Impact Factor (SJIF): 3.632

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

(Special Issue for ITECE 2016)

A SURVEY PAPER ON AUDIO TAMPERING DETECTION TECHNIQUES

Hiteshri N Patel[1], Ravi K Sheth[2]

[1]M.Tech (Cyber Security), Raksha Shakti University, Ahmedabad
[1]hiteshripatel265@gmail.com
[2]Assistant Professor (IT), Raksha Shakti University, Ahmedabad
[2]rks.it@rsu.ac.in

Abstract*— Digital Forensic is a process of multimedia signals such as images, audio or video it is used to recover the evidences set of scientific techniques. These techniques are used to reveal or identify the history of digital evidences, such as, 1) to identify the acquisition device that produced the data, 2) retrieve information from the digital signals and 3) to validate the integrity of the digital contents. In multimedia important vulnerability is copyright infringement. It means the integrity and authenticity issue. Now a day, in lawsuit, civil and criminal proceedings, use of digital audio and video evidence content are increasing. As per the literature survey enough research has been done on image and video but there is a lack of study in digital audio file. There is some method available to check integrity and authenticity of audio file. In this paper we have discussed various methods to detect forgery in audio file and their pros and cons.

***Keywords*— Authentication, Chirp Coding, Double Compression, ENF, Frame Offsets, Multimedia, Tampering**

I. INTRODUCTION

In modern era, digital multimedia is used to deliver information more accurately and attractively. It is used for evidences, entertainment and other public interest. So there is a chance for copyright infringement or tampering the digital data. So the raw content of the digital data must contain some authentic information as it captures in real time situation so it can be useful at the time of investigation for forgery detection.

In this paper mainly focuses on digital audio file. And we have learnt various methods to detect tampering in digital audio file for MP3 audio file; because MP3 is widely used compression audio format and it can be easily tampered or doctored very easily. But double compression of audio file are also widely used for forgery and forger get more benefit from this technique in this process wav file format is also used for tampering and MP3 is used with wav for double compression [5][9]. As the technical progress is increasing so editing and changing is becoming easier and faster in picture and film recordings [3]. The difference between digital sampling and conventional pirated copy is that in digital sampling sample is used for editing of the original work. Different digital sampling methods make the technical analysis and the legal classification more difficult. If effective evidence of an unauthorized use of sampling cannot be produced then the proof of this process is useless in the legal process [9].

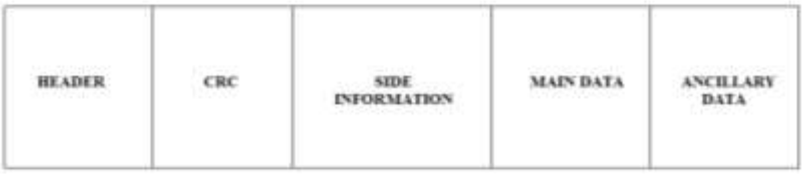
As per the literature review still there is a lack of study in digital audio content. We have studied some methods of forgery detection in digital audio file for MP3 version.

1.1 Audio File Format

1.1.1 MP3 File Format

In MP3 file format, Frame layout consists of five parts: Header, CRC, Side Information, Main Data and Ancillary Data [7].

Fig. 1 File Format of MP3 audio File



In MP3 File Format, Header Part contains 32 bits information which is synchronization word with the description of the frame [7]. Use of synchronization word is receiver can fix the signal of the stream at any point so it is stored in the each frame at the beginning. CRC is for Cyclic Redundancy Check 16 bytes and possibly check the most sensitive data for transmission errors [7]. Side Information has 17 bytes bit stream if it is single channel otherwise 32 bytes are allocated [7]. The Main Data part of the frame contains Huffman coded bits and the scalefactors. Scalefactors reduces the quantization noise. Description of decoding of Huffman coded bits is given in the side information part. In Ancillary Part the availability of number of bits are not easily given and it's also an optional [7].
For Watermarking in MP3 audio file we can embed any data or tamper file in main data part of the file format.

