Implementation of the DSSS method in watermarking digital audio objects

Jarosław Wojtuń, Zbigniew Piotrowski*, Piotr Gajewski†

*Military University of Technology,
Kaliskiego 1,01-476 Warsaw, Poland*

Abstract

The paper presents the results of implementation in the Matlab environment for watermarking embedder and extractor based on the Direct Sequence Spread Spectrum (DSSS). A block diagram of watermarking system, an analysis of watermarked signal reproduced as well as watermarking system robustness to degrading factors: lossy compression, signal-to-noise ratio (SNR) as well as a change in sampling frequency, were shown.

1. Introduction

The digital audio signal watermarking technique is more and more frequently used in copyrighting audio files, checking a consistency of original audio recording or monitoring commercial radio channels. In military telecommunications the watermarking systems based on voice signals can play an important role, for instance, providing hidden authentication during a voice link session. The watermarking technique consists in hiding additional information represented by the so-called watermark by embedding it in the original signal. This additional information should be transparent (inaudible) to potential signal listeners. Watermarking digital signals is one of the most dynamically developing

---
*E-mail address: Zbigniew.Piotrowski@wat.edu.pl
†E-mail address: Piotr.Gajewski@wat.edu.pl

branches of digital processing implemented especially in both wired and wireless telecommunication technologies. The rapid Internet development involves certain dangers leading to non-ethical behaviours, including stealing, copying or transferring different electronic items. Watermarking provides a solution aimed to prevent such threats.

2. An implementation of digital watermarking using the DSSS method

2.1. Introduction

The concept of hiding one item of information while transferring the other has been long known as steganography. The term digital watermarking was used for the first time in 1993, when A. Z. Tirkel presented a method of hiding information in images [**1**]. At the beginning the watermarking algorithms covered hiding information in images. This is because audio signal watermarking is more difficult than doing this in the case of digital videos and images, due to higher sensitivity of human auditory system to noise than human visual system.

The following watermarking methods based on DSSS are known, for example: [**2, 3, 4**], the authors of this paper have verified the watermarking system robustness to basic degrading factors.

While developing a watermarking algorithm it is good to meet certain requirements. The most important ones include [**5**]: watermark inaudibility, its robustness to destruction and the data payload. The inaudibility of watermark is associated with the above-mentioned properties of human auditory system. The watermark robustness to destruction can be defined as an ability to obtain the watermark after signal modification (lossy compression, analog-digital conversion, change in signal sampling, etc.). Information capacity defines the quantity of information bits per signal frame. Depending on the algorithm target there are different requirements for a minimum information capacity. For example, if the algorithm is to be used to protect copyrights, its minimum information capacity necessary to include author's identification data is 0.5 bps (bits per second). Establishing proper ratios between watermark inaudibility, its robustness to destruction and the amount of hidden information is always a compromise for any watermark system developer. In addition, when developing an algorithm we should consider the additional requirements, including: algorithm calculation complexity as well as the fact whether it is necessary to have original signal to recover the watermark (no blind watermark extraction) or it is unnecessary (blind watermark extraction).
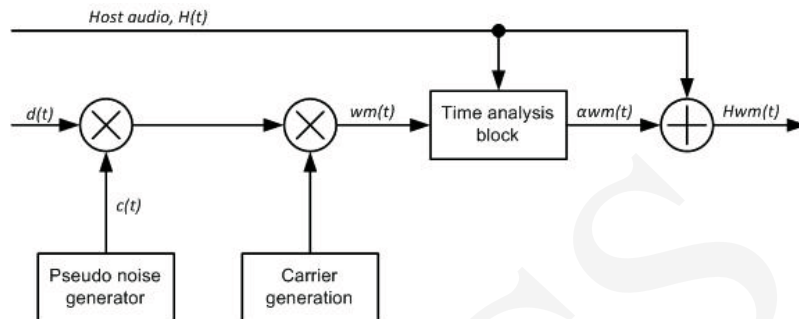
Fig. 1. Watermark embedder block diagram.

