

Anti-Forensics of Double Compressed MP3 Audio

Biaoli Tao, Ningbo University, China

Rangding Wang, Ningbo University, China

Diqun Yan, Ningbo University, China

Chao Jin, Ningbo University, China

ABSTRACT

The widespread availability of audio editing software has made it easy to create acoustically convincing digital audio forgeries. To address this problem, more and more attention has been paid to the field of digital audio forensics. There has been little work, however, in the field of anti-forensics, which seeks to develop a set of techniques designed to fool current forensic methodologies. The compression history of an audio sample can be used to provide evidence of audio forgeries. In this work, we present a simple method for distinguishing the MP3 compression history of an audio sample. We show the proposed anti-forensics method to remove the artifacts of MP3 double compression by destroying the audio frame structure. In addition, effectiveness of the proposed method is verified by three double compression detection methods. The experimental results show that the proposed method can effectively resist detection from three methods.

KEYWORDS

Anti-Forensics, Audio Forensics, Doubles MP3 Compression, Frame Structure

INTRODUCTION

With the popularization of portable recording equipment and the rapid development of multimedia technology, malicious tampering can be implemented to digital audio easier and easier. For example, it may obstruct the justice if the tampered audios are taken as evidences in court. Therefore, digital multimedia technology makes our lives convenient, and the application field of digital audio authenticity and integrity bring serious security issues. In addition, the content protection and integrity authentication of digital audio has become a hot research topic. Digital audio forensics (Jin, Wang, Yan, Ma, & Zhou, 2016) is to detect and identify the digital audio's integrity by analyzing the statistical features of the detecting audio signal.

Various effective forensic methods have been proposed in recent years. Ikram and Malik (2010) proposed a method to determine the audio integrity based on correlation similarity measure. In Ly et al. (2013), the MFCC and logarithm MFCC were considered as the discriminative features to identify the recording environment. Malik and Zhou (2012) and Hua et al. (2016) examined the integrity of

DOI: 10.4018/IJDCF.2020070104

This article, originally published under IGI Global's copyright on July 1, 2020 will proceed with publication as an Open Access article starting on January 27, 2021 in the gold Open Access journal, International Journal of Digital Crime and Forensics (converted to gold Open Access January 1, 2021), and will be distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

audios based on the electric network frequency (ENF) signal existed in the given audios. As for the MP3 audio forensics methods presented in Yang et al. (2008) and Yang et al. (2012), it was found that frame structure of MP3 audios will be destroyed when the audios were tampered.

Since audio tampering would inevitably produce double compression. It has been attention to study for audio double compression. Yu et al. (2014) put forward a method of detecting fake-quality audios, which utilized the number of MDCT coefficient as the features. The results show that their scheme was effective for detecting double compressed audios. Liu et al. (2010) proposed to detect the fake-quality audio by using Huffman table index. The experimental results demonstrated that the method was effective. Yang et al. (2010) found that the MDCT coefficients whose values fall into $(-1, 1)$ of the double compressed audio were less than that of the single compressed audio, as well as the frame offset to detect whether the audio processed with double compression. The results show that the method has higher detection rate for the MP3 audio from low bit rate to high bit rate, but its performance became poorer for the cases of high bit rate to low bit rate and the same bit rates. Liu et al. (2010) integrated the number of non-zero MDCT coefficients, the average distance between non-zero MDCT coefficients and the distribution of zero and non-zero coefficients as a feature vector to distinguish the single compressed and double compressed audios. Compared with the high bit rate to low bit- rate, the low bit rate to high bit rate achieves higher detection accuracies.

The competition between forensics technology and its countermeasure anti-forensics (Chen, Xiang, & Huang, 2016; Milani, Piazza, & Bestagini, 2014) has escalated over the past few years. In the field of audio anti-forensics, the work (Chuang, Garg, & Wu, 2012; Chuang, Garg, & Wu, 2013) present the first anti-forensic scheme and its countermeasures for ENF-based audio authentication. Chuang et al. proposed various anti-forensic operations to resist tampering with the underlying ENF signal while preserve the quality of the host signal. Zhao and Malik (2013) and Zhao et al. (2016) put forward a method to remove audio carry environmental characteristic, which aims to cover the audio splicing tampered by environmental noise.

In order to conceal/disguise the trace of tempering, the modified audio probably be recompressed with the same bit rate after the tempering. This will inevitably leave the cues of double compression. The forensic examiner can pick up the questionable audios by detecting the double compression. The goal of this work is to remove the artifacts generated by the double compression. Eventually, an anti-forensics method of MP3 audio tempering is proposed, which eliminates the double compression artifacts of tempering process by adding zero samples to the tempered audio.

