Research of the Stegosignal Propagation through the Acoustic Environment

L G Statsenko¹, U A Durneva², N A Klescheva³, M V Bernavskaya²,⁴

¹Professor, Department of Electronics, Telecommunications, and Instrumental Engineering, Far Eastern Federal University, p. Ajaks, 10., o. Russkij, Vladivostok, 690922, Russian Federation
²Master degree student, Department of Electronics, Telecommunications, and Instrumental Engineering, Far Eastern Federal University, p. Ajaks, 10., o. Russkij, Vladivostok, 690922, Russian Federation
³Professor, Department of General and experimental physics, Far Eastern Federal University, p. Ajaks, 10., o. Russkij, Vladivostok, 690922, Russian Federation
⁴Assistant Professor, Foreign Languages department, Peter the Great Saint-Petersburg Polytechnic University, ul. Politehnicheskaja, 29, Sankt-Peterburg, 195251, Russian Federation

E-mail: durneva.ua@dvfu.ru

Abstract. The paper considers two main applications of stegosystems: as a digital watermark, for copyright protection, and as a method for secretly transmitting information through an acoustic environment. A stegosystem is studied theoretically and experimentally when information is transmitted over different distances in an acoustic medium in the presence of extraneous noise. Nowadays great attention is paid to the protection of intellectual property. Illegal distribution of media content, including copies of musical compositions, is widespread and reaches unprecedented proportions. Protecting audio files from unauthorized copying and using becomes increasingly important. Digital watermarks are a promising method for solving this problem. The purpose of the article is to develop an effective method for protecting audio files from illegal use by embedding a digital watermark (DW). After detection the illegal distribution of their product, the copyright holder will be able to track the presence of the digital watermark and apply to the court for damages.

1. Introduction

Hiding information is a general concept of embedding data in content. Steganography and digital watermarking are two important subsections of information hiding [12]. Steganography seeks ways to make communication invisible by hiding data in the message, while labeling arises from the need to protect content copyright.

Digital watermarks are an important copyright protection method for multimedia content. The purpose of watermarking systems is developing systems resistant to all possible deliberate attacks. Digital watermarks must also be resistant to any transformations such as channel filtering, noise and compression [1].
There are many methods for embedding additional information in audio: amplitude modulation, adding pseudo-random noise, using wideband signals, and others. This paper will consider a method based on echo embedding: single echo-hiding.

Adding echoes to music can make it sound more natural, like a real concert hall. With correctly designed amplitudes and delays (offsets), echoes are perceived as reverberation for the received audio signals and do not create uncomfortable noise [9]. Since preserving the quality of an audio file is one of the main concerns when embedding digital watermarks, methods based on embedding reverberation are promising.

The algorithm is implemented in the MATLAB environment, evaluated in terms of stealth, reliability and security. Both watermark embedding and key detection (delay time and echo amplitude) are performed. The influence of different parameters of the stegosystem on security and the probability of errors during data extraction are considered.

2. Embedding process

Figure 1 shows a way to create watermarks for audio. The data is embedded in the cover message by adding delayed signal copies [2-4]. The data is represented by single echoes with delays for each bit.

The host signal is divided into segments. Then each frame can be considered as an independent signal containing the desired bit [1,3,5]. N – number of bits to hide and L the length of the segments. Then L should be chosen so that \( N \cdot L \) does not exceed the length of the audio signal.

![Figure 1. Impulse response of echo kernels. (a) “One” kernel. (b) “Zero” kernel](image)

A watermarked signal consists of all independent segments [6,7]. Information is embedded in the signal by repeating the original signal with one of the two echoes, and extraction consists in detecting the distance between them.

A watermarked signal \( S_w(n) \) is generated by the convolutional between host signal \( S_0(h) \) and echo kernel \( h(n) \) [2,8,9]. This process can be shown as a filter with impulse response in MATLAB. A basic echo-hiding scheme applies a signal echo kernel (1):

\[
h(n) = \delta(n) + \alpha \delta(n - d),
\]

where \( \alpha \) – echo amplitude; \( \delta(n) \) – Dirac delta function; \( d \) – delay.

