Robust Audio Steganography using Direct-Sequence Spread Spectrum Technology

(Fall 2007)

Wei Qin Cheng, Fei Han, Man Juon Tung, Kai Xu

Abstract

Watermarking has been a copyright protection technology for a while. Spread Spectrum watermarking is widely used in this area for audio copyright protection. Compared to other existing technologies, the robust character of Spread Spectrum is the reason why it was selected. Other than copyright watermarking, audio Spread Spectrum technology is also good for providing robust cover channel service. But, unfortunately, less implementation is based on audio cover channel. The increasing internet security problems provide the potential opportunity for this technology since online audio sharing is very popular right now. Users can use the audio cover channel to send security information without being detected and attacked by hackers easily, significantly reducing vulnerability to snooping and other attacks. Traditional LSB method is too simple to be invulnerable from detection. Here, we implemented a simple Dynamic Linked Library [DLL] by using managed C++ and Microsoft .NET framework. It is implemented by Direct Sequence Spread Spectrum [DSSS] method on data block base. This component provides 3rd party developers a way to implement some security mechanism for improving online or cross-platform security. A simple tool, STAGAMAN, was also implemented to demo the component.

Introduction

With the popularity of internet services today, thousands of millions people are using the internet daily for communication and playing. Harvested from the super-fast new chips and other new computer-related technologies, the bandwidth of the internet is growing wider and wider, and the price is becoming cheaper and cheaper. Consequently, the security risks are highly raised. Based on statistics, "Total costs averaged $182 per lost customer record, an increased of 30 percent over 2005 results. The average total cost per reporting company was $4.8 million per breach and ranged from $226,000 to $22 million. Since early 2005, more than 150 million personal records have been exposed". [1] Somehow, we need to rely on the internet to transfer confidential information. But, as we know, there are no existing technologies that can guarantee we are not attacked by information thieves. Many existing encryption algorithms are proved to be fragile indeed. Even with protection by modern encryption technologies, some people still can break the ciphers by the way of brute force attack with dictionary or rainbow tables. We might want to transfer information discreetly for avoiding those attacks. Cover channel can solve the problem. The benefit of the cover channel is not only simply encrypting information into a carrier media file but also confusing attackers while the user is transferring sensitive information. It will make the encrypted information not directly exposed to attackers. Because online media is so popular, this plan is achievable. Here, we are trying to improve security by using some wave media as a cover channel. For making the cover channel robust, we introduced the Spread Spectrum digital water marking steganography. We also tried to use managed C++ to implement the function instead of traditional C in Microsoft .NET 2.0 framework. That wrapped DLL file will make the component easy to be integrated with other internet technologies.

Background

A. Steganography History

The first use of steganography dates back to the ancient Greeks. Herodotus tells how a message was passed to Greeks about Xerses' hostile intentions underneath the wax of a writing tablet, and describes a technique of dotting successive letters in a cover text with secret ink due to Aeneas the Tactician. In ancient China, people were using a technology: embedding a code ideogram at a prearranged position in a dispatch. During WWII the grille method or some variants were used by spies. In the same period, the Germans developed microdot technology to print a clear, good quality photograph shrunk to the size of a dot. In the current industry market, with the advent of digital communication and publishing, one important issue is copyright enforcement, which is commonly implemented with watermarking.

B. Spread Spectrum Introduction

Spread Spectrum techniques are widely used in data communication, such as the CDMA mobile communication. The first patent was published by Hedy Lamar and George Antheil in 1941 for providing secret communication for military purposes. Spread Spectrum techniques are methods by which energy generated in particular bandwidths is deliberately spread in the frequency domain, resulting in a signal with a wider bandwidth. [2] A couple of technologies exist in this branch.

[i] DSSS stands for Direct Sequence Spread Spectrum. Data to be transmitted is divided into small pieces and each piece is allocated to a frequency channel across the spectrum. Transmitter utilizes a phase varying modulation technique to modulate each piece of data with a higher data rate bit sequence. [Figure 1]

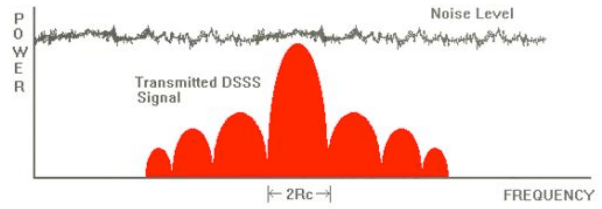

Figure 1

[ii] FHSS stands for Frequency-Hopping Spread Spectrum. It is a method of transmitting signals by rapidly switching a carrier among many frequency channels, using a pseudorandom sequence known to both the transmitter and receiver. [Figure 2]

