

New Audio Secret Sharing Schemes With Time Division Technique*

CHEN-CHI LIN, CHI-SUNG LAIH AND CHING-NUNG YANG[†]

*Cryptography and Network Security Laboratory
Department of Electrical Engineering
National Cheng Kung University
Tainan, 701 Taiwan*

[†]*Department of Computer Science and Information Engineering
National Dong Hwa University
Hualien, 974 Taiwan*

An Audio Secret Sharing (ASS) scheme is a special type of secret sharing scheme [3], which the shares of embedded messages use music as cover sound. Desmedt et al. firstly introduced the (2, 2) ASS scheme with one cover sound and also the generalized (2, n) ASS scheme with $\lceil \log_2 n \rceil$ different cover sounds. No only will more cover sounds overburden the human hearing system but also may become difficult for people to distinguish the secret bit correctly. Thus, their scheme is not practical when n is large.

In this paper, we will propose two new (2, n) ASS schemes, which carefully employ the technique of time division with only one cover sound. Comparing with the first scheme, the second scheme has the advantage of flexible improvement in relative contrast as needed. To test the acoustic result, we implemented these two proposed (2, n) ASS schemes for small n using one wave-type cover sound and then obtained near expected results.

Keywords: audio secret sharing (ASS) scheme, secret sharing schemes, threshold schemes, key protection, key management

1. INTRODUCTION

In 1998, Desmedt, Hou, and Quisquater [4] first proposed the (2, 2) audio secret sharing (ASS) scheme. We abbreviate it as the DHQ (2, 2) ASS scheme. The goal of the DHQ ASS scheme is to embed a binary secret message by cover sound, such as harmonic sound or high quality music. For example, in the DHQ (2, 2) ASS scheme, the human "ears" can decode the concealed message if one plays two shares simultaneously.

Desmedt et al. also proposed the generalized DHQ (2, n) ASS scheme based on their (2, 2) ASS scheme using $\lceil \log_2 n \rceil$ different cover sounds. Let S as the secret value, one has $S = s_0^i \oplus s_1^i$, where $1 \leq i \leq \lceil \log_2 n \rceil$ and \oplus is exclusive-or. Participant j , 0 to $n - 1$, receives shares s_1^i if the i th bit of the binary representation of the integer j is 1, else s_0^i . Therefore, more cover sounds are needed when the number of participants n increases and this will overburden the human hearing system. In other words, the more cover sounds there are, the less the human interprets them correctly.

Received March 25, 2002; accepted December 16, 2002.

Communicated by Shih-Pyng Shieh.

* The preliminary version of this paper has been presented at Information Security Conference 2002, Taichung, Taiwan, May 2002.

In this paper, we will propose two new $(2, n)$ ASS schemes, which carefully employ the technique of time division with only one cover sound. Comparing with the first scheme, the second proposed scheme has the advantage of flexible improvement in relative contrast as needed. This paper is organized as follows. We first briefly describe the DHQ $(2, n)$ ASS scheme. In section 3, the model of the new $(2, n)$ ASS is explained. In section 4, we propose two constructions of the $(2, n)$ ASS with only one cover sound. In the final section, we implement these two proposed $(2, n)$ ASS for small n using the wave-type cover sound, discuss the experimental results and make a conclusion about our new schemes.

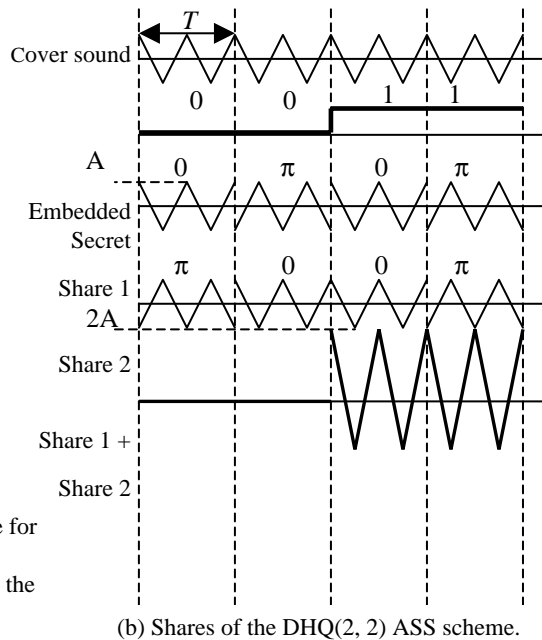
2. PRELIMINARIES

The DHQ $(2, 2)$ ASS scheme uses a triangle wave as the cover sound, and the result is shown in Fig. 1(b). The following parameters are used in this example:

