Hierarchical Selective Encryption for G.729 Speech Based on Bit Sensitivity

Zhaopin Su1,2, Jianguo Jiang1, Shiguang Lian3, Guofu Zhang1, Donghui Hu1
1School of Computer and Information, Hefei University of Technology, China
2Postdoctoral Research Station for Management Science and Engineering, Hefei University of Technology, China
3France Telecom R&D Beijing, Beijing, China
{szp, jgjiang}@hfut.edu.cn, shiguang.lian@ieee.org, {zgf, hudh}@hfut.edu.cn

Abstract

Speech encryption is becoming increasingly important as the increasing relevance of multimedia applications such as video conferencing and news broadcasting. However, multimedia encryption and decryption are often computationally demanding and impractical for power-constrained devices and narrow bandwidth environments. It is a challenge to obtain a good tradeoff between high security and low computational cost. To tackle this situation, a group of hierarchical selective encryption schemes are presented based on the computational complexity analysis of G.729 standard speech in this paper. In the proposed schemes, speech bit streams are partitioned into two parts according to the bit sensitivity, i.e., one is encrypted by a strong cipher and the other by a lightweight cipher. The contrast experiments are addressed in signal inspection and objective distortion measure. Additionally, we discuss the schemes’ cryptanalysis is investigated. The results demonstrate the effectiveness of the hierarchical selective encryption schemes for power-constrained devices and narrow bandwidth environments.

Keywords: Hierarchical selective encryption, G.729 speech, Bit sensitivity, Chaotic map.

1 Introduction

Speech encryption was first proposed to satisfy the demand of military, and always plays an important role in military use [1-2]. Due to the increasing relevance of multimedia applications in people’s daily lives, content protection and customer privacy are becoming more and more important. Since the encryption can effectively prevent eavesdropping, its use is widely advocated in many areas [3-8].

Early popular encryption techniques focused on analogue signal [9] and contained four main categories: frequency domain scrambling, time domain scrambling, two-dimensional scrambling combining the frequency domain scrambling with the time domain scrambling, and amplitude scrambling. Sridharan et al. [10-12] used the discrete Fourier transform, discrete cosine transform, Walsh-Hadamard transform, Karhunen-Loeve transform, and discrete prolate spheroidal transform to develop analogue speech encryption techniques. Lin et al. [13] proposed an encryption approach integrating a modified time domain scrambling scheme with an amplitude scrambling method, which masks the speech signal with a random noise by specific mixing. Andrade et al. [14] presented a two-dimensional scrambling method combining the frequency domain scrambling with the time domain scrambling for AMR (Adaptive Multi-Rate) speech. These analogue speech encryption techniques are simple, practical and of high quality. Besides, the bandwidth occupied is very small. However, these techniques do not change the redundancy of speech greatly, which lead to the intelligibility of the encrypted analog signal. And thus, analogue encryption has poor security [15-16].

With the development of digital signal transmission, speech can be transmitted in digital forms, and some digital speech encryption schemes are reported, which can be divided into two categories, i.e., full encryption [17-20] and selective encryption [21-23].

For full encryption, at the transmitter, multimedia signals (such as audio, image, video and so on) as input signals are encrypted in its entirety. After transmitted, multimedia signals are decrypted completely at the receiver. Full encryption can offer a high level of security, effectively prevent unauthorized access, and is widely used nowadays, such as, Fan et al. [18] viewed audio signal as ordinary binary bit streams and proposed an encryption method based on cat map and logistic map. Khan et al. [19] used Hidden Markov chain to analyze the speech recognition in case of fully encryption. However, full encryption decryption are computationally demanding, and not effectively applied to mobile and portable devices [20], where power consumption needs to be reduced as much as possible. So, the complexity of encryption and decryption algorithms should be reduced for power-constrained, real-time multimedia applications.

For selective encryption, only a subset of multimedia signals is subject to encryption instead of entire signals. At the transmitter, input signals are partitioned into two parts according to a certain strategy. One part is encrypted, and the other is unprotected. For selective encryption, the
difficulty or key pot is how to select the sensitive part to be encrypted. It is in close relation with the scheme’s security and efficiency. Selective encryption was first proposed for images and videos, and many algorithms have been developed. For example, efficient encryption algorithms for images and videos are proposed in [21-23] and [23-25], respectively.