As shown in Figure 1, echo kernels are generated with different delays (d1 and d0) to represent «1» and «0». Usually allowed delays for 44.1 kHz sampled audio is about 100 ~ 150 samples (2.3 ~3.4 ms). Each bit interval can be represented as follow (2):

\[
\forall n=1,\ldots,N: \quad S_w(n)=S(n) \ast h(n), \quad b \in \{0,1\},
\]

where \( S(n) \) – host signal; \( n \) – number of bits; \( h(n) \) – impulse filter response; \( \ast \) – convolution; \( N \) – number of samples.
3. Extracting process
Direct Fourier transform is employed for frequency domain integration $S(n)$ and $h(n)$. Then the logarithm is performed for the linear representation [2,10]. The inverse Fourier transform is used to separate hidden message from the host signal. This scheme can be described as followed (3):

$$\hat{S}_w(n) = F^{-1}\{\log(F[S_w(n)])\},$$

where $F$ – direct Fourier transform; $F^{-1}$ – inverse Fourier transform

Hidden bit is determined by comparing $\hat{S}_w(d_0)$ and $\hat{S}_w(d_1)$ (4):

$$w_e(i) = \begin{cases} 
1, & \text{if } (d_1) \geq \hat{S}_w(n), \\
0, & \text{if } (d_1) < \hat{S}_w(n),
\end{cases}$$

where $w_e(i)$ – watermark bit.

4. Matlab implementation
For the experiment we used audio recordings with a sampling rate of 44.1 kHz, a duration of 30 s and a bit rate of 16 bits per second. The oscillogram of the original signal is shown in Figure 2.

In the MATLAB, the kernel is formed using the function (5):

$$h_0 = [\text{zeros}(d_0, 1); 1] \cdot \alpha$$

where $h_0$ – “zero” kernel, $d_0$ – delay for “zero” bit, $\alpha$ – echo amplitude.

Figure 4 shows a “zero” echo implemented in MATLAB using a filter (6):

$$k_1 = \text{filter}(h_0, 1, s),$$

where $h_0$ – “zero” kernel, $s$ – host signal, $k$ – “zero” echo.

Integrated with the host signal, a watermarked signal is shown in Figure 5.
The best results of extracting are observed with echo amplitude of 1.2. For musical compositions of the rock genre, the optimal reverberation time is short; the use of large amplitudes is undesirable. When listening to audio files with embedded echoes with an amplitude of more than 0.6, the composition turns out to be too boomy, the difference between the original track and the resulting one as a result of adding a digital watermark is clearly audible. The main task when adding a watermark is to preserve the quality of the audio recording. Therefore, the peculiarities of each musical genre should be taken into account.

High information capacity is one of the main requirements for an audio system (the possible length of a hidden message should be maximum).

Table 1 shows the results of an experimental study of the bit error probability depending on the amount of information inserted.

| Audio file | Number of bits | BER, % |
|------------|----------------|--------|
| Track 1    | 136            | 0      |
|            | 155            | 0      |
|            | 189            | 0      |
|            | 210            | 0      |
| Track 2    | 136            | 0      |
|            | 155            | 0      |
|            | 189            | 0      |
|            | 210            | 7.29   |
| Track 3    | 136            | 0      |
|            | 155            | 0      |
|            | 189            | 0      |
|            | 210            | 0      |
| Track 4    | 136            | 0      |
|            | 155            | 5.73   |
|            | 189            | 4.62   |
|            | 210            | 4.16   |

Table 2 presents the results of an experimental study of the bit error rate after adding 20 dB noise.
The results of an experimental study (Table 2) prove that this method cannot be used for transmission through an acoustic medium, since echoes introduced by multiple reflections are a significant source of interference.

It should be noted that the main requirement for watermarks is system reliability. Therefore, it is necessary to conduct a study on the stability of the system to such transformations as compression.

This method is quite resistant to compression attacks, despite the increase in the number of errors during extraction (Table 3).

### Table 2. Noise adding results.

| Audio file | Echo amplitude | BER, % |
|------------|----------------|--------|
| Track 1    | 0,6            | 42,4   |
|            | 1,2            | 41,3   |
|            | 1,4            | 51,1   |
| Track 2    | 0,6            | 52,7   |
|            | 1,2            | 47,3   |
|            | 1,4            | 39,7   |
| Track 3    | 0,6            | 51,1   |
|            | 1,2            | 49,5   |
|            | 1,4            | 49,4   |
| Track 4    | 0,6            | 40,7   |
|            | 1,2            | 40,8   |
|            | 1,4            | 41,3   |
| Track 5    | 0,6            | 50,5   |
|            | 1,2            | 39,7   |
|            | 1,4            | 40,2   |

### Table 3. Results after file compression.

| Audio file | Echo amplitude | BER, % |
|------------|----------------|--------|
| Track 1    | 0,6            | 3,52   |
|            | 1,2            | 1,64   |
|            | 1,4            | 14,34  |
| Track 2    | 0,6            | 4,64   |
|            | 1,2            | 1,43   |
|            | 1,4            | 0,69   |
| Track 3    | 0,6            | 14,45  |
|            | 1,2            | 3,24   |
|            | 1,4            | 2,94   |
| Track 4    | 0,6            | 8,64   |
|            | 1,2            | 9,35   |
|            | 1,4            | 25,3   |
| Track 5    | 0,6            | 14,9   |
|            | 1,2            | 11,33  |
|            | 1,4            | 3,53   |
| Track 6    | 0,6            | 33,07  |
|            | 1,2            | 35,32  |
|            | 1,4            | 31,13  |
| Track 7    | 0,6            | 14,62  |
|            | 1,2            | 10,19  |
|            | 1,4            | 12,13  |