The rest of this paper is organized as follows. In section 2, we will introduce the artifacts of MP3 audio compression. A method to remove the double compression artifacts of MP3 audio is put forwarded in section 3. Section 4 demonstrates the validity of the algorithm based on experimental comparison. The proposed method is summarized in section 5.

MP3 COMPRESSION ARTIFACTS

MPEG-1 Audio Layer is one of the MPEG coding standards. The MPEG was established by IOS (International Organization for Standardization) and IEC (International Electro Technical Commission) of the joint technical committee, which occupies a dominant position in the international general multimedia coding standards. Its coding standard included MPEG-1, MPEG-2, MPEG-2.5, MPEG-4, MPEG-7 and MPEG-21. MP3 is the third layer structure in MPEG-1 standard. It is an important milestone in audio coding, and it is widely used in the market and research institutions.

MP3 encoding mainly includes the following parts: sub-band filter, MDCT, FFT, acoustic model and iterative loop quantization and Huffman coding, etc. PCM signal is divided into 32 sub-bands by the analysis filter bank, and then for each sub-band with MDCT in frequency line segment into 18 (long windows) or 6 (short windows) line frequency. Psychoacoustics model is used to estimate masking threshold for the maximum quantization error that can be introduced into the audio signal while still maintaining perceptually unimpaired signal quality. The spectral values are quantized

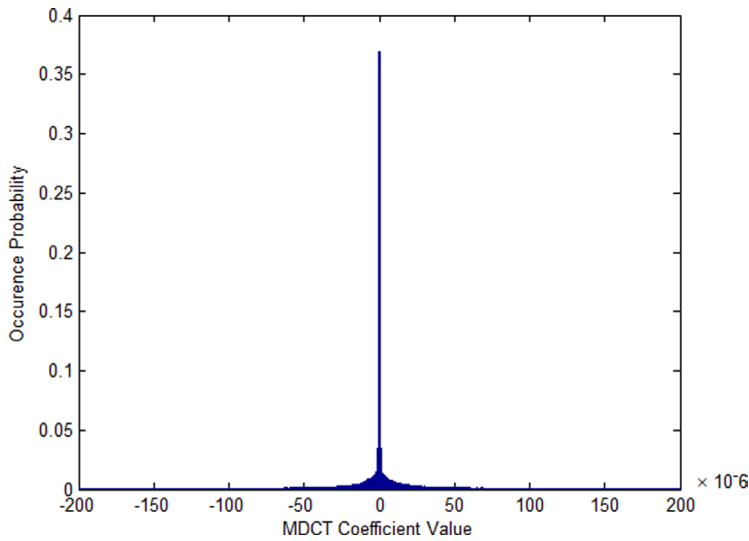
according to the masking thresholds. Spectrum values quantify comparing the quantization error in the iterative cycle. Finally, Huffman coding will be used to encode for quantized spectral values.

In MP3 audio compression forensics technology, researchers often use statistical characteristics of MDCT coefficients to obtain evidence. As shown in Figure 1, The MDCT coefficient of distribution of compressed MP3 audio and uncompressed audio is significantly different. Histogram of MDCT coefficients has some gap phenomenon, which means that the MDCT coefficients concentrate on some values. The MP3 encoding process of quantitative link, the quantitative formula is as follows:

$$ix(i) = n \text{int} \left(\left(\frac{|xr(i)|}{\frac{\text{stepsize}}{2^{\frac{4}{4}}}} \right)^{\frac{3}{4}} - 0.0946 \right) \quad (1)$$

where $xr_Q(i)$ represents the i^{th} quantized MDCT coefficient, and $xr(i)$ is the i^{th} MDCT coefficient. $n\text{int}()$ Denotes the nearest integer function. Integer operation will lead to some similar value into the same value, but the process is irreversible, which results in the phenomenon of MP3 compression artifacts, as shown in Figure 1.

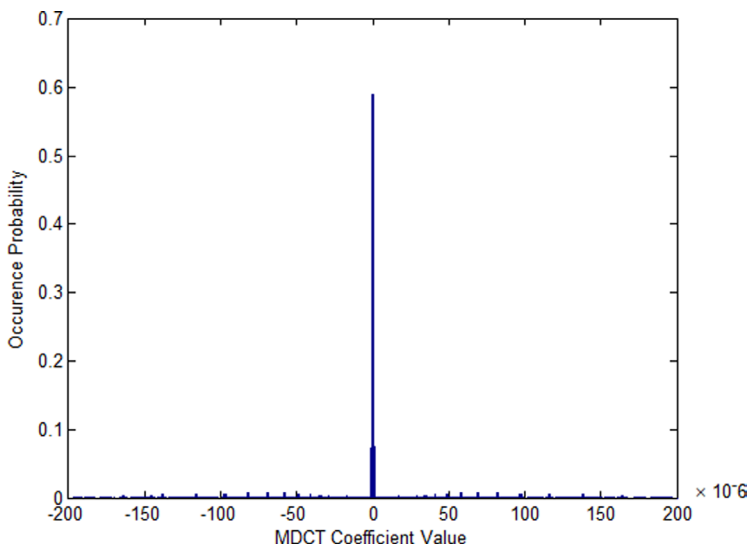
Figure 1a. Histogram of MDCT coefficients from an uncompressed audio



