International Journal of the Physical Sciences Vol. 6(16), pp. 3837-3850, 18 August, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.577
ISSN 1992-1950 ©2011 Academic Journals

Review

A review of audio based steganography and digital watermarking

M. L. Mat Kiah[1], B. B. Zaidan[2,3,4], A. A. Zaidan[2,3,4]*, A. Mohammed Ahmed[1] and Sameer Hasan Al-bakri[1]

[1]Department of Computer System and Technology, Faculty of Computer Science and IT, University of Malaya, 50603 Kuala Lumpur, Malaysia.
[2]Faculty of Engineering, Multimedia University Jalan Multimedia, 63100 Cyberjaya, Selangor, Malaysia.
[3]Predictive Intelligence Research Cluster, Sunway University, Selangor, Malaysia.
[4]Institute of Postgraduate Studies/ Research and Development Group/Al-Madinah International University, Malaysia.

Accepted 07 June 2011

With the increasing usage of digital multimedia, the protection of intellectual property rights problem has become a very important issue. Everyday, thousands of multimedia files are being uploaded and downloaded. Therefore, multimedia copyrights become an important issue to protect the intellectual property for the authors of these files. In this paper, the domains of digital audio steganography, the properties of H.A.S, the audio and the digital representation transmission environments, and its software metric, are discussed. The main purpose of this paper is to provide a proper background on the usage of audio file for the purpose of implementing new approaches and techniques in digital watermarking and steganography.

Key words: Digital audio, steganography, data hidden domains, H.A.S, copyright, intellectual property, audio environments, digital representation, watermarking and transmission environment, software metrics.

INTRODUCTION

Security is defined as the degree of protection against danger, damage, loss, and criminal activity (Chandra and Khan, 2008; Alanizi et al., 2010b; Jayakumar and Thanushkodi, 2008; Mohammed et al., 2011a; Mohammed et al., 2011b). Particularly when a sensitive message is to be delivered to a destination, authentication and confidentiality are required (Al-Frajat et al., 2010; Wang et al., 2010; Raad et al., 2010). Providing security for electronic documents is an important issue (Zaidan et al., 2010h; Alanizi et al., 2010a). In information security, confidential information or confidential data must only be used, accessed, disclosed or copied by users who have the authorization, and only when there is a real need (Nabi et al., 2010). While integrity means that data cannot be modified without authorization (Abu Ali et al., 2010), non- repudiation provides the accountability service, that is a receiver cannot deny having received the data nor can the other party denies having sent a data (Naji et al., 2009; Abomhara et al., 2010a, b; Zaidan et al., 2010f).

The term “Security through Obscurity” or “Security by Obscurity” is the belief that a system of any sort can be secure so long as nobody outside of its implementation group is allowed to find out anything about its internal mechanisms (Shihab et al., 2010; Zaidan et al., 2011a; Zaidan et al., 2011b). Data hidden considered as “Security by Obscurity” systems (Zaidan et al., 2010e). Numbers of techniques have been implemented towards improving secure data hidden approaches. They tried to overcome two main problems, which are the amount of data hidden and the secrecy of the data against the attackers. (Ping et al., 2010; Zaidan et al., 2010i)

Several packages now exist for hiding data in audio files (Medani et al., 2011), such as MP3Stego, which not only effectively hides arbitrary information, but also claims to be a partly robust method of watermarking MP3 audio files (Noto, 2001). The windows wave format lets users hide data using Steghide, it alters the least significant bits (LSB) of data in the carrier medium (Artz, 2001). All steganography techniques have to satisfy two

*Corresponding author. E-mail: aws.alaa@gmail.com.

basic requirements:

1. The first requirement is perceptual transparency or noticeable perceptual distortion which means the cover or carrier (that is, object not containing any additional data) and stego object (that is, object containing secret message) must be perceptually indiscernible (Anderson and Petitcolas, 1998);
2. The second requirement is high data rate of the embedded data.

Research objectives

Commonly, data hidden has two general techniques, which are, digital watermarking and steganography. According to the researchers, data hidden approaches have two main limitations, the size of the hidden data and the robustness of the watermark techniques. In this research we will try to achieve the following objectives:

1. To analyze the features of audio file that can be used to implement the high rate data hiding;
2. To investigate the approaches used in audio watermarking domains, audio environment for implementing a secure, robust and high rate data hiding in the audio files;
3. To carry out intensive literature reviews of the existing techniques and illustrate the advantage and the disadvantage of each technique;
4. To identify the software metrics used to evaluate the audio watermarking approaches in data hiding.

Literature review

Audio watermarking or audio steganography started consider later as attractive area that have viable applications and space for development (Zhang et al., 2010a, b; Abdulfetah et al., 2010a, b). In the past few years, several techniques for data hidden in audio sequences have been presented. All of the developed techniques take benefit of the perceptual properties of the human auditory system (HAS)

The main challenge in digital audio watermarking and steganography is that if the perceptual transparency parameter is fixed, the design of a watermark system cannot obtain high robustness and a high watermark data rate at the same time (Cvejic, 2004; Yang et al., 2009). To achieve any of data hidden goals, we need to select a proper cover, domain, and take into the account the challenges of data hidden approaches.

Arnold (2000) has tried to improve the performance of the original patchwork algorithm. Arnold's algorithm is a landmark in the area of watermarking research, especially for patchwork algorithm. Moreover, the performance of this algorithm in terms of inaudibility and robustness has been shown to be satisfactory by many researchers such as (Yeo and Kim, 2003). They have derived mathematical formulations that help to improve robustness. The core idea of the improved scheme is called the Modified Patchwork Algorithm (MPA) which can enhance the power of the original patchwork algorithm considerably.

Large work has been carried out in audio watermarking using spread spectrum technology and is presented in several key publications like (Bender et al., 1996), (Cox et al., 2002) and (Cvejic, 2004). The first method of spread spectrum into watermarking was in (Cox et al., 1997). Xu et al. (1999) proposed a multiple echo technique. Rather than embedding one large echo into the host audio signal, they use multiple echoes with different offsets. Oh et al. (2001) introduced the positive-negative echo hiding scheme. Their echo kernels comprise positive and negative echoes at nearby locations. Since the frequency response of a negative echo is the inversed shape having similar ripples as that of a positive echo, the frequency response of the positive and negative echoes has the smooth shape in the low frequency band. By employing positive and negative echoes, one can thus embed multiple echoes to allow that the host audio quality is not apparently deteriorated. Kim and Choi (2003) presented an echo hiding scheme with backward and forward kernels. The theoretically-derived results show that the amplitude of the cepstrum coefficient at the echo position from the backward and forward kernels is bigger than that from the backward kernel only when the embedded echoes are symmetric. Therefore, the backward and forward kernels can improve the robustness of echo hiding scheme.

Ko et al. (2005) went further to propose the time-spread echo kernel. With the use of pseudo-noise sequence, an echo is spread out as numerous little echoes in a time region. When the embedded data of watermarked audio signals are extracted, the pseudo-noise sequence functions like a secret key. Without obtaining the pseudo-noise sequence used in the embedding process, extracting the embedded data would be harder.

In order to add a watermark into a host signal in a perceptually transparent manner, a wide range of embedding techniques are proposed going from simple least significant bits (LSB) scheme or Low-bit encoding, Phase coding, Spread spectrum, Patchwork coding, Echo coding and noise gate technique. In the Table 1, we summarized each approach with their advantage and disadvantage

METHODOLOGY

According to Chandra and Khan (2008), we have adapted a general methodology for researcher whom concern about doing research on steganography and digital watermarking (Figure 1). According to (Zaidan et al., 2010a, b, c), steganography discuses different issues such as size of data hidden, the secrecy of the information, the available attackers to the stego files and the visibility of the noise in the stego-object, while digital watermarking concern about,

Table 1. The summary of literature.