5. Evaluation

Experiments have shown that the single echo embedding technique can be applied for protection audio files from unlicensed use. The error rate for digital watermark extraction is low, provided the optimal performance is used. The amount of embedded information per minute of audio recording is approximately 300 bits.
The basic method of embedding echoes is simple and quite reliable, but it has a number of disadvantages: poor security, visibility.

The weak security of the algorithm is explained by its simplicity: the delays \( d_0 \) and \( d_1 \) serve as the key to decoding information. By guessing these values, an attacker can easily extract the digital watermark.

By adding echoes of any amplitude, the difference between the original and modulated audio files is obvious to the listener. With the optimal fit, the 5-member control group noted an improvement in performance: adding depth and volume to the sound. However, to improve the robustness of the algorithm, high energy echoes increasing audible distortion are required.

It should also be noted that this method is not suitable for instruments with a short optimal reverberation time (such as a bass guitar), since even a small echo amplitude adds strong audible distortion.

6. Summary
In the work, the theoretical foundations of steganography, watermark algorithms for digital audio, as well as the features of systems based on the echo. The main problems and tasks of protecting audio files from unauthorized use and distribution by means of immersing a digital watermark are considered.

Experiments have confirmed the main disadvantage of the basic method - weak security. Since the received stegosignal is not resistant to extraction attacks, this method of echo embedding has limited application. The method of embedding single echoes also has advantages: reliability, simplicity, a good assessment of the audio file with a watermark by the listeners.

Experimental studies of the method have shown that with an optimal selection of stegosystem parameters, reverberation is perceived as an additional color. However, the echo information embedding did not go unnoticed by the control group of listeners, even when the reverberation amplitude was reduced to unacceptable from the point of view of reliability.

Modeling an audio signal with embedded hidden information using reverberation confirmed the prospects for the development of stegosystems based on adding echo signals and the possibility of using the methods considered in this work for copyright protection. Smart addition of a digital watermark can improve the quality of the original audio file.

The paper considers the effect of noise addition attacks. Since it was not possible to decode the message after adding noise, the transmission of an audio signal through an acoustic environment with external noise leads to the loss of embedded information. This makes it impossible to use a stegosystem based on an embedding of reverberation for the covert transmission of information.

In the future, it is planned to conduct a study of the OFDM-based stegosystem in order to obtain the best decoding characteristics when the signal passes through the acoustic environment.

7. References
[1] Arend J M, Pörschmann C 2016 Audio watermarking of binaural room impulse responses Audio Engineering Society Conference, AES International Conference on Headphone Technology 219-223
[2] Korzhik V, Morales-Luna G, Fedyanin I 2013 Audio watermarking based on echo hiding with zero error probability International Journal of Computer Science and Applications 10 1 1–10
[3] Korzhik V, Alekseev V, Morales-Luna G 2018 Audio watermarking system resistant to removal attacks by dereverberation International Journal of Computer Science and Applications 15 1 1–15
[4] Imoto K, Ono N 2017 Spatial cepstrum as a spatial feature using a distributed microphone array for acoustic scene analysis Transactions on Audio, Speech, and Language Processing 6 1335–1343
[5] Chardon G, Nowakowski T, Daudet L 2015 A blind dereverberation method for narrowband source localization IEEE Journal of Selected Topics in Signal Processing 5 815–824
[6] Fridrich J J, Goljan M and Soukal D 2006 Wet paper codes with improved embedding efficiency IEEE Transactions on Information Forensics and Security 1 1 102–110

[7] Durneva U A, Statsenko L G 2020 Investigation of the watermark system for an audio signal based on reverberation Youth and Science: actual problems of fundamental and applied Research. Materials of the III All-Russian National Scientific Conference of Students, Postgraduates and Young Scientists 250-253.

[8] Ambhaikar D A S 2012 Audio Steganography using RPrime RSA and GA Based LSB Algorithm to Enhance Security International Journal of Science and Research 12 79-83

[9] Hu H T, Chen W T 2014 A dual cepstrum-based watermarking scheme with self-synchronization. Signal Processing 92 4 1109-1116

[10] Chaitanya L V R 2017 MATLAB Implementation of Audio Steganography for Secure Data Transmission International Journal of Scientific Engineering and Technology Research 6069-6079