However, the research for speech or audio is lacking. For example, Gnanajeyaraman et al. [26] proposed an audio encryption algorithm based on higher dimensional chaotic map, which can obtain much higher security. But their encryption is full encryption, that is, all bits are subject to encryption, whose computational complexity is very high. Lan et al. [27] proposed a perception-based scalable encryption approach for AVS (Audio Video coding Standard) audio, which specifies different encryption modes by utilizing the perception classification of audio stream and provides multiple security levels by encrypting different audio coding layers. However, in their each mode (except Full encryption mode), bits of higher perception classification are encrypted, and the others are unprotected. Servetti et al. [28] proposed two partial encryption schemes of G.729 speech, and partitioned bits into two classes based on perception. Class 1, the mostly perceptually relevant bits, are to be encrypted, and Class 2, the other bits, are to be left unprotected. For these encryption approaches, there are still remaining some comprehensible bit streams structures which might leak some information for attackers and reduce security to an extent.

In most multimedia applications, speech and audio are essential, and their service is the basis in telephony industry, video conference and news broadcasting. Thus, the encryption of speech or audio is very significant. Since, the existing algorithms can not obtain a good tradeoff between security and computational cost, better encryption schemes are expected.

In this paper, novel hierarchical selective encryption schemes for speech are presented based on the widely used speech coding algorithm, i.e., the ITU-T G.729 standard [29] that has been widely used in video conference, multimedia communications, personal mobile communications, and so on.

The rest of the paper is organized as follows. In Section 2, some related works are presented, including the G.729 speech standard, and its bit sensitivity. In Section 3, the existing typical G.729 speech encryption algorithm is evaluated. The hierarchical encryption scheme is presented in Section 4. In Section 5, the schemes’ parameters are selected by theoretical analysis and signal inspection experiments. In Section 6, performances of the proposed schemes are evaluated in detail. Finally, the conclusions are drawn and future work is given in Section 7.

2 Related Work

In this Section, we will introduce G.729 speech standard, and analyze its bit sensitivity.

2.1 G.729 Speech

The ITU-T G.729 standard [29-30], using algebraic code-excited linear-prediction technique, is a widely-used speech coding algorithm that compresses voice audio in packets of 10 milliseconds duration and has been mostly used in Voice over IP (VoIP) applications for its low bandwidth requirement. G.729 provides toll quality at 8 kb/s, but there are extensions, which provide rates of 6.4 kb/s and 11.8 kb/s for marginally worse and better speech quality respectively.

A frame of G.729 is shown in Figure 1, and the corresponding bit allocation is shown in Table 1. Each frame size is 10ms (80 samples).

![Figure 1 A Frame of G.729](image)

2.2 Bit Sensitivity of G.729

Swaminathan et al. [31] categorized the G.729 codec’s bit streams into four groups:

1. LSP parameters: L0, L1, L2, L3.
2. Pitch parameters: P1, P2.
3. Gain parameters: GA1, GB1, GA2, GB2.
4. Codebook parameters: C1, S1, C2, S2.

In each group, bit errors produce a different type of distortion. Bit sensitivity measurement for bits in each group was carried out using segSNR (segmental signal-to-noise ratio) [32] with the unimpaired speech as the reference. The measurement prioritized the bits within each category, and identified at least 36 bits that are sensitive for people perceptivity. The 36 bits are tabulated in Table 2.
Hierarchical Selective Encryption for G.729 Speech Based on Bit Sensitivities

