaur** is a final year undergraduate student in the Department of Information and Communication Technology, Manipal Institute of Technology, Manipal University, India. Her research interests include Natural Language Processing, Artificial Intelligence, Machine Learning and Deep Learning. She is an active Student Member of ACM.

**Chandrakala C B** is currently working as Associate Professor in the Department of Information and Communication Technology, Manipal Institute of Technology, MAHE,Manipal, India. Her research areas are Mobile Ad Hoc Networks, Distributed Computing, Software Engineering. She has around 20 years of teaching experience.



*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*