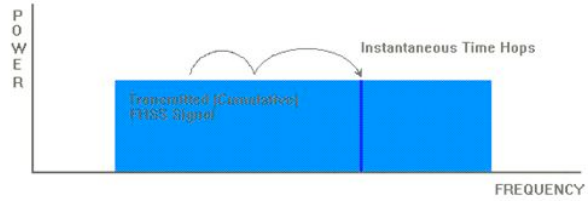

Figure 2

[iii] THSS stands for time-hopping Spread Spectrum. Short information bursts [chirps] are transmitted with pseudorandom pulse durations, or transmitted in random positions. Generally, there are two methods to implement this. Method 1: The chirp interval is determined by the PN code generator [Figure 3.1]. Method 2: The chirp goes at the same time in each bit period, the PN generator changes its duration [Figure 3.2]. Here, a chirp is a signal in which the frequency increases or decrease with time. *[iv]* General Spread Spectrum application by using chirps is also called Chirp Spread Spectrum [CSS]. *[v]* Quite often, the researcher combines these technologies together to produce some new patterned Spread Spectrum technology.

C. Wave File Format Introduction

Audio file format is a container format for storing audio data and metadata on a computer system. There are three major groups of audio file format nowadays: *[i]* Uncompressed audio formats, [e.g., WAV] *[ii]* Formats with lossless compression, [e.g., WMA] and *[iii]* Formats with lossy compression. [e.g., MP3]

Our design is based on uncompressed WAV audio format. WAV file format is a subset of Microsoft's RIFF specification for the storage of multimedia files. A RIFF file begins with a file header followed by a sequence of data chunks. In fact, a WAV file is a RIFF file with a signal "WAV" chunk. The signal "WAV" chunk is subdivided into two sub-chunks: [a] "fmt" chunk specifies the data format with some audio metadata information, such as number of channels, sample frequency rate, byte rate, and block alignment. [Figure 4] [b] "data" chunk specifies the data sample.