REMOVE AUDIO COMPRESSED ARTIFACTS

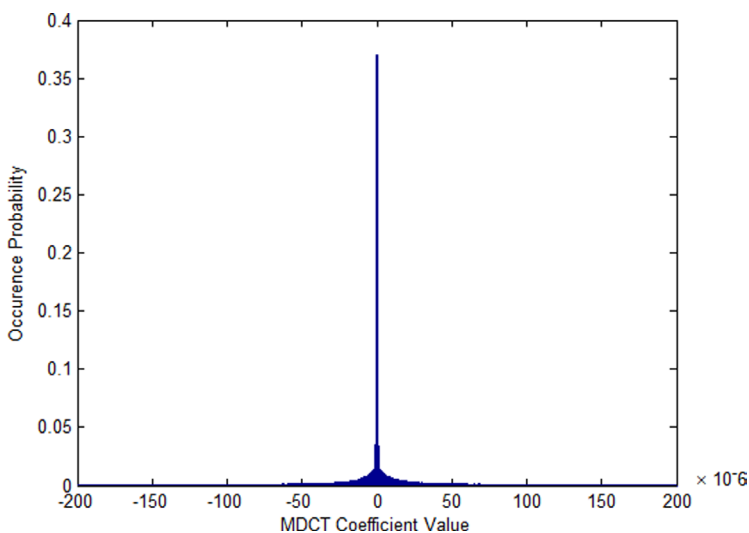
The attacker needs to remove the artifacts of double compression so that the forensic examiner cannot confirm whether the given audio is double compressed or not. For an original MP3 file, the attacker decodes it into WAV audio, tampers its content, and removes the artifacts of compression operation. Then the audio is compressed into MP3 file with the same bit rate of the original one. The detector tends to think that the tampered MP3 audio only suffered single compression instead of double compression. So, the tampered audio was detected as the credible audio. In order to achieve this goal, we put forward a method of anti-forensics by adding a zero sample to the starting position of WAV

Figure 1b. Histogram of MDCT coefficients from the same audio by MP3 compressed



audio. As show in Figure 2(a), the MP3 audio (bit rate is 64 kbps) was decoded into WAV audio, and a zero sample was added in the starting position of the decoded audio. Then the WAV audio was encoded again with the same bit rate. It can be seen that the MDCT coefficient distribution of single

Figure 2a. Adding zero samples for anti-forensics



compressed audio is similar with uncompressed audio, while completely different with compressed audio. Furthermore, the method can effectively remove the artifacts of compression. The reason is that the MP3 encoding is implemented based on the unit of frame. The frame structure will not be changed when the audio is compressed for the second time, and it can leave double compression

artifacts. however, If the frame structure was broken during the audio double compressed, the double compression artifacts will disappear. And adding zero samples is to change the structure of all frames.

In addition, we also try to test its performance with adding the zero samples with different numbers. The test results, as shown in Figure 2(b)(c)(d), are the MDCT coefficient distributions

