Blind Audio Watermarking schemes: A Literature Review

Harleen Kaur
Student Masters of Technology,
Department of Computer Science,
Sri Guru Granth Sahib World University
Fatehgarh Sahib, India

Usvir Kaur
Assistant Professor,
Department of Computer Science,
Sri Guru Granth Sahib World University
Fatehgarh Sahib, India

Abstract: **This paper presents the review of audio watermarking algorithms based on blind watermark detection techniques which present the watermarking scheme these algorithms use and results obtained. Most of the algorithms focused on improving robustness and imperceptibility, but newly proposed algorithms also focus on other parameters like embedding capacity and security.**

Keywords: audio watermarking; blind; imperceptibility; robustness.

I. INTRODUCTION

Digital watermarking schemes currently are the most interesting and popular research area for copyright protection of the multimedia data, where a lot of work has been done and is still being developed for finding more improved methods to enhance the security, robustness and quality of watermarked data. Various techniques has been used to apply watermarks to the audio signals which are generally based on two approaches: (a) *non-blind (or informed) watermarking*: This approach requires the original signal information to recover watermark from the watermarked signal and (b) *blind watermarking*: This approach does not require original signal information to recover the watermark from the watermarked signal. In recent years more research is focused towards developing blind watermarking techniques to avoid the use of original signal information to extract watermark information as the original information is always available to us.

II. PERFORMANCE EVALUATION MEASURES

An algorithm's performance is measured in terms of different metrics to check for different requirements they satisfy or not. For e.g. to check for perceptual audio quality assessment, metrics such as Signal-to-Noise ratio(SNR) and segmental Signal-to-Noise ratio (segSNR), Cepstral Distance (CD), Log-Likelihood Ratio (LLR) for objective evaluation test. Under subjective listening tests measures like MOS(Mean opinion Score), SDG(Subjective Difference Grade) are used. SNR [35] is a statistical difference metric which is used to measure the perceptual similarity between the undistorted original and the distorted watermarked audio signal and SegSNR is defined as the average of the SNR values of short segments of the watermarked signal and used to give an objective indication of the watermarking effect on the original audio. The CD [35] measures cepstral coefficients computed from the original and watermarked signals on a frame-by-frame basis, and this has the capability for accurately predicting subjective listening scores. LLR based on maximum likelihood detection used for detecting manipulations done on signal must lie on zero or close to zero which indicates good imperceptibility. SDG [35] is the difference between the absolute quality grade assigned to the distorted and the original signals, both measured on the five-grade impairment scale defined in ITU-R BS.562. If all listeners identify the hidden reference correctly, the SDG is identical to the grade on the five-grade impairment-scale assigned to the processed signal. A similar measure specified by ITU-T recommendation P.800 is MOS. To check robustness, measures such as Bit error rate (BER) and Normalized cross-correlation (NC) are used. BER [26] is used to evaluate the watermark detection accuracy after signal processing operations and a value of 0 (very close to zero) is required for good robustness. NC [26] is used to evaluate the similarity between the original watermark and the extracted watermark.

Choice of performance parameters chosen depend on the type of application on which algorithm depends. Generally robustness and imperceptibility are the main requirements for most of the applications.

III. LITERATURE SURVEY:

Keeping in view the importance of blind watermarking, this section presents the review of papers, which were based on blind detection mechanism of watermark from the watermarked signal. Earlier audio watermarking techniques were not very robust and watermarked signal doesn't have good imperceptibility level and progress was very slow which is discussed here:

A robust audio watermarking scheme using blind watermark detection approach earlier was presented by Bassia et. al. [1] in the time domain which was based on direct modification of the amplitude values in such a way that it does not produce any perceptual difference. A watermark key was used for generating the watermark to be embedded into the audio signal which is only known to copyright owner. Robustness could be increased if watermark with high amplitude values was used but with this a perceptible distortion was introduced. Watermarking scheme presented was statistically imperceptible and it was able to resist MPEG compression, rescaling, filtering and requantization.

Huang et. al. [2] presented blind watermarking algorithm based on synchronization codes and error correcting code (e.g. BCH) applied to watermark to lower the bit error rate. Embedding of the synchronization codes in time domain while the watermark in block-based DCT coefficients, made the searching of the synchronization codes easy and the watermark robust. The imperceptible watermark was found to be robust against attacks like additive noise, MP3 coding, and cropping.

Hsieh et. al. [3] proposed blind watermarking approach in cepstrum domain based on energy feature in the time domain to solve the issue of synchronization problem of losing location of watermark in case of attacks like cropping, shifting etc. The scheme was robust against the Mp3 attack and all kinds of distortion attacks such as pitch-shifting, and cut samples, but still the scheme was inefficient in case of delay attacks.