Approach	Summary	Advantage and Disadvantage
Low–bit Encoding	Low-bit encoding considered as the earliest techniques implemented in the information hiding of digital audio. It is the simplest technique to embed data into other data structures such as data of audio in image file or data of image in audio file. Low-bit encoding, can be done by replacing the LSB of each sampling point by a coded binary string (hidden data)	The major advantage of Low-bit encoding are: 1. High watermark channel bit rate 2. Low computational complexity of the algorithm compared with others techniques 3. No computationally demanding transformation of the host signal, therefore, it has very little algorithmic delay The major disadvantage is that the method are: 1. Low robustness, due to the fact that the random changes of the LSB destroy the coded watermark 2. it is unlikely that embedded watermark would survive digital to analogue and subsequent analogue to digital conversion
Phase Coding	Phase Coding watermarking works by substituting the phase of an initial audio segment with a reference phase, this phase represents the hidden data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments	The major advantage of Phase Coding are: 1. Basic technique The major disadvantage is that the method are: 1. Phase coding method is a low payload because the watermark embedding can be only done on the first block. 2. The watermark is not dispersed over the entire data set available, but is implicitly localized and can thus be removed easily by the attackers
Spread Spectrum Technique	Spread spectrum (SS) is technique designed to encode any stream of information via spreading the encoded data across as much of the frequency spectrum as possible. even though, there is interference on some frequencies, SS allows the signal reception,	The major advantage of Spread Spectrum are: 1. Difficult to detect and/or remove a signal 2. Provide a considerable level of robustness The major disadvantage is that the Spread spectrum are: 1. Spread spectrum technique used transform functions (e.g. Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), or Discrete Wavelet Transform (DWT)) with appropriated inverse transform function, which can cause a delay. 2. Spread spectrum is not a visible solution for real time applications
Patchwork Coding	Patchwork Coding considered as one of the earliest generation for digital watermarking schemes. Patchwork Coding can be done via embedding the watermark in the audio using time domain or frequency domain. In the literature, several approaches of Patchwork Coding have been proposed on	The major advantage of Patchwork Coding are: 1. Patchwork based watermarking scheme has been confirmed as an valuable to those common signal processing operations, such as low-pass filtering, image/audio compression, and so on. The major disadvantage is that the Patchwork are:

Table 1. Contd.

	frequency domain using linear transformations, such as Discrete Wavelet Transform (DWT), Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT). Frequency or time domain watermarking schemes directly tinker with sample amplitude of audio to embed the watermark	1. An attack called “curve-fitting attack” has been successfully implemented for patchwork watermarking scheme. 2. Patchwork watermarking scheme is sensitive to various synchronization attacks
Echo technique	Echo technique embeds data into a host audio signal by introducing an echo; the hidden data can be adjusted by the two parameters: amplitude and offset, the two parameters represent the magnitude and time delay for the embedded echo, respectively. The embedding process uses two echoes with different offsets, one to represent the binary datum “One” and the other to represent the binary datum “Zero”.	The major advantage of Echo are: 1. The main advantage of echo hiding is that the echo detection technique is easy to implement. The major disadvantage is that the echo hiding technique are: 1. More complicated computation is required for echo detection. 2. Echo hiding is also prone to inevitable mistakes, such as the echo from the host signal itself may be treated as the embedded echo. 3. If the echo added has smaller amplitude, then the cepstrum peak would be covered by the surrounding peaks to make the echo detection an arduous task to perform. A larger echo may increase the accuracy rate of detection but it also easily exposes the system to deliberate attacks, which then affects the sound quality
Noise Gate Technique	Noise gate technique is designed to be an alternative solution for the weakness in the previous approaches, this technique implanted in the time domain. This technique maintains a high quantity of data hidden side by side with robustness. Noise Gate Technique involve two steps approach, the first step, noise gate software logic algorithm has used to obtain a desired signal for embedding the secret message of the input host audio signal. In the second step, standard i^{th} LSB layer embedding has been done for this desired signal by simply replaces the host audio signal bit in the i^{th} layer with the bit from the watermark bit stream, if 16-bit per audio sample used, where (i=1,...,16).	The major advantage of Noise Gate Technique are: 1. High watermark channel bit rate 2. Low computational complexity of the algorithm compared with others techniques 3. No computationally demanding transformation of the host signal, therefore, it has very little algorithmic delay 4. Add level of complexity against Stego-Only Attack and Known Message Attack The major disadvantage is that the method are: 1. Fair robustness 2. Noise Gate technique is weak against Known Cover Attack, Known Chosen Cover or Chosen Message and Known Stego Attack

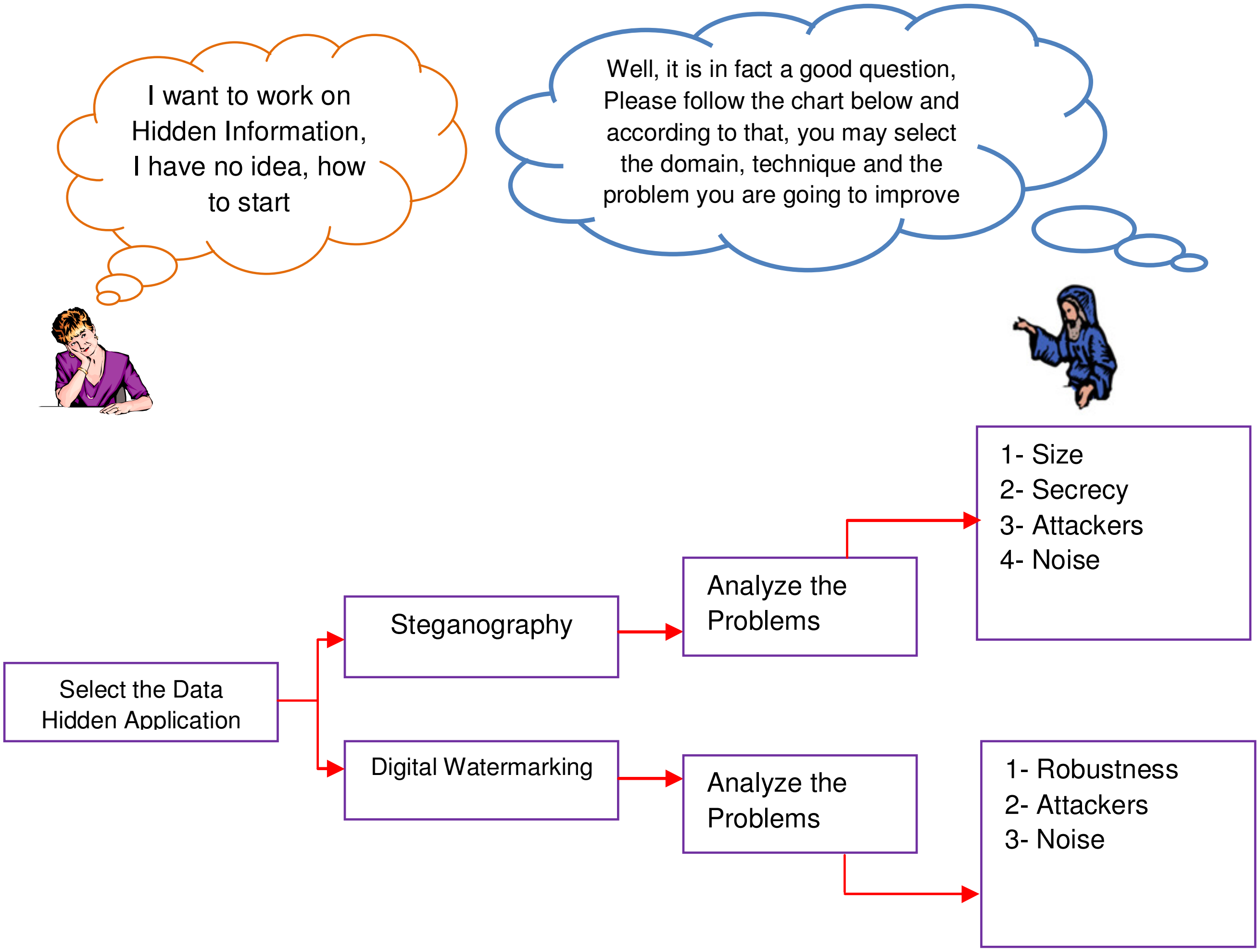

I want to work on Hidden Information, I have no idea, how to start
Well, it is in fact a good question, Please follow the chart below and according to that, you may select the domain, technique and the problem you are going to improve
Select the Data Hidden Application
Steganography
Digital Watermarking
Analyze the Problems
Analyze the Problems
1- Size
2- Secrecy
3- Attackers
4- Noise
1- Robustness
2- Attackers
3- Noise

Figure 1. Research methodology of doing research in steganography and digital watermarking.

robustness, attackers and the noise. Both (steganography and digital watermarking) approaches are at risk of stego-analysis such as stego-only attack, known cover attack, known message attack, known chosen cover or chosen message, known stego attack and other type of attack (Sameer et al., 2011).

