Research Article
Dual Key Speech Encryption Algorithm Based Underdetermined BSS

Huan Zhao,1 Shaofang He,1,2 Zuo Chen,1 and Xixiang Zhang1

1 School of Information Science and Technology, Hunan University, Changsha, Hunan 410082, China
2 Science College, Hunan Agricultural University, Changsha, Hunan 410128, China

Correspondence should be addressed to Zuo Chen; chenzuo@hnu.edu.cn

Received 10 March 2014; Accepted 27 April 2014; Published 14 May 2014

Academic Editor: Fei Yu

Copyright © 2014 Huan Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

When the number of the mixed signals is less than that of the source signals, the underdetermined blind source separation (BSS) is a significant difficult problem. Due to the fact that the great amount data of speech communications and real-time communication has been required, we utilize the intractability of the underdetermined BSS problem to present a dual key speech encryption method. The original speech is mixed with dual key signals which consist of random key signals (one-time pad) generated by secret seed and chaotic signals generated from chaotic system. In the decryption process, approximate calculation is used to recover the original speech signals. The proposed algorithm for speech signals encryption can resist traditional attacks against the encryption system, and owing to approximate calculation, decryption becomes faster and more accurate. It is demonstrated that the proposed method has high level of security and can recover the original signals quickly and efficiently yet maintaining excellent audio quality.

1. Introduction

As speech communications in our daily life become more and more common, the importance of providing a high level of security is sharply increasing. For that reason, a series of speech encryption methods have been proposed. Among which, the analogue encryption is one of the most popular encryption techniques widely used in speech communication. Generally, there are four categories of cryptographic algorithms in speech communication: frequency-domain scrambling (e.g., the frequency inverter and the band splitter), time-domain scrambling (e.g., the time element scrambling), amplitude scrambling (also known as the masking technique that covers the speech signal by the linear addition of pseudorandom amplitudes), and two-dimensional scrambling that combines the frequency-domain scrambling with the time-domain scrambling [1, 2]. In addition, there are many other analogue speech encryption algorithms in the transform domain, for example, discrete cosine transform, fast Fourier transform, wavelet transform, and so forth [3–5]. Up to date, many new speech encryption algorithms including BSS-based [2, 6], chaotic cryptosystem [7–10], and encryption using circulant transformations [11] have been developed. Due to the fact that the great amount data of speech communications and real-time communication has been required, it is not suitable to utilize traditional encryption methods directly for speech communication encryption. As such, to explore speech encryption methods that have a high level of security and efficiency and high speed in decryption while retaining excellent audio quality is an urgent issue.

Blind source separation (BSS) is used to recover unknown source or signals that are independent mutually of their observed mixtures without knowing the mixing coefficients. So, it is also known as independent component analysis (ICA). Recently, signals encryption has been more applied on the image cryptosystems [12–14] but less on the speech encryption. Speech encryption method based BSS, of which the security dependent on the difficulty of solving the underdetermined BSS problem where the number of the observed mixed signals is less than that of the source signals. The sufficient condition for constructing the underdetermined mixing matrix for encryption is presented based on the source inseparability of BSS [15].

One-time pad [16] is a simple and completely unbreakable symmetric cipher, and it has two perfect characteristics: the key is random and has the same length as the message.
The key space is large enough to resist brute-force attacks as long as the key is long enough. So, the message is secure as long as the key is protected. Motivated by the randomness and initial conditions’ sensitivity of chaotic signals, we present an underdetermined BSS-based dual key speech encryption scheme in this paper. The dual key of which are random signals (one-time pad) and chaotic signals, namely, key signals I and II, respectively. The main purpose of the algorithm is to mask the original speech signals by mixing the original speech with key signals I and II. In the decryption process, approximate calculation method is used to recover the original speech signals. The underdetermined blind source separation is a significant challenge in blind source separation (BSS) where the number of the source signals is greater than that of the mixed signals. In addition, the using of the key signals I and II (one-time pad and chaotic signal) ensures high security of the algorithm. Both extensive computer simulations and performance analysis results show that the proposed method has high level of security while retaining excellent audio quality.