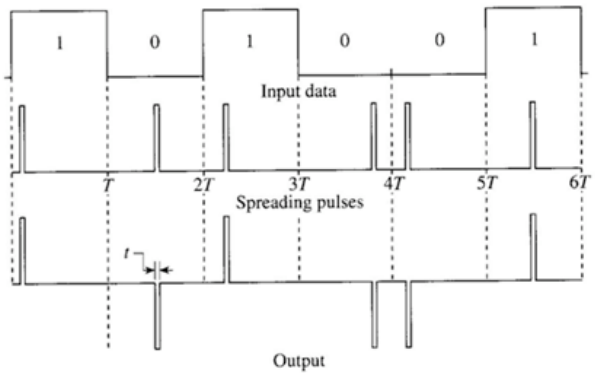

Figure 3.1

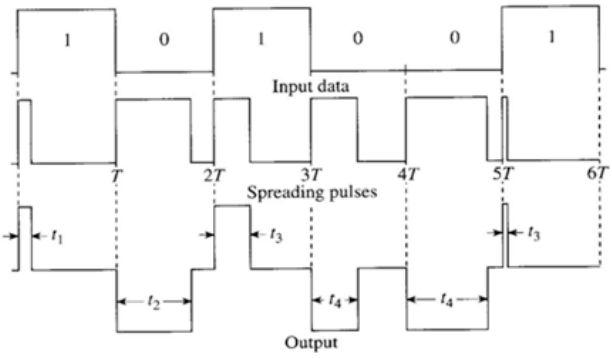

Figure 3.2

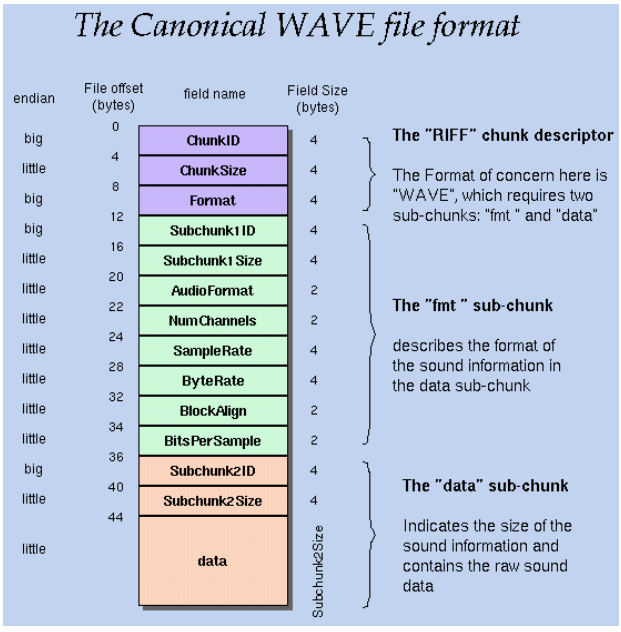

Figure 4

D. Audio Steganography Technology

Audio steganography is the technology of embedding information in an audio channel. It is used for digital copyright protection. Watermarking is the technique which hides one piece of information [message] in another piece of information [carrier]. It is widely used for applications such as 2-D image, audio clip, video clip, and even 3-D mash. More research is based on 2-D image watermarking. Most research focuses on digital watermarking for copyright protection. There are five audio steganography categories:

1. LSB Coding: the simplest way to embed information in digital audio files. It is coding the least significant bit to carry the information. The method is very simple. However, the technique is less used in the real industry because it is not robust. [Figure 5]

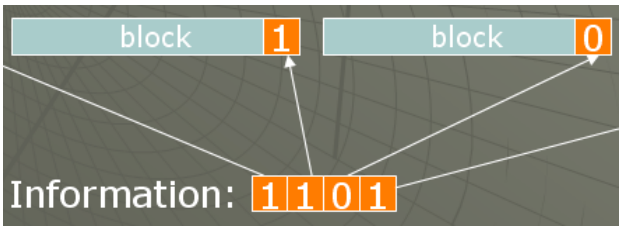

Figure 5

2. Parity Coding: The method breaks a signal down into separate regions of samples and encodes each bit for the secrete message in a sample region's parity bit. However, attackers can still find the pattern to break or remove the information from the carrier media. [Figure 6]

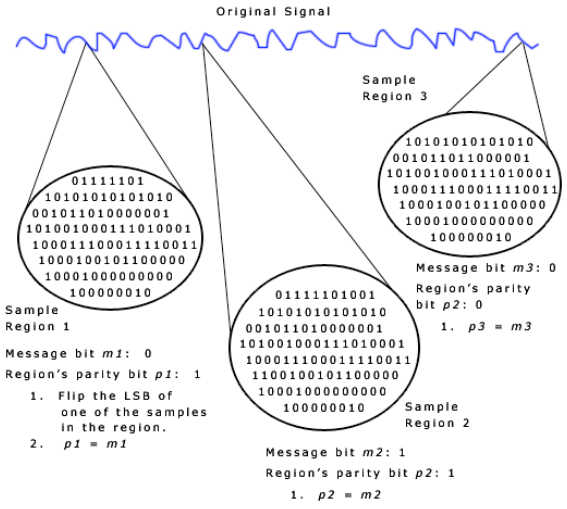

Figure 6

3. Phase Coding: It is used for addressing the disadvantage of the noise-inducing methods of audio steganography. Phase coding relies on the fact that the phase components of sound are not as perceptible to the human ear as noise is. The disadvantage is its low data transmission rate. [Figure 7]

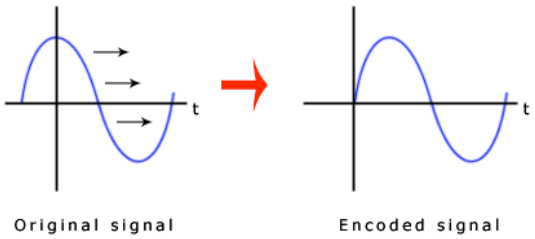

Figure 7

4. Echo Hiding: Information is embedded in a sound file by introducing an echo into the discrete signal. It allows for a high data transmission rate and provides superior robustness when compared to the noise inducing methods. [Figure 8]

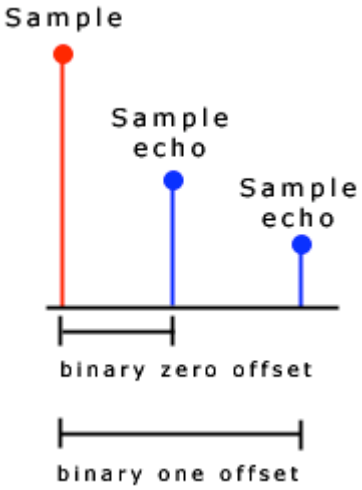

Figure 8

5. Spread Spectrum: It is the most robust watermarking technique so far. Spread Spectrum attempts to spread secret information across the audio signal's frequency spectrum as much as possible. The weakness of this technology is cost. Moreover, it is also extremely hard to implement since some expensive algorithms are involved, such as Fourier Transform. However, it is the most robust method.

