Audio and Optical Cryptography

Yvo Desmedt[1,2⋆], Shuang Hou[2*], and Jean-Jacques Quisquater[3]

[1] Center of Cryptography, Computer and Network Security,
CEAS, University of Wisconsin – Milwaukee, and
Dept. of Mathematics, Royal Holloway, University of London, UK
desmedt@cs.uwm.edu,
http://www.uwm.edu/~desmedt
[2] Department of EE & CS, University of Wisconsin, Milwaukee
P.O. Box 784, WI 53201 Milwaukee, U.S.A.
hou@cs.uwm.edu
[3] Dept of Electrical Eng., Microelectronic laboratory
Université Catholique de Louvain, Place du Levant, 3
1348 Louvain-la-Neuve, Belgium
Quisquater@dice.ucl.ac.be,
http://www.dice.ucl.ac.be/crypto/jjq.html

Abstract. In visual cryptography the additive property of light is used. Also the shares are random and therefore suspect to a censor. In this paper we present two new cryptographic schemes which use music and the wave properties of light. Both schemes are also secret sharing schemes in which shares are music or images and are not suspect to a human censor. Our scheme guarantees perfect privacy as well as high quality. To decrypt the message, one just plays two shares on a stereo system. There are two decryption methods which are either based on the interference property of sound or based on the stereo perception of the human hearing system. In optical cryptography, we use pictures as covers and the wave interference property of light. The privacy is perfect and the modified images are non-suspicious. The Mach-Zehnder interferometer is used as the decryption machine.

1 Introduction

Traditional hiding and steganography methods, e.g., [4,5] have the disadvantage that once their method is known, anyone can find the embedded message.

Visual cryptography [6] is secure in this prospect. Visual cryptography is a perfectly secure encryption scheme in which both the ciphertext and the key are pixels, with 1 bit depth, printed on transparencies. The decryption is done by stacking the key transparency on top of the ciphertext transparency and does not require any computer. But both ciphertext and key consist of random pixels and hence are suspect to censors.

A reason for not using computers is that in some countries high technology equipment is suspect. Also, computers may not be trustful. Indeed, Goldberg

⋆ A part of this research has been supported by NSF Grant NCR-9508528.

K. Ohta and D. Pei (Eds.): ASIACRYPT'98, LNCS 1514, pp. 392–404, 1998.

and Wagner just found that at least 10 digits out of 64 bits keys in GSM system were actually zeroes [7]. Not only is it dangerous to trust software, trusting hardware is also not recommended. Today's Intel Pentium Pro microprocessor contains more than 5.5 million transistors and therefore it is easy to install a Trojan horse.

More recently cerebral cryptography [3] embeds a message in images and uses human brains to decrypt the ciphertext. It is also a perfect secret sharing schemes. It uses high quality real life images as cover images and generates two shares which maintain high quality. But it requires the cover image to have a large high frequency component, i.e., enough variation. Hence, it only allows very limited bandwidth. Also, decryption in cerebral cryptography is not so easy as the authors in [3] seem to claim. Some people have problems with 3-D perception.

In this paper we first present *audio cryptography* which uses music to embed messages. We base our scheme on the inference property of sound and phase perception of the human hearing system. Our scheme has similar features to that of cerebral cryptography. The privacy is perfect and a human censor is not able to detect that a single share is suspect. So, playing a single channel of the music, sounds as normal music. By playing two channels' sounds at the same time we can listen to the secret, i.e. the embedded message.

We then present our idea of *optical cryptography* which is based on the interference property of light waves and which uses images to hide information. This approach is completely different from the one used to obtain visual cryptography [6]. It achieves the same goal on privacy and no computer is necessary to do the decryption. As in cerebral cryptography [3], the shares are not suspicious. The privacy of our scheme is perfect and the stego-images (i.e., the modified images) are of high quality. Using a Mach-Zehnder interferometer [8] on two shares, we can see the embedded image. The scheme has the advantage of providing larger bandwidth over cerebral cryptography.

The organization of the paper is as follows. We first explain a model in Section 2. In Section 3 we discuss audio cryptography. In Section 4 we present the basic idea of optical cryptography. We conclude in Section 5.

2 Model

Before we present our schemes we introduce the model on which our cryptosystems have been built.

There are two agents (or in general n) that transport some secret message from one country to another country. Each agent carries one (or in general m) pictures/music, in which the secret message is embedded. They can not use computers.