2.2. Watermark embedder

One of the paper components was to carry out a watermark embedder program in the Matlab environment. The embedder principle of operation uses the previously described Direct Sequence Spread Spectrum (DSSS) method. The block diagram showing the embedder principle of operation is presented in Figure 1. The embedder program described herein allows us to embed a watermark in audio files at a sampling frequency of 8kHz, 44.1kHz and 48kHz. The application supports mono (single channel) and stereo (double channel) sound systems. The algorithm was based on solving perfectly the following system of equations:

$$
\begin{cases}
nL_b = N \\
mL_c = N \\
T_b L_b = \frac{N}{f_p} \\
T_c L_c = \frac{N}{f_p}
\end{cases}
\tag{1}
$$

where:

n, m – natural numbers;

L_b – number of information bits per signal frame;

L_c – number of spread sequence chips per signal frame;

N – number of samples per signal frame;

T_b – duration of single bit;

T_c – duration of spread sequence chip;

f_p – audio signal sampling frequency.

In addition, the following relations govern the values of (1) equation:

$$
B_0 = \frac{1}{T_b}
\tag{2}
$$

$$
B_r = \frac{1}{T_c}
\tag{3}
$$

$$z = \frac{B_r}{B_0} = \frac{T_b}{T_c} = \frac{L_c}{L_b} \tag{4}$$

$$z[dB] = 10log_{10}(z) \tag{5}$$

where:

B_0 – watermark bandwidth;

B_r – watermark bandwidth following pseudo random sequence spread;

z – processing gain.

Considering the watermark requirements described in the introduction, while solving the (1) system of equations it was necessary to find a compromise between:

1. Number of hidden information bits;

2. Processing gain, value of which determines the watermark inaudibility and affects its robustness;

3. Number of samples per signal frame that affects the calculation complexity of algorithms and its execution time;

4. Maximum spread band available for a given sampling frequency;

The following values taken from the (1), (2), (3), (4) and (5) equations have been assumed in our paper:

Table 1. List of values necessary to develop a watermarking algorithm.

$f_p[kHz]$	$\frac{f_p}{2}[kHz]$	L_b	$T_b[ms]$	$B_0[Hz]$	L_c	$T_c[\mu s]$	$2B_r[kHz]$	N	z	$z[dB]$
8	4.00	5	200	5.0	1600	625	3.20	8000	320	25.05
44.1	22.05	6	36	27.6	1920	113	17.64	9600	320	25.05
48	24.00	6	33	30.0	1920	104	19.20	8000	320	25.05

The pseudo random sequence generator presented in Figure 1 acts based on the Gold sequence generator described in the previous chapter. The same spread sequence is generated for each sampling frequency based on the following primitive polynomials:

$$[5000]_8 = [101000000000]_2 \Leftrightarrow x^{11} + x^9 + 1 \tag{6}$$

$$[5500]_8 = [101101000000]_2 \Leftrightarrow x^{11} + x^9 + x^8 + x^6 + 1 \tag{7}$$

The initial status for both registers is the same and amounts to 11111111111. A Gold sequence generator for this case is presented in Figure 2. Some examples of timing curves for data signals at sequence generator output d(t) and pseudo random generator c(t) for the sampling frequency of fp=48000Hz are presented in Figure 3. Figure 4 presents amplitude/frequency based images of hidden information, spread sequence, hidden information following the spreading and modulating by carrier wave. Another functional block presented in the watermark embedder schematic diagram is a time analysis block. Within the
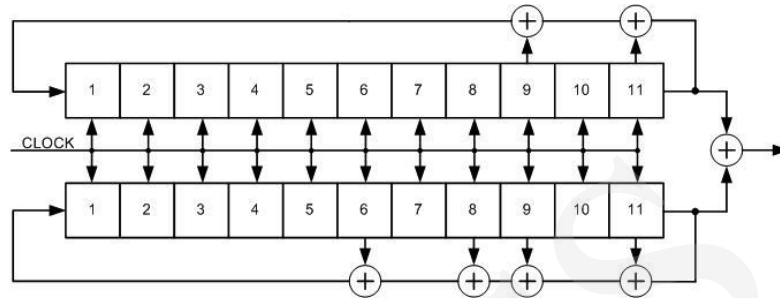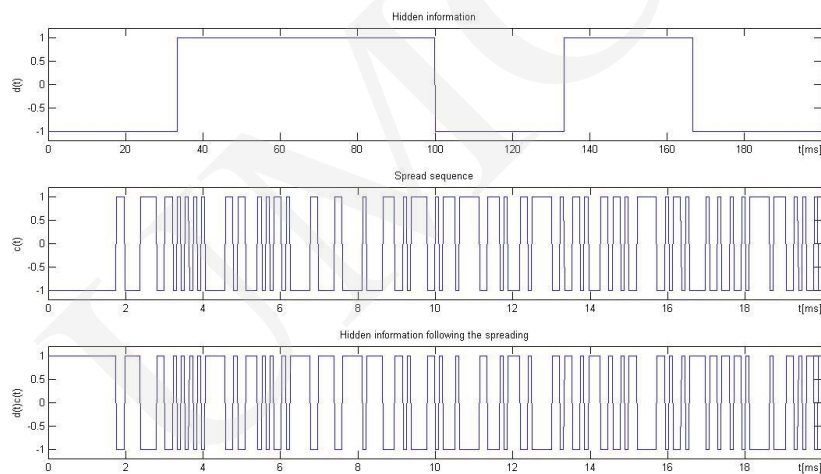
Fig. 2. Pseudo random from Figure 1.

