Integer Wavelet Transform Based Watermarking System for PCM Signals

Mario González-Lee[1], Luis J. Morales-Mendoza[1], Rene F. Vázquez-Bautista[1], Celia Calderon-Ramon[1], Héctor Pérez-Meana[2], and Mariko Nakano-Miyatake[2]

[1] Facultad de Ingenieria en Electronica y Comunicaciones, Universidad Veracruzana
Av. Venustiano Carranza S/N, Col. Revolucion.
Poza Rica Veracruz, Mexico
{mgonzalez01,javmorales,favazquez}@uv.mx
[2] Seccion de Estudios de Posgrado e Investigacion
Instituto Politecnico Nacional
Av. Santa Ana No. 1000, Col. San Francisco Culhuacan
Mexico D.F.

Abstract. In this paper, a watermarking system based on the integer wavelet transform is proposed. A set of different wavelets was chosen for this propose so their performance as watermark transmission channel can be analyzed. The watermark is embedded in the second decomposition level of the wavelet transform of a PCM signal in a block wise approach, the coefficients of the wavelet transform are assumed to be accurately modeled as a Laplacian channel. A set of attacks were performed in order to test system's robustness, we found that this approach is robust against several attacks such as additive white noise, low pass filtering, cropping, among other attacks; the best suited wavelet class for developing watermarking systems was identified too.

Keywords: Watermarking, Integer wavelet transform, Optimal detector, Laplacian channel.

1 Introduction

A very handy approach to watermarking is to establish analogies to the very strong field of the theory of communications, in this context, we can think of a watermark as a signal that propagates through a communications channel.

The watermarking model is shown in Fig. 1, we can identify the main input variables: the cover, which is an audio signal that will carry the watermark, a user's key which is used to generate a pseudo-random signal and the embedding gain which is related to the embedding energy of the watermark.

In this work, a watermark is a binary signal $\mathbf{W} = [w_i]$ with $w_i \in \{-1, 1\}$ and is zero mean with variance 1. This watermark is embedded in the cover $\mathbf{X} = [x_i]$ so we get the watermarked signal $\mathbf{Y} = [y_i]$; Ideally, the cover does not interfere the watermark, however in practice this is not true, in consequence, we model the effects of the cover within the channel block, and attacks to the watermark are modeled as noise in the channel during the propagation of the watermark.

J.A. Carrasco-Ochoa et al. (Eds.): MCPR 2013, LNCS 7914, pp. 185–194, 2013.

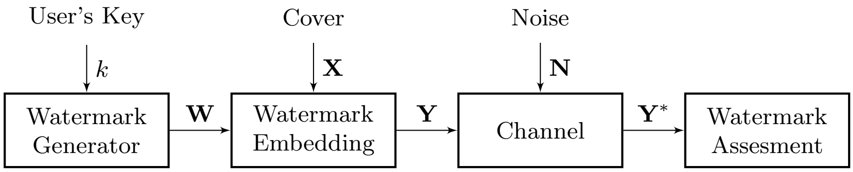

Fig. 1. Watermark propagation model

Once the watermark reaches the detector, it has to assess the presence of the watermark, usually by computing two statistics, one that measures the presence of the watermark in the possibly watermarked audio signal and the other is a threshold. If the computed statistics surpasses the threshold value, the watermark is detected, otherwise, the watermark is considered to be absent. The computed statistics are often known as the decision variable d and the decision threshold Th respectively.

The remaining of this paper is organized as follows in Sect. 1.1 related works are discussed, then in Sect. 1.2 we will introduce the integer wavelet transform, in Sect. 2 the proposed system is discussed in detail and the corresponding experimental results are presented in Sect. 3; finally the conclusions in Sect. 4 and references.

1.1 Related Works

There are some works in the field of audio watermarking that make use of the Laplacian channel model, for example, in [1] a method that uses MPEG 1 Layer 3 compression as a reference to determine where and how the watermark must be embedded. The results show that their watermarking scheme is robust against the attacks of the audio Stirmark benchmark.

The authors of [2] change the length of the intervals between salient points of the audio signal to embed data. Their results suggest that the algorithm is robust to common audio processing operations e.g. MP3 lossy compression, low pass filtering, and time-scale modification. The major drawback of this proposal is its low bit embedding rate.

