Exploration of Simple Audio Cryptography Schemes in Time Domain and Frequency Domain

Rizky Maulana Nugraha - 13508083[1]
Informatics Engineering
Schools of Electrical Engineering and Informatics
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
[1]r_maulana_n@students.itb.ac.id; if18083@students.if.itb.ac.id; lana_pcfre@yahoo.com; lana.pcfre@gmail.com

***Abstract*—Visual cryptography is a cryptographic technique which allows visual information to be encrypted in such a way that the decryption can be performed by the human visual system, without the aid of computers. Similarly, audio cryptography is a cryptographic technique which allows audible information to be encrypted in such a way that the decryption can be performed by the human auditory system, without the aid of computers. This idea is an extension of visual cryptography. Several paper has been found to implement Audio Secret Sharing (ASS) schemes. Audio Secret Sharing in 2-out-of-2 has been widely used, by using a cover sound. This paper will explain about the author's exploration of another possible simple audio cryptography schemes in 2-out-of-2 schemes.**

***Index Terms*—Audio Cryptography, Visual Cryptography, One-time-pad Cipher.**

I. Introduction

Based on Simple Audio Cryptography (Adriansyah, 2010), we can easily create a secret sharing schemes of audio signal in time domain without a cover sound. We can further extend this idea to create a secret sharing schemes in frequency-domain. We could also add a cover sound. Based on Audio and Optical Cryptography (Y. Desmedt, 1998), we could add a harmonic cover sound. We could change the harmonic sound with noise as cover sound.

The basic of visual cryptography study is to create a secret sharing schemes that do not rely on computers or other hardware to perform decryption process. In other word, the decryption process can be done by Human Auditory System. In visual cryptography, decryption process can be done by overlapping all the visual shares. Similarly in audio cryptography, decryption can be done by overlapping all the audible shares, or in other words, by playing the shares simultaneously, for example by using stereo system.

In throughout the paper, the author uses MATLAB to compute, and to plot the sound wave.

II. Basic Theory

Fourier Analysis

Sound is a signal. Just as other signal, we could apply Fourier Analysis to the signal. Fourier analysis is used to decompose the signal into a number of individual signal with certain frequency. We could use Fourier analysis to decompose the sound as secret sharing scheme. In order to decompose discrete signal, we must transform the signal into frequency domain. We do this by applying Discrete Fourier Transform (DFT). Because DFT is cost-expensive, we use the efficient DFT algorithm namely Fast Fourier Transform (FFT), to compute the transform.

The result of FFT is an array of complex numbers that represent the signal in frequency domain. Because it is a complex number, it has a magnitude and phase information within each frequency. For example, if we create a harmonic sound with certain frequency f. The resulting Fourier Transform will have a sudden peak in certain position of x (frequency) axis in the frequency domain (if we plot the frequency vs magnitude of the transform). That is because the signal only contain one frequency distinct frequency, that is f. If the signal is a discrete signal, or any complex discrete signal (for example, music), the resulting transform will gave us information of all the frequencies and phases the signal uses.

Superposition Principles

Sound is mechanical wave. Just like another mechanical wave, superposition and interference principle can be applied to sound. Superposition principle in time domain and freqency domain can be described in the following figures. We will look at two sound, sound A and B. First we take a look at time domain representation of sound A.

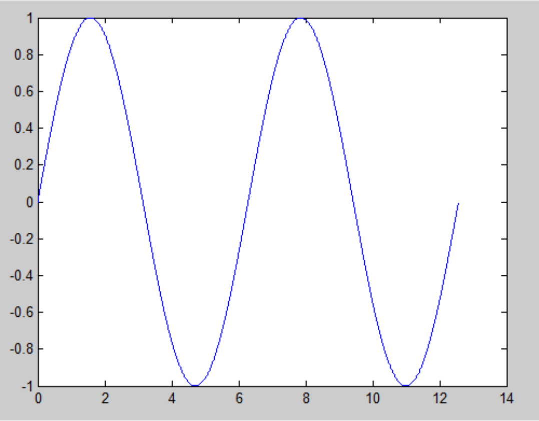

Figure 1 Time-domain representation of sound A

Below is frequency domain representation of sound A. We obtain the frequency domain by applying Fast Fourier Transform on the time domain representation. We plot the frequency domain in the complex plane.