Liu et. al. [4] proposed a scheme in which watermark were represented by sinusoidal patterns consisting of phase-modulated sinusoids by elements of pseudorandom sequences.. The watermark detection is completed by computing the correlations between watermarked signal and pseudorandom sequences using FFT properties, and comparing the correlations with a threshold to determine whether watermark is present in the signal. Experiment results demonstrated that sinusoidal patterns generated based on pseudorandom sequences keep the same correlation properties with those of pseudorandom sequences. The effectiveness both in audibility and robustness of the proposed method using sinusoidal patterns was tested and The tested results indicated that the sinusoidal pattern is an effective format for using spread-spectrum technology in audio watermarking. However, technical issues for some signal manipulations involving nonlinear transforms were unsolved.

The scheme based on Peak Point Extraction for synchronization was proposed for achieving the blind watermark extraction mechanism and use of error correcting codes to improve robustness by Wei et. al.[5] The scheme used a three stage process for embedding watermark into the original signal comprising the encoding of signature (watermark) into specific binary format, finding location for synchronization points(which is the last point of every peak point group) and embedding repeatedly in discrete cosine transform (DCT) and a three stage process to extract watermark from the watermarked signal comprising of finding synchronization locations, extracting the bit information and recovering the symbols representing the watermark. Good Imperceptibility of the added watermark was achieved as the scheme utilized the property of less sensitivity to phase change. The proposed scheme was found robust against attacks, especially cropping and re-sampling, but was not good enough for additive noise, low-pass filtering and mp3 compression.

Soon Blind watermarking schemes started gaining importance and progress in this area started with more paces, more and more techniques especially based on transform domain methods were developed. Here are discussed some important techniques which were developed during this phase:

Audio watermarking scheme embedding synchronization codes and hidden informative data into the low frequency coefficients in DWT (discrete wavelet transform) domain to achieve robustness against common signal processing such as resampling, requantization, cropping, MP3 compression and gaussian noise corruption was proposed by Wu. et. al. [6]. Performance was measured using SNR and BER calculations which revealed that embedding strength is greatly depended on the type and magnitudes of the original audio signals.

Wang and Zhao [7] proposed algorithm which utilized the properties of DWT and DCT for watermark embedding. The algorithm divides the original audio signal into segments and cuts each segment into two sections. Barker code used as synchronization code is embedded into first section using mean value of samples. DWT and DCT transformation are applied to second section and watermark is embedded into low frequency components by quantization. Inverse operations are applied to get watermarked signal. Experimental results obtained from calculation of NC, BER and SNR over two different length signals revealed that Algorithm was inaudible and robust to common signal manipulations such as noise, resampling, requantization, cropping and MP3 compression but not to pitch invariant time scale modification attack.

In the method proposed by Z. Li et. al[8], Statistic characteristics of host audio are used for generating private key and which helped in blind watermark detection. Watermark is embedded into high energy coefficients of DWT with specific intensities. Independent component analysis (ICA) is done to extract watermark and to realize true blind detection mechanism. Also multi-embedded watermark can be detected and solve the problem of multi-embedded watermark problem using time tag. SNR and NC calculations revealed watermarked signal has good imperceptibility and is robust against common signal processing such as cropping, low pass filter, noise, compression, scaling, shift etc.

The method by Z. Xu et. al. [9] uses a remote sensing image (like satellite image) as watermark and adopts the half toning error diffusion method in the preprocessing step for converting watermark into suitable format . Chaotic sequence is embedded along with watermark information to make the scheme self- synchronized. Algorithm quantizes the lowest frequency coefficients of DWT of audio signal so that watermark detection can be done in a blind way. Bit error code rate (BRE) and SNR (for attacks performed by stirmark audio attacking software) calculations revealed that algorithm robust against noise, MP3 compression etc. and familiar attacks, but it was not robust to high-pass and resample.

Blind watermarking scheme based on convolution code and embedding domain with lifting wavelet was applied by D. Xu et. al. [10]. Watermark information is encoded with convolution codes and interleaved to resist disturbance and is embedded into the low-middle frequency coefficients in the lifting wavelet. During watermark detection soft decision viterbi decoder is used to improve robustness of watermarking system. Also to improve security, watermark is encrypted using chaotic sequence. SNR was found to be greater than 20 db for this scheme. Experimental results with BER and NC values proved that soft decision decoding is better than hard decision decoding for watermark extraction.

In the scheme presented by X. Wang et. al. [11], template information and watermark signal is embedded into low – middle DCT coefficients of DWT- decomposed original signal by adaptive quantization according to local audio correlation and human auditory masking. For watermark extraction, firstly corresponding features of template and watermark are extracted from watermarked audio and are selected as training sample to train SVR. Actual outputs are predicted according to corresponding feature and thus digital watermark is recovered using well-trained SVR. NC, PSNR, distortion ratio calculations revealed that scheme was inaudible and robust to some common signal processing operations. Its performance was better than SVM watermarking schemes; involve easy calculations and easy implementation, which enhances practicality and application value.

The algorithm proposed by H. Liu et. al. [12] embeds the watermark information into absolute average of audio low frequency coefficients in the DWT and error correcting code is applied to lower the bit error rate of watermark in extraction. SNR, NC, BER calculations shows that value of SNR is 23.342 for watermarked signal and method is robust to common signal processing attacks and especially robust to de-synchronization attacks. Also if error rate is greater than 20% watermark cannot be extracted correctly, so repetition code shouldn't be applied in that case.