Figure 2b. Adding two zero samples for anti-forensics

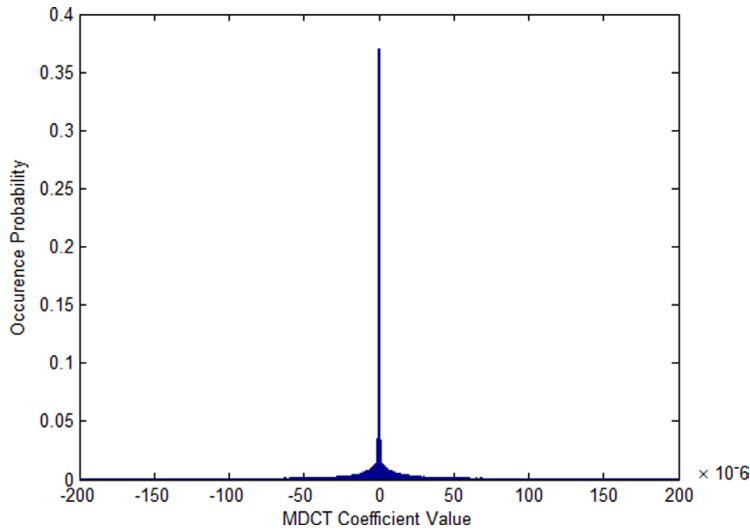
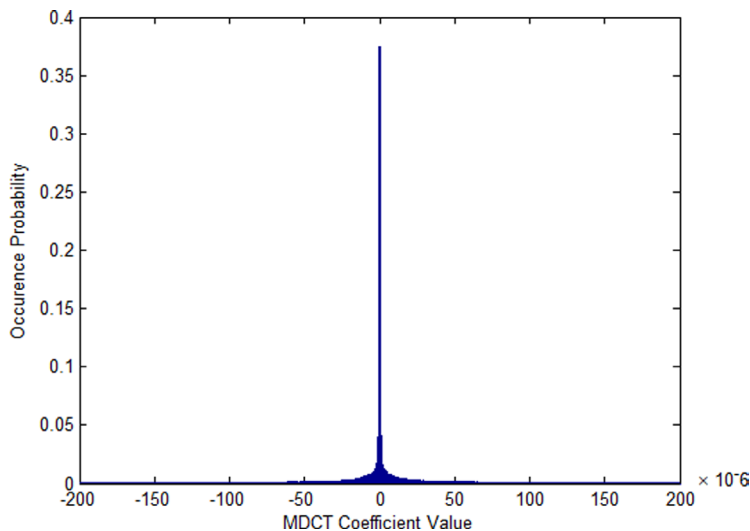
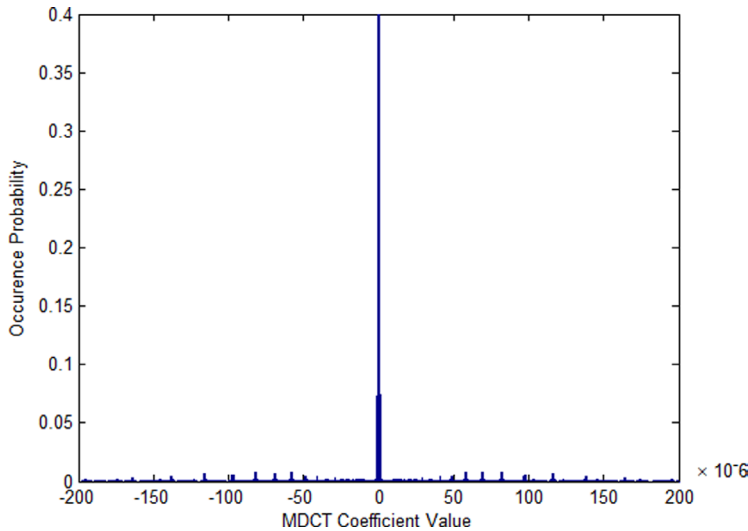


Figure 2c. Adding 100 zero samples for anti-forensics



with adding 2, 100 and 576 zero samples, respectively. It can be observed that the MDCT coefficient distributions are highly consistent when added 1, 2, 100 zero samples, and similar to the distribution of the uncompressed audio. Whereas, the MDCT coefficient distribution of the audio added by 576 zero samples is similar to distribution of the compressed audio. It implies that adding multiple

Figure 2d. Adding 576 zero samples for anti-forensics



zero samples also can achieve the goal of removing the compressed artifacts. However, the anti-forensics method will fail to detect the audio added with 576 zero samples. Because MP3 encoding is compressed granule by granule, and each granule contains 576 samples. When audio was added 576 zero samples, the 576 samples will be independent as a granule to encode, and the frame structure will not be changed from the original frame structure. The original data according to the previous structure will be encoded again. Therefore, it keeps the artifacts of the double compression. Although adding multiple zero samples can also achieve the effect of anti-forensics for audio, the anti-forensics operation must keep concealment and integrity of the original data. Therefore, our anti-forensics method is to add a zero sample instead of multiple zero samples.

EXPERIMENTAL RESULTS

In this section, three methods of detecting double compression are used to test the proposed anti-forensics method, respectively, Yang's method (2010), Yu's method (2013) and Ma's method (2014). Yang's method points out that the first digits statistical distribution of MDCT coefficients obeys to Benford's law when the audio has not been compressed. However, MDCT coefficients of the double compressed audio do not follow this law. Then Yang put forward to use the first digits of quantized MDCT coefficients as a 9-dimensional feature vector, and using the SVM classifier to classify. The results show that their method is effective to detect double compression. Yu's method shows that the number of quantized MDCT coefficients in two adjacent compressions is significant different. Then they put forward a method to use the differences as the characteristics, and use the SVM classifier to classify whether the audio experienced double compression or not. Ma et al. analysis the statistical characteristic of scale factor when the audio was compressed with the same bit rate. They found that the number of scale factor of adjacent compression have certain difference, so the statistical characteristics of scale factor are used as a feature set by using SVM classifier to classify double compression. Experiments show that the three methods can effectively detect double compressed MP3 audios with the same bit rate.