1.1.2 WAV File Format

A WAV file format contains a header and data. A WAV file contains a header and the raw data, in time format. Information stored in file format is in bit size. Bit size works with the amplitude. Now a days file format should be in 16bit, 8 bits file is less in size but it has less resolution. In 16 bit, a total of 65,536 amplitude levels are available [1]. Dynamic range of the file increase as resolution of the file will increase. CD-Audio uses 16 bit samples. No of samples per seconds defines the sample rate. CD-Audio has a sample rate of 44,100 [1]. This means that 1 second of audio has 44,100 samples [1]. The highest frequency can be considered to be ½ of the sample rate [1]. There is always two parts in the file format, first one is a header part and second one is a data part. The beginning of the WAV file is header part. Header part provides information like file type, sample rate, size of the file and sample size and also provides the overall size. The header of a WAV file is 44 bytes long. So for watermarking process we can add watermark in to wav audio file after 44 bytes.

III. DIFFERENT METHODOLOGY FOR FORGERY DETECTION

Here, we have discussed various methods to detect tampering in audio signals.

Method-I: Using Frame offsets

The method to detect forgery in MP3 file is by checking the frame offset of the signals. Audio samples are divided into frames to encode, and after encoding process each frame has its own encoding. By breaking the frame grids of audio signals forger can generate a tampered audio file. So the offsets of audio signals frame are used to indicate location of forgeries, and the detection of the frame offsets can be possible using the quantization characteristics of the audio signals. According to this research during the encoding process some spectral coefficients are completely masked by other components so that many spectral coefficients are usually quantized to zero value. The increase of zero spectral coefficients is a quantization characteristic of MP3 coding. The absolute value of quantized spectral and its unquantized spectral vary from the 10^{-5} to 10^{-1} [8]. Hence, quantized and its unquantized spectral is not easily visible in their real value form but they are visible in a logarithmic representation [8].

In this way, we can get the location of doctored positions of audio signals automatically. The ratio of detection of tampered audio signals is above 94% for different bitrates [8]. This technique is very useful for both audio and speech and can counter a forger who makes insertions or deletions at frame boundaries. But this technique is not giving the result on additive noise or on double compressed audio files. In this two techniques are there. One is number of coefficients and differentiated sorted spectral. The number of active (non-zero) spectral coefficients of frame offset can be used as a function for automatic identification of the characteristic spectral [8].

Method-II: Based on singularity analysis with wavelet packets

Second Method is for Detection of audio file using singularity analysis with wavelet packets and it can also locate the forgery position in digital audio file. With the help audio forensics techniques are used to detect and analyse digital audio forgeries with respect to time domain, it performs discrete wavelet packet decomposition and scans the singularity points of audio signals. According to this analysis there are some samples which are closed to the tampering position, then there is possibility of decreasing and breaking of the correlation property and it will generate new singular points. In analysis of detecting and locating forged position of wavelet packet where singular points are given in the time domain. But the disadvantage is it causes serious interference at the time of detection due to some singularity points stays in the form of group. It is quite different for the sample rate of the digital audio file and for analysis process of forged and inherent singularity point. Hence, it can be helpful for developing a method for detection of forged position accurately and also locating it. This method can give up to 70% accurate result [3].

Method-III: Using chirp based robust watermarking

Another technique is blind tamper detection to detect digital audio file forgeries using chirp based robust watermarking. Use of chirp coding is without generating intuitive degradation of audio quality to embed a watermark [6]. Using energy based features; watermark can be derived from the audio signal. Chirp coding is used for embedding the watermark in MP3 audio data. In chirp coding technique we can recover watermark from the watermarked signal and also we can derived it from the original audio signal. It enables the blind recovery of the watermark. It also provides the benefit that we can derive two independent watermark extraction process. From that we can assure the authentication of audio data and if both watermark are mismatched then we can say that data of audio file may have been tampered.

The ability of a chirp coded watermark is that it can be detect in a high noise background is used for embedding in an MP3 audio signal [6]. Signal dependent watermark sequence helps to detect blind tampering of the audio signal. This scheme is used to found attacks like scaling, resampling interpolation, compression and decimation [6].

Method-IV: Based on max offsets for cross correlation between ENF (Electric Network Frequency) and Reference signal

Another method is Audio forgery detection based on max offsets for cross correlation between ENF(Electric Network Frequency) and Reference signal. When electronic recording devices are connected to electric power lines then the Electric Network Frequency (ENF) is embedding into the audio signals [11].