Therefore, we need to understand each particular fact before going further on the research. Moreover, identify the range of research, define the problems, select the appropriate domain and technique to solve the problem and finally select the environment of the test and evaluation; can help to put the researcher in the right way (Wang et al., 2011; Zeki and Manaf, 2011).

Summary

Hiding data in audio can be done a number of ways like: Phase coding, Spread spectrum, Echo data hiding, Patchwork coding, Low–bit encoding and Noise gate. The analysis of these techniques shows:

1. Low-bit encoding technique has the highest watermark channel bit rate but with low robust.
2. Noise gate technique can carry more data with fair robustness

Inaudibility, the watermark data rate and robustness to attacks are in the corners of the magic triangle (Figure 2). The magic triangle (Johnson et al., 2001) has displayed simplest requirements of information hiding in digital audio.

This model is suitable for a visual representation of the required trade-offs between the capacity of the watermark data and the robustness to specific watermark attacks, while keeping the perceptual quality of the watermarked audio at an acceptable level. It is not possible to get high robustness to signal modifications and high data rate of the embedded watermark at the same time. Hence, if a high bit rate of the embedded watermark is required from the audio steganography technique, the robustness will be low and vice versa (Ahmed et al., 2010). In additional, we have to take into the account some of the attacks as in Table 2.

No real development for these attackers has appeared in the literature, and therefore, if the researcher involve in the area of stego-analysis or watermarking analysis, they should come with non- standard module.

Audio steganography and audio watermarking domains

Researchers who work in the area of data hidden know that steganography and digital watermarking are using the same concepts and techniques. Steganography and digital watermarking

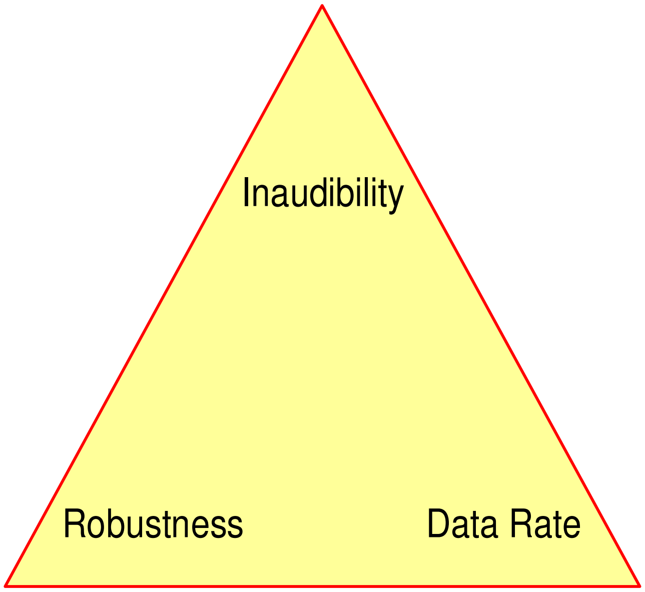


Figure 2. Magic triangle-three contradictory requirements of audio steganography.

Table 2. The attack and the environment of attacking.

Attack	Environment of attacking
Stego-only attack	If the attacker catch only the stego file that contains the hidden data, In this case, attacker try to analyse this stego file. Analysis is done by trial and error
Known cover attack	Once the attacker knows the cover and the original file before embedding data, in this case the attacker will match both and extract the differences that lead to the hidden file
Known message attack	Here the attacker may know the complete hidden message. Thus, the attacker can analyze the file that carries the hidden information, compare it with what it is similar to, and extract the real cover, which probably can be used in the future to extract new hidden information/data
Known chosen cover or chosen message	Here the attacker has part of the real cover or the real message, thus he/she will use the partial matching method with trial and error method to analysis and extract the data
Known stego attack	The goal is known as well as the algorithm of Steganography system, and this is the most dangerous type of attack, because attacker directly applies the algorithm to reconcile the concealed message

techniques are used to protect information, address digital rights management, and conceal secrets. Information hiding techniques provide an interesting challenge for digital forensic investigations. Research into steganalysis techniques aims to analyze and discover the hidden information, moreover, steganalysis techniques lead research toward improved methods for hiding information. Data hidden can be classified according to the domain where the watermarking or steganography has been applied. The following

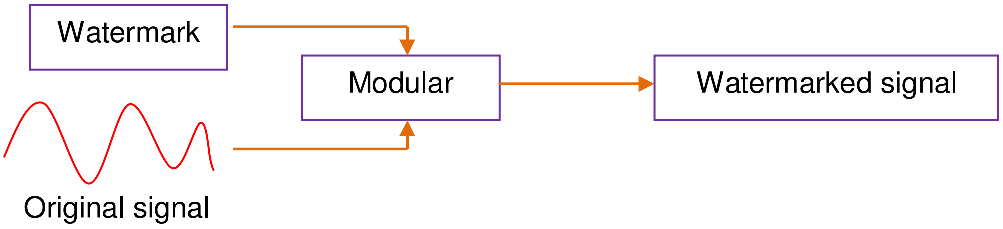


Figure 3. Time domain audio steganography.

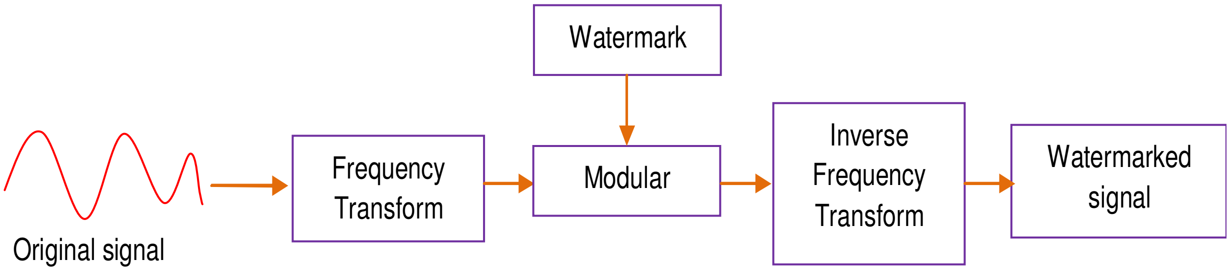


Figure 4. Frequency domain audio steganography (Alsalami and Al-Akaidi, 2003).

sections discus these domains and classify them to four categories (Alsalami and Al-Akaidi, 2003; Ahmed et al., 2010; Sheikhan and Asadollahi, 2010) as thus described.

Time domain audio steganography and digital watermarking

In time domain steganography techniques, watermark is directly embedded into audio signal, where no domain transform is required in this process. Watermark signal is shaped before embedding operation to ensure the robustness (Figure 3).

The existing time domain steganography approaches insert the watermark into audio signal by adding the watermark to the signal. Hiding the watermark into time domain engage several challenges related to robustness and inaudibility.

Shaping the watermark before embedding enables the system to maintain the original audio signal audibility and renders the watermark inaudible. Concerning to the robustness, the approaches in the time domain steganography systems use different techniques to improve the robustness of the watermark (Alsalami and Al-Akaidi, 2003). As an example of audio steganography technique in this domain is Low-bit encoding.