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0</td>
<td>Switch MA predictor index of LSP quantizer</td>
<td>1</td>
</tr>
<tr>
<td>L1</td>
<td>1st stage vector of LSP quantizer</td>
<td>7</td>
</tr>
<tr>
<td>L2</td>
<td>2nd stage lower vector of LSP quantizer</td>
<td>5</td>
</tr>
<tr>
<td>L3</td>
<td>2nd stage higher vector of LSP quantizer</td>
<td>5</td>
</tr>
<tr>
<td>P1</td>
<td>Pitch delay 1st subframe</td>
<td>8</td>
</tr>
<tr>
<td>P0</td>
<td>Parity bit for pitch delay</td>
<td>1</td>
</tr>
<tr>
<td>S1</td>
<td>Signs of fixed-codebook pulses 1st subframe</td>
<td>4</td>
</tr>
<tr>
<td>C1</td>
<td>Fixed codebook 1st subframe</td>
<td>13</td>
</tr>
<tr>
<td>GA1</td>
<td>Gain codebook (stage 1) 1st subframe</td>
<td>3</td>
</tr>
<tr>
<td>GB1</td>
<td>Gain codebook (stage 2) 1st subframe</td>
<td>4</td>
</tr>
<tr>
<td>P2</td>
<td>Pitch delay 2nd subframe</td>
<td>5</td>
</tr>
<tr>
<td>S2</td>
<td>Signs of fixed-codebook pulses 2nd subframe</td>
<td>4</td>
</tr>
<tr>
<td>C2</td>
<td>Fixed codebook 2nd subframe</td>
<td>13</td>
</tr>
<tr>
<td>GA</td>
<td>Gain codebook (stage 1) 1st subframe</td>
<td>3</td>
</tr>
<tr>
<td>GB2</td>
<td>Gain codebook (stage 2) 1st subframe</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1 Bit Allocation for G.729

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Bits</th>
<th>Sensitive bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>L2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>P1</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>P2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>GA1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>GB1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>GA2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>GB2</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2 Sensitive Bits of G.729

3 Security Analysis of Servetti’s Selective Encryption of G.729 Speech

Servetti et al. [28] took Swaminathan’s work [31] as the reference to propose two partial encryption schemes of G.729 speech.

First, speech signals are partitioned into two classes based on perception. Firstly, Class 1, the mostly perceptually relevant 36 bits, are to be encrypted, and Class 2, the other 44 bits, are to be left unprotected.

Then, two partial encryption techniques, a high-protection scheme (Scheme-H) and a low-protection scheme (Scheme-L), are developed. The Scheme-H high-protection scheme (See Figure 2[a]), based on the encryption of 36 sensitive bits, can achieve content protection comparable to that obtained by full encryption of the bit stream, as verified by signal inspection in both the time and frequency domains, means of objective distortion measures and means of formal listening tests. The Scheme-L low-protection scheme (See Figure 2[b]), the encryption of 24 bits, can eliminate intelligibility of bit streams, but it leads to 66% success rate for random guess.

Figure 2 Scheme-H and Scheme-L

(a) High-protection scheme

(b) Low-protection scheme
4 The Proposed Hierarchical Selective Encryption Scheme for G.729 Speech

In order to enhance the security of encryption and not increase the computation complexity, a group of novel hierarchical selective encryption schemes are proposed. We partition signals into two hierarchies based on bit sensitivity. The perceptual relevant bits (less than 36 bits) are encrypted with a strong cipher, and the others with a lightweight cipher. Here, the strong cipher is often of high security and also with high computational cost, while the lightweight cipher is often efficient in computing. We compare our schemes with Servetti’s Scheme-H and full encryption scheme in signal inspection and objective distortion measures. In the following content, we propose the encryption algorithms based on cat map and logistic map respectively, which will be used to encrypt the partitioned bits. Additionally, the hierarchical encryption schemes will be presented.

4.1 The Encryption Algorithm Based on Cat Map

Cat map, first proposed by Arnold [33-34] in the 1960s, is a chaotic map and can be presented as (1).

\[
\begin{bmatrix}
    x_{n+1} \\
    y_{n+1}
\end{bmatrix} = A \begin{bmatrix}
    x_n \\
    y_n
\end{bmatrix} (\mod N), \quad A = \begin{bmatrix}
    1 & 1 \\
    1 & 2
\end{bmatrix}
\]  
(1)

In order to use cat map for encryption, cat map is first extended to two-dimensional \(N \times N\) matrix and discretized as (2), where \(x, y \in \{0,1,2,\ldots,N-1\}\).

\[
\begin{bmatrix}
    x_{n+1} \\
    y_{n+1}
\end{bmatrix} = A \begin{bmatrix}
    x_n \\
    y_n
\end{bmatrix} (\mod N)
\]  
(2)

Then, encryption parameters can be obtained by changing the elements of \(A\). In normal condition, \(A\) is shown as (3) where \(|A| = 1\) and \(a, b\) are positive integers.

\[
A = \begin{bmatrix}
    ab + 1 & a \\
    b & 1
\end{bmatrix}
\]  
(3)

Thus, the confusion key of cat map is composed of the parameter \(a\) and \(b\). For a \(n\)-bit signal, the computational complexity of cat map is

\[
C_{\text{cat}}(n) = n^2 \cdot \lceil \frac{n}{2} \rceil \cdot \lceil \frac{n}{2} \rceil \cdot 4
\]  
(4)

We use cat map to encrypt a 10-second speech signal, and compute the segSNR between original and encrypted signals. Suppose \(I_{cp}\) is the iteration number of encryption, the relation between \(I_{cp}\) and segSNR is shown in Figure 3.