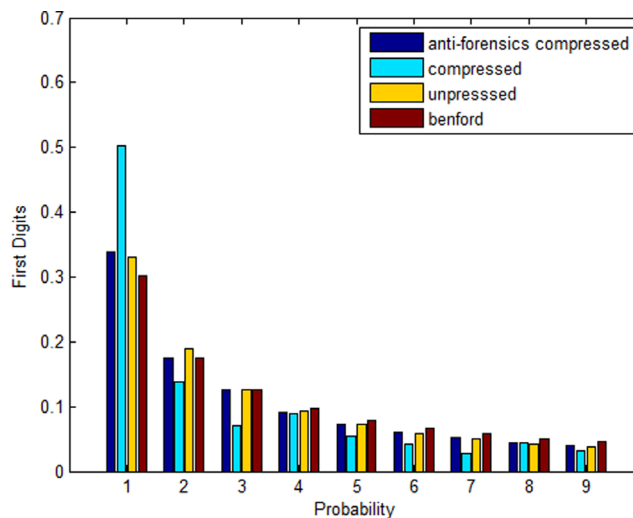
Experimental Environment

A dataset of 300 audio clips has been constructed in this paper. The format information of all clips are WAV, 5-second-long, 44.1 KHz sampling rate, 16-bit quantization and mono. All the clips were coded by LAME-3.99.5 with 6 different compression ratios: 32, 64, 96, 112, 128 and 192 kbps. The 1800 single compressed audios denote the sample library 1(SL1). Then the single compressed MP3 audios were decoded into WAV audios which were processed to generate another two libraries. One library, defined as samples library 2(SL2), consists of 1800 double compressed MP3 audios encoded by those WAV audios with the same bit rate of the single compressed audios. The other one named library 3(SL3) is also composed of 1800 double compressed audios, but these audios were added with a zero sample in starting position before suffered the second compression.

Results of Removed Artifacts

We use the samples of 64 kbps compression for experiment analysis. In Yang's method, the authors point out that the characteristics of the first digits statistical distribution of MDCT coefficients approximate to Benford's law distribution for the uncompressed WAV audio, while the MDCT coefficients of the compressed MP3 audio will not follow this law. The MDCT coefficients were extracted based on sample libraries 1, 2 and 3, respectively. The statistical distributions of the first digits shown in Figure 3. It can be observed that the first digits statistical distributions of the audios without

Figure 3. The first digits statistical distribution of MDCT coefficients from compressed and uncompressed audio and anti-forensics compressed



compression and the anti-forensics processed audios are both approximately identical to Benford's law distribution, but the compressed audios' first digits statistical distribution has obvious differences. It indicates that the proposed anti-forensics method is able to remove the double compression artifacts.

Yu's method points out that the number of different quantized MDCT coefficients between single and double compressions is significantly more than that of the double and third compressions. According to the results of the analysis, we test the 3 samples libraries for experiments. As shown in Figure 4a, Yu's feature can obviously distinguish the single compressed audios from the double compressed audios. However, when the double compressed audios are processed by the proposed anti-forensics method, it turns to be difficult to separate the single compressed audios from these

