World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:8, 2011

A Semi- One Time Pad Using Blind Source Separation for Speech Encryption

Long Jye Sheu, Horng-Shing Chiou and Wei Ching Chen

Open Science Index, Computer and Information Engineering Vol:5, No:8, 2011 publications.waset.org/6210.pdf

Abstract**—**We propose a new perspective on speech communication using blind source separation. The original speech is mixed with key signals which consist of the mixing matrix, chaotic signals and a random noise. However, parts of the keys (the mixing matrix and the random noise) are not necessary in decryption. In practice implement, one can encrypt the speech by changing the noise signal every time. Hence, the present scheme obtains the advantages of a One Time Pad encryption while avoiding its drawbacks in key exchange. It is demonstrated that the proposed scheme is immune against traditional attacks.

Keywords**—**one time pad, blind source separation, independent component analysis, speech encryption.

I. Introduction

As speech communications become more and more widely used and even more vulnerable, the importance of providing a high level of security is a major issue. Up to date, many speech encryption techniques have been proposed. In general, there are four categories of cryptographic algorithms widely used in speech communication, namely time domain [1], transform domain [2-4], amplitude scrambling, two-dimensional mixed scrambling methods [3, 4]. Recently, some new speech encryption methods including circulant transformation [5] and underdetermined blind source separation [6] have also been developed.

Blind source separation (BSS) techniques are applied to recover unknown signals or sources from their observed mixtures. If the number of the original sources is larger than that of the observed mixtures, there poses a significant difficulty of separation. The problem is called as underdetermined blind source separation (UBSS). However, the intractability of UBSS has motivated researchers to study whether it could replace other intractable problems (e.g. integer factorization) in the construction of cryptographic algorithms. Recently, Lin et al. [6] introduce the concept of UBSS for image and speech encryption. The quality of the decrypted speech/images is excellent. Unfortunately, these schemes have been found to be insecure against known-/chosen-plaintext attack and chosen-ciphertext attack [7].

A one-time pad [8] is a very simple yet completely unbreakable symmetric cipher. To use a one-time pad, you need two copies of a “pad” or key which is a block of truly random data. If the key is truly random, an XOR-based one-time pad is perfectly secure against ciphertext cryptanalysis. A pad is only used once and discarded, hence the name one-time pad. In a perfectly secure OTP cipher, the difficulty lies on the key exchange between the encryption and decryption.

In this letter, we propose a speech encryption using blind source separation. The original speech is mixed with key signals which consist of the mixing matrix, chaotic signals and a random noise. The present scheme obtains the advantages of a One Time Pad encryption while avoiding its drawbacks in key exchange. It is demonstrated that the proposed scheme to be immune against traditional attacks. The design also provides a new perspective toward secure communication.

II. Blind Source separation

The Blind source separations are techniques to recover n independent sources, $\mathbf{S}(t)=[s_1(t),s_2(t),..,s_n(t)]^T$, from their mixtures, $\mathbf{X}(t)=[x_1(t),x_2(t),..,x_m(t)]^T$, which are linear combination of the independent sources by an unknown matrix, **A**. For the sake of simplicity, we assume the number of mixed signals is the same as the number of independent sources, i.e. $m=n$. This is a simplifying assumption that is not completely necessary. Then, the mixed vector can be written as

$$\mathbf{X}(t)=\mathbf{AS}(t)\,; \tag{1}$$

where

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \tag{2}$$

is the mixing matrix giving the mixing weights. Matrix **A** is generally assumed to be unknown. The source signals, s_i, are unknown as well. When only mixed signals, $x_i(t)$, are known, the BSS algorithms are designed to separate the estimated independent sources, $\hat{\mathbf{S}}(t)=[\hat{s}_1(t),\hat{s}_2(t),...,\hat{s}_n(t)]^T$, such that

$$\hat{\mathbf{S}}(t)=\mathbf{BX}(t)=\mathbf{BAS}(t)\approx\mathbf{S}(t)\,, \tag{3}$$

Long-Jye Sheu is with the Dept. of Mechanical Engineering, Chung Hua University, HsinChu, Taiwan (corresponding author, e-mail: ljsheu@chu.edu.tw).

Horng-Shing Chiou is with the Dept. of Electrical Engineering, Technology and Science Institute of Northern Taiwan, Taipei, Taiwan (e-mail: hschiou@tsint.edu.tw).