We can determine whether the audio signal was manipulated or edited by comparing the max offsets based on a block by block method. We can generate max offset by find cross correlation between the reference signal and the extracted ENF. To indicate the boundaries of the edited region we can use the change of the max offsets [11]. In this they have used spatial domain method so that it doesn't need DFT (discrete Fourier Transform) because it relies on the computation accuracy of the ENF phase. The ENF often contains other frequency components from the background noise and the voice by the band pass filter from a practical audio signal [11]. These components vary because of the accuracy of estimation of the ENF signal can be vary. In this method computational load is lighter than other methods. Fluctuation of the practical ENF frequency caused it also interferes with the accuracy of estimation of the ENF phase by real electric networks [11].

Method –V: detection of double compression using content independent method

To detect double compression of MP3 audio file using Content Independent Feature in this improvised technique according to researcher, double compression is nothing but the manipulation of the audio file created for malicious activity. Double compression is achieved by decompression of audio file and again recompress at different compression ratio. In this research work, researcher have proposed an idea to detect up transcoded MP3 audio file as well as down transcoded MP3 audio file. The meaning of up transcoded compression is the first compression used lower quality than the second compression and the meaning of down transcoded is the first compression used higher quality than the second one [4]. Up transcoded audio files are used to show the fake quality of an audio file cause customer choose online stores for downloading music and in that they download music according to the bitrates forger can make more profit on up transcoded files [4]. For malicious manipulation like insertion, splicing, deletion, substitution, down transcoding is used to achieve more sophisticated memory. If we compress audio files second time using lower quality then it will generate noise in MP3 audio files.

Using quantized MDCT coefficients and their derivatives we can extract statistical patterns and based on that it can reveal the real quality of compression. Firstly minimize the false prediction caused by individual characteristics of diversified audio clips [4]. To measure the difference between signal based features and reference based signal features they have generated reference audio signals by recompressing and calibrating the audio files. For binary and multiclass classifications dynamic evolving neural fuzzy interference system and support vector machine (SVM) were applied on that [4]. To provide a content independent analysis and detection reference audio signals are generated. For binary and multi class classification, support vector machine with radial basis function (RBF) kernel and dynamic evolving neural fuzzy interference systems (DENFIS) to predict quantitative and qualitative approach [4]. This can detect MP3 double compression and also can reveal the history of double compression.

To create a MP3 double compression firstly decompresses MP3 audio into temporal domain and then recompresses it using a different bit rate [4]. It will reset the quantization levels in audio file is called additional distortion. Experimental result of this work shows in detection of double compressed MP3 file and gives accurate result. Still down transcoded needs further investigation when the first quality is high.

Method-VI: Exposing The Double Compression In Mp3 Audio By Frequency Vibration

In detection of double compressed MP3 audio file with the analysis of double compression by frequency vibration effect on MDCT (Modified Discrete Cosine Transform) coefficients in MP3 audio File [10]. To measure the vibration caused by double compression on audio file is called Frequency Vibration Value (FVV) [10]. In this research work authors shows a better method than the existing double compression method. And they have also calculated the original bitrate for a double compressed MP3 file.

In this paper, authors have presented an approach to detect MP3 double compression for both down transcoded and up transcoded audio and calculated the original bitrate of first compression. The disadvantage to detect double compression in both type of MP3 audio is to develop some effective features for both audios which are discriminative with the help of survey, The MDCT coefficients are different from the normal coefficients because of double quantization. Double compression of MP3 audio file contains double quantization on the MDCT coefficients. In normal MP3 audio file the coefficients executes a Laplace distribution [10]. This distribution can't be ensured after the double compression. And it possible to distinguish a normal MP3 audio from a double compressed audio. In very few cases we can say that whether a coefficient distribution execute a Laplace distribution of normal MP3 audio file.

In this case, they have considered to describe the coefficient distribution in the form of transform domain because if an orthogonal transform can re-distribute the energy then good properties can occur. Here, two kinds of the quantized MDCT coefficients are obtained; first one is normal and second is double compressed one with Huffman decoding. After that MDCT coefficient distribution is computed and it has transformed into a DFT(Discrete Fourier Transform) method.

Recognizing the properties of MP3 double compression, that is based on the frequency vibration value. Authors have implemented an algorithm for this process. Here, steps of the algorithm are given: (1) To get the quantized MDCT coefficients firstly decompress the MP3 audio file: (2) transform decompress audio file by DFT(Discrete Fourier Transform); (3) then after Frequency Vibration Value(FVV) has been calculated; (4) double compression of audio file can be determined [10].