Charfeddine Maha et. al. [13] proposed a method based on Human psychoacoustic model (HPM), DWT, Neural Network (NN) and error correcting code. Frequency perceptual watermarking property is utilized to make watermark inaudible. To assure watermark embedding and extraction, neural networks is used to memorize the relationships between a wavelet central sample and its neighbors. To increase robustness watermark is encoded using Hamming error correction code and embedded into DWT transformed audio signal. The experiments results demonstrate that using Hamming error improves the NC value. In fact, detection errors introduced by some attacks were refined. The Hamming code permit to correct one error in 8 bits and help then to overcome the part of the corruption of the watermark, thus the NC values of the watermark is higher than the one without the error correcting code.

Bhat K et. al. [14] proposed a method utilizing Cepstrum transform and DWT transform properties and watermark encoded using BCH coding. In this scheme, original audio is segmented into number of frames equal to size of BCH-coded watermark, transformed into cepstrum domain and decomposed into 2-level wavelet transform and watermark is embedded. This technique takes advantage of error correcting code and cepstrum transform to lower the bit error rate of the extracted watermark. Results indicate that scheme achieves better results than non-BCH techniques

The method proposed by Y. Wang et. al. [15] segments the original audio into pieces and each piece is divided into 3 parts and each part is embedded with synchronous code, serial number and chaotic-encrypted watermark respectively in DCT domain in low frequency components using quantization. BER and NC calculations revealed that algorithm was robust to attacks such as adding white noise, LP filtering, requantization and random cutting, but algorithm involved huge quantity of calculation and was not robust to resampling attack.

To solve the problem of de-synchronization attacks, scheme based on utilizing statistic characteristics of audio and synchronization code techniques was proposed by X. Wang et. al [16]. Audio is segmented and each segment is cut into two sections and synchronization code is embedded I first section using spatial techniques and second is transformed using DWT and watermark is embedded into statistic average value of low frequency components. BER and PSNR calculations revealed that scheme was inaudible and robust against common signal manipulations and de-synchronization attacks such as random cropping, amplitude variance, pitch-shifting, time-scale modification and jittering.

A novel approach based on DFRST was proposed by Fan and Wang [17] which utilizes chaotic sequences to improve the security. To Obtain the real watermarked audio DFRST was implemented FFT was chosen, middle frequency components were chosen for applying DFRST to have good imperceptibility and from these largest amplitude values were chosen to embed scrambled watermark bits. Tests were performed with NC, BER, SNR calculations on several audio pieces such as bagpipe music, classical, country, dance, jazz and rock music and also embedding capacity, false alarm and false rejection analysis and scheme was found to be secure, imperceptible and robust to many attacks but it was not robust against random cropping and time scale modification attacks.

A new multiple-audio watermarking algorithm applying DS-CDMA was proposed by Y. Chen et. al [18]. An orthogonal gold sequence set is generated using principle of DS-CDMA communication used to spread spectrum of two different watermarks. These CDMA-encrypted messages are combined and, then mixed and encrypted by linear instantaneous mixing model and then embedded into DWT approximate coefficients of the original audio. Blind Extraction is carried out using Fast ICA and secret keys. All watermarks can be embedded without considering embedding order. Method achieves satisfactory security and has certain robustness to against audio signal manipulations. Watermarks can be detected with low bit error rate.

Algorithm based on watermark embedding in MP3 compressed signal and using psychoacoustic model and Gaussian distribution analysis to adaptively control watermark parameters was proposed by B. Chen et. al. [19]. Watermark is embedded adaptively and inaudibly after MDCT and before quantization. The experimental results proved that this new watermarking algorithm was robust against MP3 compression when the compression bit rate is better than 48 kbps and can survive most common attacks.

A scheme based on embedding watermark bit stream into on-silent samples of stereo signals by replacing least significant bits or bits in higher layer of original audio layers was proposed by W.Cao et . al. [20]. SNR and robustness measures at different extraction rates against different attacks for mono, stereo and speech signals were calculated and it was found that this scheme was not robust to MP3 compression and re-sampling attack if re-sampling frequency was not integer multiple of original sampling frequency, However this scheme was found robust to attacks than mono-signal watermarking scheme.

X. Liu et. al. [21] proposed a methodology utilizing features of DWT transform in which audio signal was decomposed into frames and each frame is decomposed with 3-level DWT, then energy and the maximal peaks of all its sub-bands were extracted as local features. SVR technique was utilized to model relationship between local features and embedding strength of audio frame to make scheme adaptively control the embedding strength. Due to good leaning ability of SVR, watermark can be exactly recovered unless attacked severely. It was found robust to filtering, lossy compression, re-sampling, re-quantization etc.