In [3], the analysis filterbank decomposition, the psychoacoustic model and the empirical mode decomposition (EMD) techniques are used. This algorithm embeds the watermark bits in the final residue of the subbands in the transform domain. The authors claim that the scheme is robust against MP3 compression and Gaussian noise attacks. A drawback is that it might not be robust to common attacks such as band-pass filtering and cropping.

The adaptive tabu search (ATS) has been explored in order to develop watermarking systems, in [8], the Daubechies wavelet decomposition is used for watermark embedding. The optimal intensity of watermark is searched by using the ATS. Experimental results show that watermark is inaudible and robust to many digital signal processing, such as resamplig, cropping, low pass filter, additive noise and MP3 compression.

In [4,5], the authors propose to use the Laplacian channel model for audio signal in temporal domain, they derive an optimal detector and the threshold equation, their system is semi-fragile, but has several advantages, first, since no transform domain is involved, the resulting system has very low computational complexity, second, memory requirements could be easily satisfied, and finally, the system performs very well for attacks such as additive white noise, cropping and echo attacks.

In this paper, propose a system in the Integer Wavelet Transform domain assuming they can be accurately modeled as a Laplacian communication channel, so the optimal detector equation and the threshold equation proposed in [4,5] can be used for detecting watermarks. Several Wavelets were tested in order to select the best suited wavelet for a robust watermarking system. Once identified, a watermarking system was derived, the resulting scheme has low complexity and it proved to perform remarkably well, furthermore, it proved to be unaffected for low pass filtering attack and additive white noise and phase inverting attack whilst it is almost unaffected for cropping attack and echo attacks.

1.2 Integer Wavelet Transforms

An Integer Wavelet Transform (IWT) maps a set of integers into another set of integers, this property is well suited for signal coding for example in lossless image compression [7]. In our case, integer coefficients reduce damage to embedded watermark caused by quantization, in addition, the well know advantages of the IWT can be exploited for robust watermarking systems design.

The following equations are mathematical definitions of several IWT used in this work, here the notation IWT(m,n) is used to denote an IWT with m and n vanishing moments in the analysis and synthesis high pass filters respectively ([6,7]).

IWT(1,1):

$$d_n = x_{2n+1} - x_{2n} \quad , \tag{1}$$

$$s_n = x_{2n} + \left\lfloor \frac{d_n}{2} \right\rfloor \quad . \tag{2}$$

IWT(2,4):

$$d_n = x_{2n+1} - \left\lfloor \frac{1}{2}(x_{2n} + x_{2n+2}) + \frac{1}{2} \right\rfloor \quad , \tag{3}$$

$$s_n = x_{2n} + \left\lfloor \frac{19}{64}(d_{n-1} + d_n) - \frac{3}{64}(d_{n-2} + d_{n+1}) + \frac{1}{2} \right\rfloor \quad . \tag{4}$$

IWT(4,2):

$$d_n = x_{2n+1} - \left\lfloor \frac{9}{16}(x_{2n} + x_{2n+2}) - \frac{1}{16}(x_{2n-2} + x_{2n+4}) + \frac{1}{2} \right\rfloor \quad , \tag{5}$$

$$s_n = x_{2n} + \left\lfloor \frac{1}{4}(d_{n-1} + d_n) + \frac{1}{2} \right\rfloor \quad . \tag{6}$$

IWT(6,2):

$$d_n = x_{2n+1} - \left\lfloor \frac{75}{128}(x_{2n} + x_{2n+2}) - \frac{25}{256}(x_{2n-2} + x_{2n+4}) + \frac{3}{256}(x_{2n-4} + x_{2n+6} + \frac{1}{2}) \right\rfloor \quad , \tag{7}$$

$$s_n = x_{2n} + \left\lfloor \frac{1}{4}(d_{n-1} + d_n) - \frac{1}{2} \right\rfloor \quad . \tag{8}$$

In this paper, we denote the coefficient sequence ant the *k-th* decomposition level of any given signal as $\mathbf{S}_{(k)} = [s_{(k)i}]$ and its related detail sequence as $\mathbf{D}_{(k)} = [d_{(k)i}]$. The operator $\lfloor \cdot \rfloor$ denotes the floor operation.

The IWT equation set just introduced is going to be used for computing the IWT for the proposed watermarking system discussed right away.