Wei-Ching Chen is with the United Distribution Co. (email: wc137@hotmail.com)

Open Science Index, Computer and Information Engineering Vol:5, No:8, 2011 publications.waset.org/6210.pdf

where **B** is the demixing matrix,

$$\mathbf{B}=\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} \approx \mathbf{A}^{-1}. \qquad (4)$$

Many algorithms exist for calculating the demixing matrix, **B**. Among them, Independent component analysis (ICA) [9] is a faithful, easy and efficient method. In general, the estimated elements of matrix **B** differ from those of $\mathbf{A}^{-1}$. The components of $\hat{\mathbf{S}}(t)$ separated by the ICA reveal opposite phases and unequal amplitudes with the components of the original source, **S**(t).

III. Proposed Scheme

The core idea of this Letter is to utilize UBSS with partial One Time Pad for speech encryption. The block diagram of proposed BSS-based speech encryption is shown in Fig. 1.

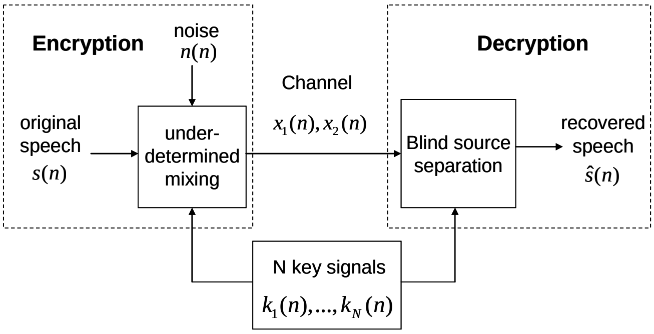

Fig. 1 Encryption scheme based on blind source separation

Encryption: the original speech $s(n)$ is mixed with a random noise $n(n)$ and N key signals $\mathbf{k}=[k_1(n),...,k_N(n)]^T$ into two mixed ciphertexts $\mathbf{x}=[x_1(n),x_2(n)]^T$ by using an underdetermined mixing matrix $\mathbf{A_e}$. Specifically, given source matrix $\mathbf{s}=[s(n), n(n), \mathbf{k}^T]^T$, the encryption can be represented by the following equation:

$$\mathbf{x} = \mathbf{A_e s} = [\mathbf{a_s} \ \mathbf{a_n} \ \mathbf{a_k}]\mathbf{s} \qquad (5)$$

where the $\mathbf{a_s} \ \mathbf{a_n} \ \mathbf{a_k}$ represent the weightings of original speech, random noise and key signals in the ciphertexts, respectively. Obviously, Eq. (1) constructs an UBSS problem since there are (N+2) signals in the two ciphertexts. The two ciphertexts are transmitted to the receiver through the public channel.

Decryption: To recover the original speech, at least (N+2) signals are required as inputs of the BSS in the decryption. Once the two decrypted ciphertexts are received from the public channel, the N key signals are regenerated by the secret seed I$_0$ to provide the rest of the N inputs of the BSS. Hence, the inputs of the BSS can be written as

$$\mathbf{x_d} = \begin{bmatrix} \mathbf{x} \\ \mathbf{k} \end{bmatrix} = \mathbf{A_d s} \qquad (6)$$

where

$$\mathbf{A_d} = \begin{bmatrix} \mathbf{A_e} \\ \mathbf{0} \ \mathbf{I} \end{bmatrix} \qquad (7)$$

In Eq. (3), **0** is a $N \times 2$ zero matrix, **I** is a $N \times N$ identity matrix. It is noted that $\mathbf{A_d}$ is a square matrix of full rank. When $\mathbf{x_d}$ is feed into the BSS, the independent component analysis technique is used to recover the estimate of the original speech $\hat{s}(n)$ as shown in Fig. 1.