There are human censors at each custom office who check each passenger's baggage. They cannot use a computer, either. We allow for two types of censors. Some that only censor suspicious pictures/music (then two pictures/music are

sufficient). The other type of censor will randomly destroy pictures/music (then we need n agents).

There is also some counterintelligence who may intercept one suspect picture/music. They have unrestricted computer power, but we assume they will never obtain two shares.

Our goal is that at least two agents can enter the other country successfully and finally meet each other. They put their shares together and they can decrypt the message without using any computer.
Our model has the following security properties:

- Unconditional privacy, i.e., the counterintelligence has infinite computer power.
- Censors can only use human computation.

Note: modern cryptography has three levels of computation powers, i.e.,

- infinite computer power
- quantum computer power
- polynomial time (Turing machine) computer power.

We have extended this to include human computation power.

In this paper, we use "embedded message" to refer to the plaintext, "cover" to represent the original image or music which is used to encrypt the plaintext and "stego-" to refer to the modified image or music which is transported by agents.

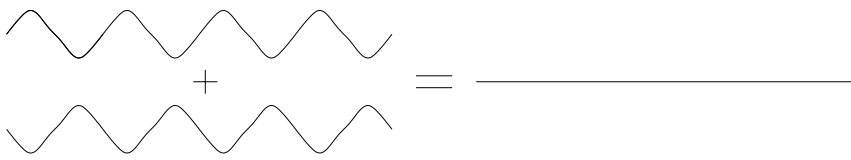
3 Audio Cryptography

One of our approaches to hide information is based on the interference property of sound waves. The other is based on the fact that the human hearing system is capable of observing phase differences. The two methods only differ in the way to decrypt the ciphertext. In Section 3.1 we will first give a simple explanation of both concepts. In Section 3.2 we will construct our basic scheme using a harmonic sound and then we will extend it to regular music. In Section 3.3 we demonstrate our results.

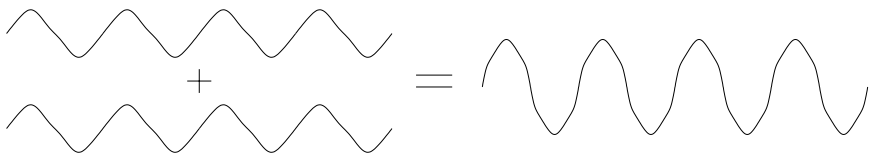
3.1 Two Concepts

Interference of Sound Sound is a pressure wave traveling through air, water or any other media. Interference occurs when two sound waves encounter each other while traveling. A sound wave is a moving series of sompressions (high pressure) and rarefactions (low pressure). If the high-pressure part of one wave lines up with the low-pressure of another wave, the two waves interfere destructively and there is no more pressure fluctuation (no more sound). On the other hand, if the high-pressure part of one wave meets the high-pressure part of another wave, it results in an intensified high-pressure. Note that the matching must occur in both

space and time [9]. As shown in Figure 1a, if two simple harmonic sound waves are of the same frequency and amplitude, and if they are superimposed upon one another out of phase (with a 180 difference in phase), then they will destroy each other completely. While in Figure 1b, if they are superimposed upon one another with 0 difference in phase, the resulting wave has an amplitude which is twice of the original one.



a. Destructive interference.



b. Constructive interference.

Fig. 1. Interference illustration.

This property has been applied to active noise control [12] where active attenuation of noise is obtained by using artificially generated acoustic waves mixed with the unwanted sound so that when the waves are in anti-phase, then destructive interference results.

We observe that the interference principle acts like a *not-exclusive-or* operation, which gives 1 (corresponding to an amplitude of 2) only when the two operands are of equal value.

Stereo Conception We can localize the direction from where the sound originate. As shown in Figure 2, the sound Source 1 has the same distance from both ears and the Source 2 is on the right side of the person in the figure. The waves from Source 1 arrive at the two ears with the same amplitudes and same phases. But the waves from Source 2 travel a little longer to get to the left ear compared to the right ear. This means that the waves striking two ears are of different amplitudes and most likely of different phase. As a result, the human hearing system can observe whether the sound comes from Source 1 or from Source 2.

The aspect of observing the phase differences is used in one of our decryption methods.