Frequency domain audio steganography and digital watermarking

In the Frequency Domain, The input signal should transform to frequency domain in first stage, and then the watermark can embedded. To get the watermarked signal, the inverse frequency transform should be applied (Figure 4)

Transforming audio signal from time domain to frequency domain enables steganography system to embed the watermark into perceptually significant components. According to (Zhang et al., 2010a, b; Cox et al., 1997) this technique offers high level of robustness, due to that any attempt to remove the watermark will result in introducing a serious distortion in original audio signal fidelity.

Moreover, there are several different frequency domains, each defined by a different mathematical transformation, which are used to analyze signals. The most common transforms used and the fields in which they are used in digital audio steganography are: Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT). Examples of techniques for this domain are Phase coding, Spread spectrum, Echo data hiding.

Compressed domain audio steganography and digital watermarking

Compressed domain audio steganography has removed the perceptually irrelevant parts of the audio and makes the audio signal distortion inaudible to the human ear. MPEG audio compression is a lossy algorithm and uses the special nature of the HAS, these type of systems are suitable for “pay audio” scenario, where the provider stores audio contents in compressed format. During download of music, the customer identifies himself with his unique customer ID, which therefore is known to the provider during delivery. In order to embed the customer ID into the audio data using a steganography technique, a scheme is needed that is capable of steganography compressed audio on the fly during download(Alsalami and Al-Akaidi 2003; Ahmed et al., 2010) (Figure 5).

MPEG encoding process has the following steps:-

1. The audio samples pass through mapping filter to divide the audio data into subsamples of frequency;
2. Audio samples pass through MPEG psychoacoustics model at the same time. This process creates a masking threshold of audio signal. Masking threshold is used by quantization and coding step to determine how to allocate bits that minimize the quantization noise audibility;

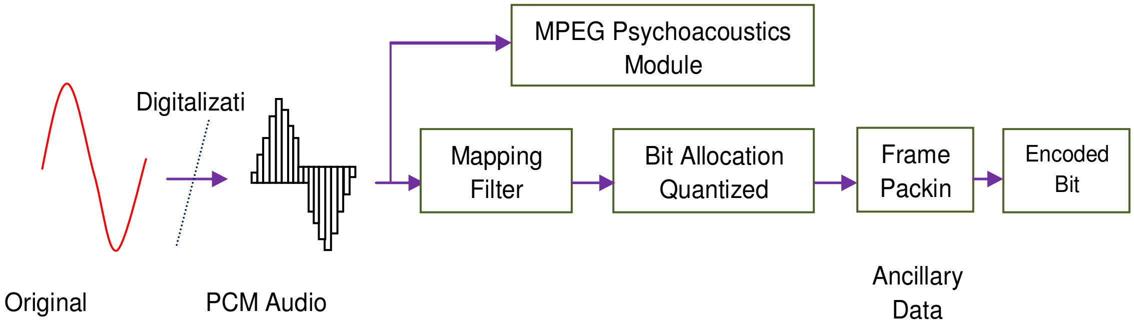
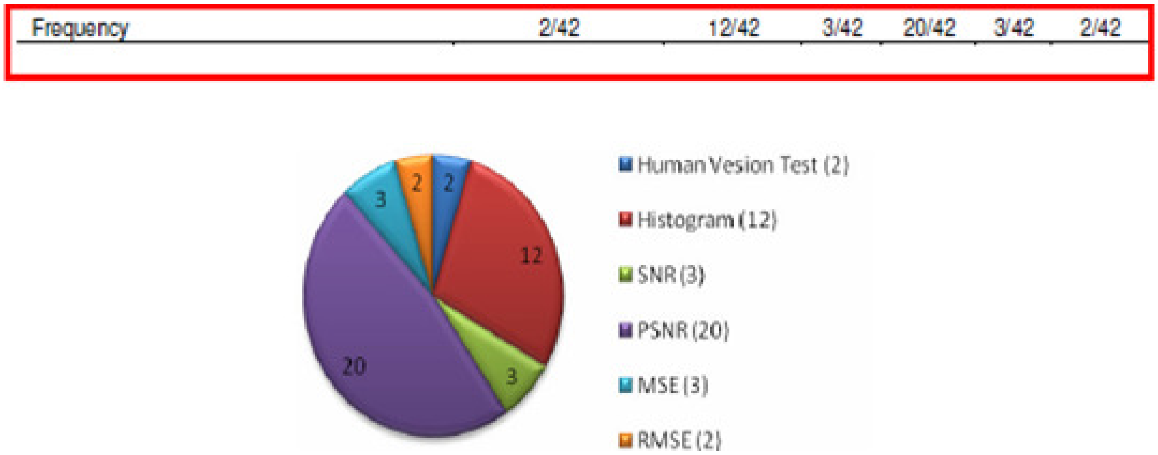


Figure 5. MPEG audio encoder structure.

Header	CRC	Bit Allocation	Scale Factors	Encoded Samples	Ancillary Data

Figure 6. Frame format of MPEG audio.



Frequency	2/42	12/42	3/42	20/42	3/42	2/42

Figure 7A. The frequency of the metrics for sample of 42 research articles (Hmood et al., 2010b).

3. In the final stage, the quantized subsamples are packed into frames (coded stream). (Figure 3) shows the basic structure of an MPEG audio encoder (Alsalami and Al-Akaidi 2003; Ahmed et al., 2010).

The filter divides the input audio signal into 32 equal-width of subsamples, subsequently, the number of bits used in quantization is determine upon masking threshold to minimize the audibility of possible distortion maybe introduced by quantization. (Figure 6) Frame is the smallest unit which can be decoded individually. Each frame contains audio data, header, CRC (Cyclic Redundancy Code), and ancillary data. In frame, each subsample has three groups of samples with 12 samples per group. The encoder can use a different scale factor for each group. Scale factor was determined upon masking threshold and used in reconstruction of audio signal. The decoder multiplies the quantizer output to reconstruct the quantized subsamples.

As in shown Figure 7B, MPEG audio decoding process is reverse of the encoding process. The decoding takes the encoded bit stream as an input, unpacks the frames, reconstructs the frequency samples (subsamples) using scale factors, and then inverses the mapping to re-create the audio signal samples(Ahmed et al., 2010).

Wavelet domain audio steganography

Wavelet transform can be used to decompose a signal into two parts, high frequencies and low frequencies (Shahad et al., 2011). The low frequencies part is decomposed again into two parts of high and low frequencies. The number of decompositions in this process is usually determined by application and length of original signal. The data obtained from the above decomposition are called the DWT (Discrete wavelet transform) coefficients. Moreover, the original signal can be reconstructed form these coefficients. This

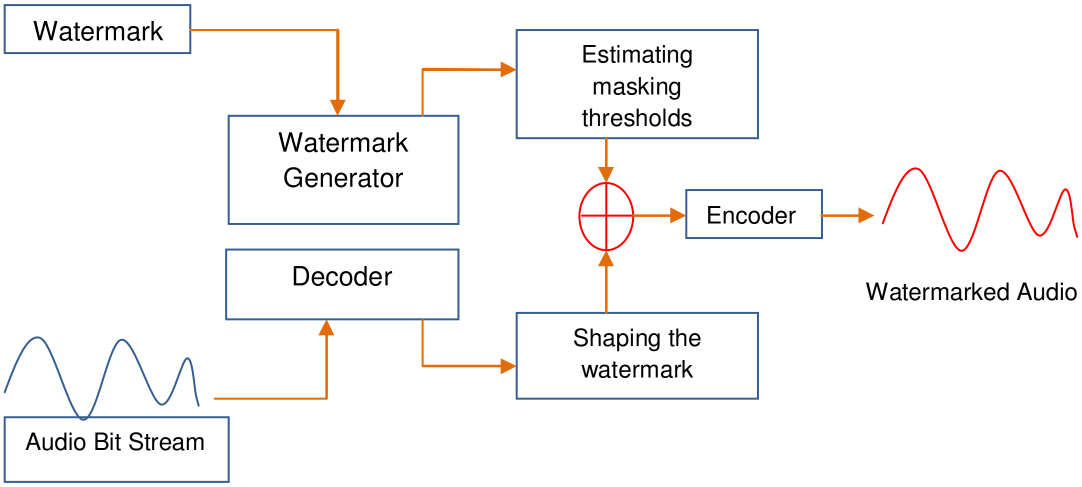


Figure 7B. MPEG audio decoding process.

reconstruction is called the inverse DWT (Ahmed et al., 2010). An example method of audio signal watermarking in wavelet domain uses patchwork algorithm (Kim and Choi, 2003). This method provides a fast synchronization between the watermark embedding and detection parts without original audio signals.