Advantages of this scheme: The proposed design offers a higher level of security. We would like to discuss the role of noise signal n(n) in encryption. It is noted in Fig. 1 that n(n) is not a key signal since no information about the noise n(n) is necessary in the decryption. In encryption process, the noise signal n(n) plays the same role as the key signals $\mathbf{k}=[k_1(n),.......,k_N(n)]^T$ to interfere the original speech. One can encrypt the speech by changing n(n) every time. This means that n(n) could work like an one-time pad (OTP) cipher to provide the security of this scheme. A one-time pad [8] is a very simple yet completely unbreakable symmetric cipher where the same key is used for encryption and decryption of a message. To use an OTP, you need two copies of a “pad” or key which is a block of truly random data. If the signal is truly random, an OTP is perfectly secure against ciphertext cryptanalysis. A pad is only used once and discarded, hence the name OTP. In a perfectly secure OTP cipher, the difficulty lies on the key exchange between the encryption and decryption. In the present design, the information exchange of n(n) between encryption and decryption is not necessary. In fact, n(n) could be randomly generated to interfere the original speech and then deleted immediately after underdetermined mixing is constructed. Hence, n(n) is used in the present scheme to obtain the advantages of an OTP encryption while avoiding its drawbacks.

IV. Result

The next step is the evaluation of the scheme to encrypt and decrypt a speech. The simulation is to securely transmit a speech file recording a person saying the digits “one” to “ten” in English with a little audible music [10]. The speech signal is sampled at 16 kHz and 6.875 seconds long (totally 110000 samples) as shown in Fig. 2(a). In this study, two key signals (N=2) are generated to encrypt the speech. Here, the chaotic signals are used as the key signals. It is known that the chaotic signals are very sensitive to parameters and initial conditions. The random-like behavior of chaotic signals provides potential to mask the original signals. We choose two key signals from the the Chen-Lee system [11]:

Open Science Index, Computer and Information Engineering Vol:5, No:8, 2011 publications.waset.org/6210.pdf

$$\text{Chen-Lee system:}\quad \begin{aligned} \dot{z}_1 &= -z_2 z_3 + \alpha z_1 \\ \dot{z}_2 &= z_1 z_3 + \beta z_2 \\ \dot{z}_3 &= (1/3) z_1 z_2 + \gamma z_3 \end{aligned} \quad , \tag{8}$$

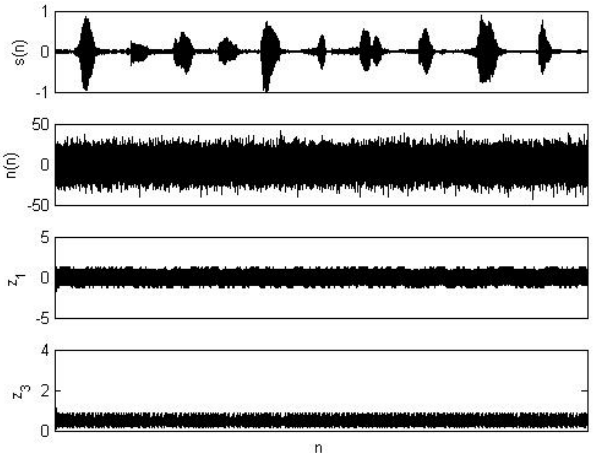

Fig. 2 Original signals

The parameters and initial conditions are chosen to be, $(\alpha, \beta, \gamma) = (5, -10, -3.8)$,and $[z_1(0), z_2(0), z_3(0)] = [0.2, 0.2, 0.2]$. We generate z$_1$ and z_3 with time intervals 0.01sec as the key signals. The random noise n(t) is generated by Marsaglia's Ziggurat pseudorandom number generator (PRNG). It is known that Marsaglia's Ziggurat PRNG generates Gaussian floating-point values with a very huge period , good statistical performance and fast speed. Figure 2b~2d shows the random noise and the key signals z$_1$, z$_3$, respectively.

The underdetermined mixing matrix used for simulation is

$$\mathbf{A_e} = \begin{bmatrix} 0.7342 & \text{-}0.8612 & 0.1014 & 0.7662 \\ 0.2995 & 0.9409 & 0.8779 & 0.2957 \end{bmatrix} \tag{9}$$

By using Eq. (5), two ciphertexts are deduced. Figure 3 shows the two ciphertexts and their power spectra. It is seen that the original speech has been well masked by the key signals and random noise. With the mixed sets of random noise and chaotic signals, the spectra of the ciphertexts are very plat which makes them almost indistinguishable from random noise during transmission.