2 Proposed System

In this section, we introduce the proposed system, as we stated early in this paper, the system embeds a watermark in the second decomposition label of the IWT, first, the signal is divided in non overlaping blocks of length 16 times the sampling frequency, then the IWT of the blocks is computed, the detail sequences $\mathbf{D}_{(1)}$ and $\mathbf{D}_{(2)}$ are stored for later use, and the wavelet coefficients $\mathbf{S}_{(2)}$ is watermarked, the audio signal is reconstructed by reversing the transformations as shown in Fig. 2. At the receiver, the system computes the second level IWT decomposition in order to detect the watermark as shown in Fig.3.

The embedding rule and detection variables are discussed in next section.

2.1 Embedding Algorithm

The embedding algorithm is the multiplicative embedding rule since it exhibits several desirable properties, for example, one the most important is the masking effect that allows greater embedding strength while imperceptibility holds. The multiplicative watermark embedding rule is given as:

$$s'_{(2)i} = s_{(2)i}(1 + gw_i), \tag{9}$$

where $s'_{(2)i}$ is the *i-th* watermarked coefficient, w_i is the *i-th* watermark bit and g is the watermark embedding gain, which controls the watermark energy.

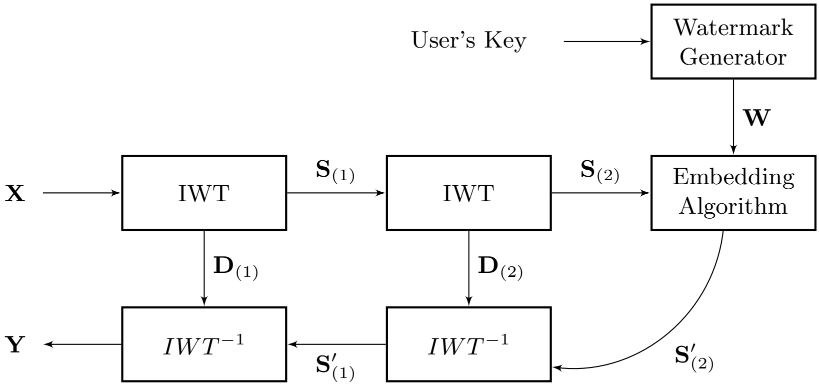

Fig. 2. Block diagran for the proposed watermarking system. IWT is the forward integer wavelet transform whilst IWT^{-1} is the inverse of IWT.

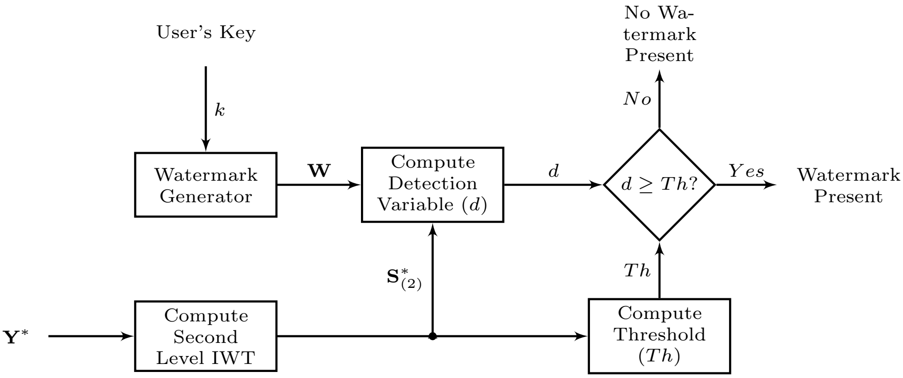

Fig. 3. Block diagram for a generic watermark detection system

2.2 Optimal Watermark Detector for Laplacian Channel

The coefficients of the IWT, and thus the channel, are considered to be statistically modeled using a Laplacian PDF, so the optimal detector variable is computed as [4,5]:

$$d = \hat{g} = \frac{1}{N\alpha_y} \sum_{i=1}^{N} |s'_{(2)i}| w_i \quad , \tag{10}$$

where,

$$\alpha_y = \frac{1}{N} \sum_{i=1}^{N} |s'_{(2)i}| \quad . \tag{11}$$

Equation (10) can not be used "as is" since the IWT coefficients must to be normalized first. We found that as one goes deeper in wavelet decomposition, the resulting coefficients must to be scaled to ensure they hold the Laplacian channel assumption, our experiments showed that in order to enforce the IWT coefficients to hold the Laplacian assumption, (10) should be scaled this way:

$$d = k\frac{1}{N\alpha_y}\sum_{i=1}^{N}|s'_{(2)i}|w_i \quad , \tag{12}$$

where $k = \frac{1}{32}$. One can normalize individual samples, however for a block size of L, L divisions must to be made, whilst by scaling (10) it takes just one.