Fig. 3. Examples of timing curves for data signals d(t) pseudo random
sequence c(t).

time analysis the block watermark and signal power are calculated based on
the following equation:

$$P = \sum_{k=1}^{N} x^2(i) \tag{8}$$

$$P[dB] = 10log_{10}(P) \tag{9}$$

Theoretical considerations in chapter one show that the processing gain value
in dB is equivalent to the usable signal to noise ratio at which the signal is still
properly received. In the watermarking system described the usable signal is the
watermarking signal, whereas the noise signal is the original audio signal. During the analysis of block output we obtain corrected watermark signal $\alpha wm(t)$.
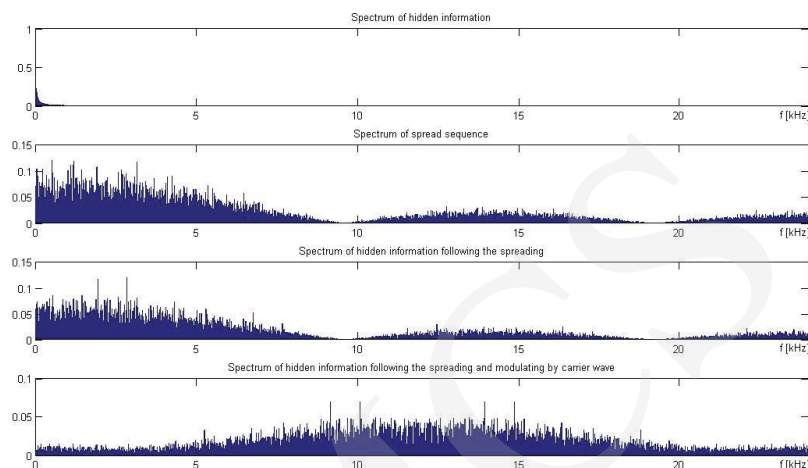
Fig. 4. Amplitude/frequency signal spectra d(t), c(t), d(t)c(t) and wm(t).

The α correction signal is selected separately for each frame calculated based on audio signal and watermark power to provide correct watermark detection at the reception side and at the same time to make the watermark inaudible against the audio signal background. Figure 5 presents the examples of original and watermarked audio signals and their amplitude – frequency based spectra.