In this study, the joint approximate diagonalization of eigenvalues for real signals (JADER) algorithm [12] was applied to calculate the demixing matrix, **B** as

$$\mathbf{B} = \begin{bmatrix} 0.0736 & 0.0677 & 0.0681 & \text{-}0.0034 \\ 0.1026 & \text{-}0.0122 & \text{-}0.0002 & \text{-}0.0061 \\ 0.0433 & 0.0408 & \text{-}0.0022 & \text{-}0.3033 \\ 12.4550 & 11.4001 & \text{-}0.5637 & \text{-}0.6486 \end{bmatrix} \tag{10}$$

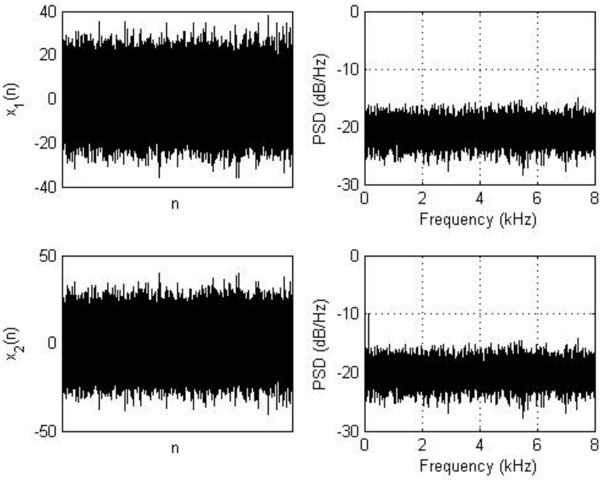

Fig. 3 Two encrypted signals

and the estimated original source signals are derived from Eq. (3). Figure 4 illustrates the recovery of the original speech using this scheme. It is clearly seen that faithful recovery of the message signal is possible. Comparing both waveforms between the original speech and the recovered speech, it can be seen that the original speech has been decrypted faithfully. Listening tests also indicate a very high quality decryption.

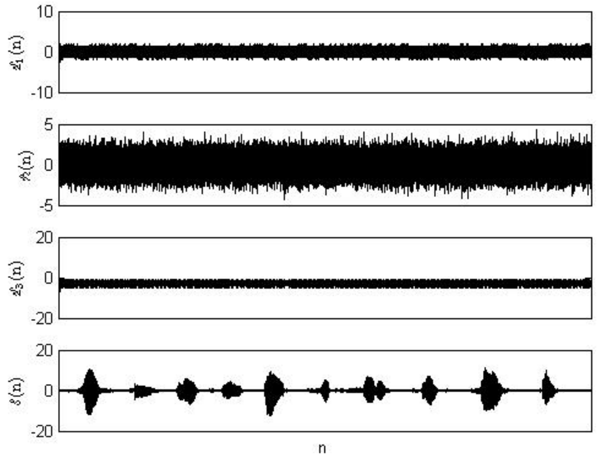

Fig. 4 Decrypted signals

V. Security Analysis

This section presents a cryptanalysis study of the new scheme. The analysis focuses on determination of the key space, key selection rules and analysis of security.

In the present scheme, the key signals, $z_1(t)$ and $z_3(t)$, are generated from the chaotic Chen-Lee system with the parameters (α, β, γ) and initial conditions $[z_1(0), z_2(0), z_3(0)]$. The parameters and initial conditions are the secret keys in this scheme. Hence, the secret key consists of six numbers $(\alpha, \beta, \gamma, z_1(0), z_2(0), z_3(0))$. Since these six numbers could be real numbers, the space of the keys will be a

Open Science Index, Computer and Information Engineering Vol:5, No:8, 2011 publications.waset.org/6210.pdf

6-dimensional space. The space is nonlinear since all of the keys are not equally strong. In the subspace where the fractional derivative orders or parameters of the fractional Lorenz system originate periodic orbits, the sub-key space is degenerative because it is relatively easy to break. Values of $(\alpha, \beta, \gamma, z_1(0), z_2(0), z_3(0))$ which give rise to periodic windows should be avoided since chaotic bands are preferred for encryption.

The key space for a encryption scheme should be large enough to resist the brute force attack. If the precision for each key is 10^{-10}, the key space size is 10^{60}. The key space size is large enough to resist all kinds of brute force attacks. In order to demonstrate the sensitivity of our communication system to keys, we consider an estimate of keys, say $(\alpha, \beta, \gamma) = (-5, -10, -3.8)$ and $(z_1(0), z_2(0), z_3(0)) = (0.2 + 10^{-10}, 0.2, 0.2)$ in which there is a slight mismatch with the real keys in $z_1(0)$. Fig. 5 shows the sensitivity of present secure communication scheme to slight mismatch of keys. It is noted that the recovered plaintexts with wrong keys is of random behavior and totally different from the original speech.