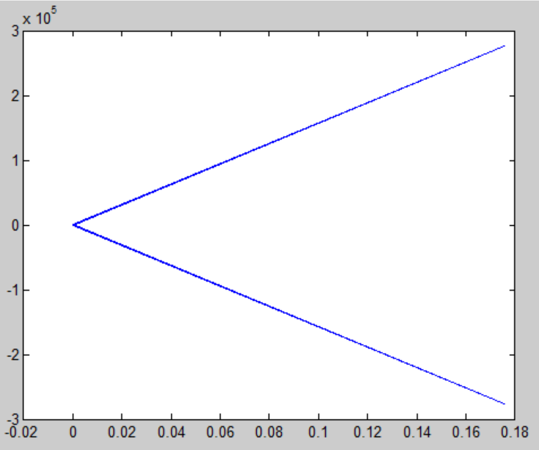

Figure 2 Frequency-domain representation of sound A

Next, we look at sound B in time domain, which have a higher frequency than A.

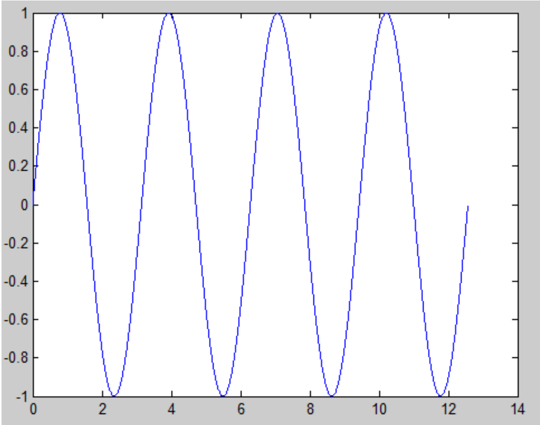

Figure 3 Time-domain representation of sound B

Below is frequency domain representation of sound B. It is also obtained by applying Fast Fourier Transform. The frequency domain also displayed in complex numbers plane. Notice the x axis of the plot. It has a wider range than sound A

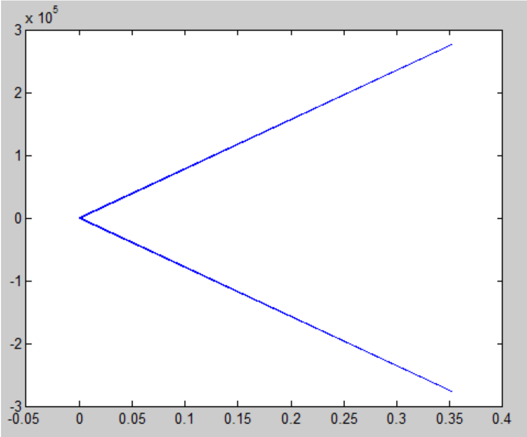

Figure 4 Frequency-domain representation of sound B

Next, we look at sound A “adds” B, where “adds” means using superposition principle in A and B. The figure below is the time domain representation of A “adds” B.

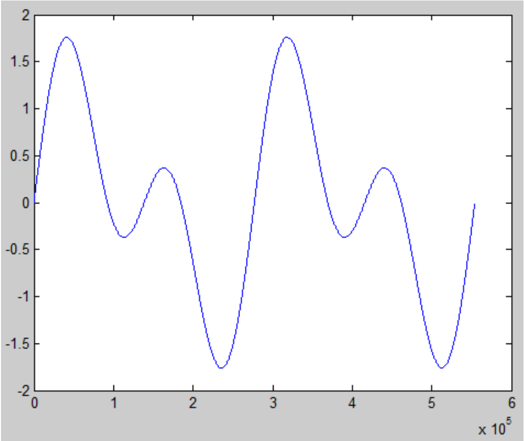

Figure 5 Time-domain representation of A + B

The figure below is the frequency domain representation of A “adds” B in the complex plane.