Figure 4a. The distribution of QMDCT coefficients from Yu's method

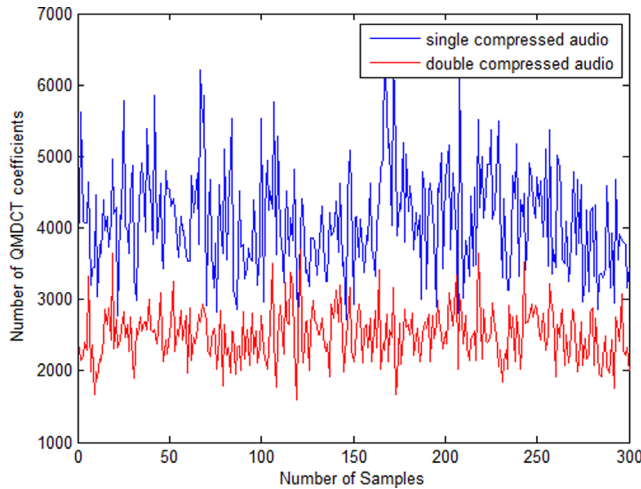
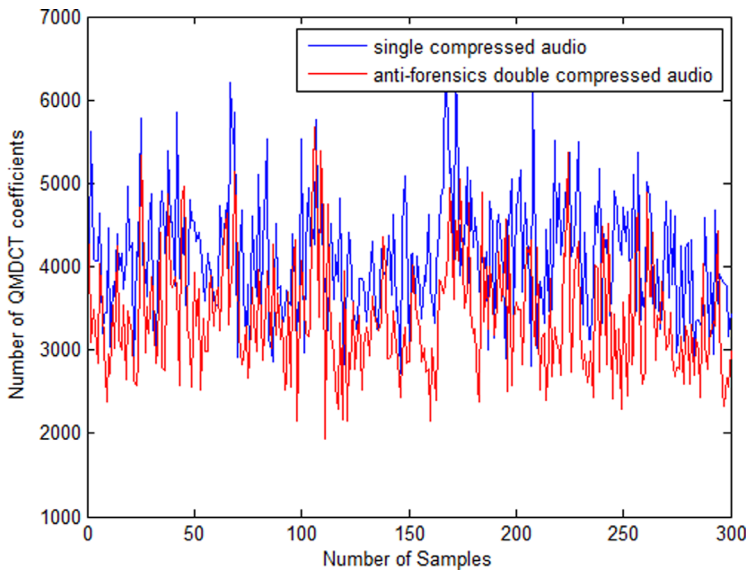


Figure 4b. The distribution of QMDCT coefficients from Yu's method with anti-forensics



audios. The experimental results show that Yu's method will fail to detect the double compressed MP3 audio after anti-forensics processing. On the contrary, it reveals that the proposed method is effective to Yu's method.

Ma's method points out that the number of different scale factor coefficients between single and double compressions is significantly more than that of the double and third compressions. We also test it on the 3 samples libraries for experiments. Experimental results, as shown in Figure 5a, demonstrate that Ma's feature can obviously distinguish the single compressed audios from the double compressed audios. However, when the double compressed audios are processed by the proposed anti-forensics method, it turns to be difficult to separate the single compressed audios from these audios. From the result, we can infer that Ma's method unable to identify the double compression

Figure 5a. The distribution of scaling factor coefficients from Ma's method

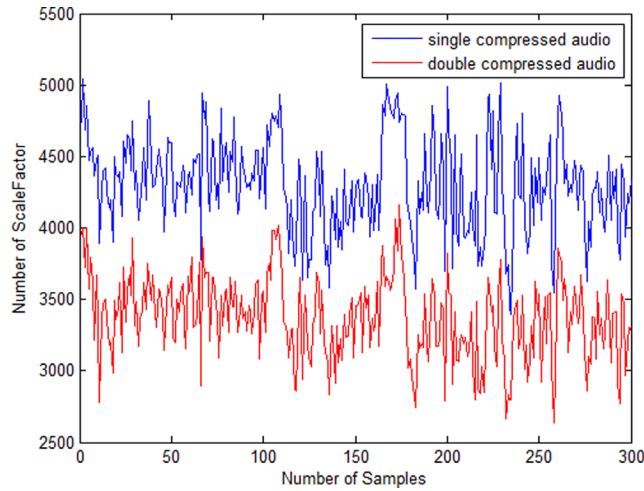
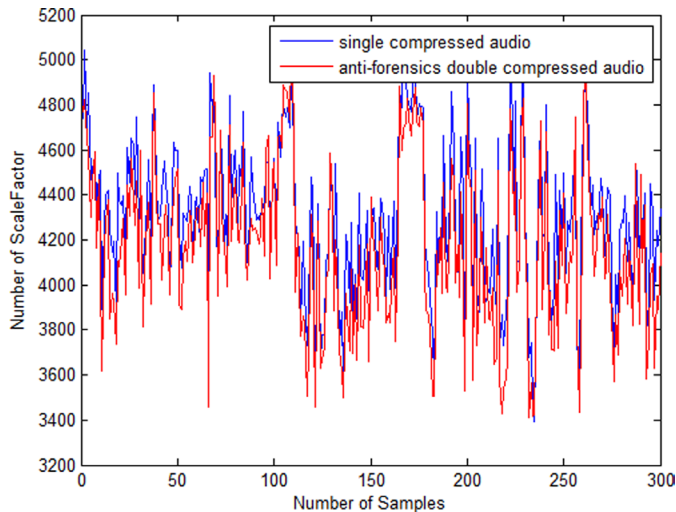


Figure 5b. The distribution of scaling factor coefficients from Ma's method with anti-forensics



when the audio was operated by the proposed method, which indicates that the proposed anti-forensics method is effective for Ma's method.