Summary

The main purpose of data hiding, are the secrecy of the hidden message, robustness of the approach and data hidden size. Several audio steganography and audio watermarking approaches have been developed in literature using different domains like time domain, frequency domain, and wavelet domain to achieve the above purposes. The process of selecting the domain depends on the purpose of developing the approach, for example, the target of the developer is to achieve high rate data hidden, and in this case they need to use time domain or compressed domain.

Properties of the H.A.S

The operation of hiding data in the audio signal is a particular challenge because the Human Auditory System (HAS) works dynamically in a wide range of frequencies, which falls between (20Hz - 20000Hz), therefore this system is very sensitive to add random noise, the perturbations in a sound file can be detected as low as on part in ten million (80dB below ambient level).

Embedding more and additional information into audio sequences is a more tedious task than that of images, due to the dynamic supremacy of the (HAS) over human visual system (Cvejic, 2004; Bender et al., 1996). In addition, the quantity of data that can be embedded in the video frames is higher than the quantity of data that can be embedded transparently into audio samples upon the fact that, audio signal has less dimension then video. On the other hand, many malicious attacks are against image and video watermarking algorithms (e.g., geometrical distortions and spatial scaling) cannot be implemented against audio watermarking schemes (Ahmed et al., 2010). However, there are some “holes” available which need to be addressed. While the (HAS) has a large dynamic range, it has a pretty small differential range. As a result, loud sounds tend to mask out quiet sounds. Additionally, the (HAS) is unable to perceive absolute phase, only relative phase. Finally, there are some environmental distortions so common as to be ignored by the listener in most case (Bender, 1996).

Two attributes of the (HAS) dominantly used in watermarking algorithms are: frequency (simultaneous) masking and temporal masking. The concept using the perceptual holes of the (HAS) is taken from wideband audio coding (e.g., MPEG Compression 1, Layer 3, usually called MP3) (Noll, 1993). In the compression algorithms, the holes are used in order to decrease the amount of the bits needed to encode audio signal, without causing a perceptual distortion to the coded audio. Along with that, the information hiding scenarios, masking properties are used to embed additional bits into an existing bit stream, again without generating perceptible noise in the audio sequence used for data hiding (Cvejic, 2004).

Audio environments

When developing a data hiding method for audio, one of the first considerations is the likely environments the sound signal will travel between encoding and decoding. There are two main areas of modification which we consider (Bender et al., 1996):

1. The storage environment, or the digital representation of the signal that will be used;
2. The transmission passageway the signal might travel.

Digital representation

There are two critical parameters to most digital audio representations:

1. Sample quantization method. The most popular format for representing samples of high-quality digital audio is (16-bit linear quantization), e.g., Windows Audio-Visual (WAV) and Audio Interchange File Format (AIFF). Another popular format for lower quality audio is the logarithmically scaled 8-bit μ-law. These quantization methods introduce some signal distortion, somewhat more evident in the case of 8-bit μ-law. Popular temporal *sampling rates* for audio include 8 (kilohertz), 9.6, 10, 12, 16, 22.05, and 44.1 kHz;
2. Temporal sampling rate. Sampling rate impacts data hiding in that it puts an upper bound on the usable portion of the frequency

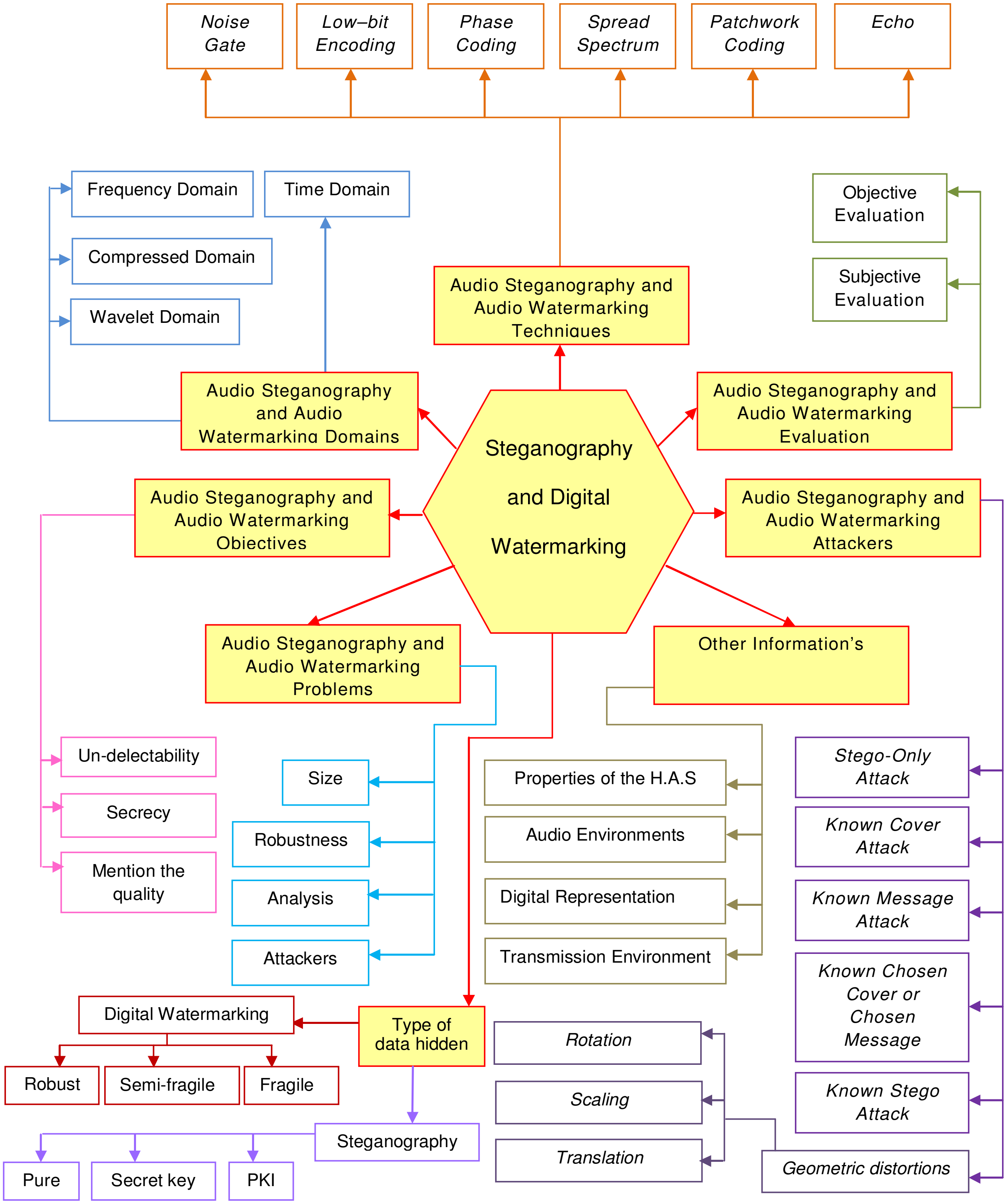


Figure 8A. Data hidden environment.

spectrum (if a signal is sampled at ~8 kHz, we cannot introduce modifications that have frequency components above ~4 kHz). For most data-hiding techniques we have developed, usable data space increases at least linearly with increased sampling rate.