Form Figure 3, the value of segSNR is bigger when \(2 \leq I_{cp} \leq 7\), and when \(I_{cp} = 3\), the value of segSNR is the biggest which is -1.6364dB. When \(I_{cp} > 42\), the value of segSNR is less than 1, which shows that the effects of selective encryption decrease and the performance of security is degraded. So, \(I_{cp}\) should satisfy \(1 \leq I_{cp} \leq 42\) when cat map is used for encryption.

![Figure 3 Cat Map -- segSNR](image)

4.2 The Encryption Algorithm Based on Logistic Map

Logistic map [35-36] is represented as (5), where \(x\) is an input floating number in the interval \([0,1]\), and \(i\) plays the role of discrete time. It is well-known that the logistic map with \(\mu = 4\) is chaotic, and its confusion key is the input floating number \(x_0\).

\[
x_{i+1} = \mu \cdot x_i \cdot (1 - x_i)
\]  
(5)

For a \(n\)-bit signal, the computational complexity of logistic map is

\[
C_{\text{log}}(n) = n^2 \cdot \lceil \frac{n}{2} \rceil \cdot \lceil \frac{n}{2} \rceil \cdot 4
\]  
(6)

Similar to cat map, we use logistic map to encrypt the same speech signal, and compute the segSNR between original and encrypted signals. Suppose \(I_{lp}\) is the iteration number of encryption, the relation between \(I_{lp}\) and segSNR is shown in Figure 4.

From Figure 4, when \(2 \leq I_{lp} \leq 6\), the value of segSNR are -9.7277dB, -10.6132dB, -9.7690dB, -9.5162dB, -9.4719dB respectively. When \(I_{lp} > 6\), the value of segSNR fluctuates between -9.5162dB and -9.4719dB. So, when \(2 \leq I_{lp} \leq 6\), the algorithm using logistic map can obtain a good effect of encryption.
4.3 The Hierarchical Selective Encryption Schemes

For existing researches, only one part of bit streams is subject to encryption, and the other part is unprotected, which greatly reduces the performance of security. In this section, we take Swaminathan and Servetti’s work as the references, partition signals into two hierarchies (See Figure 5): H1, the most sensitive bits, are encrypted by a strong cipher and H2, the remaining bits, are encrypted by a lightweight cipher, and propose a group of encryption schemes.

In this paper, the strong encryption means that the bits are encrypted by both cat map and logistic map for $x$ times, while lightweight encryption is only by cat map for $x'$ times, where $x' < x$.

Suppose that $\langle u, v \rangle$ is an encryption scheme, where $u$ is the number of bits subject to encryption by a strong cipher, and $u \leq 36$ according to literature [28] and Table 2; $v$ is the number subject to encryption by a lightweight cipher, and $u + v = 80$. Our purpose is to find a group of encryption schemes $\langle u, v \rangle$ whose computational complexity are lower than Scheme-H, and security performance are better than Scheme-H.

Literature [28], for Scheme-H, only pointed out that 36 bits are subject to encryption, but did not refer any encryption method. In order to compare $\langle u, v \rangle$ with Scheme-H, 36 bits in Scheme-H are encrypted by a high degree of encryption, and the other 44 bits are unprotected.

5 Parameter Selection

In this section, we first obtain a group of schemes with low computational complexity by the theoretical analysis, and then obtain the final group of schemes with high security and low cost by contrast experiments.