2.3. Watermark extractor

The schematic diagram that illustrates the watermark extractor principal of operation is presented in Figure 6. The watermark system presented in this paper takes place without any participation of original (unmarked) audio signal, which is referred to as blind watermark extraction. As it has been already mentioned in chapter one, for the spread spectrum system to operate correctly, pseudo random sequences generated by the transmitter and receiver must have the same form and be mutually synchronized. In the watermarking system described synchronization takes place in its synchronization block and is carried out by synchronizing watermarked audio signal with the pseudo random sequence. The first frame of watermarked audio signal is multiplied in the synchronization block by the pseudo random sequence and carried wave signal values. Then, the power spectral density for such a frame is calculated. It is defined as the Fourier transform for the autocorrelation function, which for
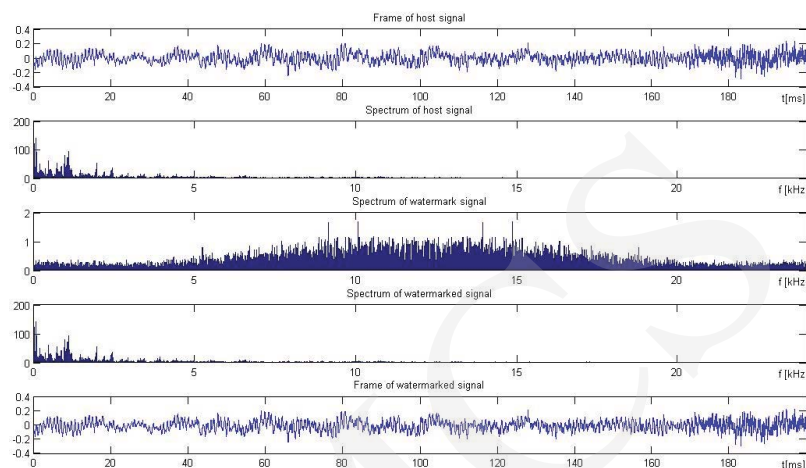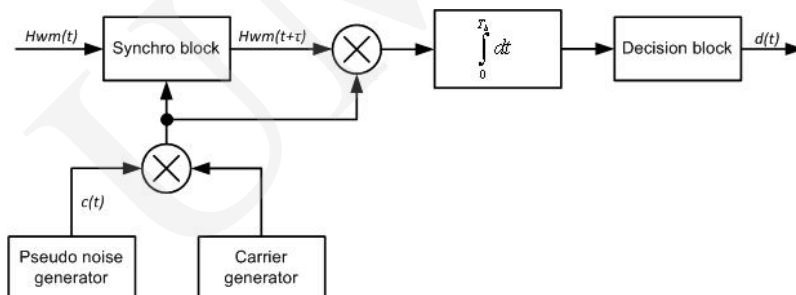
Fig. 5. Examples of signal curves H(t) awm(t) and Hwm(t).

Fig. 6. Schematic diagram for the watermark extractor.

discrete signals is expressed as follows [**6**]:

$$P_{xx}(f) = \sum_{m=-\infty}^{\infty} R_{xx}(m)e^{j2\pi(\frac{f}{f_p})m}. \tag{10}$$

Then, the audio signal frame is transferred by one sample and the actions repeat. The cycle repeats N times (Table 2.1). In the synchronization block the time shift τ is determined, which corresponds to the number of samples the audio signal should be shifted to synchronize it with the pseudo random sequence and carrier wave signal. Proper τ value is determined by defining the number of shift samples to obtain maximum local values out of previously calculated power spectral densities. The examples of power spectral densities are presented in Figure 7. For the power spectral densities presented in Figure 7,
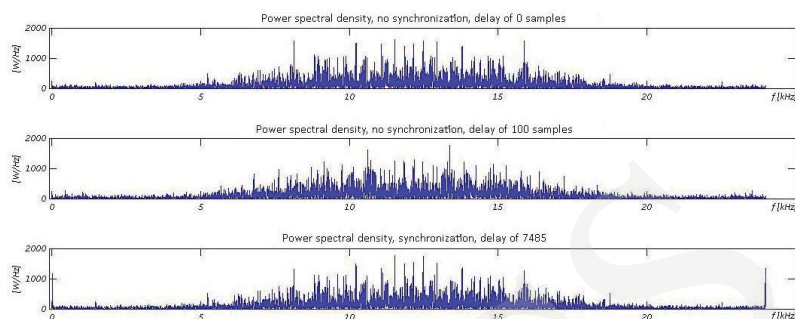
Fig. 7. Examples of power spectral densities calculated in synchronization block.

we can observe clearly that local peak for the watermarked signal is shifted by 7485 samples. The peak is found in the constant component and for the frequency component equaling to one half of the sampling frequency. As a synchronization criterion in this paper we assumed the peak found at the Nyquist frequency. Once the watermarked signal is synchronized with the pseudo random sequence and carrier wave signal, the system multiplies each signal frame by the pseudo random sequence and carrier wave signal values. Then, the resulting signal is transferred to the integrator, where the process of integrating takes place for the single bit duration period. Finally, the extractor outputs the signal from the integrating system to the decision-making system, where it reproduces the final watermark signal form. The examples of signal curves at particular branches of the extractor schematic diagram are shown in Figure 8.