A last representation to consider is that produced by lossy, compression algorithms, such as the International Standards Organization Motion Pictures Expert Group—Audio (ISO MPEG-AUDIO) perceptual encoding standard. These representations

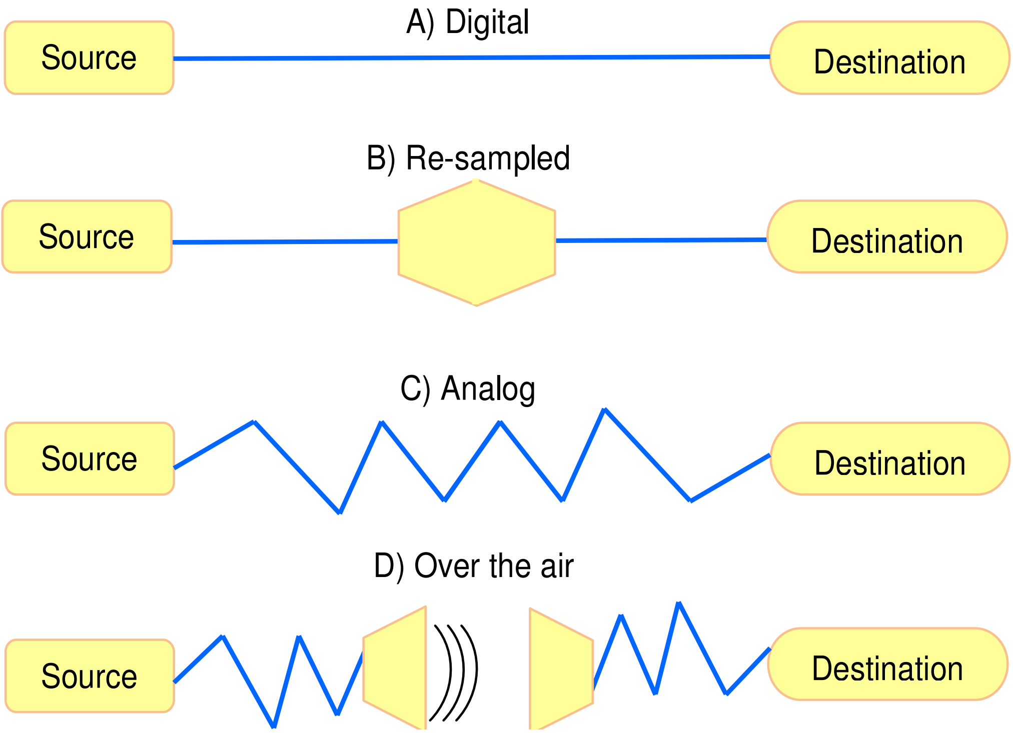


Figure 8B. Transmission environments (Bender et al., 1996).

drastically change the statistics of the signal; they preserve only the characteristics that a listener perceives (that is, it will sound similar to the original, even if the signal is completely different in a least squares sense) (Bender et al., 1996).

Transmission environment

There are many different transmission environments that a signal might experience on its way from encoder to decoder. We consider four general classes for illustrative purposes (Figure 8B) (Bender et al., 1996).

The first is the digital end-to-end environment as in Figure 8(B). This is the environment of a sound file that is copied from machine to machine, but never modified in any way. As a result, the sampling is exactly the same at the encoder and decoder. This class puts the least constraints on data-hiding methods. The next consideration is when a signal is resample to a higher or lowers sampling rate, but remains digital throughout *as* in Figure 8(B).

This transform preserves the absolute magnitude and phase of most of the signal, but changes the temporal characteristics of the signal. The third case is when a signal is “played” into an analog state, transmitted on a reasonably clean analog line and re-sampled as in Figure 8(B), absolute signal magnitude, sample quantization, and temporal sampling rate are not preserved. In general, phase will be preserved. The last case is when the signal is “played into the air” and “re-sampled with a microphone” as in Figure 8(B). The signal is subjected to possibly unknown nonlinear modifications resulting in phase changes, amplitude changes, drift of different frequency components, echoes, etc. Signal representtation and transmission pathway must be considered when choosing a data-hiding method. Data rate is very dependent on the sampling rate and the type of sound being encoded. A typical value is 16 bps, but the number can range from 2 bps to 128 bps.

RESULTS

In the field of software engineering, the process of measuring software quality or some of its specifications is called software metrics. Like other software, steganography and digital watermarking approaches are evaluated in the literature using subjective (that is, listing or viewing) and objective (that is, PSNR, PSNR, MSE and RMSE) tests. Away from the subjective and the objective tests, there is another test for data hiding called the histogram, on which, the researchers compare between the histograms before and after hiding the data. However, these metrics have been criticized in the literature.

Subjective evaluation

Subjective listening tests by human’s auditory perception, the subjects are asked to discriminate the differences between the watermarked signal and original audio clips. The watermarked signal is graded with respect to the host signal according to five-grade scale (Table 3) defined in ITU-R BS.562. According to Arnold (2000), the five-grade scale called Subjective Difference Grade (SDG), which is the difference between the subjective ratings given individually to the watermarked signal and the original signal

Subjective listening tests are indispensable and essential toward perceptual quality evaluation, due to the

Table 3. Subjective difference grade (SDG) (Arnold, 2002).

Description of impairments	Difference grade
Very annoying	1
Annoying	2
Slightly annoying	3
Perceptible but not annoying	4
Imperceptible	5

ultimate judgment that is made by human perception and the unreliability of the objective test. However, carrying out such listening tests is quite difficult and also not enough for manufacturing. Therefore, objective evaluations are also useful to provide a convenient, consistent and fair measurement (Lin and Abdulla, 2008).

Objective evaluation

The aim of objective evaluation tests is to facilitate the implementation of subjective listening tests. To achieve its goal, results of objective evaluation should relate well with SDGs. Recently, the commonly used objective evaluation is to assess the perceptual quality of audio data via a stimulant ear.

The objective metrics of the hidden information are mainly used to measure the distortion level in the steganographic object. The metric should demonstrate the possibility of any alteration to the perceptual layout of the audio, image or video. However, the used metrics in the literature have a number of limitations. The main full-reference objective tests for image, audio and video quality metrics that have been appears in the literature are:

1. Mean squared error (MSE);
2. Peak signal-to-noise ratio (PSNR);
3. Root means squared error RMSE;
4. Signal-to-noise ratio SNR.

Regardless, PSNR is widely used because it is simple to calculate, has clear physical meanings, and is mathematically easy to deal with for optimization purposes (Figure 7A). However, these metrics have been widely criticized as well for not correlating well with perceived quality measurement (Hmood et al., 2010; Alam et al., 2010)

DISCUSSION

This paper described the digital audio properties, audio steganography and watermarking domains, audio quality evaluation, audio steganography and digital watermarking techniques. This review might help the researchers to design, develop, and establish new methods, modules, algorithms, further analysis on steganography and digital watermarking. Moreover, several problems, approaches and techniques from the literature were discussed in this paper. The main challenges of steganography and digital watermarking are: the survival of the Watermark; the main challenge of data hidden is the survival against all types of attacks. Secondly, protection of the watermark: a multiple layers watermark when layers aim to protect each other from being analyzed, “The more robust and reliable the implementation is, the longer it will last”. According to (Zaidan et al., 2010i ; Yang et al., 2011; Al-Azawi and Fadhil, 2010; Hmood et al., 2010a,c; Shirali-Shahreza, and Shirali-Shahreza, 2008; Rabah, 2004; Al-Hamammi and Al-Hamadani, 2005; Al-Jaber and Aloqily, 2003; Luo et al., 2007; Li et al., 2011; Fiaidhi, and Mohammed, 2003;Hong et al., 2010; Wang et al., 2008; Wang et al., 2011;Liang et al., 2011; Liu et al., 2010; Phadikar et al., 2007; Prasannakumari, 2009; Khan et al., 2008; Shao et al., 2008; Hu and Niu, 2010; Xiao et al., 2009; Eltahir et al., 2009; Othman et al., 2009; Zaidan et al., 2010d) we have translated the data hiding environmental into the Figure 8.