5.1 Analysis of Computational Complexity

Suppose $C_{\langle u, v \rangle}$ and $C_{\text{Scheme-H}}$ are the computational complexity of $\langle u, v \rangle$ and Scheme-H, they are shown as (7) and (8) respectively.

\[
C_{\langle u, v \rangle} = [C_\phi(u) + C_\psi(v)] \cdot x + [C_\phi(v)] \cdot x' \tag{7}
\]

\[
C_{\text{Scheme-H}} = [C_\phi(36) + C_\psi(36)] \cdot x \tag{8}
\]

If

\[
C_{\langle u, v \rangle} < C_{\text{Scheme-H}}, \tag{9}
\]

we obtain (10) by substituting (7), (8) into (9).

\[
[C_\phi(u) + C_\psi(u)] \cdot x + [C_\phi(v)] \cdot x' < [C_\phi(36) + C_\psi(36)] \cdot x \tag{10}
\]

Then substituting (4), (6) into (10), we can obtain (11). And (11) can be changed as (12).

\[
\left[ u^2 \cdot \left( \frac{1}{v^2} \cdot \frac{4 + u^2}{u^2} \right) \cdot x + \left( \frac{1}{v^2} \cdot \frac{4 + u^2}{u^2} \right) \cdot x' \right] \cdot x' < \left[ 36^2 \cdot \left( \frac{1}{36^2} \cdot \frac{4 + 36^2}{36^2} \right) \right] \cdot x \tag{11}
\]

\[
\frac{x'}{x} < \frac{10 - 36^2 \cdot \left( \frac{4 + u^2}{u^2} \right) - \frac{4 + u^2}{u^2}}{\left( \frac{4 + u^2}{u^2} \right)} \tag{12}
\]

From section 4.1 and 4.2, $2 \leq I_p \leq 6$, and $1 \leq I_c \leq 7$, so, $x = I_p$, that is $2 \leq x \leq 6$. Since $x' < x$, $1 \leq x' \leq 5$. Thus, $\frac{1}{6} \leq \frac{x'}{x}$, then we can obtain (13).
\[ \frac{1}{6} \leq 10^{36^2} - \left( u^2 \cdot \left[ u^1 \right]^2 + 4 + u^3 \cdot \left[ u^1 \right]^2 \right) \left( v^2 \cdot \left[ v^1 \right]^2 \cdot 4 \right) \]  
(13)

That is
\[ \left[ 10^{36^2} - \left( u^2 \cdot \left[ u^1 \right]^2 + 4 + u^3 \cdot \left[ u^1 \right]^2 \right) \right] 6 \]
\[ - \left( v^2 \cdot \left[ v^1 \right]^2 \cdot 4 \right) \geq 0 \]  
(14)

Suppose
\[ f(u) = \left[ 10^{36^2} - \left( u^2 \cdot \left[ u^1 \right]^2 + 4 + u^3 \cdot \left[ u^1 \right]^2 \right) \right] 6 \]
\[ - \left( v^2 \cdot \left[ v^1 \right]^2 \cdot 4 \right) \]  
(15)

Since \( u + v = 80 \), when \( 1 \leq u \leq 33 \), \( f(u) > 0 \), that is \( C_{\text{enc}} < C_{\text{Scheme-H}} \). However, when the number of bits encrypted is less than a value \( u' \), the security performance can not been guaranteed. Thus, we have to analysis of \( 1 \leq u \leq 33 \) by contrast experiments in the following content, including signal inspection and segSNR, with the original speech as the reference, and to obtain the final group of encryption schemes.

5.2 Analysis of Signal Inspection

In this section, according to the bit sensitivity of G.729 in Table 2, combinations of \( <u, v> \) in Table 3 are tested. And in our signal inspection experiments, \( x = 6 \) and \( y = 3 \).

The test speech materials consisting of 7 paragraphs, (1, 2, 3 are male speech, and 4, 5, 6, 7 are female speech), each 10 seconds long, is first encoded using G.729, and then protected with full encryption, Scheme-H and \( <u, v> \) in Table 3 respectively.

Different selective encryption schemes degrade the speech signals in different degree. Taking the second paragraph for an example, the specific features of encrypted signals in the time domains are tested.