3. Analysis of watermarking system quality and robustness

3.1. Introduction

According to Figure 1, the parameters that characterize the system include: information capacity, watermark inaudibility, watermark robustness to destruction. The inaudibility of watermark integrated in the original signal can be defined using subjective or/and objective signal quality assessment methods. The signal quality assessment using a subjective method requires playing the signal in real time. Objective methods can be divided into the ones adopting parametric models and methods using signal parameterization. For the needs of this paper a test of watermarked signal quality subjective assessment was performed. The robustness of a watermarking system is defined as a capacity to recover the watermark after modifying the watermarked signal. The basic attacks aimed at destroying the watermark embedded in the audio sequence
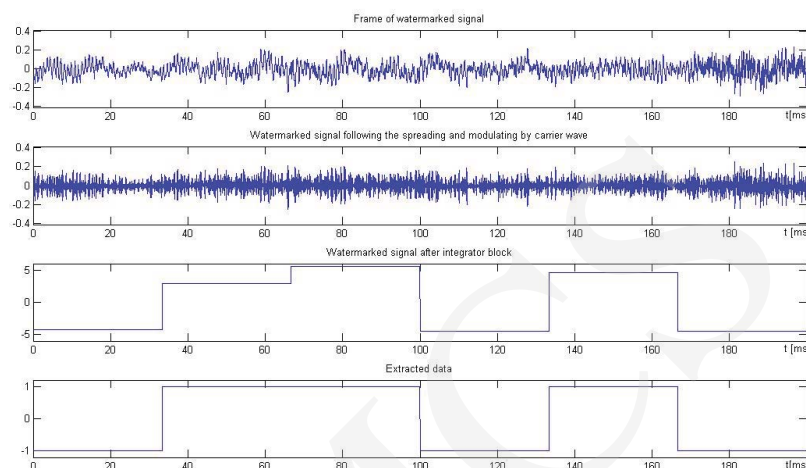
Fig. 8. Signal curves in the watermark extractor.

include: filtration, adding noise to signal, changing sampling frequency, lossy compression, watermarked signal modulation (amplitude modulation, adding a choir effect – reverberation, adding vibrations), de-synchronization. The paper presents an analysis of watermarking system robustness to some of the above–listed attacks. The parameter used in this paper to quantitatively determine the watermarking system robustness is a bit error rate defined by equation 11:

$$BER[\%] = \frac{1}{100} \sum_{i=1}^{L_b} \begin{cases} 1, & d^*(i) \neq d(i) \\ 0, & d^*(i) = d(i) \end{cases} \tag{11}$$

where:
$d(t)$ – hidden watermark sequence;
$d^*(t)$ – reproduced watermark sequence.

3.2. Watermarked signal quality assessment

The watermark embedder is a program designed to convert the input data stream H(t) into the watermarked signal Hwm(t), containing a watermark. The level of watermarked signal deformation as compared with the original one can be presented using a 5–level MOS (Mean Opinion Score) scale, shown in Table 2.

The quality assessment test covered a group of 10 officer cadets who previously read Table 2. 5 audio files different in dynamics and frequency characteristics were selected for assessment. All files featured a sampling frequency of fp=48kHz. The test consisted in listening to the original signal, then to the watermarked one, then the test participants were requested to define the

Table 2. Level of watermarked signal deformation in the MOS scale.

MOS	Degree of distortion
5	Imperceptible
4	Perceptible but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

Fig. 9. Process of adding noise to the watermarked signal.

level of noise. The test results are presented in Table 3. They show that wa-

Table 3. MOS scale based assessment of watermarked signal noise compared with the original one

Test signal	Average MOS
ABBA	4.44
Omega	4.67
Queen	4.22
Vangelis	4.44
Suzanne Vega	3.67
FINAL	4.29

termark is audible by the listener, but its presence does not compromise the watermarked signal quality vs. the original signal quality. Only in the case of the latter track under analysis (Suzanne Vega), the watermarked embedded in the original signal weakens its quality.