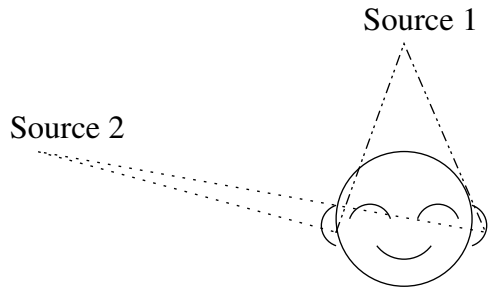


Fig. 2. Illustration of observing sound sources.

3.2 Schemes

Our goal is to use shares of the embedded message that are high quality music. We therefore start from some high quality music as the cover. We then want to produce two piece of stego-music of high quality, such that if one plays these, one obtains the embedded message. For convenience, we will refer to this scheme as a 2-out-of-2 secret sharing scheme.

The basic idea of audio cryptography is as follows. One generates the share s_1 based on random coin flips b and the second share s_2 based on $\overline{b \oplus S}$, where S is the secret bit we want to hide and $\oplus$ is the exclusive-or. It is clear from the properties of the *one-time pad* [13,11] that such schemes guarantee perfect secrecy.

In the following, we first use harmonic sound as the embedded message. We then generalize it by using music to obtain our anti-censor goal.

Harmonic Scheme Our basic scheme using harmonic sound is presented first.

The setting

S: a plaintext message which is a binary string.
L: the length of the embedded message which represents how many bits are in S.
T: a parameter which represents how many seconds of sound are used per secret bit. So, we need a total of $T \times L$ seconds of sound in order to encrypt a secret message of L bits.
B: a cover sound which is a single frequency signal lasting $T \times L$ seconds.

Procedure:

- Generate the first share s_1 as follows: Initialize s_1 to B. For every T seconds data from s_1, flip a coin b. If b is 1, multiply the corresponding T second data with -1, implying a 180 phase change. Otherwise leave them unchanged. So, one has randomly chosen to flip the T seconds data to its opposite phase or not.

– Generate the second share s_2: Initialize s_2 to B. For every T seconds data from s_2, compute $b' = \overline{b \oplus S}$. If b' is 1, multiply the corresponding T second data with -1, implying a 180 phase change. Otherwise leave them unchanged. In other words, if the secret bit is 1 then the corresponding T seconds sound from s_2 has the same phase as that from s_1. If the secret bit is 0 then the corresponding T seconds sound from s_2 has the opposite phase as that from s_1.

Two Decryption Methods There are basically two ways that can be used to decrypt the ciphertext in order to get the embedded message. They are either based on the interference property of waves or based on the stereo conception property of the human hearing system.

In the first method, we put two speakers very close and face to face. Then, we send share 1, s_1, to one speaker and send share 2, s_2, to the other speaker. We can clearly notice the effect of volume changing, in which louder represents secret bit 1 and more silent represents secret bit 0. The cancellation is not complete due to the incomplete destructive interference, the reflection from the wall, etc.

In the second method, we move one speaker to our left side and the other speaker to our right side. Then, as in Method 1 we play two shares from two speakers respectively. We can observe that the sound sources move from sides to center and from center to sides, which is due to the phase differences in two channels. If both signals from two channels are of the same phase, which encodes secret bit 1, we observe only one source which is from the center. If two signals are out of phase, which corresponds to secret bit 0, we observe two sources, one from left and one from right.

In a variant of the second method, we use a set of headphones instead of two speakers. We play one share in each ear, we obtain the same effect as in Method 2 due to the phase conception property.

Testing on Harmonic Sound We have tested our scheme on three harmonic sounds which have frequencies 300Hz, 500Hz and 1000Hz respectively. All the decryption methods worked pretty well. But, each share is suspicious. We heard some clicks at each phase changing point as shown in Figure 3. This is because the modified signal is not of a single frequency any more and the added frequencies make the click very recognizable in the pure tone environment.

Music Scheme We extend our basic scheme of using a harmonic sound to a more general one of using music. We modify the algorithm described in the harmonic method only by using a piece of music instead of a harmonic sound as the sound B uses to hide the share. Nothing else need to be changed.