In order to properly detect the watermark, the decision variable d must be compared to a threshold, a watermark is present if $d \geq Th$, the general threshold equation derived from the Neyman-Person criterion in [4,5] is:

$$Th = erfc^{-1}(1 - 2p_{fp})\sqrt{\frac{2}{N}}, \tag{13}$$

where $erfc^{-1}(\cdot)$ is the inverse error function complementary.

Since (13) does not depend upon the samples, then no scaling is needed for this equation. Equations (12) and (13) are meant to be used for detecting watermarks in the IWT domain, the corresponding experimental results are shown in next section.

3 Computer Simulations

All test were carried out under the following scenario: the watermark was embedded in non overlapping blocks with length of 16 times the sampling frequency of an audio signal using (9). Detection is made in the same block wise approach, d and Th are computed for each block using (10) and (13) and the responses for each block are accumulated and averaged. We let $P_{fp} = 10^{-6}$ and the embedding gain was set to 0.5. All audio signal used for our tests were uncompressed 16-bit stereo WAV files with 48000 Hz sampling rate.

We repeated the same attacks using the wavelets defined in section 1.2, so we can identify which of them has the best performance for practical watermarking system design. The performance of the system is presented in next section.

3.1 Main Results

In this section we present simulation results that validate that (12) and (13) provide an accurate watermark detection model. Several audio were used in the following test, however, we show our worst case.

First we tested detector performance; we present in Fig.4 the response for IWT(4,2). Given that watermark number 500 is the watermark that was embedded in an audio signal. One can verify that the response is way bigger that any other watermark, furthermore, the response to any watermark different to the one embedded is far from crossing the decision threshold which confirms the Laplacian channel model is accurate for IWT(4,2) coefficients.

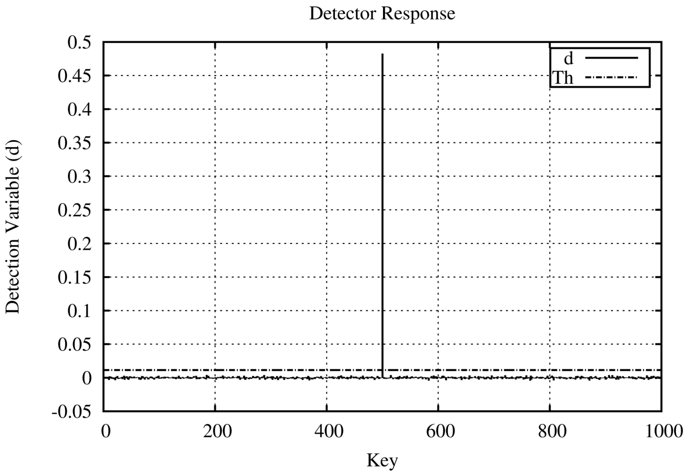

Fig. 4. Detector characteristics. It is shown that the system performs remarkably for IWT(4,2) even for our worst case host signal.

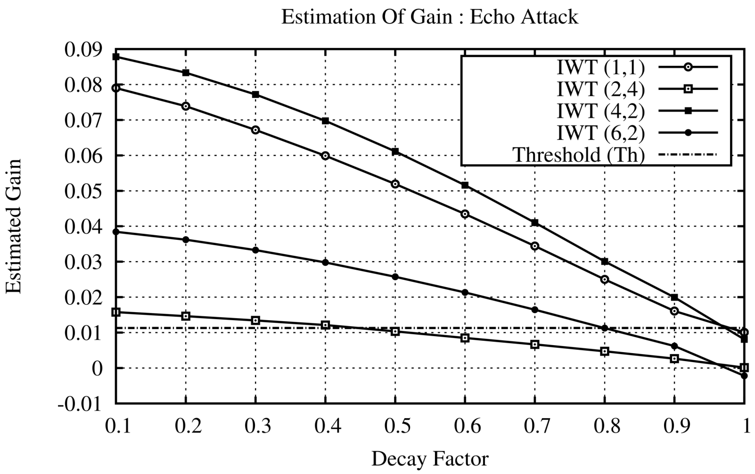

Fig. 5. Watermark detection for different IWT for various values of the decay factor

In the second test, we added a echo signal, The echo signal has a delay of 1 second for a given value of decay factor which determines how fast the echo fades away. In Fig.5 the response for various values of the decay factor is shown, it is clear that IWT(4,2) perform better whilst IWT(2,4) performed poorly.