H. Peng et. al. [22] proposed a method in which the audio signal was partitioned into audio frames, and the watermark was embedded in quantified low frequency coefficients of wavelet domain (DWT with 3 levels). For each audio frame, the energy and the maximal peaks of its all sub-bands were extracted as the local features, and SVR was used to model the relationship between the local features and the embedding strength of the audio frame in order to adaptively control the embedding strength of the audio frame. Due to the good learning ability of SVR, the watermark can be correctly extracted under several different attacks. Due to the good learning ability of SVR, the watermark could be exactly recovered unless the watermarked audio signal was attacked severely. The experimental results show that the proposed method possesses significant robustness to be against the various attacks.

Multi-watermarking scheme was proposed by Xu and Yang [23] in which watermark is scrambled and scrambling parameters are used for generating private key which is utilized for watermark extraction. Scrambled audio was segmented and on each frame 2-level DWT was applied and then on approximate coefficients DCT. Largest coefficients of DCT were chosen for Watermark embedding. After performing inverse of above operation watermarked audio is treated as new original file and same procedure is followed using new scrambling key. Tests were performed with transparency measures: SNRseg, ODG (Objective Difference Grade) and robustness measures: NC and BER. Subjective listening test indicate that there is no perceptible distortion in watermarked audio with multiple watermarking signal. Extracted watermarks have low BER and high NC and scheme was found robust to various signal manipulations as well as malicious substitutions.

M. K. Dutta et. al. [24] presented a novel scheme based on extracting biometric features to generate watermark key which automatically addresses the issue of authentication and copyright protection was proposed. This Bio-key is embedded into thus high energy bands of wavelet. SNR, NC and BER, MOS(Mean opinion score) calculations on different music samples such as Flute, classical, Blues, Pop, Country etc. revealed that this method was not suitable for vocal or classical music as it does not involve percussion instruments. NC of the extracted bio-key was found to be more than 0.9 for the signal processing attacks.

Algorithm utilizing Time –Domain and FFT domain proposed by D. Megias et. al. [25] embeds synchronization marks in time domain and watermark information into frequency domain taking into account tuning factors. Capacity, Tuning parameters, Robustness and Transparency measures (with sound quality assessment material, full songs and human voice) in terms of ODG were used in calculations. Algorithm has fast execution time and good transparency while achieving robustness against typical attacks making it suitable for real-time applications like broadcast monitoring.

In the scheme proposed by V. Bhat et. al. [26], synchronization code is embedded into first part watermark was embedded into SVs of the segmented approximate wavelet coefficients after applying wavelet transform to second part of 2-part divided audio file based on QIM. Only quantization parameters are used during watermark extraction. MOS, Subjective grade (SG), SNR, NC and BER, security measures on different audio signals (like classic, pop etc.) revealed that Scheme was found robust to MP3 compression, cropping, LP filtering, additive noise, resampling, requantization, echo addition, and denoising. False negative and false positive error probabilities are very low.

D. WU et. al.[27] proposed a technique in which BCH encoded watermark sequence is embedded into coefficients of DWT-decomposed signal (decomposed performed by Haar wavelet bases). The technique is advantageous as it utilizes the error correction capacity of BCH code to increase robustness and consequently diminishes the intensity of watermarking. SNR value of 42.878 for watermarked data indicate that the watermarked audio has very high perceptual quality. The embedded watermark was found robust against various types of attacks such as noise, re-sampling, re-quantizing and MP3 encoding, cutting etc by using the proposed technique. However, it was not excellent algorithm for the asynchronous attacks.

Algorithm utilizing three key techniques polyphase filterbank decomposition, the psychoacoustic model and the empirical mode decomposition (EMD) was proposed by L. Wang et. al. [28]. The analysis filterbank was used to decompose the host audio signal into multiple subbands, and each of the subbands was embedded with a unique watermark message. Within each of the subbands, the signal was firstly segmented and EMD is applied to each of the segments. The watermark bits are embedded into the final residue extracted by the EMD process. The inaudibility of the watermarks is guaranteed with the use of the psychoacoustic model. The system performance (in terms of the bit error rate) is highly related with the watermark strength, which depends on the masking threshold for the particular audio segment. 4 watermarks were embedded by the author. Listening test on 20 persons and SDG, BER calculations revealed that when embedded with multiple watermarks, the subjective quality evaluation could not provide promising results (the resulting watermarking system would produce a noticeable distortion while 4 watermark messages were embedded). The proposed blind audio watermarking scheme is proved to be robust against MP3 compression and adding Gaussian noise attacks. However this method may not be robust to some other attacks such as band-pass filtering and cropping.

Y. Bai et. al. [29] proposed a methodology in which an area of the entire audio signal was selected for watermark embedding and is designated by adaptive adjustment. This area was segmented, FFT was applied to every segment and finally watermark was embedded into dc coefficients (has good HAS properties). Synchronization code was introduced to confirm the initial position of the watermark. Watermark was scrambled using Arnold transform to increase robustness. BER and NC were found to be 0 and 1 respectively for attacks of requantization, resampling, additive noise, LP filtering, echo and for mp3 compression (except at 64, 48, and 32 kbps). Experimental results with SNR, listening test on 5 persons, BER and NC revealed that the watermarked audio signal is imperceptible and robust against various attacks.