The rest of this paper is organized as follows. In Section 2, firstly, we introduce the BSS mixing model and the underdetermined BSS problem briefly, secondly, the details of speech encryption and decryption are described, and finally, we analyze the feasibility of approximate calculation in the decryption process. Sections 3 and 4 conduct computer simulations to illustrate and analyze the performance of the method. We conclude this paper in Section 5.

2. Proposed Method

2.1. BSS Mixing Model and Underdetermined Problem [2]. Suppose that \(s_1(t), s_2(t), \ldots, s_M(t)\) is \(M\) independent source signals and \(N\) observed mixtures of the source signals are \(x_1(t), x_2(t), \ldots, x_N(t)\) \((M < N)\). The linear BSS mixing model is represented as follows:

\[
x(t) = As(t),
\]

where \(s(t) = [s_1(t), s_2(t), \ldots, s_M(t)]^T\), which is \(M \times 1\) column vector collected from the source signals, similarly, \(N \times 1\) column vector \(x(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^T\) collects the observed signals, and \(A\) is an \(N \times M\) mixing matrix that contains the mixing coefficients. The aim of BSS is to find a \(M \times N\) demixing matrix \(W\) such that output vector:

\[
u(t) = Wx(t) = WAs(t) = PDs(t),
\]

where \(P \in R^{M \times M}\) is a permutation matrix and \(D \in R^{M \times M}\) is a diagonal scaling matrix. When the number of the mixed signals is less than that of the source signals; that is, \(M > N\), BSS becomes a difficult case of the underdetermined problem, in which the complete separation of the source signals is impossible.

2.2. Encryption. The main idea of the proposed algorithm is to construct the intractable underdetermined BSS problem in encryption, and in decryption it can only be solved with the dual key. The block diagram of the underdetermined BSS-based speech encryption scheme is shown in Figure 1.

Two main steps in the encryption process are the segment splitter and the underdetermined mixing. Suppose that the original speech is divided into frames, and every frame is encrypted, respectively; \(q\) is the frame pointer. The frame \(q\) is encrypted as follows.

1. Segment splitter: the segment splitter first partitions the frame \(q\) into \(P\) segments \(s_1(t), s_2(t), \ldots, s_p(t), t = 1, \ldots, T\), where \(T\) is the segment length.

2. Underdetermined mixing: the source signals are composed of three parts, they are original speech signals, key signals I generated by pseudorandom number generator (PRNG) with secret seed \(I_0\), and key signals II from chaotic system. \(s(t) = [s_1(t), s_2(t), \ldots, s_p(t)]^T\) denotes original speech, key signals I are \(k(t) = [k_1(t), k_2(t), \ldots, k_p(t)]^T\), and \(h(t) = [h_1(t), h_2(t), \ldots, h_p(t)]^T\) is key signals II. Therefore, \(3p \times 1\) column vector of the source signals is \([s^T(t), k^T(t), h^T(t)]^T\). A \(p \times 3p\) underdetermined mixing matrix \(A_e = [B \ aB \ \beta B]\) for encryption is first generated randomly, where \(B\) is a \(P \times P\) matrix of full rank, which is pseudorandomly generated with normal distribution between \(-1\) and \(1\). \(\alpha, \beta \ll 2\) are scalar values to make the original speech be covered well by the dual key signals. The encryption equation can be represented as follows:

\[
x(t) = A_e[s^T(t), k^T(t), h^T(t)]^T
\]

\[
= [B \ aB \ \beta B]\begin{pmatrix} s(t) \ k(t) \ h(t) \end{pmatrix}
\]

\[
= Bs(t) + aBk(t) + \beta Bh(t),
\]

where \(x(t) = [x_1(t), x_2(t), \ldots, x_p(t)]^T\) is the observed signals. (Parameters \(P, T, I_0,\) initial condition of chaotic system, and scalar \(\alpha, \beta\) are inserted into the head data of the encryption speech in a definite format for transmission.)