Techniques and Implementation

No free implementation of audio cover channel can be found on the internet. Many research papers and interests are focused on digital watermarking for copyright protection. More resources are found for image spread spectrum than for audio cover channel. The possible reasons are:

a) Major marketing demand comes from copyright concern.

b) Compared to 2-D images, the implementation of Spread Spectrum cover channel in audio music is much harder, because of the characteristics of human voice.

c) The precision of message decoding is also a problem.

d) Cost is expensive due to the lower embedding information bandwidth.

e) Implementation is hard and complex.

f) The decoding process is time consuming.

g) Much code is based on C and MatLab and is not well-organized.

We are not attempting to solve all these problems in our implementation because of the limited time and resources. We implemented a Microsoft .NET 2.0 based DLL component for audio .wav files by referring to "Spread-Spectrum Watermarking of Audio Signals" [3] and the project by Darko Kirovski and Henrique S.Malvar in the Microsoft Research Lab. The DLL implementation is based on the two above projects. We rewrote the code using managed C++ in Microsoft .NET 2.0 Framework and Visual Studio 2005 and converted the function to a cover channel for embedding encrypted information.

A. General Watermarking Model

The following diagram [Figure 9] shows the general concept of a blind detector. Blind detector simply means the detector can recognize the embedded message without knowing the original cover information. The detector does not depend on knowing the original cover message. It only depends on the watermark key for encoding. The blind watermarking model fits our design since it is impossible to keep the original .wav music file confidential.

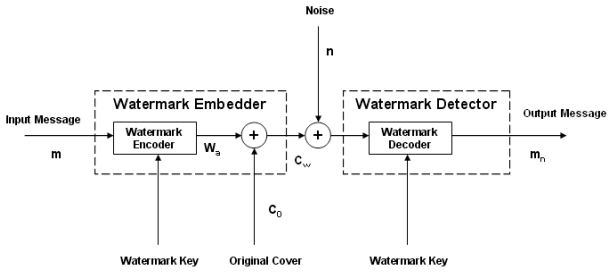

Figure 9

B. Selected Spread Spectrum Technique

We will briefly describe the major components in the design of DSSS.

Audio watermarking schemes rely on the imperfections of the human auditory system [HAS] [4]. The data hiding techniques explore the fact that the HAS is insensitive to small amplitude changes either in time or frequency domain [3]. The advantage of SS is that watermark detection does not require the original recording, and it is difficult to extract the hidden data using optimal statistical analysis under certain conditions.[11]

The DSSS method can be expressed using the following diagram. [Figure 10]

A modulated complex lapped transform [MCLT] is applied to transfer the audio information into the frequency domain. Information is embedded into the signal at frequency domain by applying a certain algorithm. Please refer to [3]. Finally, the inverted MCLT [iMCLT] is applied to transfer the signal back to the time-domain samples.

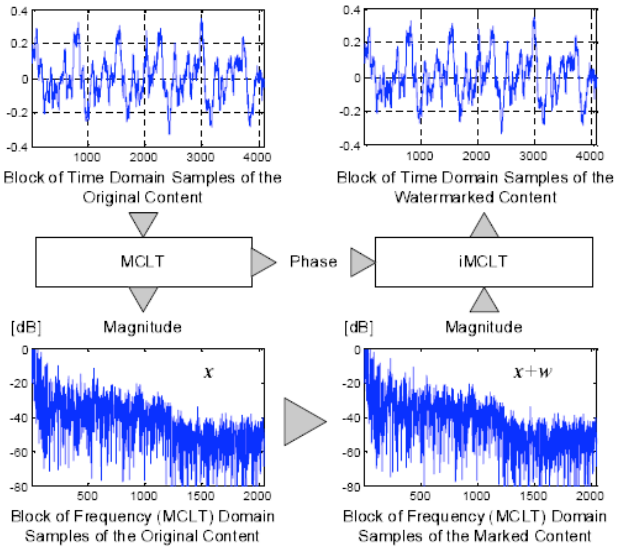

Figure 10

The following diagram [Figure 11] shows how the information is encoded and decoded. A pseudo-random R is applied here to generate the watermark chirps for Block i.