Message \ Shares	L		H		
	S1	0	π	0	π
	S2	π	0	0	π

(a) The relation of the phrase change for each share.

H(L): High(Low) volume to represent the plaintext message 1(0).



(b) Shares of the DHQ(2, 2) ASS scheme.

Fig. 1. The DHQ $(2,2)$ ASS scheme with a triangle cover sound.

- $S = (0\ 0\ 1\ 1)$: a plaintext message.
- $L = 4$: the length of the embedded message.
- T_b (seconds): the length of the secret bit.
- B : a cover sound (triangle harmonic sound).

The procedure for generating shares is described below. The dealer first initializes

two shares, s_1 and s_2 , to be the cover sound. Share s_1 is changed according to random coin flips, c , for every T_b second of sound. If c is 1 or 0, this makes a change of 180 (π) phase or zero (0) phase for the corresponding T_b second of cover sound. Share s_2 based on $c' = \overline{c \oplus S}$ is generated for every T_b second of sound. If c' is 1 or 0, this also makes a change of 180 (π) phase or zero (0) phase for the corresponding T_b second of cover sound. The relation of the phase change is shown in Fig.1a. According to the embedded message, each share has fixed maximum amplitude A and has four times change of phase in $T \times L$ period. Playing one share sounds like normal music; however, while playing two shares simultaneously, one can hear the secret, i.e., the embedded message.

Two decryption methods are generally used in the DHQ (2, 2) ASS scheme, which are based either on the interference property or on the stereo perception of human ears. In the decryption method of the interference property, two speakers are put very close together, and face to face. The listener can clearly notice the change of the volume because loud volume represents secret bit 1 and low volume represents secret bit 0. In the decryption method of the stereo perception, the listener can detect only one source, which is from the source if the secret bit is 1. Otherwise, the listener can detect two sources, one from the left and the other from the right, if the secret bit is 0.

Our scheme mainly depends on the property of sound interference. If two sounds with the same frequency and amplitude are played simultaneously, then there will be two results in the theoretical environment. One is constructive interference because the two shares are in the same phase. The other is destructive interference for the two shares are out of phase. The interference principle is similar to an “or”ed operation symbolized by “+”. Therefore, when share1 “or”ed share2, this means that the different shares are played simultaneously from two stereo channels or two speakers. Then, one will hear the change of the volume, $2A$ or silent amplitude, and perceive the embedded message 1 or 0 for each T time slot. Finally, one can get the secret, S .

3. THE PROPOSED MODEL

The drawback of the DHQ (2, n) ASS scheme is that it needs $\lceil \log_2 n \rceil$ different cover sounds. Our goal is to propose one model to construct the (2, n) ASS scheme with only one cover sound. Using the subpixel concept of visual cryptography [5], the proposed model utilizes the technique of time division to produce the sub time slot in each share.

This model is described as follows. Either a harmonic sound or high quality music is chosen as the cover sound lasting $T \times L$ seconds, where each secret bit occupies T seconds for an L bit binary secret. The dealer firstly initializes n shares to be all the cover sound. In the period of each binary bit, the dealer logically divides T seconds of each share into m sub-slots (or $m \times r$ sub-slots, $r \in \mathbb{N}$) with the same time intervals. The m sub-slots in each secret bit are defined as an m -vector V . In the DHQ ASS model, the processed time unit is the period of one secret message bit (T seconds). However, in our model, the processed time unit is one sub-slot, equal to T / m (or $T / (m \times r)$ seconds), where r is the repeated cycle of m -vector V in the period of one secret bit.