A fragile watermarking scheme based on SVD and DWT transform was proposed by W. Jiang [30]. DWT transform was applied to every frame of audio. Chaos encrypted watermark was embedded in the descending dimension in the singular values obtained after applying SVD decomposition to the approximate coefficient values of wavelet. SNR Results demonstrate that algorithm was fragile and could be used for checking the authenticity and integrity of information; however this algorithm cannot remove temper. SDG grades were found to be zero which validates watermarked audio inaudibility.

Algorithm was designed by Huang and Jiang [31] to be fragile and adaptive. Watermark image was encrypted by chaotic sequence, modulated by logistic map to increase security and secrecy. MDFT is applied to selected frames of audio signal chosen by pseudorandom sequence. Watermarking bits were embedded into N/2th Modified DFT coefficients making the scheme strongly adaptive to signal energy. Unitary transform energy conservation property of MDFT makes this method adaptive. NC, SNR and SDG measures revealed that scheme was fragile and was suitable for checking the signal integrity.

The method proposed by Maha et. al. [32] chooses a set of set of random blocks of audio signals for watermark embedding. Hamming correction codes are embedded along with watermark to increase robustness. Watermark embedding and extraction are based on neural network architecture which establishes relationships between central sample and its 8- neighbours. Chosen watermark could be image or text. SNR level achieved was 43.10 which indicate good performance level. NC, listening test on 10 persons, SNR tests revealed that Algorithm has good imperceptibility and was robust to various attacks like MP3 compression with different compression rates and stirmark attacks (e.g. rc_highpass, amplify, compressor), but method was not able to resist echo and fft_invert attack.

Methodology based on reduced SVD by J. Wang et. al. [33] mainly focuses on preserving audio fidelity through threshold distortion technique and which is further supplemented by distortion suppression utilizing psychoacoustic principles. Audio signal is decomposed into frames, each frame’s spectral magnitudes is organized into 2D form and RSVD is applied to obtain a unitary matrix and watermark information is embedded into it utilizing a spectrum distortion control and minimization process. Perceptual distortion suppression is performed and signal is reconstructed using overlap-add procedure into time domain to obtain the watermarked signal. SDG, SNR, mean and standard deviation of precision against attacks tests were performed and results demonstrated that the proposed scheme has a high data payload and is imperceptible and also robust against MP3 compression at various bit rates as well as other selected attacks.

In current times, some more improved and new algorithms have been developed which provides a good direction towards new approaches for audio watermarking and suggest some was to improve the proposed algorithms:
Algorithm based on combining moment invariance features and wavelet features having excellent characteristics such as ability to capture local information, robustness against common signal processing operations and linear relationship between signal and its wavelet moments etc. was proposed by X. Wang et. al. [34]. Audio is segmented and each segment is cut into two parts, with synchronization code embedded in the first utilizing spatial domain technique and converting the second part into the 2D form and then watermark bits are embedded into average value of modulus of the low-order of wavelet moments. PSNR, Perceptual Audio Quality Measure(PAQM)- listening test on 10 persons, SG, BER tests indicated that watermarking algorithm has good auditory quality and reasonable resistance against most common signal manipulations and desynchronization attacks. DA/AD conversion (a common signal processing operation) and the TSM (Time-scale Modification) attacks were still challenging issues.

Algorithm based on DCT and SVD domain by B. Y. Lie et. al. [35] embeds the bit information in the high frequency coefficients of the blocks of SVs obtained by applying SVD transformation to the audio blocks. Adaptive frequency masking properties are utilized to control the embedded local watermark. Chaotic sequences were used as synchronization codes to increase the security and to find the location of watermark. Objective evaluation test using measures such as segmental SNR, Cepstral Distance (CD), Log-Likelihood Ratio (LLR), Subjective Listening test measures: Subjective Difference Grade(SDG), MOS for checking watermarked audio quality, BER and NC (against standard stirmark attacks.) for checking robustness, checking of Embedding capacity in bits per second was performed. Different audio quality test revealed that this scheme has good imperceptibility. Approximate SNR level achieved was 32.53. True capacity calculated was found to be 43 bits/s. Synchronization increases the robustness against cropping and shifting attacks. Embedding capacity is satisfactory and can be further improved in future.

H. Yang et. al. [36] presented in which firstly, the wavelet de-noising is performed on the original host audio, the de-noised digital audio is segmented, and then each segment is cut into two parts. Secondly, with the spatial watermarking technique, synchronization code is embedded into the statistics average value of audio samples in the first part. And then, the higher-order statistics are obtained by using the Hausdorff distance. Finally, the digital watermark scrambled using Arnold transform is embedded into the original audio signal in wavelet domain by using the higher order statistics (effective tool for analyzing multichannel signals). Simulation results show that the proposed watermarking scheme is not only inaudible and robust against common signal processing such as MP3 compression, noise addition, resampling, re-quantization, etc., but also robust against the desynchronization attacks such as random cropping, amplitude variation, pitch shifting, jittering, etc. Future work will focus on improving robustness against some serious TSM attacks.