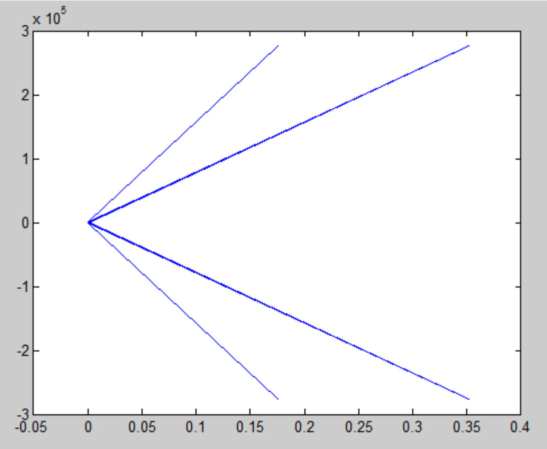

Figure 6 Frequency-domain representation of A+B

You will notice that superposition principles also can be applied in frequency domain. Therefore, we could use superposition principles in frequency domain and still affect the resulting wave in time domain.

Sound Interference

Decryption process in audio cryptography use interference principles. We must play the sound simultaneously in stereo system or in two different speaker, in order the sound to interfere each other. If both the signal have opposing phase, then the signal will be destroyed (destructive interference). This interference will destroy unnecessary signal and will resulting the secret sound we hide. We must note that complete signal destruction is impossible due to the environmental condition. We should align the stereo system to create maximum destructive interference to the unnecessary signal.

Digital Representation of Sound

Throughout the paper, the author will assume that the sound is represented digitally. The author assume the sample rate of the sound is 44100 Hz. For more information about digital representation of sound, Adriansyah (Adriansyah, 2010) also explain this on his paper, or you can refer to one of my paper (Nugraha, 2011).

III. Encryption Schemes

Encryption Scheme in Time Domain

Simple Audio Cryptography (Adriansyah, 2010) uses simple superposition scheme. Please refer to the paper for more information. This scheme did not use cover sound. If we heard each individual share, we could only heard scrambled noise. This scheme will break each sample into a valid combination of share's sample. If we "adds" this share's sample, it will result the original sample signal. Therefore, this simple scheme do not need any cover sound to hide it's content. This scheme manipulate the samples in time domain. The content information can be any audible sound.

Encryption Scheme in Frequency Domain

Audio and Optical Cryptography (Y. Desmedt, 1998) uses simple interference scheme. Please refer to the paper for more information. This scheme uses cover sound, for example a harmonic sound to hide its secret sound in each shares. This scheme takes advantages of Human Auditory System. Human capable of distinguishing frequency and volume (amplitude) in audible sound, but incapable of distingushing phase changes. So, if the sound had phase changes in the same frequency, we can not detect it. The idea is to encrypt information represented in High Volume and Low Volume bit with a cover sound. It is different with the scheme proposed by Adriansyah where the content information can be any audible sound. The author included this scheme in frequency domain category, because this scheme manipulate the signal phase. In this scheme, the cover sound phase is manipulated in such a way so that in the decryption process, the opposing phase in each shares will create destructive interferences or constructive interference depending on the desired volume changes in encrypted information.

Explored Schemes

Based on the explanation above, the author categorize the Audio Cryptography Schemes based on the manipulated domain. If we manipulate the amplitude in the given time domain, then it is time-domain encryption. If we manipulate the phase in the given frequency domain, then it is frequency-domain encryption. The author is interested in frequency-domain encryption and explored or proposed several possible schemes. In the following subsection, the author will describe the 3 explored frequency-domain encryption schemes.

A. Phase Sharing Encryption

This schemes do not need a cover sound. The idea is to transform the sound into frequency domain representation, and manipulate the sound's phases. It shares the sound's phases information into two shares. We could use further sharing algorithm, but in this paper we use simple even and odd sharing to demonstrate the scheme.