The m -vector V of those n shares is realized by using an $n \times m$ Boolean matrix $D = [s_{ij}]$, where $s_{ij} = \pi$ if the j th piece of one m -vector V in the i th share has a phase change of 180 degree compared with the original cover sound, or $s_{ij} = 0$ if the j th piece share of one

m -vector V in the i th share remains unchanged. When those shares i_1, \dots, i_t are played simultaneously, the loudness level of the combined result is proportional to the operation weight $w(V)$, which is the sum of the m -vector r_{i_1}, \dots, r_{i_t} , where r_{i_1}, \dots, r_{i_t} are the corresponding rows i_1, \dots, i_t of D . In other words, the operation weight $w(V)$ represents the number of “increasing” interference values. One “increasing” interference value is not a zero value from the “or”ed operation of the sub-slots of the qualified shares. Therefore, the maximum number of $w(V)$ has m or $m \times r$ values of “increasing” interference in the period of a secret bit.

For this model, let D_H and D_L be two collections of $n \times m$ Boolean matrices. The dealer randomly chooses one of the matrices in $D_H(D_L)$ to process secret bit 1(0). One row of the Boolean matrices $D_H(D_L)$ can be called a high-level (low-level) vector. The chosen matrix defines the phase change of the m sub-slots in each of the n shares. $A(t, n)$ ASS scheme is considered valid if it satisfies two conditions $A1$ and $A2$, where $A1$ is called acoustic contrast and $A2$ is called security.

- A1. For any $D \in D_L$, the sum of rows r_{i_1}, \dots, r_{i_t} satisfies $w(V) = l \leq d - \alpha(m) \times m$; whereas for any $D \in D_H$, it turns out that $w(V) = h \geq d$. The relative difference, $\alpha(m)$, is the ratio of the “increasing” interference numbers between the high-level and the low-level vectors. The relative contrast $\alpha(m)$ can be defined as $(h - l)/m$ to represent the loss of fidelity. The value of the relative contrast $\alpha(m)$ should be at least $1/m$ and designed as large as possible.
- A2. For any subset $\{i_1, \dots, i_p\}$ of $\{1, 2, \dots, n\}$ with $p < t$, the two collections of $p \times m$ matrices D_t for $t \in \{L, H\}$ obtained by confining each $n \times m$ matrix in D_t (where $t = L, H$) to rows r_{i_1}, \dots, r_{i_p} are indistinguishable in the case where they contain the same matrices with the same frequencies.

The first condition is related to the acoustic contrast. When a set of qualified users play their shares simultaneously, they can correctly detect the embedded message. The second condition is called security since it implies that nothing can be learned about the embedded information by inspecting fewer than t shares.

As Naor and Shamir stated about visual cryptography [5], the visual effect of a black subpixel in one of the transparencies cannot be undone by the color of that subpixel in another transparency which is laid over it. However, in audio cryptography [4], the audio effect of a sub-slot of one share can be undone by an anti-phase sub-slot of another share when they are played simultaneously in an ideal environment. In fact, our model pays more sub time-slots to get the benefit of one cover sound.

4. TWO PROPOSED $(2, n)$ ASS SCHEMES

It can be ruled out that the m -vector V of the n cover shares is constructed by two collections of Boolean matrices (D_H and D_L). Thus, the human hearing system can distinguish the embedded string by averaging the contributions of the “increasing” interference numbers from the qualified participant set. Therefore, we have to use a threshold d and relative difference $\alpha > 0$ to distinguish the acoustic loudness. Based on the above model, we can develop two different constructions for $(2, n)$ ASS schemes with only one cover sound.

4.1 Method-1 (2, n) ASS Scheme

In the construction Method-1, we simply use the following collections of $n \times m$ matrices:

$$D_L = \left\{ \text{all the matrices obtained by permuting the column of } \begin{bmatrix} \pi & 0 & 0 & \dots & 0 \\ 0 & \pi & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \pi \end{bmatrix} \right\},$$

$$D_H = \left\{ \text{all the matrices obtained by permuting the columns of } \begin{bmatrix} \pi & 0 & \dots & 0 \\ \pi & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \pi & 0 & \dots & 0 \end{bmatrix} \right\}.$$

Theorem 1 Method-1 (2, n) is an ASS scheme with the parameter $m = n$ and the relative contrast $\alpha(m) = 2/n$. It achieves perfect privacy.