Algorithm designed by Y. Zhao [37] was to resist air channel resistance along with common signal processing manipulations. Barker code was utilized for synchronization purpose and for the complete embedding low frequency coefficients of DCT adjacent to each other were chosen to be modified. BER calculations for specific testing environment (e.g. with different distances like 0.5 m, 1.5m, 3 m ,volumes with 30% and 70% and in different directions with 0, 30, 60, 90 degrees etc.) were performed. Low error rates were found in different testing environment. This method was able to effectively resist attacks cropping, Mp3 compression, and 30% of resample and interference with the air channel. Embedding capacity is also large for this algorithm.

B. Lei et. al. [38] proposed a method utilizing the features of LWT, SVD, synchronization code technique and QIM (Quantization Index Module). This method not only focuses on the robustness, imperceptibility, but also on security, payload and time complexity and achieves all these requirements satisfactorily. Initially LWT decomposition is performed on the audio segments and approximate coefficients are obtained which are divided into non-overlapping blocks and watermark sequence is embedded into low-frequency subband. SVD is performed on each block and random chaotic sequence encrypted watermark is embedded into SVs using QIM technique. The performance analyses and comparison results indicate that the proposed watermarking scheme maintains good audio quality and high robustness against various attacks and achieves a satisfactory compromise between robustness, payload, imperceptibility and time complexity.

Methodology combining the robustness of vector norm and the approximation components of the discrete wavelet transform (DWT), achieves a blind and adaptive audio watermarking algorithm proposed by X. Wang et. al. [39]. In order to improve the robustness and imperceptibility, a binary image encrypted by Arnold transform as watermark is embedded in the vector norm of the segmented approximation components, the count of which depends on the size of the watermark image, after DWT of the original audio signal through quantization index modulation (QIM) with an adaptive quantization step selection scheme. Moreover, a detailed method has been designed to search the suitable quantization step parameters. Experimental results indicate that with this scheme capacity achievable is high, up to 102.4 bps, and still algorithm is able to maintain good quality of the audio signal and tolerate a wide class of common attacks such as additive white Gaussian noise (AWGN), Gaussian Low-pass filter, Kaiser Low- pass filter, resampling, requantizing, cutting, MP3 compression and echo, but does not perform well in case of amplitude scaling attack, so this algorithm needs to be improved to tolerate this attack.

A watermarking scheme jointly exploiting the Discrete Wavelet Packet Transform, DCT, QIM and perceptual entropy was proposed by Hu and Li [40]. To achieve a balance between the robustness and imperceptibility, an adaptive QIM has been developed under the guidance of the human auditory masking effect for frame synchronization and watermark embedding. The quality of the watermarked audio signal was evaluated using the SNR along with perceptual evaluation of audio quality (PEAQ) using ODG and results guarantee the imperceptibility and robustness against common signal manipulations. Payload capacity was found to be near 140 bps. Author suggests that Payload capacity can be further improved by accommodating more critical bands in each frame or by embedding more bits in each band.

Methodology by X. Wang et. al [41] adopting logarithmic quantization index modulation based on μ-Law companding is utilized to transform vector norm of the segmented wavelet approximation components of the original audio signal, and the binary watermark image scrambled by chaotic sequence is embedded in the transformed domain with a uniform quantization scheme. Experimental results demonstrate that watermarked audio has good perceptual quality, and capacity of the scheme achieved is high, up to 102.4 bps. Moreover, with almost the same imperceptibility and capacity results provided, the algorithm is able to achieve good robustness against common attacks, especially amplitude scaling attack and also reduce computational complexity, which is useful for real-time applications. Author suggests that developing a synchronization audio watermarking algorithm combined with this scheme will be a good direction which we will place great emphasis on in the further research.

IV. CONCLUSION

This paper presents the analysis of algorithms of watermarking audio signals using blind detection approach which does not need original signal information to recover watermark from the watermarked signals at the receiver side. Various watermarking techniques have been adopted to embed watermarks and performance of these algorithms differ at the levels of robustness, perceptual quality and other parameter. Previously techniques focused on only improving robustness against common signal manipulations and improving perceptual quality, but recent techniques also focus on desynchronization attacks or others along with common attacks and on satisfactorily compromising the requirements of security, embedding capacity, processing complexity along with robustness and perceptual quality. Survey of current techniques presents future directions to improve the proposed methods.

REFERENCES

[1] P. Bassia and I. Pitas, robust audio watermarking in the time domain, IEEE transactions on Multimedia, Vol. 3, Issue: 2, pp. 232 –241, 2001.

[2] Jiwu Huang, Yong Wang, Yun Q. Shi, 2002, A Blind AudioWatermarking Algorithm With Self-Synchronization, IEEE International Symposium on Circuits and Systems, ISCAS 2002, Vol. 3, pp. 627 – 630.

[3] Ching-Tang Hsieh, Pei-Ying Tsou, Blind Cepstrum Domain Audio Watermarking Based On Time Energy Features, 14th International Conference on Digital Signal Processing, DSP 2002, vol.2, pp. 705 – 708.