Let s is the sound we want to encrypt in time-domain representation. The scheme is described in the following step:

1. Transform s into its frequency-domain representation S using Fast Fourier Transform.

$$S = fft(s)$$

2. Split phase information in S into two shares, namely odd shares and even shares. Odd shares and even shares must satisfies the following rule. Let N be the number of sample in s, thus the number of element in S. Also, let p defined as

frequency partition. We could use

$$k = \{0,1,2,\ldots,N-1\}$$
$$odd_k = 0 \quad for\ every\ even\ floor(k/p)$$
$$odd_k = S_k \quad for\ every\ odd\ floor(k/p)$$
$$even_k = 0 \quad for\ every\ odd\ floor(k/p)$$
$$even_k = S_k \quad for\ every\ even\ floor(k/p)$$

3. Thus we have shared phase's information into two shares. The final step is to transform even and odd shares into its time-domain representation. Note that ifft means inverse Fast Fourier Transform

$$Odd\ Sound = ifft(odd)$$
$$Evem\ Sound = ifft(even)$$

B. Phase Changing Encryption

In this phase changing encryption, we change the phase's information of sound. So, w don't need any cover sound. We create a random phase with a certain magnitude. We put this phase's information into the first share. Let S_1 be the phase's information in the first share. Then we calculate the phase's information in the second share in order the shares to satisfies the following equation:

$$k = \{0,1,2,\ldots,N-1\}$$
$$S_k = S_{2k} + S_{1k}$$

Let s is the sound we want to encrypt. The scheme is described in the following step:

1. Transform s into its frequency-domain representation S using Fast Fourier Transform.

$$S = fft(s)$$

2. Create a random phase's information with fixed magnitude m. Let N be the number of sample in s, thus the number of element in S. Let m be the complex magnitude of the random phase, and S_1 be the first share.

$$k = \{0,1,2,\ldots,N-1\}$$
$$phi = \{phi | 0 \le phi < 2\pi\}$$
$$S_{1k} = m \cdot \cos(phi) + m \cdot \sin(phi) \cdot i$$
$$for\ every\ k, phi\ is\ random\ radian\ value$$

3. Let S_2 be the second share. Calculate the phase's information in the second share, with the following formula:

$$k = \{0,1,2,\ldots,N-1\}$$
$$S_{2k} = S_k - S_{1k} \quad for\ every\ k$$

4. Thus, we have changed the phase's information with random complex value. The final step is to transform into its time-domain representation.

$$First\ share = ifft(S_1)$$
$$Second\ share = ifft(S_2)$$

C. Noise Encryption

In this scheme, we add noise to the shares. So, actually it is lika adding a cover sound to the shares in the form of noise. We could add the noise in time-domain. However, we can not be sure how strong the noise we must provide to cover the secret sound. So, we must add the noise in frequency domain with certain noise magnitude m with a random phase. It is important to note, that the noise magnitude must be large enough to break and damaged the secret signal we want to encrypt. This is important so that the resulting share will not contain any secret signal.

Let s is the sound we want to encrypt. The scheme is described in the following step that is quite similar with phase changing encryption:

1. Transform s into its frequency-domain representation S using Fast Fourier Transform.

$$S = fft(s)$$

2. Create a random phase's information with fixed magnitude m. Let N be the number of sample in s, thus the number of element in S. Let m be the complex magnitude of the random phase.

$$k = \{0,1,2,\ldots,N-1\}$$
$$phi = \{phi | 0 \le phi < 2\pi\}$$
$$noise_k = m \cdot \cos(phi) + m \cdot \sin(phi) \cdot i$$
$$for\ every\ k, phi\ is\ random\ radian\ value$$

3. Then, add the signal with the noise to compute the first share. The second share is computed by adding the signal with the noise with opposing phase. Let S_1 be the first share, and S_2 be the second share. For every k, calculate with the following formula:

$$k = \{0,1,2,\ldots,N-1\}$$
$$S_{1k} = S_k + noise_k \quad for\ every\ k$$
$$S_{2k} = S_k - noise_k \quad for\ every\ k$$