Proof: We will show that D_H and D_L conform to the contrast and security rule described for the above model. In the following discussion, we assume r equals one, i.e., uses one cycle of the m -vector V . For the contrast condition, any two high-level vectors added together result in constructive operations of n sub-slots in the period of secret bit 1. In contrast, any two low-level vectors operated together result in constructive operations of $n - 2$ sub-slots in the period of one secret bit 0. In other words, there are n increasing interference numbers in the period of one secret bit 1 and there are $n - 2$ increasing interference numbers in the period of one secret bit 0. The relative contrast between these two cases is $2/n$. For each low-level (or high-level) vector in D_L or D_H , there is a random choice of $n - 1$ “zero phase change” sub-pieces and one “anti-phase change” sub-piece. Hence, it yields the same distribution from a random permutation of the columns in D_L or D_H regardless of which row was chosen. Thus, the “security” in Method-1 (2, n) ASS scheme holds.

Example 1 Let us consider for Method-1 construction. One cover sound of a triangle wave is shown in Fig. 2. If we use one cycle of the m -vector V , i. e. r equals to one, to process the phase change in each secret bit, there will be only three phase-changes in the period of one secret bit. For any two participants, there are three constructive interferences in the period of secret bit 1 and there is one constructive interference in the period of secret bit 0. In other words, the number of “increasing” interference is three in secret bit 1 and one in secret bit 0. The relative contrast of this example is $2/3$. The 3×3 basic low-level matrix D_L and high-level matrices D_H for (2, 3) ASS scheme are constructed as follows:

$$D_L = \begin{bmatrix} \pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{bmatrix} \quad D_H = \begin{bmatrix} \pi & 0 & 0 \\ \pi & 0 & 0 \\ \pi & 0 & 0 \end{bmatrix}$$

4.2 Method-2 (2, n) ASS Scheme

In Method-1 construction, the larger the share number n is, the lower the relative contrast will be. Hence, it is not easy to detect the secret message when n is large. In Method-2 construction, we use the constant weight codes to design the Boolean matrices. The constant weight codes have been well studied [1, 2] and can be defined as $A(x, y, z)$, where x is the length of maximum number of binary vectors, y is the minimum number of Hamming distance and z is the constant weight. In Method-2 construction, we apply the analog for 1 and 0 in Hamming weight terminology to π and 0 symbols used in our model. The author of [2] surveyed the known techniques for constructing constant weight codes. Thus, we use the lower bounds in the $A(x, y, z)$ tables in [2] to construct the low-level matrix D_L . We can obtain $D_L(n, m)$, where n is the table entry and m is equal to x if we choose z to $\lceil x/2 \rceil$ from lower bounds in the $A(x, y, z)$ tables. Therefore, each row in D_L constructed from the lower bounds in the $A(x, y, z)$ table [4] has $\lceil x/2 \rceil$ “ π ”