Proposition 1. when \( 16 \leq u \leq 33 \), \( <u, v> \) has lower computational complexity and higher security performance than Scheme-H.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>The number of bits encrypted</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>2 3 3 4 4 5 5 6 6 6 6 6 6</td>
</tr>
<tr>
<td>L2</td>
<td>2 2 2 3 3 3 3 4 4 4 4 5 5</td>
</tr>
<tr>
<td>P1</td>
<td>3 3 4 4 4 5 5 5 5 5 5 5 5 5</td>
</tr>
<tr>
<td>P2</td>
<td>2 2 3 3 3 3 3 3 3 3 3 3 3 3</td>
</tr>
<tr>
<td>GA1</td>
<td>1 1 2 1 2 2 2 2 2 2 2 2 3 3 3</td>
</tr>
<tr>
<td>GB1</td>
<td>2 2 2 2 2 2 2 2 2 2 2 2 3 3 3</td>
</tr>
<tr>
<td>GA2</td>
<td>1 2 1 1 1 1 2 2 2 2 2 2 3 3 3</td>
</tr>
<tr>
<td>GB2</td>
<td>2 2 2 2 2 2 2 2 2 2 2 2 3 3 3</td>
</tr>
<tr>
<td>( u )</td>
<td>15 16 18 20 22 24 26 28 30 32 34 36 38 40 42</td>
</tr>
<tr>
<td>( v )</td>
<td>65 64 62 60 58 56 54 52 50 48 47</td>
</tr>
</tbody>
</table>

Figure 6 shows the signals in time domain. As can be seen, Scheme \( <15, 65> \) makes the waveform of signals encrypted have little change compared to the original signal, and its security performance is less than Scheme-H. When \( u \geq 16 \), the encrypted effects are better than Scheme-H. Especially, Scheme \( <33, 47> \) can have the same performance as full encryption. So, we can obtain the proposition as follows.

Proposition 1. when \( 16 \leq u \leq 33 \), \( <u, v> \) has lower computational complexity and higher security performance than Scheme-H.
6 Performance Analysis for the Proposed Encryption Scheme

In order to confirm Proposition 1, in this section, the contrast experiments are addressed in objective distortion measure (segSNR). Additionally, schemes’ cryptanalysis is investigated.

6.1 Analysis of SegSNR

We use segSNR to objectively assess the performance of \(<u, v>\), and compute segSNR for 7 paragraphs encrypted by full encryption, Scheme-H and \(<u, v>\) in Table 3, respectively. Table 4 gives the result of the 2\textsuperscript{nd} and 5\textsuperscript{th} paragraphs.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(2\textsuperscript{nd})</th>
<th>(5\textsuperscript{th})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full encryption</td>
<td>-8.8953</td>
<td>-12.8630</td>
</tr>
<tr>
<td>Scheme-H</td>
<td>-2.8760</td>
<td>-3.2753</td>
</tr>
<tr>
<td>(&lt;15, 65&gt;)</td>
<td>-2.1769</td>
<td>-2.7518</td>
</tr>
<tr>
<td>(&lt;16, 64&gt;)</td>
<td>-5.0717</td>
<td>-3.7772</td>
</tr>
<tr>
<td>(&lt;18, 62&gt;)</td>
<td>-7.4117</td>
<td>-5.4983</td>
</tr>
<tr>
<td>(&lt;20, 60&gt;)</td>
<td>-4.3314</td>
<td>-3.4068</td>
</tr>
<tr>
<td>(&lt;22, 58&gt;)</td>
<td>-4.1470</td>
<td>-5.6888</td>
</tr>
<tr>
<td>(&lt;24, 56&gt;)</td>
<td>-4.8223</td>
<td>-6.2077</td>
</tr>
<tr>
<td>(&lt;26, 54&gt;)</td>
<td>-6.4074</td>
<td>-5.0645</td>
</tr>
<tr>
<td>(&lt;28, 52&gt;)</td>
<td>-4.4014</td>
<td>-3.9904</td>
</tr>
<tr>
<td>(&lt;30, 50&gt;)</td>
<td>-4.6820</td>
<td>-8.9009</td>
</tr>
<tr>
<td>(&lt;32, 48&gt;)</td>
<td>-5.9986</td>
<td>-9.1442</td>
</tr>
<tr>
<td>(&lt;33, 47&gt;)</td>
<td>-8.8879</td>
<td>-9.4834</td>
</tr>
</tbody>
</table>

From Table 4, for the two paragraphs, the performance of scheme \(<15, 65>\) (2.1769dB, 2.7518dB respectively) are lower than Scheme-H (2.8760dB, 3.2753dB respectively). And when \(u \geq 16\), the performance of \(<u, v>\) are higher than Scheme-H. Especially, for the 2\textsuperscript{nd} paragraph, the performance of scheme \(<33, 47>\) (8.8879dB) can almost be the same as that of full encryption (8.8953dB).