2.3. Decryption. Once the mixture signals \(x(t) = [x_1(t), x_2(t), \ldots, x_p(t)]^T\) are received, the key signals I are regenerated by the secret seed \(I_0\) and the key signals II are produced by the chaotic system using the initial conditions. Usually, BSS is then performed [2,17] to recover original signals. But in this paper, we employ approximate calculation to recover original signals.

2.3.1. The Approximate Calculation for Decryption. Multiply \(k^T(t)\) at both sides of (3), and we get equation:

\[
x(t)k^T(t) = Bs(t)k^T(t) + aBk(t)k^T(t) + \beta Bh(t)k^T(t).
\]
Similarly, multiply \( h^T(t) \) at both sides of (3), and then get equation
\[
x(t) h^T(t) = B s(t) h^T(t) + \alpha Bk(t) h^T(t) + \beta B h(t) h^T(t)
\]
denoted by
\[
R_{sk} = x(t) k^T(t), \quad R_{kk} = s(t) k^T(t),
R_{sh} = h(t) k^T(t), \quad R_{hh} = s(t) h^T(t),
R_{kh} = k(t) h^T(t), \quad R_{hh} = h(t) h^T(t).
\] (6)
Equations (4) and (5) can be represented by (7) and (8), respectively,
\[
R_{sk} = B R_{kk} + \alpha B R_{kh} + \beta B R_{hh},
R_{sh} = B R_{sh} + \alpha B R_{kh} + \beta B R_{hh}.
\] (7)
(8)
Since the original speech signals are independent statistically of the key signals I and II, we have \( R_{sk} \ll R_{kk}, R_{sh} \ll R_{hh} \), that is, \( BR_{sk} \ll \alpha B R_{kk}, BR_{sh} \ll \beta B R_{hh} \), \( 1 \ll \alpha, \beta \ll 2 \); therefore, (7) and (8) are represented approximately by (9) and (10):
\[
R_{sk} \approx \alpha B R_{kk} + \beta B R_{hh},
R_{sh} \approx \alpha B R_{kh} + \beta B R_{hh}.
\] (9)
(10)
Plus (9) and (10), we can get equation:
\[
R_{sk} + R_{sh} \approx B (\alpha R_{kk} + \beta R_{hh} + \alpha R_{kh} + \beta R_{hh} ).
\] (11)
So we obtain an estimate for \( B \) as follows:
\[
\hat{B} = (R_{sk} + R_{sh}) (\alpha R_{kk} + \beta R_{hh} + \alpha R_{kh} + \beta R_{hh})^{-1}.
\] (12)
Substituting \( B \) in (3) with (11), original signals \( s(t) \) can be estimated as
\[
\hat{s}(t) = (\alpha R_{kk} + \beta R_{hh} + \alpha R_{kh} + \beta R_{hh})^{-1} x(t) - \alpha k(t) - \beta h(t).
\] (13)

2.3.2. The Steps of Calculating Original Signals for Decryption.
Parameters \( P, T \) and scalar \( \alpha, \beta \) are transmitted together with the secret seed \( I_0 \) and initial condition of chaotic system in the head data of the encryption speech. The original speech signals can be decrypted with very high quality by employing the decryption equation (13). Supposing that the mixed signals are received and the double key signals are regenerated, the original signals can be calculated as follows:

(a) calculate \( R_{sk}, R_{kk}, R_{sh}, R_{kh}, R_{hh} \) respectively;
(b) calculate \( (\alpha R_{kk} + \beta R_{hh} + \alpha R_{kh} + \beta R_{hh}) \) and \( (R_{sk} + R_{sh})^{-1} \);
(c) calculate \( \hat{s}(t) \) using (13).