In the next test, additive white noise was added to the carrier signal, the detection characteristic for several values of amplitude of the noise is shown in Fig.6. It can be seen that again IWT(4,2) outperforms the other wavelet decompositions and once again IWT(2,4) has the worst performance.

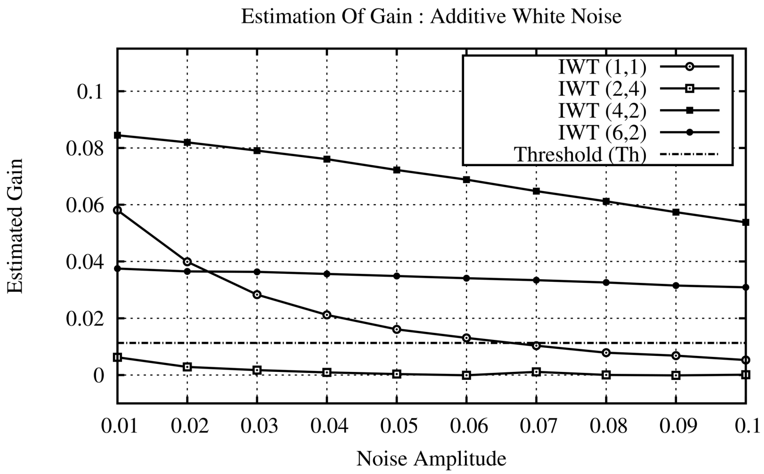

Fig. 6. Detector characteristics for additive white noise at various amplitudes

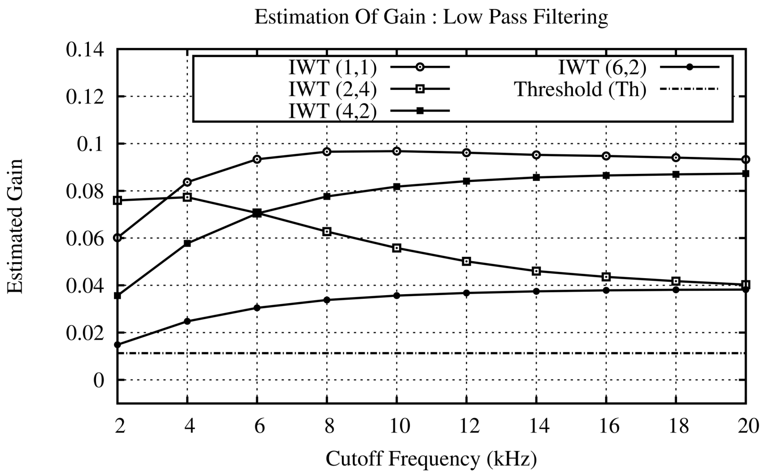

Fig. 7. Low pass filtering attack for various cutoff frequencies

A low pass filter was applied and detector response was measured for various cutoff frequencies, the results are shown in Fig.7, in this test all wavelets exhibit very good performance against low pass filtering with IWT(1,1) being the best and IWT(6,2) was the worst.

Another attack is cropping, Fig.8 shows detection characteristics as the cropping percentage is varying. In this case IWT(4,2) performed best again and once more IWT(2,4) performed poorly.