So, \(<u, v>\) (16 \(\leq u \leq 33\)) have higher security performance than Scheme-H.

6.2 Analysis of Security

The encryption schemes’ security also depends on the adopted ciphers. Here, the proposed encryption schemes \(<u, v>\) (16 \(\leq u \leq 33\)) adopt the ciphers based on chaotic maps, whose keys are controlled by several sensitive parameters.

(1) Parameters \(a, b\) of cat map: \(a, b\) are natural numbers.

Since G.729 provides toll quality at 8 kb/s, \(a, b\) should be 0 \(< a < 256\), 0 \(< b < 256\).

(2) Initial input number \(x_0\) of logistic map: \(x_0\) is a floating number, and 0 \(< x_0 < 1\). In our experiments, precision of logistic map is 1/10000.

(3) Times of encryption \(x\) and \(x'\): 2 \(\leq x \leq 6\), and 1 \(\leq x' \leq 5\).

(4) The number \(u\) of bits encrypted: 16 \(\leq u \leq 33\).

Thus, the total number of secret keys is 255 \(\times 255 \times 10000 \times 6 \times 5 \times 18 \approx 10^{11}\). Suppose that crackers attack our system at the speed of 10\(^6\) per second, they need 10\(^5\) seconds. For speech applications in power-constrained devices and narrow bandwidth environments, it is difficult to be broken. So, our work can effectively prevent eavesdropping.

7 Conclusion and Future Work

The typical selective G.729 speech encryption algorithm proposed by Servetti is evaluated, whose disadvantages are analyzed. Based on this, a group of hierarchical selective encryption schemes are presented for ITU-T G.729 standard speech. First, we discuss the schemes’ computational complexity compared to Scheme-H, and obtain a group of encrypted schemes \(<u, v>\) (1 \(\leq u \leq 33\)) with low complexity. And then we propose two encryption algorithms based on chaotic maps, i.e., cat map and logistic map, and give contrast experiments addressed in signal inspection, objective distortion measure, and get the final group of schemes \(<u, v>\) (16 \(\leq u \leq 33\)). Additionally, we discuss the schemes’ security. The results demonstrate that the good tradeoff between security and cost can be obtained by the proposed hierarchical selective encryption schemes, and they are suitable for narrow bandwidth environments and power-constrained devices.

In future work, the properties against transmission errors will be investigated, and the encryption of speech/ audio standards supporting scalable streams will be considered.

References


Hierarchical Selective Encryption for G.729 Speech Based on Bit Sensitivities


Biographies

Zhaopin Su received the BS and PhD degrees in computer science from Hefei University of Technology, Hefei, China, in 2004 and 2008. Currently, she is a Lecturer with the School of Computer and Information, Hefei University of Technology. Her research interests include multimedia and communication security, chaotic algorithms.

Jianguo Jiang received the MS degree in computer science from Hefei University of Technology, Hefei, China, in 1989. He is currently a Professor and PhD supervisor in the School of Computer and Information, Hefei University of Technology. He is Head of the Texas Instruments-Hefei University of Technology DSPS Laboratory. His research interests include multimedia security, automatic control, image processing, and software engineering. He has published over 80 research papers and two books, and holds four patents in these fields.

Shiguo Lian received the PhD degree in multimedia security from Nanjing University of Science and Technology, Nanjing, China, in 2005. He is currently a Research Scientist, France Telecom R&D (Orange Labs) Beijing, China. He has published over 80 research papers, contributed 15 book chapters and held 14 filed patents. His research interests include information and communication security, intelligent multimedia processing, and ubiquitous communication and services.

Guofu Zhang received the BS and PhD degrees in computer science from Hefei University of Technology, Hefei, China, in 2002 and 2008. He is currently a Lecturer with the School of Computer and Information, Hefei University of Technology. His major research interests are in soft computation and information security.

Donghui Hu received MS degree in computer science and technology from University of Science and Technology of China, China, in 2004. He is currently pursuing the PhD degree in information security at the department of computer science, Wuhan University, Wuhan, China. His research interests include information security, digital image forensic and steganalysis.