2.3.3. Analysis of the Approximate Calculation for Decryption.
Key signals I and II are generated by pseudorandom number generator (PRNG) and chaotic system, respectively. They are both independent statistically of the original signals. For illustrating the feasibility of the approximate calculation, we compute values of \( R_{sk}, R_{kk}, R_{sh}, R_{kh}, R_{hh} \) in the example. The original signals of digital "1" in English are regarded as a frame, which is divided into two segments, that is, \( P = 2, T = 16000 \), the original signals \( s(t) = [s_1(t), s_2(t)]^T \), correspondingly key signals I and II are \( k(t) = [k_1(t), k_2(t)]^T \), \( h(t) = [h_1(t), h_2(t)]^T \). The key signals I and II are generated by PRNG with the secret seed \( I_0 = p \times T = 32000 \), and choosing initial condition \( (a, b, c) = (35, 3, 28), [h(1), y(1), z(1)] = [\ldots] \),...
Finally, we calculate the values of $R_k$, $R_{kk}$, $R_{sh}$, and get the average values of diagonal elements value of $R_k = s(t)k^T(t)$, $R_{kk} = k(t)k^T(t)$, $R_{sh} = s(t)h^T(t)$, and $R_{hh} = h(t)h^T(t)$ are $1.3775, 5365.35, −0.5212$, and $876.29$, respectively.

In the example, the original signals are also split into other different numbers of segments, which are $p = 4, 8$, corresponding to $T = 8000, 4000$, use the same key signals I and II, and compute $R_k$, $R_{kk}$, $R_{sh}$, and $R_{hh}$, respectively. Table 1 shows the results of the average value of diagonal elements and upper-triangular for comparison of $R_k$, $R_{kk}$, $R_{sh}$, and $R_{hh}$ in three different cases.

From Table 1 we can see that the average values of diagonal elements of $R_k$ are considerably much smaller than those of $R_{kk}$, and the average values of upper-triangular elements of $R_{sh}$ are also much smaller than those of $R_{hh}$, and obviously, $R_k \ll R_{kk}$, $R_{sh} \ll R_{hh}$, therefore, using approximation calculation for decryption is feasible, which means that the decryption method in this paper not only have the characteristic of computing simply and quickly but also maintaining excellent audio quality.

### 3. Computer Simulations

In order to illustrate the feasibility of the proposed scheme, we carry out extensive computer simulation. In common experiments, recorded audio files in wave format are adopted and transmitted within local area network. In our experiment, we use the speech file recording a man saying the digit “1” in English. The speech signals are sampled at 16 KHz, as shown in Figure 2(a). For the purpose of simulating encryption process, we regard the speech signals as one frame directly and separate them into two segments; that is, $p = 2, T = 16000$, as shown in Figure 2(a). The key signals I are generated by PRNG with the secret seed $I_0 = p \times T = 32000$, and Figure 2(b) is the split wave of key signals I. Choosing initial condition $(a, b, c) = (35, 3, 28)$ and $[h(1), y(1), z(1)] = [0, 1.001, 0]$, the key signals II $(h(t)/40)$ are generated by Chen-Lee chaotic system [18]:

$$\dot{h} = −yz + ah,$$

and the split wave of which is showed in Figure 2(c). Choosing $α = β = 2$, the underdetermined mixing matrix $A_e = [B \; αB \; βB]$ used for simulation is

$$A_e = \begin{pmatrix}
0.9501 & 0.6068 & 1.9003 & 1.2137 & 1.9003 & 1.2137 \\
0.2311 & 0.4860 & 0.4623 & 0.9720 & 0.4623 & 0.9720
\end{pmatrix}$$

using (3), and two cipher texts are deduced quickly. Figure 2(d) shows the two cipher-text segments. Obviously, the original speech signals are well covered with the mixed sets of key signals I and II. In the decryption process, the mixed signals are received and the double key signals are regenerated; we can regain the original speech signals according to the steps of approximate calculation decryption method. The recovered signals are showed in Figure 2(e).