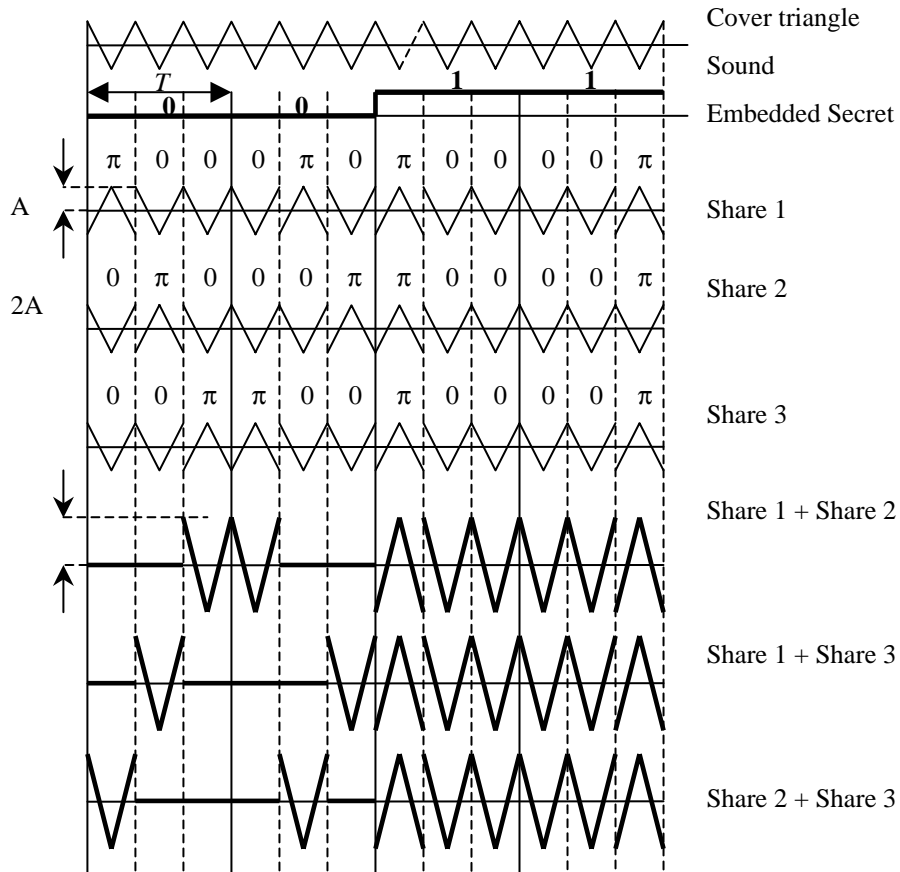


Fig. 2. Method-1 (2, 3) ASS scheme.

and $m - \lceil x/2 \rceil$ “0”; any two rows exclusive-ored has y hamming distance. In addition, each row of the high-level matrix D_H is constructed by $\lceil x/2 \rceil$ successive “ π ” and $m - \lceil x/2 \rceil$ successive “0”. The collections of $n \times m$ matrices are all the matrices obtained by permuting the columns of the Boolean matrices D_L and D_H .

Theorem 2 The above Boolean matrices D_L and D_H are used to construct the method-2 $(2, n)$ ASS scheme. The scheme has relative contrast $\alpha(m) = y / m$ and the perfect privacy.

Proof: In the following discussion, we assume r equals one. In this scheme, each row of the low-level matrix $D_L(n, m)$ constructed from the lower bounds in the $A(x, y, z)$ table [2] has $\lceil x/2 \rceil$ “ π ” and $m - \lceil x/2 \rceil$ “0”; any two rows exclusive-ored has y hamming distance. In addition, each row of the high-level matrix D_H consists of $\lceil x/2 \rceil$ successive “ π ” and $m - \lceil x/2 \rceil$ successive “0”. For the contrast condition, any two low-level rows added together always produce y sub-slot destructive operations in the period of secret bit 0 because the Hamming distance is at least y apart in these tables, and any two high-level vectors operated together have zero sub-slot destructive operation in the period of one secret bit 1. The relative contrast between these two cases is y/m .

For any low-level (or high-level) vector in D_L (or D_H), there is a random choice of $m - \lceil x/2 \rceil$ “zero phase change” pieces and $\lceil x/2 \rceil$ “anti-phase change” pieces. Therefore, each share constructed by D_L or D_H has the phase change $\{0, \pi\}$ with the same probability. Thus, the “security” of Method-2 $(2, n)$ ASS scheme holds.