4. The final step is to transform the shares into its time-domain representation.

$$First\ share = ifft(S_1)$$
$$Second\ share = ifft(S_2)$$

IV. Decryption Scheme

There is not much we can cover in this chapter. The objective of audio cryptography is to enable the decryption by using Human Auditory System (HAS). We could perform decryption by using stereo system or, playing both shares simultaneously. Let us assume that both shares will be played in a different machine simultaneously. Then each share will output in the different speaker. In order to hear the secret sound, we must put the speaker in the position so that each share will interfere each other. The people who will decrypt the secret sound, let us say this the decryptor, must hear the resulting interference in the right position. The distance of the speaker with the decryptor will affect the incoming share's signal phase in the ear. If the phase is not the same, the decryption can not happen. So, it is important to place the decryptor in the position so that the distance of each speaker to the decryptor will be the same, thus the incoming signal will have the same phase.

Physical Decryption

Practically, physical decryption can be done in the following step:

1. Put both speaker as close as possible facing one single direction (to the decryptor).
2. Place the decryptor in front of both the speaker.
3. Play both the shares simultaneously.
4. The decryptor can reveal the secret sound or message by using only one ear.

Computer Aided Decryption

The above step conclude the decryption scheme physically. If the decryptor insist in decrypting the message using computer (computer aided), we could use audio mixing software. One of the available free audio mixing software is Audacity, and can be downloaded via internet. Decryption can be done by mixing both shares in the same channel (mono channel). The resulting signal is the secret sound. The example guide of using Audacity to decrypt the shares can be described in the following step:

1. Open Audacity
2. Drag both shares (audio files) into Audacity's window. This will load each shares into different track in Audacitys window.
3. Combine the track into a single track. You can see the screenshoot for more detail.

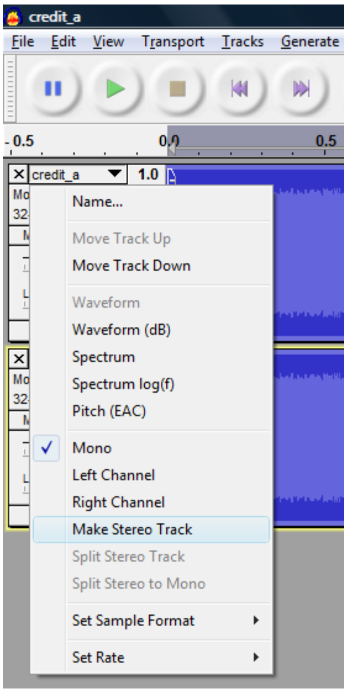

Figure 7 Menu to combine into stereo track in Audacity

4. Mix both channels by selecting the track and click the menu, Track > Stereo Track to Mono.

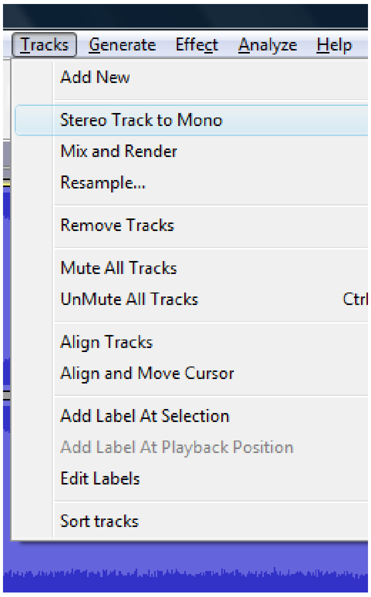

Figure 8 Menu to mix stereo track into mono track in Audacity

5. The decryption process is done. Now you can play the resulting sound to reveal the secret message.

V. Experiment Result

The author test and compare each encryption scheme mentioned in this paper. The encryption scheme tested are, Adriansyah's (Adriansyah, 2010), Desmedt's (Y. Desmedt, 1998), and the author's three explored schemes that is Phase Sharing Encryption (PSE), Phase Changing Encryption (PCE) and Noise Encryption (NE). The author compared the resulting share based on the following parameter, the description of the encrypted sound, the information the scheme can contain, the domain the encryption algorithm used, the use of cover sound, the cryptographic strength of the algorithm, computational cost and complexity, disadvantages of the algorithm, and the advantages of the algorithm. The comparison table can be found in the appendix page.