### 4. Performance Analysis

#### 4.1. Signal-to-Noise Ratio Computation. For the purpose of quantifying the performance of the proposed method, we calculated the signal-to-noise ratio (SNR) index of original signals segments in each encrypted signal segments and decrypted signals segments. Particularly, the SNR index of original segments in the decrypted segments is represented as follows:

$$\text{SNR (dB)} = 10 \log \frac{\sum_{t=0}^{T-1} s^2(t)}{\sum_{t=0}^{T-1} \tilde{s}^2(t)},$$

where $s(t)$ is original signals and $\tilde{s}(t)$ is decrypted original signals that are calculated by the approximate calculation method. If $\tilde{s}(t)$ is replaced by $x(t)$, which denotes encrypted signals, we can obtain the SNR index of encrypted signals. Employing the data in computer simulations, we can get SNR of two original signals segments in two encrypted segments and two decrypted signals segments. Table 2 shows the results.

These SNR indexes show that in the encrypted segments dual key signals have well masked the original segments, and the original signals in the decrypted segments that are recovered by approximate calculation method have excellent quality.
Figure 2: (a) The original speech signals and two segments $s_1(t), s_2(t)$; (b) two segments of key signals $k_1(t), k_2(t)$; (c) two segments of key signals $h_1(t), h_2(t)$; (d) two encrypted segments $x_1(t), x_2(t)$; (e) two recovered segments by approximate calculation $\hat{s}_1(t), \hat{s}_2(t)$.

Table 2: SNR (dB) of two original signals segments in two encrypted segments and two decrypted signals segments.

| Original signals segments | Encrypted signals segments | Decrypted signals segments |
|--------------------------|---------------------------|---------------------------|
| $s_1$                    | $x_1$                     | $\hat{s}_1$              |
|                          | 81.4021                   | -94.595                   |
| $s_2$                    | $x_2$                     | $\hat{s}_2$              |
|                          | 105.586                   | 25.2364                   |


4.2. Security Analysis. We take advantage of the underdetermined BSS problem to propose a dual key encryption algorithm in this paper. There are three aspects to ensure the security of the algorithm. Firstly, the intractability of the underdetermined BSS problem can be ensured by the mixing matrix for encryption. Secondly, the key signals I that as long as the original speech signals have the prefect property of the one-time pad cipher, which is statistically independent and non-Gaussian characteristics.

Finally the key signals II are generated from the Chen-Lee chaotic system that has the characteristic of power randomness and high sensitivity to initial condition. In order to illustrate the sensitivity of the proposed encryption algorithm to secret keys, we choose an estimate of the initial condition that is $(a, b, c) = (35, -3, 28), [h(1), y(1), z(1)] = [0.001, 1, 0.001]$, in which there is a slight mismatch with the real initial condition and use the same key signals I with secret seed $I_b = 32000$. Figure 3 shows two segments of the recovered signals utilizing the approximate calculation for decryption. Obviously, the recovered signals with the wrong secret key are totally different from the original speech signals. In short, the proposed method is sensitive to secret keys and immune against the ordinary attacks on cryptosystems, for example, cipher-text-only attack, known-plaintext attack, chosen-plaintext attack, and brute-force attract.

5. Conclusions

In this paper, we proposed a new dual key encryption scheme based on the underdetermined BSS problem. Since the mixing matrix for encryption ensures the intractability of the underdetermined BSS problem, and the key signals I approximately have the prefect property of the one-time pad cipher, and the key signals II are generated from the chaotic system that has the characteristic of power randomness and high sensitivity to initial condition. In the decryption process, using approximate calculation decryption method can recover the original signals quickly and efficiently yet maintaining high level audio quality. The design of this encryption method has five merits: (1) it is impossible to recover the original signals without the parameters in the head data of the encryption speech signals; (2) the approximate calculation method used in decryption process ensures the recovery of the original signals efficiently yet maintaining excellent speech quality; (3) the key signals I approximately has the prefect property of the one-time pad cipher, in which the length of the key is the same as the original speech signals; hence, the space of the keys is so large that all brute-force attacks against the system are infeasible; (4) the key signals II generated from chaotic system that provides the present scheme having the property of cipher text are very sensitive to secret keys, and (5) it can resist all kinds of traditional attacks against cryptosystems.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was sponsored by the National Nature Science Foundation of China (6117306), Specialized Research Fund for the Doctoral Program of Higher Education of China (2010016120021), the Natural Science Foundation of Hunan Province of China (14JJ3062), and Young Teacher’s Growth Program of Hunan University.