The problem which exists with the harmonic method does not exist in our general scheme. We could hardly hear such clicks. When playing only one stego-sound, either share 1 or share 2, we get very good quality music which sounds just as the original one. It is hard to tell any difference. When playing both, we

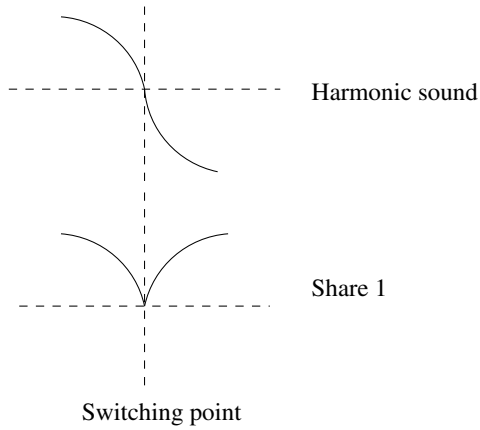


Fig. 3. Illustration of how clicks are generated in a harmonic method.

can observe the volume changing if using the decryption Method 1 and we can observe the switching of the sound sources if using the decryption Method 2 or its variant. All methods provide correct decryption.

Doing some spectrum analysis, we can see how close the two curves in Figure 4 are, one for the original music and one for the signal after being randomly phase changed. The music is rich in frequencies and therefore the added noise, which is also distributed flatly among a wide range of frequencies, has little impact on human ears.

If the volume of the music goes up and down dramatically and frequently and the cancellation is not complete by using two speakers, it may be difficult to make the right decryption using Method 1. But, in such circumstances, one can always use the methods which are based on the phase conception property, i.e., Method 2 and in particular its variant.

2-out-of-n Schemes To generalize our previous 2-out-of-2 to 2-out-of-n, we use the secret sharing scheme discussed in [2] and use $\lceil \log_2(n) \rceil$ different pieces of music as covers.

We remind the reader that the 2-out-of-n secret sharing scheme in [2] is based on $\lceil \log_2(n) \rceil$ many 2-out-of-2 sharing schemes executed independently. So if k is the secret key, one has $k = r_0^i \oplus r_1^i$, where $1 \leq i \leq \lceil \log_2(n) \rceil$. When numbering the participants from 0 to $n-1$, participant j receives share r_0^i if the ith bit of the binary representation of the integer j is 0, else r_1^i.

So, in our context, one uses $\lceil \log_2(n) \rceil$ pieces of music as covers. For practical purposes they are different. For each of the $\lceil \log_2(n) \rceil$ pieces of music one creates shares R_0^i and R_1^i as in our previous 2-out-of-2 audio cryptosystem. A participant j receives the audio channel R_0^j when the ith bit of the binary representation of j is 0, else receives R_1^i.

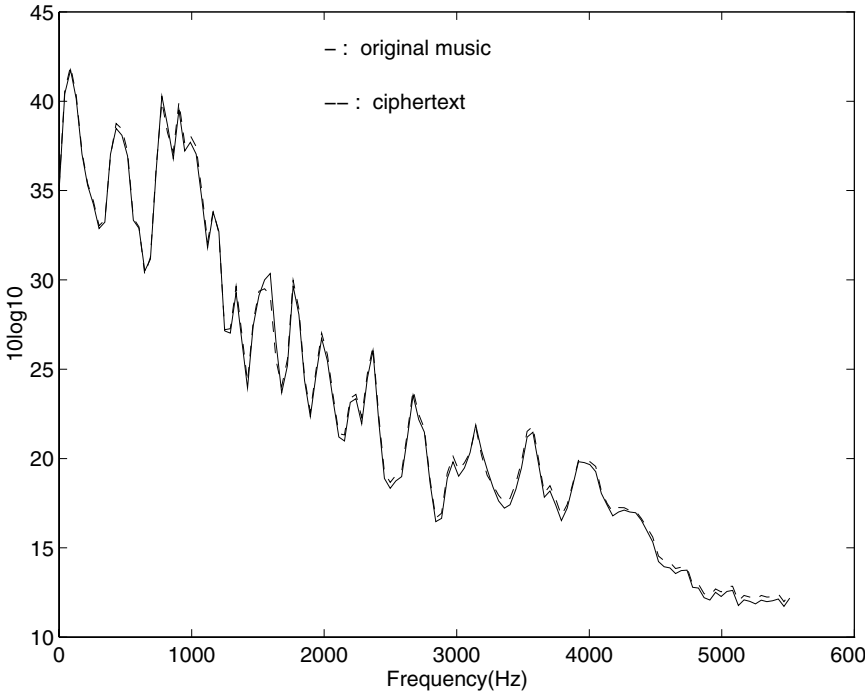


Fig. 4. Spectrum comparison of original music with ciphertext.