Example 2 The advantage of Method-2 construction is that we can design the relative contrast as needed. We can use $(4, 6)$, $(7, 7)$, $(14, 8)$ and so on as the size of the (n, m) Boolean matrix if we construct the D_L matrix using the lower bounds in the $A(x, 4, z)$ table [2]. In the $D_L(4, 6)$ example, the lower bounds in the $A(6, 4, 3)$ table can give four partition sets of all binary vectors of length 6, weight 3 and distance 4. In this example, we choose the first partition as matrix D_L in this example. However, we can also use any other partitions as matrix D_L . The matrices D_L and D_H are shown below.

$$D_L = \begin{bmatrix} 0\pi 00\pi\pi \\ 0\pi\pi\pi 00 \\ \pi 00\pi 0\pi \\ \pi 0\pi 0\pi 0 \end{bmatrix} \quad D_H = \begin{bmatrix} \pi\pi\pi 000 \\ \pi\pi\pi 000 \\ \pi\pi\pi 000 \\ \pi\pi\pi 000 \end{bmatrix}$$

In the above example, the relative contrast is $2/3$ in Method-2 construction. If we want better relative contrast, we can use other tables such as the lower bounds in $A(x, 6, z)$, $A(x, 8, z)$, and so on. Thus, these tables in [2] can provide larger columns to construct the Boolean matrix and better relative contrast resolution. This scheme offers the advantage of flexible improvement in relative contrast as needed. The more the blocklength m increases, the better the relative contrast we can obtain.

4.3 Comparison of the Three $(2, n)$ ASS Schemes

For clarity, we give a comparison of the two types of $(2, n)$ ASS schemes listed in Table 1 which include the DHQ and Method-2 $(2, n)$ ASS schemes. Concerning relative contrast, Table 2 shows a comparison between Method-1 and Method-2 $(2, n)$ ASS schemes, which Method-2 uses the lower bounds in the $A(x, 4, z)$, $A(x, 6, z)$ and $A(x, 8, z)$ tables, where n is at least 166.

Table 1. Comparisons of two $(2, n)$ ASS schemes.

Scheme	Cover sound Number	Decryption Effect	Relative Contrast Improvement
DHQ $(2, n)$ ASS scheme	$\lceil \log_2 n \rceil$	The more n increases, the worse the decryption performance.	No
Method-2 $(2, n)$ ASS scheme	1	It is not affected by size n .	Yes

Table 2. A comparison of relative contrast for $n = 166$.

Construction	n	m	h	l	$\alpha(m) = (h-l)/m$
Method-1 $(2, n)$ ASS scheme	166	166	166	164	Note: $\alpha(m) = 2/n \ 1/83$
Method-2 $(2, n)$ ASS scheme with $A(x, 4, z)$	166	13	13	9	Note: $\alpha(m) = 2/n \ 4/13$
$A(x, 6, z)$	166	17	17	11	Note: $\alpha(m) = 2/n \ 6/17$
$A(x, 8, z)$	166	20	20	12	Note: $\alpha(m) = 8/n \ 2/5$

5. SOME REMARKS ON THE IMPLEMENTATION RESULTS AND CONCLUSIONS

In the experiment, we chose one rock-type music as the cover music with T be 4 seconds, channel be stereo and sample rate be 44100 (hz). We also chose different repeated cycles to implement those two proposed $(2, n)$ ASS schemes. In fact, we will obtain interesting results with the proposed schemes from above two choices.

In the demonstrated examples of session 4.1, we used the repeated cycle (symbolized by r) equal to one in Figs. 1 and 2. This meant that each sub-piece was $T/3$ sec-

onds long in period T of one secret bit when the repeated cycle r was equal to one. If there are 10 repeated cycles of m -vector V in T seconds, the m -vector V will occur ten times in T seconds. Thus, each sub-slot will occupy only $T / 30$ seconds in the period of one secret bit.

In our implementation, we could obtain different acoustic effects from different repeated cycle values r . The listener could hear a smoother secret message in the range $r \in (1176, 2940)$ than in the range $r \geq 11760$. For a very large r value (≥ 11760), the listener can only hear some intermittent bits of rock music in the period of the secret bit "0" when two shares were played simultaneously, and the combined effect in the secret bit "1" is louder than the original cover sound. When the value r decreased, the sub-piece with the change of 180 phases lasted longer and the destructive time of two shares is longer. The magnetic field of the speaker pulled backwardly if the phase had a change of 180 phases. Therefore, for a small r value (≤ 588), the listener could hear noticeable chafing or vibration when each share was played or any two shares were played together.