Conclusion

The main goal of steganography and digital watermark is to be unsuspected by the human eyes or human ear. For instance audio watermarking or audio steganography is a great example for data protection and intellectual property. Therefore, design, development and implementation of new methods or techniques required a burly background on signal processing. In this paper, we summarize the domains on which digital audio steganography and watermarking are implemented. Moreover, properties of the Human Auditory System (HAS), the dynamic of HAS, audio environments, digital representation transmission environment, audio watermarking techniques, the available stego-analysis attackers and audio quality assessment are reviewed in this paper.

ACKNOWLEDGEMENTS

This research has been funded in part from University of

Malaya under No. UM.C/625/1. The authors would also like to acknowledge the Multimedia University (MMU) and Sunway University as the collaborator for this research.

REFERENCES

Abdulfetah AA, Sun X, Yang H (2010). “Robust adaptive video watermarking scheme using visual models in DWT domain”. Inf. Technol. J., 9(7): 1409-1414

Abdulfetah AA, Sun X, Yang H, Mohammad N (2010b). “Robust adaptive image watermarking using visual models in DWT and DCT domain”. Inf. Technol. J., 9(3): 460-466.

Abomhara M, Khalifa OO, Zakaria O, Zaidan AA, Zaidan BB, Alanazi HO (2010). "Suitability of Using Symmetric Key to Secure Multimedia Data: An Overview." J. Appl. Sci., 10(15): 1656-1661.

Abu Ali AN, Alnaimat AK, Abu-Addose HY (2010). “Evaluating the vulnerability and the security of public organizations websites in Jordan”, J. Appl. Sci., 10: 2447-2453.

Ahmed MA, Kiah MLM, Zaidan BB, Zaidan AA (2010). "A Novel Embedding Method to Increase Capacity and Robustness of Low-bit Encoding Audio Steganography Technique Using Noise Gate Software Logic Algorithm". J. Appl. Sci., 10(1): 59-64.

Alam GM, Kiah MLM, Zaidan BB, Zaidan AA, Alanazi HO (2010). “Using the features of mosaic image and AES cryptosystem to implement an extremely high rate and high secure data hidden: Analytical study”. Sci. Res. Essays, 5(21): 3254-3260

Alanazi HO, Jalab HA, Alam GM, Zaidan BB, Zaidan AA (2010a)." Securing electronic medical records transmissions over unsecured communications: An overview for better medical governance”. J. Med. Plants Res., 4(19): 2059-2074 .

Alanizi H, Kiah O, Zaidan MLM, Zaidan BB, Zaidan AA, Alam GM (2010b). “Secure topology for electronic medical record transmissions”. Int. J. Pharmacol., 6 (6): 954-958.

Al-Azawi AF, Fadhil MA (2010). Arabic text steganography using kashida extensions with huffman code. J. Appl. Sci., 10(5): 436-439.

Al-Frajat AK, Jalab HA, Kasirun ZM, Zaidan BB, Zaidan AA (2010). "Hiding Data in Video File: An Overview ". J. Appl. Sci., 10(15): 1644-1649.

Al-Hamammi A, Al-Hamadani MH (2005). Proving poverty of steganography system. Inf. Technol. J., 4(3): 284-288.

Al-Jaber A, Aloqily I (2003). High quality steganography model with attacks detection. Inf. Technol. J., 2(2): 116-127

Alsalami MAT, Al-Akaidi MM (2003). “Digital Audio Watermarking: Survey”, 17th European Simulation Multi-conference, UK.

Anderson RJ, Petitcolas FAP (1998). “On The Limits of Steganography”. IEEE J. Selected Areas Commun., 16(4): 474-481.

Arnold M (2000). “Audio watermarking: Features, applications and algorithms”, In Proceedings IEEE Int. Conf. Multimedia Expo., 2: 1013-1016

Artz D (2001). “Digital steganography: hiding data within data” Internet Computing. IEEE, 5(3): 75-80.

Bender W, Gruhl D, Morimoto N, Lu A (1996). “Techniques for data hiding”. IBM Syst. J., 35: 313-336.

Chandra S, Khan RA (2008). “Object oriented software security estimation life cycle-design phase perspective”. J. Softw. Eng., 2: 39-46.

Cox IJ, Kilian J, Leighton FT, Shamoon T (1997). “Secure spread spectrum watermarking for multimedia”, IEEE Trans. Image Process., 6(12): 1673 -1687

Cox J, Miller ML, Bloom JA (2002). Digital watermarking. Academic Press.

Cvejic N (2004). “Algorithms for audio watermarking and steganography”, Department of Electrical and Information Engineering, Finland, University of Oulu.

Eltahir ME, Kiah LM, Zaidan BB, Zaidan AA (2009). "High Rate Video Streaming Steganography", Int. Conf. Info. Manage. Eng., icime2009, pp. 550-553.

Fiaidhi JAW, Mohammed SMA (2003). Towards developing watermarking standards for collaborative e-learning systems. Inf. Technol. J., 2(1): 30-34.

Hmood AK, Jalab HA, Kasirun ZM, Zaidan BB, Zaidan AA (2010a). "On the Capacity and Security of Steganography Approaches: An Overview ", J. Appl. Sci., 10(16): 1825-1833.

Hmood AK, Jalab HA, Kasirun ZM, Zaidan BB, Zaidan AA (2010b). “On the accuracy of hiding information metrics: Counterfeit protection for education and important certificates ”, Int. J. Phys. Sci., 5(7): 1054-1062.

Hmood AK, Zaidan BB, Zaidan AA, Jalab HA (2010c). "An Overview on Hiding Information Technique in Images " J. Appl. Sci., 10(18): 2094-2100.

Hong W, Chen TS, Lin KY, Chiang WC (2010). A modified histogram shifting based reversible data hiding scheme for high quality images. Inform. Technol. J., 9(1): 179-183.

Hu YY, Niu XM (2010). Image hashing algorithm based on robust bits extraction in JPEG compression domain. Inf. Technol. J., 9(1): 152-157

Jayakumar J, Thanushkodi K (2008). “Application of exponential evolutionary programming to security constrained economic dispatch with FACTS devices”, Asian J. Sci. Res., 1(4): 374-384.

Johnson NF, Duric Z, Jajodia S, Memon N (2001). “Information Hiding: Steganography and Watermarking Attacks and Countermeasures". J. Electron. Imaging, 10(3): 825

Khan A, Niu X, Yong Z (2008). A robust framework for protecting computation results of mobile agents. Inform. Technol. J., 7(1): 24-31

Kim HJ, Choi YH (2003). “A novel echo-hiding scheme with backward and forward kernels”, IEEE Trans. Circuits Systems Video Technol., 13: 885-889.

Ko BS, Nishimura R, Suzuki Y (2005). Time-spread echo method for digital audio watermarking. IEEE Trans. Multimed., 7(2): 212-221

Li J, Wang RD, Zhu J (2011). “A Watermark for Authenticating the Integrity of Audio Aggregation Based on Vector Sharing Scheme”. Inf. Technol. J., 10(5): 1001-1008

Liang W, Sun X, Ruan Z, Long J (2011). The design and FPGA implementation of FSM-based intellectual property watermark algorithm at behavioral level. Inf. Technol. J., 10(4): 870-876.

Lin Y, Abdulla WH (2008). “Perceptual evaluation of audio watermarking using objective quality measures”. IEEE Int. Conf. Acoust. Speech Signal Process., pp. 1745-1748

Liu Z, Sun X, Liu Y, Yang L, Fu Z, Xia Z, Liang W (2010). Invertible transform-based reversible text watermarking. Inf. Technol. J., 9(6): 1190-1195

Luo, G, Sun X, Xiang L (2008). Multi-blogs steganographic algorithm based on directed hamiltonian path selection. Inf. Technol. J., 7(3): 450-457.