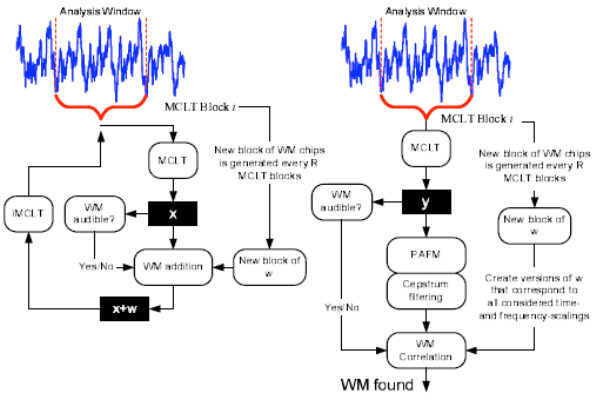

Figure 11

C. Implementation

audioSSWM.DLL:
We used managed C++ to implement the kernel DLL component audioSSWM.DLL under Microsoft .NET Framework. Managed C++ is selected to meet security concerns. It makes the component less easily attacked by malicious code. Supporting .NET platform benefits the 3rd party developer to develop other products using this component.

STAGAMAN.EXE:
It is a demonstration program written in C#. It proves the component can be used by a cross-platform application design. The functions of STAGAMAN are quite simple. Users can simply encode and decode information into the 44100Hz x 2Channel .wav file.
Here are some screen shots:

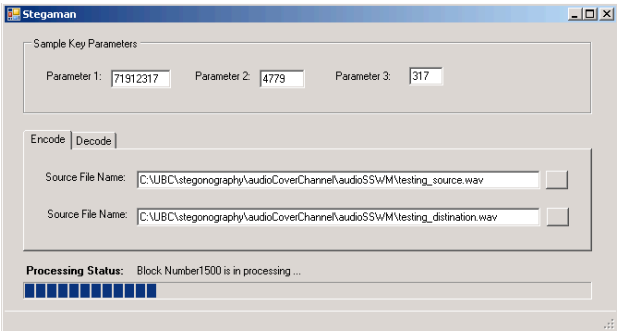

Figure 12.a

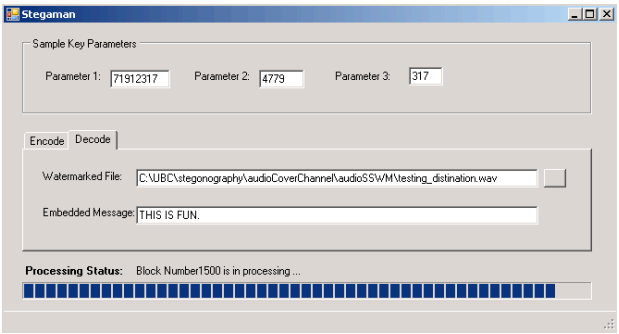

Figure 12.b

Security Analysis

The audioSSWM.DLL component is implemented by DSSS. A pseudo-random function is applied to generate the watermark key. As a secondary encryption method, the cipher space is reasonably large. Each possible attack will cost more time due to the applied method. Three parameters are used to control the sequence of the random function. Even one digit difference will make the whole information undetectable. Moreover, the window size is alterable in the component and it will also affect the output of the function. That provides more security into the design.

That does not mean the watermarked audio file can be invulnerable to all attacks. Smart users can estimate the watermark chirps from the marked signal. Since the attacker does not know if there is information embedded in the .wav file, the possibility of being attacked is significantly reduced.

Although we did not implement the modern RSA, AES or D-H algorithm in message coding, we still encourage that be done before putting information into the channel. The main purpose is to increase security by the 'Design in-depth' security principle.