[4] Zheng Liu and Akira Inoue, IEEE Transactions On Circuits And Systems For Video Technology, Audio Watermarking Techniques Using Sinusoidal Patterns Based on Pseudorandom Sequences, Vol. 13, NO. 8, pp. 801 - 812, August 2003

[5] Foo Say Wei, Xue Feng and Li Mengyuan, A Blind Audio Watermarking Scheme Using Peak Point Extraction, IEEE International Symposium on Circuits and Systems, ISCAS 2005, Vol. 5, pp. 4409 – 4412.

[6] Shaoquan Wu, Jiwu Huang, Daren Huang, and Yun Q. Shi, Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission, IEEE Transactions On Broadcasting, VOL. 51, NO. 1, pp. 69 – 76, 2005.

[7] Xiang-Yang Wang and Hong Zhao, A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT, IEEE Transactions on Signal Processing, Vol. 54, No. 12, 4835 - 4840, December 2006.

[8] Li Zhang, Li-Min Chen, Gong-Bin Qian, Self-Synchronization Adaptive Blind Audio Watermarking, 12th International Multi-Media Modelling Conference Proceedings, 2006, Beijing, ISBN 1-4244-0028-7.

[9] Zhao Xu, Ke Wang and Xiao-hua Qiao, Digital Audio Watermarking Algorithm Based On Quantizing Coefficients, Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing , IIH-MSP'06, ISBN: 0-7695-2745-0, pp. 41 – 46.

[10] Da-Wen Xu, Rang-Ding Wang, Ji-Long Bao, A Blind Audio Watermarking Algorithm Based on Convolutional Code, 8th International Conference on Signal Processing, 2006, Beijing, Vol. 4

[11] Xiangyang Wang, Wei Qi, and Panpan Niu, A New Adaptive Digital Audio Watermarking Based on Support Vector Regression, IEEE Transactions On Audio, Speech, And Language Processing, Vol. 15, No. 8, pp. 2270 - 2277, November 2007.

[12] Hai-yan Liu , Xue-feng Zheng, Ying Wang, Dwt –Based Audio Watermarking Resistant To Desynchronization, Seventh International Conference on Computer and Information Technology, pp. 745 – 748.

[13] Charfeddine Maha, Elarbi Maher and Ben Amar Chokri, A blind audio watermarking scheme based on Neural Network and Psychoacoustic Model with Error correcting code in Wavelet Domain, 3rd International Symposium on Communications, Control and Signal Processing, ISCCSP 2008, pp. 1138 – 1143.

[14] Vivekananda Bhat K, Indranil Sengupta, and Abhijit Das, Audio Watermarking Based on BCH Coding using CT and DWT, International Conference on Information Technology, ICIT '08, pp. 49 – 50.

[15] Yongqi Wang, Yang Yang, A Synchronous Audio Watermarking Algorithm Based on Chaotic Encryption in DCT Domain, 2008 International Symposium on Information Science and Engieering, ISISE'08, Vol. 2, pp. 371 – 374.

[16] Xiang-Yang Wang, Pan-Pan Niu, Hong-Ying Yang, A robust digital audio watermarking based on statistics characteristics, Pattern

Recognition, Volume 42, Issue 11, Pages 3057–3064, November 2009.

[17] Mingquan Fan, Hongxia Wang, Chaos-based discrete fractional Sine transform domain audio watermarking scheme, Computers and Electrical Engineering, Volume 35, Issue 3, May 2009, pp. 506–516.

[18] Yong-Hong Chen, Jian-Cong Chen, A New Multiple Audio Watermarking Algorithm Applying Ds-Cdma, Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009, Volume: 4 , pp. 2205 – 2210.

[19] Bingwei Chen and Jiying Zhao, An Adaptive and Secure Audio Watermarking Algorithm Robust to MP3 Compression, International Instrumentation and Measurement Technology Conference, Singapore, I2MTC 2009, pp. 629 – 632.

[20] Wei Cao, Yixin Yan, Shengming Li, Bit Replacement Audio Watermarking Using Stereo Signals, International Conference on New Trends in Information and Service Science, 2009, NISS '09 , pp. 603 – 606.

[21] Xiao-li Liu et. al., An Audio Watermarking Method of Resistance Statistics Attack Based on Psychoacoustic Model, Fifth International Conference on Information Assurance and Security 2009, Vol. 1, pp. 737 – 739.

[22] Hong Peng, Jun Wang, Weixing Wang, An Adaptive Audio Watermarking Method Based on Local Audio Feature and Support Vector Regression, 10th ACIS International Conference on Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing, 2009, pp. 381 – 384.

[23] Tingting Xu, Zhen Yang, Blind Audio Watermarking Scheme for Copyright Protection of Multi-User, International Conference on Multimedia Information Networking and Security, 2009, Vol. 2, pp. 159 – 162.

[24] Malay Kishore Dutta , Phalguni Gupta, Vinay K. Pathak, Blind Watermarking in Audio Signals using Biometric Features in Wavelet Domain, IEEE Region 10 Conference TENCON 2009, pp. 1 – 5.