Result of Anti-Forensics Detection

In order to evaluate the effectiveness of the proposed method in terms of the three detecting double compression schemes. We test the samples library in a variety of bit rate compressions. SL1 and SL2 respectively were used for testing Yang's, Yu's and Ma's methods, SL2 and SL3 for the three methods as well. The results of double compression detection for two groups of sample libraries were obtained. The classifier LIBSVM was chosen according to the experiment requirements. The higher classification accuracy shows that the method to identify audio double compression is more effective.

Experimental results are shown in Table 1. It can be seen that the detecting accuracy of Yang's method for double compression is higher in different bit rate compression. The detection accuracy

Table 1. The detection accuracy of three methods and corresponding anti-forensics methods

Bit Rate(kbps)	Yang (%)	Anti-Forensics for Yang (%)	Yu(%)	Anti-Forensics for Yu (%)	Ma(%)	Anti-Forensics for Ma (%)
32	94.67	76.33	57.00	69.00	70.00	66.67
64	100.00	54.67	94.67	65.67	91.00	52.67
96	100.00	61.33	99.00	66.67	96.00	53.67
112	100.00	59.67	97.00	56.00	98.67	48.33
128	100.00	55.33	89.00	58.00	93.67	47.33
192	99.67	71.33	50.67	45.67	89.33	60.67

of double compression obviously drops to 50% for Yang's method when we used the proposed anti-forensics method. For the lower bit rate of 32 kbps and higher bit rate of 192 kbps the detection accuracies are 70% for our proposed method, the accuracy is decreased by 25%. The results show that proposed method of anti-forensics has good ability to resist Yang's method. It can be seen from the test results that Yu's method has good detection effect only when the bit rate from 64 to 128kbps, and this method is invalid for the lower or higher bit rate cases. The detection accuracy has obvious drop when the proposed anti-forensics method was used. The final detection accuracy is close to 50-60%. It is almost impossible to detect the audio of double compressed. And the method still unable to detect the double compressed audio when the bit rates are 32 kbps and 192 kbps. Experimental results also show that Yu's method cannot detect the double compressed audio when the proposed anti-forensics method was used, which means the proposed method of anti-forensics has good resistance effects for Yu's method. The test results of Ma's method illustrate that detection accuracy has appeared a downward trend in the bit rate too low or too high, but it has certain detection effect with different bit rates. The accuracy of detection obviously dropped when using the proposed anti-forensics method, close to 50%, it almost impossible to detect the double compression. Experimental results also show that Ma's method almost impossible to detect the double compression of audio when the anti-forensics method was used. It also shows that the proposed anti-forensics method can resist Ma's method to detect double compression of audio.

In conclusion, the three methods of detecting double compressed audio are effective without utilizing the proposed anti-forensics method. And the detection accuracy dropped to 50-60%, almost impossible to detect the double compression of audio when we used the anti-forensics method. Therefore, the proposed anti-forensics method can resist detection of the three methods.

CONCLUSION

Inspired by the principle of the existing MP3 double compression detection methods, we found that the distribution of MDCT coefficients has great difference when the audio was compressed and uncompressed. Then a method for removing the compression artifacts was proposed. This method can make the MDCT coefficients distribution of the double compressed audio becoming similar to the distribution of uncompressed audio's MDCT coefficients. So, the double compressed audio after being removed artifacts will be identified as the single compressed audio. Meanwhile, we also evaluated our anti-forensics method based on three typical MP3 double compression methods, and the experimental results show that our method can effectively resist the detection of these methods.

Digital multimedia forensics technology and anti-forensics technology is restricting each other and developing mutually. The digital image and video forensics technology related to anti-forensics

work have achieved some advances. The anti-forensics work on digital audio still less, and there will be more and more scholars focus on audio anti-forensic research in the future.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61672302, 61300055), Zhejiang Natural Science Foundation (Grant No. LZ15F020010, Y17F020051), Ningbo University Fund (Grant No. XKXL1405, XKXL1420, XKXL1509, XKXL1503) and K.C. Wong Magna Fund in Ningbo University.