Three ordinary attacks on cryptosystems are (1) ciphertext-only attack; (2) known-plaintext attack; and (3) chosen-plaintext attack. If we encrypt the same plaintext twice, we will not have two identical ciphers since $n(t)$ is changed for each time. Hence, these attacks are not effective to break this scheme. This also provides the present scheme to have ciphertexts which are very sensitive to keys.

VI. Conclusions

In this paper, a new perspective on speech communication using blind source separation is proposed. The design of this scheme has many merits: (a) it provides advantages of a One Time Pad encryption while avoiding its drawbacks in key exchange; (b)The key space is large enough to resist all kinds of brute force attacks; (c) The ciphertexts are very sensitive to the secret and (d) it is immune against traditional attacks.

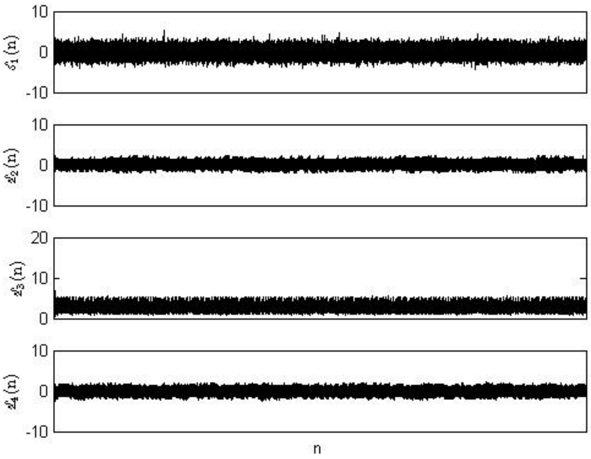

Fig. 5 Decrypted signals of present scheme with mismatch of keys.

References

[1] Grangetto, M., Magli, E., Olmo, G.: Multimedia selective encryption by means of randomized arithmetic coding IEEE Trans. Multimed. **8**, 905–917 (2006)

[2] Lee, J., Vijaykrishnan, N., Irwin, M.J., Chandramouli, R.: Block-based frequency scalable technique for efficient hierarchical coding. IEEE Trans. Signal Process. **54**, 2559– 2566 (2006)

[3] Mao, Y., Wu, M.: A joint signal processing and cryptographic approach to multimedia encryption. IEEE Trans. Image Process. **15**, 2061–2075 (2006)

[4] Ashtiyani, M., Behbahani, S., Asadi, S., Birgani, P.M.: Transmitting encrypted data by wavelet transform and neural network. In: IEEE Int. Symposium on Signal Processing and Information Technology, pp. 385–389 (2007)

[5] Manjunath, G., Anand, G.V.: Speech encryption using circulant transformations. In: IEEE Int. Conf. Multimedia and Expo., vol. 1, pp. 553– 556 (2002)

[6] Lin, Q.H., Yin, F.L., Mei, T.M., Liang, H.: A blind source separation based method for speech encryption. IEEE Trans. Circuits Syst. **53**, 1320– 1328 (2006)

[7] S. Li, C. Li, K. T. Lo and G. Chen, "Cryptanalyzing an encryption scheme based on blind source separation," IEEE Trans. Circuits Syst. 55, 1055-1062 (2008).

[8] B. Schneier, *Applied Cryptography: Protocols, Alorithms and Source Code in C.* Wiley Computer Publishing, John Wiley and Sons, Inc, 1996.

[9] C. Jutten and J. Herault, “Blind separation of sources, Part 1: an adaptive algorithm based on neuromimetic architecture,” Signal Processing, Vol. 24, pp. 1-10, 1991

[10] T. W. Lee, A. J. Bell, and R. H. Lambert, “ Blind separation of delayed and convolved sources,” Adv. Neural Inf. Process. Syst., vol. 9, pp. 758–764, 1996.

[11] Chen HK, Lee CI. Anti-control of chaos in rigid body motion. Chaos, Solitons & Fractals 2004;21:957–65.

[12] .F. Cardoso and A. Souloumiac, “Blind beamforming for non-Gaussian signals,” IEE Proceedings F, Vol. 140, pp. 771-774, 1993.