The author test the algorithm using MATLAB, matlab code, and FFT function provided by MATLAB. The code itself is not complicated, and self explanatory based on the implemented schemes.

Adriansyah's Simple Audio Cryptography (Adriansyah, 2010)

This algorithm can encrypt any sound, but the cryptographic strength of the resulting share depend on the original sound before encryption. The algorithm become less secure if the original sound contain distinct message, such as human spoken words, because the resulting share randomly contain original sound. The resulting encrypted sound is noise-like sound. The author categorized this scheme as time-domain encryption scheme. The computational complexity is relatively small, and the cost depend on the original sound length. This algorithm can be easily implemented.

Desmedt's Audio Cryptography (Y. Desmedt, 1998)

This algorithm can only encrypt volume information sound. Volume information is an analog to bit information. High volume represent bit 1 and low volume represent bit 0. This kind of information is encrypted by using a cover sound. The simplest cover sound can be used is harmonic sound. This simple idea is very secure. It can not be decrypted unless we have both shares. The encrypted sound can be heard as cover sound. Because the algorithm manipulated the cover sound's phase, the author categorize this scheme as frequency-domain encryption. The computational complexity is relatively small, and the cost depend on the volume information length and the time slot used.

Phase Sharing Encryption

This algorithm can encrypt any sound. The resulting share will contain meaningless and confusing echo of the original sound if the algorithm's parameter can be chosen carefully. The author categorized it as frequency-domain encryption. This algorithm is not very secure because it is using echo from the original sound (the result of shared phase). However, in some situation where the attacker or cryptanalyst does not know the spoken context of the original sound, the resulting share can be confusing. This algorithm do not need any cover sound, because it already use its own sound as cover. The computational complexity is considered high and expensive, because it is using Fourier Transform. As consequences of the limitation in current FFT algorithm, there is limitation in the duration of the encrypted sound. However, it is still feasible with short duration sound. The author only tried the scheme with duration below 20 seconds, which is considered sufficient.

Phase Changing Encryption

This algorithm can also encrypt any sound. The resulting share will contain noise like sound and damaged the original sound entirely if the algorithm's parameter carefully chosen. The author categorized it as frequency-domain encryption. This algorithm is very secure. However, because of the scheme's nature, one of the share will contain spectrum with the same absolute value in all the frequency in frequency-domain, let us say this is the key share. The other share will contain a similar spectrum with the original sound, let us say this is the keyhole share. Imagine some attacking scheme like following. The attacker only have the key share. The attacker can not decrypt the sound, because it doesn't even have information about the original sound. Now, let us assume the attacker only have the key hole share. The attacker now have the similar spectrum of the original sound. The attacker can also have the information about the absolute value of the key. However, the attacker must also need the phase information of the key in order to decrypt the message. This phase information can not be obtained from the keyhole, unless the attacker doing some thorough spectrum analysis which is very exhausting. This algorithm do not use cover sound. The computational complexity of the scheme is also considered high and expensive, because it is using Fourier Transform and additional complex number operation. This scheme suffer the same limitation as Phase Sharing Scheme.

Noise Encryption

This algorithm can encrypt any sound. The resulting share will contain noise like sound and damaged the original sound entirely if the algorithm's parameter carefully chosen. The author categorized it as frequency-domain encryption. This algorithm is very secure. If we use the same analogy as Phase Changing Encryption, this algorithm will produce two keyhole with the spectrum similar with the original sound. The difference between the share's phase is the key. Just like Phase Changing Encryption, the cryptographic strength of the scheme will become insecure if the attacker can guess the key spectrum. This algorithm uses cover sound, that is the noise itself, which is generated randomly. Just like the other scheme that using Fourier Transform, this scheme also suffer the same limitation, but it still feasible with short duration sound.