[25] David Megias, Jordi Serra-Ruiz, Mehdi Fallahpour, Efficient self-synchronised blind audio watermarking system based on time domain and FFT amplitude modification, Signal Processing, 2010, Volume 90, Issue 12, pp. 3078–3092.

[26] Vivekananda Bhat K, Indranil Sengupta, Abhijit Das, An adaptive audio watermarking based on the singular value decomposition in the wavelet domain, Digital Signal Processing, 2010, Vol. 20, Issue 6, pp. 1547–1558.

[27] Dongmei Wu, Shangguo Yang, Zuochen Wang, Wenhua Qin, BCH Code-Based Robust Audio Watermarking Algorithm in the DWT Domain, International Conference on Intelligent Computation Technology and Automation, 2010, Vol. 1, pp. 878 – 880.

[28] *Liang Wang, Sabu Emmanuel, Mohan S. Kankanhalli*, EMD And Psychoacoustic model Based Watermarking For Audio, Ieee International Conference On Multimedia And Expo (ICME), 2010, pp. 1427 – 1432.

[29] Yubao Bai, Sen Bai, Guibin Zhu, Chunyan You, Bowen Liu, A Blind Audio Watermarking Algorithm Based on FFT Coefficients Quantization, International Conference on Artificial Intelligence and Education (ICAIE), 2010, pp. 529 – 533.

[30] Wei-zhen Jiang, Fragile Audio Watermarking Algorithm Based on SVD and DWT, International Conference on Intelligent Computing and Integrated Systems (ICISS), 2010, pp. 83 – 86.

[31] Xiong-hua Huang, Wei-zhen Jiang, An Adaptive Fragile Audio Watermarking Based on MDFT, International Conference on Intelligent Computing and Integrated Systems (ICISS), 2010, pp. 91 – 94.

[32] Charfeddine Maha, Elarbi Maher, Koubaa Mohamed and Ben Amar Chokri, DCT Based Blind Audio Watermarking Scheme, International Conference on Signal Processing and Multimedia Applications (SIGMAP), Proceedings of the 2010, pp. 139 – 144.

[33] Jian Wangn, RonHealy,JoeTimoney, A robust audio watermarking scheme based on reduced singular value decomposition and distortion removal, Signal Processing, 2011, Vol. 91, Issue 8, pp. 1693–1708.

[34] Wang Xiang-yang, Niu Pan-pan, Lu Ming-yu, A robust digital audio watermarking scheme using wavelet moment invariance, The Journal of Systems and Software, 2011, Volume 84 Issue 8, pp. 1408-1421.

[35] Bai Ying Lei , Ing Yann Soon, Zhen Li, Blind and robust audio watermarking scheme based on SVD–DCT, Signal Processing, 2011, Volume 91, Issue 8, pp. 1973–1984.

[36] Yang Hong-ying, Wang Xiang-yang, Ma Tian-xia, A robust digital audio watermarking using higher-order statistics, International Journal of Electronics and Communications (AEÜ), vol. 65, Issue , pp. 560-568.

[37] Yanbin Zhao, Wanyi Yang, Di Chang, Wei Guo, A Robust Audio Sonic Watermarking Algorithm Oriented Air Channel, International Conference on Computational and Information Sciences(ICCIS), 2011, pp. 53 – 57.

[38] Baiying Lei, Ing Yann Soon, Feng Zhou, Zhen Li, Haijun Lei, A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition, Signal Processing, Volume 92, Issue 9, pp. 1985–2001.

[39] Xinkai Wang, Pengjun Wang, Peng Zhang, Shuzheng Xu, Huazhong Yang, A norm-space, adaptive, and blind audio watermarking algorithm by discrete wavelet transform, Signal Processing, April 2013, Volume 93, Issue 4, pp. 913–922.

[40] Hwai-Tsu Hu* and Wan-Cyun Li, A Perceptually Adaptive and Retrievable QIM Scheme for Efficient Blind Audio Watermarking, International Conference on Information Science and Applications , 2012, pp. 1 – 5.

[41] Xinkai Wang, Pengjun Wang, Peng Zhang, Shuzheng Xu, Huazhong Yang, A blind audio watermarking algorithm by logarithmic quantization index modulation, Multimedia Tools and Applications, springer, 2012, Online ISSN:1573-7721.

AUTHORS PROFILE

Harleen Kaur is persuing her Master's in Engineering in Computer Science from Sri Guru Granth Sahib World University, Fatehgarh Sahib. She is currently working on the project of audio watermarking for his research work. She has interests in subject areas like image processing, signal processing, software development, web development etc.
Email: harleen.unique@gmail.com

Usvir Kaur has persued her B.Tech from Guru Nanak Dev Engineering College, Ludhiana and received her postgraduate degree in M.TECH, computer Science from Punjabi University, Patiala. Currently she is working as Assistant Professor in Computer Science Department, Sri Guru Granth Sahib World University. She has published more than 5 research papers. Her areas of interest include data mining, search Engine Optimization, image processing etc.