3.3. Adding noise to the watermarked signal

Attempts to destroy the embedded watermark in the audio signal can be made by adding noise signal n(t) at a proper power to the watermarked signal Hwm(t). As a result, we receive the signal Hwm*(t). The schematic diagram of the process described is presented in Figure 9. Each time the power of noise signal n(t) was selected to obtain a specific value of watermarked signal-to-noise
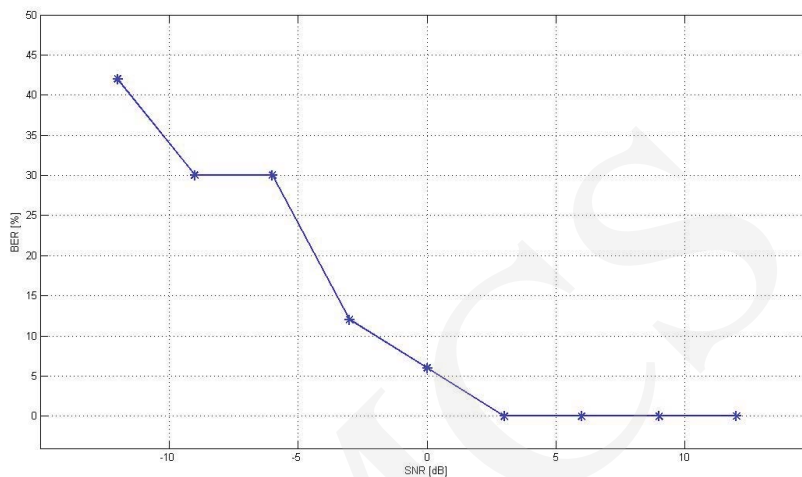
Fig. 10. Robustness of watermarking system to adding noise to the watermarked signal.

ratio according to equation (12):

$$SNR[dB] = 10log_{10} \left\{ \frac{\sum_{i=1}^{N} H_{wm}^2(i)}{\sum_{i=1}^{N} [H_{wm*(i)-H_{wm}(i)}]^2} \right\}. \tag{12}$$

As a result of decoding the signal Hwm*(t), we receive the signal d*(t). The interdependence between the bit error rate (BER) calculated according to equation 11 as a function of signal-to-noise ratio is presented in Figure 10. If the noise signal power is twice lower than watermarked signal power (SNR=3dB), then the watermarking system can decode a watermark with no errors. Adding to the watermarked signal noise with power equalling the signal (SNR=0dB) results in bit error rate (BER) at a level of 6.67%, which is a satisfying value. Increasing the noise power results finally in a situation when the noise signal has a higher power than the watermarked signal (SNR<0dB), at the same time bit error rates (BER) exceed 10%.

3.4. Change in the sampling frequency

Another attack under analysis aiming at destroying the watermark is a change in the watermarked signal sampling frequency. The watermark was embedded in the original signal at a sampling frequency of fp=48kHz. Then, using the

CoolEdit 2000 environment researchers changed the signal sampling frequency by raising it to 96kHz ("frequency change upwards") and lowering it to 44.1kHz ("frequency change downwards"). Signals modified this way were subjected to another change in sampling frequency returning to the initial frequency of 48kHz, then, it was sent to the watermark extractor input. After decoding the watermark, the bit error rate was calculated using equation (11). The test results are presented in Table 4.

Table 4. System robustness to change in the watermarked signal sampling frequency.

Analizowany sygnał, audio	BER [%]	BER [%]
	"Upsampling"	"Downsampling"
ABBA	0.00	0.00
Omega	0.00	0.00
Queen	0.00	0.00
Vangelis	0.00	0.00
Suzanne Vega	0.00	0.00

The watermarking system is fully robust to changes in the watermarked signal sampling frequency. No matter whether the frequency value was increased (from 48kHz to 96kHz), or lowered (from 48kHz to 44.1kHz), the extractor played a watermark for all tracks without any error.