3.3 Demonstration

We present some sound sample showing the original music signal, two shares and corresponding secret bit in Figure 5.

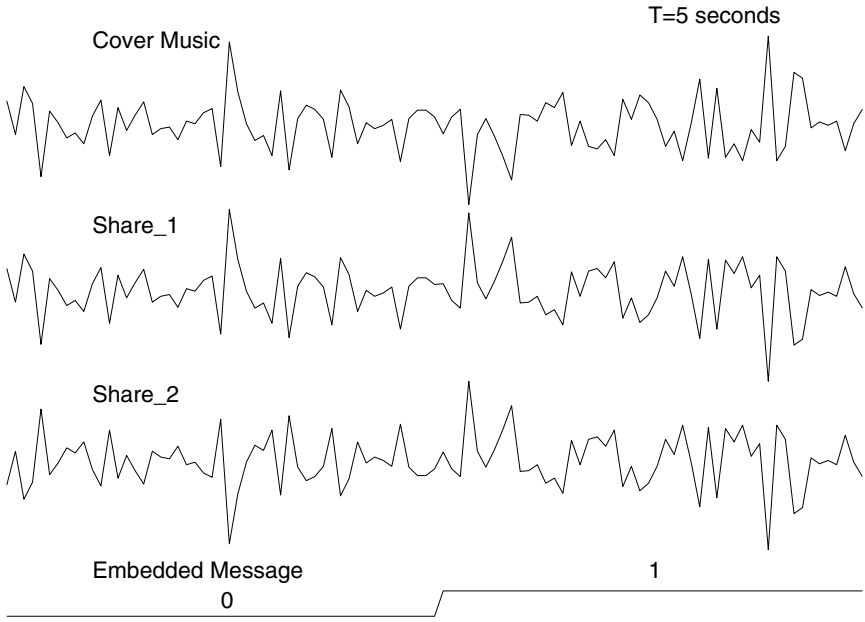


Fig. 5. Audio cryptography illustration.

We have done tests on pop music and also on classical music. These tests can be found at `http://www.cs.uwm.edu/~desmedt/audio/` Both results have shown that each share (stego-sound) is of the same quality as the cover music and the decryption is correct.

4 Optical Cryptography

Light is also a kind of wave and therefore has the interference property. If two beams of light from the same source meet out of phase, they will destroy each other and this results in total darkness. If they meet with the same phase, then they produce an intensified light.

Our idea of optical cryptography is as follows. Our plaintext is a 1 bit/pixel digital image (e.g., a blueprint). We choose a high quality n bits per pixel image which has a larger size than that of our plaintext. We pad the plaintext to make it the same size of the cover image. We generate the share 1 by randomly flipping the least mth significant bit of each pixel in the cover image. We copy share 1 to

share 2 as its initial value. Then if in the plaintext a pixel has the value 1 then we flip the least mth significant bit of this corresponding pixel in share 2. So, now the mth significant bits in the generated shares are uniformly random bits. If m is small enough then we maintain the high quality. (When n is 8, m can be 4 and the alternation is unnoticeable to a human as shown in Figure 6 and Figure 7.) The two shares only differ in the least mth significant bit. Denote the least mth significant bit from share 1, share 2 and the plaintext as s_1, s_2 and s respectively. Then, they are clearly related by $s_2 = s_1 \oplus s$ which is equivalent to $s = s_1 \oplus s_2$.

Fig. 6. Share 1 for optical cryptography scheme with n =8 bits/pixel and m =4th least significant bit.

Now we can use a machine called Mach-Zehnder interferometer [8] to reconstruct the plaintext. As shown in Figure 8, the laser beam passes some lenses and becomes a wide parallel beam. Then, it is split into two beams, beam 1 and beam 2, which take different paths. When beam 1 passes share 1, its amplitude is changed by the corresponding pixel values in share 1. The beam 2 passes share 2 and carries similar information about share 2. When finally the two beams meet out of phase, the result is the plaintext.

Fig. 7. Share 2 for optical cryptography scheme with n =8 bits/pixel and m =4th least significant bit.