Medani A, Gani A, Zakaria O, Zaidan BB, Zaidan AA (2011). “ Review of mobile short message service security issues and techniques towards the solution“. Sci. Res. Essays, 6(6): 1147-1165.

Nabi MSA, Kiah MLM, Zaidan BB, Zaidan AA, Alam GM (2010).” Suitability of Using SOAP Protocol to Secure Electronic Medical Record Databases Transmission“. Int. J. Pharmacol., 6(6): 959-964.

Naji AW, Zaidan AA, Zaidan BB (2009). “Challenges of Hidden Data in the Unused Area Two within Executable Files”. J. Comput. Sci., 5(11): 890-897.

Noll P (1993) “Wideband speech and audio coding”. IEEE Commun. Mag., 31: 34-44

Noto M (2001). “MP3Stego: Hiding Text in MP3 Files”, SANS Institute.

Oh HO, Seok JW, Hong JW, Youn DH (2001). “New echo embedding technique for robust and imperceptible audio watermarking”, In Proceeding of IEEE Int. Conf. Acoust. Speech Signal Process., 3: 1341 – 1344

Othman F, Maktom L, Taqa AY, Zaidan BB, Zaidan AA (2009). “An Extensive Empirical Study for the Impact of Increasing Data Hidden on the Images Texture”, International Conference on Future Computer and Communication, ICFCC2009, pp. 477 – 481.

Phadikar A, Verma B, Jain S (2007). Region splitting approach to robust color image watermarking scheme in wavelet domain. Asian J. Inf. Manage., 1(2): 27-42

Ping Z, Xi C, Xu-Guang Y (2010). “The software watermarking for tamper resistant radix dynamic graph coding”. Inf. Technol. J., 9(6): 1236-1240.

Prasannakumari V (2009). A robust tamperproof watermarking for data integrity in relational databases. Res. J. Inf. Technol., 1(3): 115-121.

Raad M, Yeasin NM, Alam GM, Zaidan BB, Zaidan AA (2010). “Impact of spam advertisement through email: A study to assess the influence of the anti-spam on the email marketing”. Afr. J. Bus. Manage., 4(11): 2362-2367.

Rabah K (2004). Steganography-the art of hiding data. Inf. Technol. J., 3(3): 245-269.

Rabah KVO (2005). Implementation of one-time pad cryptography. Inform. Technol. J., 4(1): 87-95.

Sameer HAl, Mat Kiah ML, Zaidan AA, Zaidan BB, Alam GM (2011).“ Securing peer-to-peer mobile communications using public key cryptography: New security strategy“. Int. J. Phys. Sci., 6(4): 930-938.

Shahad N, Mohd.Ali MA, Zaidan AA, Zaidan BB, Najah H (2011). “Computerized Algorithm for Fetal Heart Rate Baseline and Baseline Variability Estimation based on Distance Between Signal Average and α Value“. Int. J. Pharmacol. (IJP), 7(2): 228-237.

Shao LP, Qin Z, Gao HJ, Heng XC (2008). 2D triangular mappings and their applications in scrambling rectangle image. Inf. Technol. J., 7(1): 40-47

Sheikhan, M, Asadollahi K (2010). “High Quality Audio Steganography by Floating Substitution of Lsbs in Wavelet Domain”. World Appl. Sci. J., 10(12): 1501-1507

Shirali-Shahreza M, Shirali-Shahreza S (2008). High capacity persian/arabic text steganography. J. Appl. Sci., 8(22): 4173-4179.

Wang B, Sun X, Ruan Z, Ren H (2011). “Multi-mark: multiple watermarking method for privacy data protection in wireless sensor networks”. Inf. Technol. J., 10(4): 833-840

Wang B, Sun X, Ruan Z, Ren H (2011). Multi-mark: multiple watermarking method for privacy data protection in wireless sensor networks. Inf. Technol. J., 10(4): 833-840.

Wang X, Sun X, Liu Y, Liu Y (2008). Natural language watermarking using chinese syntactic transformations. Inf. Technol. J., 7(6): 904-910

Wang X, Yang L, Sun X, Han J, Liang W, Huang L (2010). Survey of anonymity and authentication in P2P networks. Inf. Technol. J., 9(6): 1165-1171.

Wu CH, Zheng Y, Ip WH, Lu ZM, Chan CY, Yung KL (2011). “Effective hill climbing algorithm for optimality of robust watermarking in digital images”. Inf. Technol. J., 10(2): 246-256

Xiao X, Sun X, Wang X, Rao L (2009). DOSM: A data-oriented security model based on information hiding in WSNs. Inf. Technol. J., 8(5): 678-687.

Xu C, Wu J, Sun Q, Xin K (1999). “Applications of digital watermarking technology in audio signals”. J. Audio Eng. Soc., 47(10): 805-812

Yang B, Sun X, Xiang L, Ruan Z, Wu R (2011). “Steganography in Ms Excel Document using Text-rotation Technique”. Inf. Technol. J., 10: 889-893.

Yang H, Sun X, Sun G (2009). “A semi-fragile watermarking algorithm using adaptive least significant bit substitution”. Inf. Technol. J., 9(1): 20-26

Yeo I, Kim HJ (2003). “Modified Patchwork Algorithm: A Novel Audio Watermarking Scheme”, IEEE Trans. Speech Audio Process., 11(4): 381-386.

Zaidan AA, Ahmed, Karim NN, Abdul H, Alam GM, Zaidan BB (2011a). “Spam Influence on the Business and Economy: Theoretical and Experimental Study for Textual Anti-spam Filtering Using Mature Document Processing and Naïve Bayesian Classifier”, African. Afr. J. Bus. Manage., 5(2): 596-607.

Zaidan AA, Zaidan BB, Alanazi HO, Gani A, Zakaria O, Alam GM (2010c). “Novel approach for high (secure and rate) data hidden within triplex space for executable file”. Sci. Res. Essays, 5(15): 1965-1977.

Zaidan AA, Zaidan BB, Al-Fraja AK, Jalab HA (2010a). Investigate the Capability of Applying Hidden Data in Text File: An Overview." J. Appl. Sci., 10(17): 1916-1922.

Zaidan AA, Zaidan BB, Al-Frajat AK, Jalab HA (2010b). An overview: Theoretical and mathematical perspectives for advance encryption standard/rijndael. J. Appl. Sci., 10(18): 2161-2167.

Zaidan AA, Zaidan BB, Taqa AY, Mustafa KMS, Alam GM, Jalab HA (2010d). “Novel Multi-Cover Steganography Using Remote Sensing Image and General Recursion Neural Cryptosystem“. Int. J. Phys. Sci., 5(21): 3254-3260.

Zaidan BB, Zaidan AA, Al-Frajat AK, Jalab HA (2010e). "On the Differences between Hiding Information and Cryptography Techniques: An Overview J. Appl. Sci., 10(15): 1650-1655.

Zaidan BB, Zaidan AA, Mat Kiah ML (2011b). “Impact of Data Privacy and Confidentiality on Developing Telemedicine Applications: Review, Participates Opinion and Expert Concerns“. Int. J. Pharmacol., 7(3): 382-387.

Zaidan BB, Zaidan AA, Taqa A, Alam GM, Kiah MLM, Jalab HA (2010f). “StegoMos: A Secure Novel Approach of High Rate Data Hidden Using Mosaic Image and ANN-BMP Cryptosystem”. Int. J. Phys. Sci., 5(11): 1796-1806.

Zeki AM, Manaf AA (2011) “ISB watermarking embedding: A block based model”. Inf. Technol. J., 10(4): 841-848.

Zeng W, Wu Y (2010). “A visible watermarking scheme in spatial domain using HVS model”. Inf. Technol. J., 9(8): 1622-1628.

Zhang Y, Lu ZM, Zhao DN (2010). “A blind image watermarking scheme using fast hadamard transform”. Inform. Technol. J., 9: 1369-1375.

Zhang Y, Lu ZM, Zhao DN (2010b). “Quantization based semi-fragile watermarking scheme for H.264 video”. Inf. Technol. J., 9(7): 1476-1482.