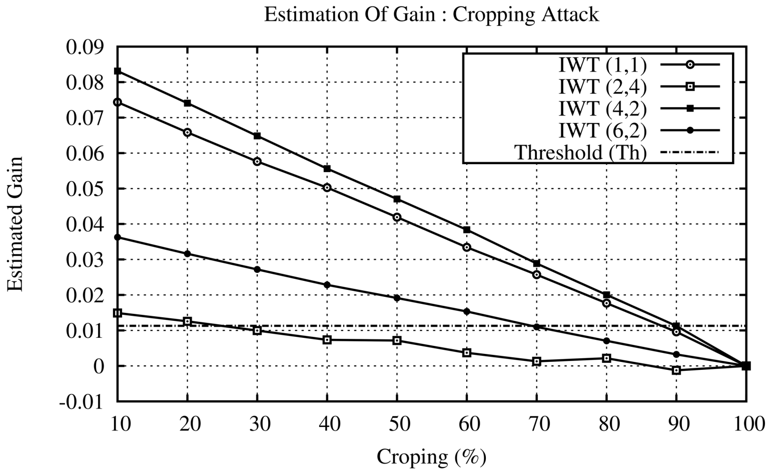

Fig. 8. Cropping attack; horizontal axis is the cropping percentage

Inverting attacks exploits the fact that the human auditory system is unable to perceive phase changes, the system response is $d = 0.0813$, $d = 0.0162$, $d = 0.0899$, $d = 0.0394$ for wavelets IWT(1,1), IWT(2,4), IWT(4,2) and IWT(6,2) respectively ,threshold value is $Th = 0.0113$. So the system is not affected by inverting attacks and once more, IWT(4,2) is clearly the winner.

These results clearly show a trend: the IWT(4,2) is best in almost all cases whilst IWT(2,4) is worst in almost all cases.

4 Conclusions

We can draw several conclusions, first, IWT based watermarking proved to be more robust than other previously proposed systems, it has the additional advantage of being a blind algorithm, so the original audio is not needed.

Second, IWT(4,2) is the best choice among the set of wavelets tested in this paper whilst IWT(2,4) performed worst in almost all test. Furthermore, a test not shown in this paper using IWT(2,2) performed so bad that we decided not to include it in the tests presented in this work, its poor performance prevented the proposed structure from detecting watermarks accurately even when no attacks were made.

The resulting system has low complexity so the detection of the watermark can be done in a couple of seconds for a full length song using a consumer laptop.

Future works include: Analyzing the performance of other classes of IWT and to investigate the influence of different decomposition levers in the wavelet transform and the proper derivation of the scaling factor of (12).

Acknowledgments. The authors want to thank the FIEC Poza Rica and the Universidad Veracruzana for their support to this work.

References

1. Megias, D., Herrera-Joancomarti, J., Minguillon, J.: A robust frequency domain audio watermarking scheme for monophonic and stereophonic PCM formats. In: Procedings of the 30th Euromicro Conference, pp. 1089–6503 (2004) ISSN 1089-6503
2. Mansour, M.F., Tewfik, A.H.: Audio watermarking by time-scale modification. In: Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001), vol. 3, pp. 1353–1356 (2001) ISSN 1520-6149
3. Wang, L., Emmanuel, S., Kankanhalli, M.S.: EMD and psychoacoustic model based watermarking for audio. In: Procedings of the 2010 IEEE International Conference on Multimedia and Expo (ICME), pp. 1427–1432 (July 2010) ISSN 1945-7871
4. Gonzalez-Lee, M., Morales-Mendoza, L.J., Vazquez-Bautista, R.F., Morales-Mendoza, E.: Optimal Detection of Multiplicative Watermarks in Laplacian Channels. In: Proceedings of the 9th International Conference on Electrical Engineering, Computing Science and Automatic Control, Mexico City, Mexico (2012)
5. Gonzalez-Lee, M., Morales-Mendoza, L.J., Vazquez-Bautista, R.F., Morales-Mendoza, E.: Semi Fragile Watermarking System in Temporal Domain. In: CERMA 2012, Cuernavaca Morelos, Mexico (2012)
6. Calderbank, R., Daubechies, I., Sweldens, W., Yeo, B.L.: Lossless Image Compression Using Integer To Integer Wavelet Transform. In: Proceedings of the International Conference on Image Processing, vol. 1, pp. 596–599 (October 1997)
7. Sheng, F., Bilgin, A., Sementilli, P.J., Marcellin, M.W.: Lossy and Lossless Image Compression Using Reversible Integer Wavelet Transform. In: Proceedings of the 1998 International Conference on Image Processing, vol. 3, pp. 876–880 (1998)
8. Sriyingyong, N., Attakitmongcol, K.: Wavelet-Based Audio Watermarking Using Adaptive Tabu Search. In: Proceedings of the 2006 1st International Symposium on Wireless Pervasive Computing, pp. 1–5 (2006)