REFERENCES

- Chen, J., Xiang, S., Huang, H., & Liu, W. (2016). Detecting and locating digital audio forgeries based on singularity analysis with wavelet package. *Multimedia Tools and Applications*, 75(4), 2303–2325. doi:10.1007/s11042-014-2406-3
- Chuang, H., Garg, R., & Wu, M. (2012). How secure are power network signature based time stamps. *Proceedings of the ACM conference on Computer and Communications Security* (pp. 428-438). ACM.
- Chuang, H., Garg, R., & Wu, M. (2013). Anti-forensics and countermeasures of electrical network frequency analysis. *IEEE Transactions on Information Forensics and Security*, 8(12), 2073–2088. doi:10.1109/TIFS.2013.2285515
- Hua, G., Zhang, Y., Goh, J., Vrizlynn, L., & Thing, L. (2016, May). Audio authentication via exploring the absolute-error-map of the ENF signals. *IEEE Transactions on Information Forensics and Security*, 5(11), 1003–1016. doi:10.1109/TIFS.2016.2516824
- Ikram, S., & Malik, H. (2010, July). Digital audio forensics using background noise. *Proceedings of the 2010 IEEE International Conference on Multimedia and Expo* (pp. 106-110). IEEE.
- Jin, C., Wang, R., Yan, D., Ma, P., & Zhou, J. (2016). An efficient algorithm for double compressed AAC audio detection. *Multimedia Tools and Applications*, 75(8), 4815–4832. doi:10.1007/s11042-015-2552-2
- Liu, Q., Sung, A., & Qiao, M. (2010). Detection of double MP3 compression [J]. *Cognitive Computation*, 2(4), 291–296. doi:10.1007/s12559-010-9045-4
- Lv, Z., Hu, Y., Li, C. T., & Liu, B. B. (2013, July). Audio forensic authentication based on MOCC between ENF and reference signals. *Proceedings of the 2013 IEEE China Summit and International Conference on Signal and Information Processing* (pp. 427-431). IEEE. doi:10.1109/ChinaSIP.2013.6625375
- Ma, P., Wang, R., Yan, D., & Jin, C. (2014). Detecting double-compressed MP3 with the Same Bit rate. *Journal of Software*, 9(10), 93–96. doi:10.4304/jsw.9.10.2522-2527
- Malik, H., & Zhao, H. (2012, March). Recording environment identification using acoustic reverberation. *Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1833-1836). IEEE.
- Yu, X., Wang, R., Yan, D., & Zhu, J. (2013). Method for detecting double MP3 compression with same bit rate. *Computer Engineering and Applications*, 49(12), 93–96.
- Milani, S., Piazza, P., & Bestagini, P. (2014). Audio tampering detection using multimodal features. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE Press. doi:10.1109/ICASSP.2014.6854466
- Yang, R., Qu, Z., & Huang, J. (2008). Detecting digital audio forgeries by checking frame offsets. *Proceedings of the 10th ACM workshop on Multimedia and Security, International Multimedia Conference* (pp. 21-26). doi:10.1145/1411328.1411334
- Yang, R., Qu, Z., & Huang, J. (2012). Exposing MP3 Audio Forgeries Using Frame Offsets. *ACM Transactions on Multimedia Computing Communications and Applications*, 8(2S), 1651–1654. doi:10.1145/2344436.2344441
- Yang, R., Shi, Y., & Huang, J. (2010). Detecting double compression of audio signal. *Proc. SPIE on Media Forensics and Security*. Academic Press.
- Yu, X., Wang, R., Yan, D., & Ma, P. (2014). Detecting fake-quality MP3 based on Huffman table index. *Journal of Software*, 9(4), 907–912. doi:10.4304/jsw.9.4.907-912
- Zhao, H., Chen, Y., Wang, R., & Malik, H. (2016). Anti-Forensics of Environmental-Signature-Based Audio Splicing Detection and Its Countermeasure via Rich-Features Classification. *IEEE Transactions on Information Forensics and Security*, 11(7), 1603–1617. doi:10.1109/TIFS.2016.2543205
- Zhao, H., & Malik, H. (2013). Audio recording location identification using acoustic environment signature. *IEEE Transactions on Information Forensics and Security*, 8(11), 1746–1759. doi:10.1109/TIFS.2013.2278843

Rangding Wang received his M.S. degree in the Department of Computer Science and Engineering from the Northwest Polytechnic University, Xian in 1987, and received his Ph.D. degree in the School of Electronic and Information Engineering from Tongji University, Shanghai, China, in 2004. He is now a professor in College of Information Science and Engineering, Ningbo University, China. His research interests mainly include multimedia security, digital watermarking for digital rights management, and data hiding.

Diqun Yan received B.S., M.S., Ph.D. degrees in Circuit and System from Ningbo University, Zhejiang China, in 2002, 2008 and 2012 respectively. He is currently an associate professor in College of Information Science and Engineering, Ningbo University. His research interests include multimedia forensics and security.

Chao Jin received his Ph.D. degree in Signal and Information Processing from Ningbo University, China, in 2017. His research interests mainly include audio steganalysis, multimedia forensics and audio/speech signal processing. His research interests mainly include audio forensics and audio/speech signal processing.