VI. Conclusion

The author proposed a way of categorizing Audio Encryption Scheme into two domain, time-domain and frequency-domain. Audio encryption scheme is categorized as time-domain encryption if the attacker only bothered by the amplitude of the share to break the encryption. Similarly, audio encryption scheme is categorized as frequency-domain encryption if the attacker not only bothered by the amplitude of the share, but also by the phase and frequency of the share. If we applying this category, Adriansyah's Scheme is considered time-domain encryption and Desmidt's Scheme is considered frequency-domain encryption. Generally, frequency-domain encryptions are much stronger, because the attacker must consider a larger variable than in the time-domain.

The author also considered Phase Changing Encryption and Noise Encryption can be very secure, depend on the algorithm's parameter. While Phase Sharing Encryption is considered less secure but can be confusing, depend on the algorithm's parameter.

When testing the encryption schemes the author experiments with different parameter and found the sufficient parameter.

1. Phase Changing Encryption uses following paramater: Complex magnitude of random phase (m) equal to 1000.
2. Noise Encryption uses following parameter: Complex noise magnitude of random phase (m) equal to 1000.
3. Phase Sharing Encryption uses following parameter: frequency partition (p) equal to 20.
4. Desmedt's Scheme uses following parameter: harmonic sound with frequency 200 Hz as cover sound, time slot length equal to 1 s.

Actually cryptographic strength in Desmedt's schemes does not depend on the parameter. But the author's schemes depend on the parameter. When testing the author's schemes, the author used spoken words as original sound.

Based on the author's experiment and opinion, audio cryptography is very secure, just like visual cryptography. It also shares the same characteristics as visual cryptography, such as it is relying on Human Auditory System in the decryption process. Audio cryptography can become an alternative of visual cryptography for people with visual disability. Audio cryptography also can be implemented to become One-time-pad Cipher. If implemented correctly, these schemes can be as strong as One-time-pad, because the scheme relying on random numbers. This characteristic is very much alike with visual cryptography.

REFERENCES

Adriansyah, Y. (2010). *Simple Audio Cryptography.* Bandung: Department of Informatics Engineering, Schools of Electronics and Informatics Engineering, Bandung Institute of Technology.

Naor, M., & Shamir, A. (1994). *Visual Cryptography.* Eurocrypt 94.

Nugraha, R. M. (2011). *Implementation of Direct-Sequence Spread Spectrum Steganography on Audio Data.* Bandung: Department of Informatics Engineering, Schools of Electronics and Informatics Engineering, Bandung Institute of Technology.

Socek, D., & Magliveras, S. S. (2005). *General Access Structures in Audio Cryptography.* Florida: CiteSeerX.

Y. Desmedt, S. H. (1998). *Audio and Optical Cryptography.* Advances in Cryptology-Asiacrypt'98.

STATEMENT

I hereby stated that this paper is written by my own and is my own work, it is not copied from other's paper, not a translation from other's paper, and not a plagiarism.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 9 Mei 2011

Rizky Maulana Nugraha – 13508083

Appendix

Table i Comparison table of explained audio encryption algorithm

Algorithm \ Parameter Name	**Adriansyah's Simple Audio Cryptography**	**Desmidt's Audio Cryptography**	**Phase Sharing Encryption**	**Phase Changing Encryption**	**Noise Encryption**
Original sound	Any sound	Volume information	Any sound	Any sound	Any sound
Resulting sound	Noise-like	Cover sound	Confusing echoed original sound	Noise-like	Noise-like (noise is the cover sound)
Encryption domain	Time-domain	Frequency-domain	Frequency-domain	Frequency-domain	Frequency-domain
The use of cover sound	Do not use cover sound	Use cover sound	Do not use cover sound	Do not use cover sound	Do not use cover sound
Cryptographic strength	Can be less secure, depend on the original sound	Very secure, independent of the original sound	Not secure yet confusing, depend on the parameter and original sound	Very secure, depend on the parameter	Very secure, depend on the parameter
Disadvantages	Depend on the original sound and random function	Can only encrypt volume information	Echoed share may reveal the message, FFT limitation	FFT limitation, can burden HAS	FFT limitation, can burden HAS
Advantages	Simple computation	Simple computation and very reliable	Can create confusion	Confusing noise-like sound, can encrypt any sound	Confusing noise-like sound, can encrypt any sound