Future Work

The audioSSWM.DLL is only the starting point of the audio watermarking component. There are several improvements that are achievable in the next stage. Some ideas are based on the information embedding and detecting process:

a) There is some latency in decoding processing. It is quite possible to improve it. The most interesting way is using GPU [Graphic Process Unit] to optimize the Fast Fourier Transform calculation. GPU is optimized for processing matrices, which is quite suitable to complex DSP cases.

b) Our project is only designed for 44100Hz x 2 Channel .wav sound files. The bandwidth of information embedded is not high and minimum window size is setup at 11.1483s. However, potentially the size can be downgraded to half. As a result, some digital transmission could be not precise. For compensating this problem, we can introduce Cyclic Redundancy Check methods to auto-fix the error bits.

c) AudioSSWM.DLL is not going to replace the existing information encryption method. It is just a novel way to protect the message from attack. We can continue working on the module to support information encryption by modern AES, RSA, or D-H algorithms.

d) Introducing more pseudo-random functions into the module to increase the watermark key space.

Conclusion

Audio watermarking technology and relative products are only commonly used for copyright protection. Not much cover channel practice based on audio signal watermarking can be found cross the internet. Computer security problems become more and more important in our daily life. That significantly increases the demand to design more secure systems. No single security approach, as we know, can protect information totally or make the system invulnerable.

Design in-depth security principle is important. That requires us to widely evaluate and select security components during the design. Direct Sequence Spread Spectrum is a robust watermarking technology. We implemented a Microsoft Framework .NET 2.0 based Audio DSSS watermarking dynamic link library [audioSSWM.DLL] to harvest the 3rd party developer for producing more secure products across the internet. We also developed a simple demonstration Embedder and Detector. STAGAMAN is implemented with C# and Microsoft .NET 2.0 Framework. It proves that the model can be functional across the network and different platforms.

The experiment proved that the audioSSWM.DLL meets the original design request. It can embed information into a 44100Hz x 2 Channel .wav music audio without losing the quality of original sounds. The information can also be well-detected using a selected watermarking key. The key space is reasonably large and the information is very sensitive to different watermarking keys. It makes the idea of cover channel achievable.

Using audioSSWM.DLL or STAGAMAN alone for transmitting security information is not our purpose. We also didn't encourage the potential user to simply rely on that. The main purpose of our audioSSWM.DLL project is providing an alternative to improve security, especially when the user wants to reduce the possibility to be attacked on certain sensitive information. Moreover, the managed C++ coding is better than the traditional DSP C coding. The DLL itself is less vulnerable to attack. Last but not least, the component based on .NET framework makes it suitable to all sorts of other web and cross-platform application developers.

References

[1] "Computer Theft & Recovery Statistics" [Online] Available: http://www.absolute.com/resources/computer-theft-statistics-details.asp 2007 Nov 13

[2] "Digital watermarking", Wikipedia [Online] Available: http://en.wikipedia.org/wiki/Digital_watermarking

[3] D. Kirovski and H. S. Malvar, "Spread-Spectrum Watermarking of Audio Signals", IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, v 3, 2001

[4] S. Katzenbeisser and F. A. P. Petit colas, Eds., "Hiding Techniques for steganography and Digital Watermarking" Boston, MAL Artech House, 2000

[5] Ingmar J. Cox, Matthew L. Miller, Jeffrey A. Bloom, "Digital Watermarking", Morgan Kaufmann Publishers, SanFrancisco, CA, 2002

[6] A. Bahramshahry, H. Ghasemi, A. Mitra, V. Morada, "Design of a Data Hiding Application Using Steganography [April 2007]," EECE 412 Term Project Reports From Previous Years, April 2007. [Online]. Available: http://www.vista.ubc.ca.

[7] A. Smith, "Spread Spectrum Watermark Estimation Through Autocorrelation, " M.S. Thesis, Dept. Comp. Sci., Univ. of Wisc. , Madison, WI, 2002. [Online]. Available: http://pages.cs.wisc.edu/~aasmith/mypapers/smith_MSthesis.pdf

[8] R. Radhakrishnan, K.Shanmugasundaram, "Data Masking: A Secure-Covert Channel Paradigm", Polytechnic University, Brooklyn, NY 11201,USA [Online] Available:http://ieeexplore.ieee.org/iel5/8561/27103/01203315.pdf

[10] N. F Johnson, S. Katzenbeisser, "A Survey of Steganographic Techniques", in S. Katzenbeisser and E Petitcoals [Eds.]: Information Hiding. Artech House, Norwood, MA, 2000.

[11] J.K.Su and B. Girod "Power-spectrum condition for energy-efficient watermarking", in Proc. Int. Conf. on Image Proc., Kobe, Japan, Oct. 1999, pp. 301-305