The experimental results for different value of r can be found at the website <http://crypto.ee.ncku.edu.tw/~lchenchi/temp>. There are two directories in the website, and each directory holds some implementation files for each scheme. The file name has three parts: the file attribute, construction method and one decimal value representing the repeated cycle r .

In conclusion, the repeated cycles and relative contrast are two main factors that affect the decryption effect of our two constructions in our implementation. Therefore, it is important to experiment with a range of repeated cycles to obtain better decryption effects for the two proposed schemes. For example, in Method-1 (2, 3) ASS scheme, the listener could hear a better quality secret message according to the experimental results if $980 \leq r \leq 5880$ is taken.

REFERENCES

1. A. E. Brouwer, F. J. MacWilliams, A. M. Odlyzko, and N. J. A., Sloane, "Bounds for binary codes of length less than 25," *IEEE Transactions on Information Theory*, Vol. IT-24, 1978, pp. 81-93.
2. A. E. Brouwer, J. B. Shearer, N. J. A. Sloane, and W. D. Smith, "A new table of constant weight codes," *IEEE Transactions on Information Theory*, Vol. 36, 1990, pp. 1334-1380.
3. A. Shamir, "How to share a secret," *Communications of ACM*, Vol. 22, 1979, pp. 612-613.
4. Y. Desmedt, S. Hou, and J. Quisquater, "Audio and optical cryptography," in *Advances in Cryptology-Asiacrypt '98*, Springer-Verlag LNCS, pp. 392-404.
5. M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology-Eurocrypt '94*, Springer-Verlag LNCS, 1995, pp. 1-12.



Chen-Chi Lin (林禎吉) was born on May 22, 1965 in Penghu Hsien, Taiwan. He received the M.S. degree in 1990 from the Department of Electrical Engineer at National Sun Yat-Sen University. Now, he is pursuing Ph.D. degree in Electrical Engineering of National Cheng Kung University. Since 1990, he has worked as a professional teacher at Tainning Institute Kaohsiung Center, Chunghwa Telecom Co., Ltd. His research interests include Cryptology, Information Security, Public Key Infrastructure and Telecommunication Systems.



Chi-Sung Laih (賴溪松) was born on June 4, 1956 in Chiayi, Taiwan, Republic of China. He received his B.S., M.S. and Ph.D. degrees all in Electrical Engineering from National Cheng Kung University in 1984, 1986 and 1990, respectively. Since September 1986, he has been on the faculty of the Department of Electrical Engineering at National Cheng Kung University, Tainan, Taiwan, and currently is a professor. From August 1993 to January 1997, he was an adjunct research fellow at Engineering and Technology Promotion Center of the National Science Council of the Republic of China. Currently, he is the director of computer and network center. From February 1997, he was the director of Project Management, office of Research and Development at National Cheng Kung University. From June 1997, he was elected as the Chairman of Chinese Cryptology and Information Security Association (CCISA). His research interests include Cryptology, Information Security, Error Control Codes and Communication Systems. Dr. Laih is a member of IEEE, ACM and IACR. He was the winner of the 1991 and 1997 Acer Long Term Award for Outstanding M.S. Thesis Supervision, the winner of Graduate Team of TI-Taiwan 1994 DSP Design Championship and the winner of 1997 and 1999 Outstanding Paper Award and 1996 of CCISA. He also obtained the 1997-1998 and 1999-2000 Outstanding Research Award of the National Science Council of the Republic of China. He received 1999 Outstanding Talent Award in Information Science, Republic of China.



Ching-Nung Yang (楊慶隆) was born on May 9, 1961 in Kaohsiung, Taiwan. He received the B.S. degree in 1983 and the M.S. degree in 1985, both from the Department of Telecommunication Engineering at National Chiao-Tung University. He received Ph.D. degree in Electrical Engineering from National Cheng Kung University in 1997. During 1987-1989 and 1990-1999, he worked at Telecommunication Lab., and Tainning Institute Kaohsiung Center, Chunghwa Telecom Co., Ltd., respectively. He is presently an assistant professor in the Department of Computer Science and Information Engineering at National Dong Hwa University. His research interests include coding theory, Information security and cryptography.