References

[1] H. J. Beker and F. C. Piper, Secure Speech Communications, Academic Press, London, UK, 1985.
[2] Q.-H. Lin, F.-L. Yin, T.-M. Mei, and H. Liang, “A blind source separation based method for speech encryption,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 6, pp. 1320–1328, 2006.
[3] B. Goldburg, S. Sridharan, and E. Dawson, “Design and cryptanalysis of transform-based analog speech scramblers,” *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 5, pp. 735–744, 1993.
[4] A. Matsunaga, K. Koga, and M. Ohkawa, “Analog speech scrambling system using the FFT technique with high-level security,” *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 4, pp. 540–547, 1989.
[5] E.-L. Ma, J. Chen, and Y.-M. Wang, “Wavelet transform-based analogue speech scrambling scheme,” *Electronics Letters*, vol. 32, no. 8, pp. 719–721, 1996.
[6] L.-J. Sheu, H.-S. Chiu, and W.-C. Chen, “A semi-one time pad using blind source separation for speech encryption,” *World Academy of Science, Engineering and Technology*, vol. 5, no. 8, 2011.
[7] K. Li, Y. C. Soh, and Z. G. Li, “Chaotic cryptosystem with high sensitivity to parameter mismatch,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, no. 4, pp. 579–583, 2003.
[8] S. M. H. Alwahbani and E. B. M. Bashier, “Speech scrambling based on chaotic maps and one time pad,” in *Proceedings of the International Conference on Computing, Electrical and Electronics Engineering (ICCEE’13)*, pp. 128–133, August 2013.
[9] L. J. Sheu, "A speech encryption using fractional chaotic systems," Nonlinear Dynamics, vol. 65, no. 1-2, pp. 103–108, 2011.
[10] Y. P. Zhang, F. Duan, and X. Liu, "The research of applying chaos theory to speech communicating encryption system," in Advances in Multimedia, Software Engineering and Computing Vol.2, vol. 129 of Advances in Intelligent and Soft Computing, pp. 197–202, Springer, Berlin, Germany, 2011.
[11] G. Manjunath and G. V. Anand, "Speech encryption using circulant transformations," in Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '02), vol. 1, pp. 553–556, 2002.
[12] W. Kasprzak and A. Cichocki, “Hidden image separation from incomplete image mixtures by independent component analysis,” in Proceedings of the 13th International Conference on Pattern Recognition, vol. 2, pp. 394–398, Vienna, Austria, 1996.
[13] C.-C. Chang, M.-S. Hwang, and T.-S. Chen, "A new encryption algorithm for image cryptosystems," Journal of Systems and Software, vol. 58, no. 2, pp. 83–91, 2001.
[14] K. D. Rao, K. P. Kumar, and P. V. M. Krishna, "A new and secure cryptosystem for image encryption and decryption," IETE Journal of Research, vol. 57, no. 2, pp. 165–171, 2011.
[15] X.-R. Cao and R. Liu, “General approach to blind source separation,” IEEE Transactions on Signal Processing, vol. 44, no. 3, pp. 562–571, 1996.
[16] B. Schneier, Applied Cryptography: Protocols, Algorithms and Source Code in C, John Wiley & Sons, 1996.
[17] D.-P. Guo and Q.-H. Lin, “Fast decryption utilizing correlation calculation for BSS-based speech encryption system,” in Proceedings of the 6th International Conference on Natural Computation (ICNC '10), vol. 3, pp. 1428–1432, August 2010.
[18] H.-K. Chen and C.-I. Lee, “Anti-control of chaos in rigid body motion,” Chaos, Solitons & Fractals, vol. 21, no. 4, pp. 957–965, 2004.