3.5. Lossy compression

The latter attack covering the watermark to be analyzed in this paper is defining a specific lossy compression signal. The lossy compression consists in lowering the number of bits necessary to express an item of information. The lossy compression algorithms are based on a psychoacoustic model. The paper used Cool Edit 2000 environment, where the lossy compression algorithm acts based on a psychoacoustic model developed by the Fraunhoffer Institute and Thompson company. The watermarked signal was subjected to lossy compression in MPEG 3 (layer–3) standard at different compression levels. Then, decompression to the original format was performed and the signal was sent to the extractor input. Once the watermark was decoded, the bit error rate was calculated using the equation (11). The curve of bit error rate depending on the compression level is presented in Figure 11. For the compression levels lower than 6.9:1 we obtain error-free watermark reproduction, whereas for the compression ratios of 8.0:1 and 9.6:1, the bit error rate amounts to 3.33%. The results obtained allow to confirm that the watermarking system is, to a large extent, robust to lossy compression for the compression ratios lower than 9.6:1.
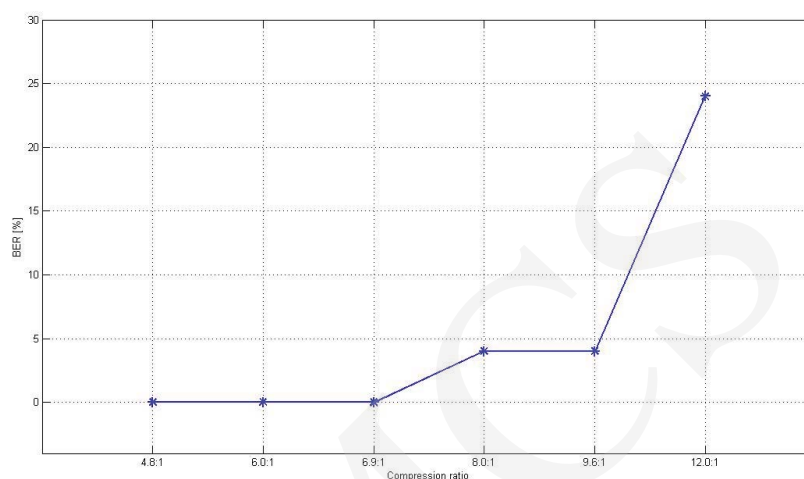
Fig. 11. Watermarking system robustness to the watermarked signal
lossy compression.

It corresponds to lowering the throughput for the watermarked audio signal
from 768kbps to 80kbps (for audio signal with a sampling frequency of 48kHz
and a resolution of 16 bits per sample). Lossy compression at a higher compression ratio produces an excessive number of erroneously decoded watermark
bits.

4. Conclusions

The above-described principle of operation for the watermark embedder and
extractor as well as the results of quality and robustness analysis of the watermarking system developed show that the DSSS method can be used as a
dedicated solution for the robust audio signal watermarking. Our further research should cover enhancing the algorithms by developing a proper transmission protocol for the coder to increase the amount of information hidden, at the
same time increasing the processing gain value. The extractor can be enhanced
by developing a more effective method to synchronize the watermarked signal
with the pseudo random sequence.

References

[1] Chun-Shien L., Multimedia Security: steganography and digital watermarking techniques for protection of intellectual property, Londyn, Idea Group Publishing (2005).

[2] Simon M. K., Omura J. K., Scholtz R. A., Levitt B. K., The Spread Spectrum Handbook, New York, McGraw-Hill (1994).

[3] Haykin S., Systemy telekomunikacyjne, Tom 2, Warszawa, WKiL (1998).

[4] Jackowski S., Telekomunikacja, Czesc 1, Radom, Wydawnictwo Politechniki Radomskiej (2005).

[5] Cox I. J., Miller M. L., Bloom J. A., Friedrich J., Kalker T., Digital watermarking and steganography, second edition, USA, Morgan Kaufman (2008).

[6] Zielinski T. P., Cyfrowe przetwarzanie sygnalow. Od teorii do zastosowan, Warszawa, WKiL (2007).