In this paper, to detect double compression based on the Frequency vibration value (FVV) of MDCT coefficients authors have present a simple method. Using Frequency Vibration Value they have estimated the effect of double MP3 compression and prove it reasonable on the quantized MDCT coefficients. The experimental result of this method shows that this method is effective comparatively typical method of this field to detect double compression [10]. Disadvantage of this method is down transcoding is an issue but it can estimate original bit rate for double compression.

IV. COMPARATIVE ANALYSIS OF DIFFERENT METHODS

Here, we have analysis the advantage and disadvantage of the different audio forgery detecting methods which are explained above:

TABLE 1: Comparisons of Different Methods

Method	Pros	Cons
Using frame offsets	Gives 94% accurate result	Spectral are not visible in real value form, cannot remove additive noise
Based on singularity analysis with wavelet packet	Locate the tampering point	Difficult to find when singularity points stays in the form of group
Based on chirp coding	Detectable in high noise background	Due to filtering audio quality is degraded
Based on max offsets for cross correlation between ENF and reference signal	Light computation load	Does not give accurate result
Detection of double compression using content independent method	Detect both up transcoding and down transcoding audio file using SVM DENFIS	Down transcoded is not accurate
Detection of double compression using frequency vibration value	Works for both up transcoded and down transcoded	Down transcoded is not accurate

V. CONCLUSION AND FUTURE WORK

The audio forensics techniques which we have discussed here are used to detect forgery and tampering in digital audio file. So basically these are used to check authenticity and integrity of digital audio file. Many robust and objective techniques are proposed by the researchers. Hence, we will propose another technique to detect forgery in digital audio file based on this survey. In this technique we will use digital signature to check integrity and authenticity. By extracting the digital signature from watermarked signal and comparing with the original signal we can check the authenticity and integrity of the digital audio file.

REFERENCES

[1] Digital Audio, Avilable via: <http://www.topherlee.com/software/pcm-tut-waveformt.html> [Accessed 4 December 2015]

[2] Jiaorong Chen, Shijun Xing, Hongbin Huang, Weiping Liu, "Detecting and Locating digital audio forgeries based on singularity analysis with wavelet packet" Springer Science Business Media NewYork, Volume 75, Issue 4, pp 2303-2325, December 2014

[3] Marking Stefan K. Braun, "Forensic Evidence of Copyright Infringement by Digital Audio Sampling" International Journal of Cyber-Security and Digital Forensics(IJCSDF), The Society of Digital Information and Wireless Communications, ISSN:2305-0012, 2014

[4] Mengyu Qiao,"Improved Detection of MP3 Double Compression using Content-Independent Features" Signal Processing, Communication and Computing (ICSPCC), IEEE International Conference, pp1-4, 2013

[5] Multimedia Forensics Available at: <http://lesc.det.unifi.it/en/node/157>_ [Accessed on 21 November 2015]

[6] O. Farooq. S. Datta, And J. Blackledge,"Blind Tamper Detection in Audio using Chirp based Robust Watermarking" whitepaper of researchgate, 2008

[7] Rassol Raissi, "The Theory Behind MP3", December 2002

[8] Rui Yang, Zhenhua Qu, Jiwu Huang, "Detecting Digital Audio Forgeries by Checking Frame Offsets" ACM digital library,pp 21-26 2008

[9] Swati Gupta, Intel Seongho Cho and C.-C. Jay Kuo, , "Multimedia in Forensics, Security, and Intelligence Current Developments and Future Trends in Audio Authentication" IEEE Multimedia. Volume: 19, Issue:1, pp 50-59, January-March 2012

[10] Tianzhuo Wang , Xiangwei Kong , Yanqing Guo, Bo Wang, "Exposing The Double Compression In Mp3 Audio By Frequency Vibration " Signal And Information Processing, IEEE China Summit And International Conference, pp 450-454, 2014

[11] Yongjian Hu, Chang-Tsun Li, Zhisheng L , and Bei-bei Liu, "Audio Forgery Detection Based on Max Offsets for Cross Correlation between ENF and Reference Signal " Springer, pp 253-266, 2013