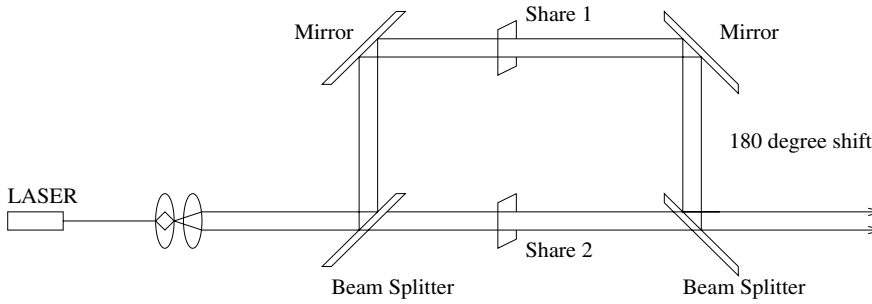


Fig. 8. Illustration of the decryption for optical cryptosystems.

Optical cryptography allows for high bandwidth encryption while still maintaining our covert property of the shares. The high bandwidth results from using the modification of each pixel.

This 2-out-of-2 perfect threshold scheme can easily be extended to a 2-out-of-n perfect threshold scheme as for audio cryptography. We use the secret sharing scheme discussed in [2] and use $\log_2(n)$ different images as cover images.

We will report about the testing results of optical cryptography in the final paper.

5 Conclusions and Open Problems

We have demonstrated that audio cryptography uses high quality music sound as shares and provides perfect privacy. Decryption is easy by playing both shares at the same time. We have discussed two decryption methods. For the decryption, we only need a stereo player and a stereo headphone (or two speakers).

We also presented optical cryptography which is different from cerebral cryptography. It has all the aspects of visual cryptography and cerebral cryptography. Only a Mach-Zehnder interferometer, a laser beam and some lenses are needed for decryption.

Both schemes can be considered as 2-out-of-2 threshold secret sharing schemes. We have shown how to generalize them to 2-out-of-n secret sharing schemes by using different cover pictures or sounds. It is not clear how to generalize them to more general t-out-of-n schemes.

Audio cryptography as well as optical cryptography do not need a digital computer to decrypt the ciphertext, however they *do* require one to encrypt the plaintext. This introduces two open questions:

- Can a cryptographic scheme be developed that does not need a digital computer or equivalent electronic hardware to encrypt plaintext and hide the share as in our schemes, and
- Can a scheme be developed that does not rely on digital computers (or electronic equipment) for encryption as well as for decryption.

References

1. Blakley, G. R.: Safeguarding cryptographic keys. In *Proc. Nat. Computer Conf. AFIPS Conf. Proc. (1979)* pp. 313–317
2. Desmedt, Y. and G. Di Crescenzo and Burmester, M. : Multiplicative non-abelian sharing schemes and their application to threshold cryptography. In *Advances in Cryptology — Asiacrypt '94*, Proceedings (Lecture Notes in Computer Science 917), pp.21–32.
3. Desmedt, Y. and Hou, S. and Quisquater, J.J.: Cerebral Cryptography. *Workshop on information hiding,* Preproceedings, April 15-17, 1998, Portland, Oregon, USA.
4. Franz, E., Jerichow, A., Möller, S., Pfitzmann, A. and Stierand, I. : Computer Based Steganography: How it works why therefore any restrictions on cryptography are nonsense, at best. *Information Hiding*, Proceedings, 1996, pp.3–21.

5. Kurak, C., McHugh, J.: A cautionary note on image downgrading. In *Proceedings of the 8th Computer Security Applications Conference (December 1992)*
6. Naor, M., Shamir, A.: Visual cryptography. In *Advances in Cryptology — Eurocrypt '94*, Proceedings (Lecture Notes in Computer Science 950) (May 9–12, 1995). A. D. Santis, Ed. Springer-Verlag pp. 1–12
7. *The New York Times*, April 14, 1998, pp.C1.
8. Nussbaum, A. and Phillips, R. A. *Contemporary Optics for Scientists and Engineers.* Prentice-Hall, 1976
9. Sears, F. W. and Zemansky, M. W.: *University Physics.* Addison-Wesley, 1964
10. Shamir, A.: How to share a secret. *Commun. ACM* **22** *(1979)* 612–613
11. Shannon, C.E.: Communication theory of secrecy systems. *Bell Systems Technical Journal*, 28 (1949), 656-715
12. Tokhi, M. O. and Leitch, R. R. : *Active Noise Control.* Oxford Science Publications 1992
13. Vernam, G.S.: Secret signaling system. *U.S. Patent # 1,310,719, 22 Jul 1919*