

一种基于 G. 729a 语音的信息隐藏方法

郭舒婷, 田 晖

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

**摘要:** 针对 VoIP 系统普遍使用 G. 729a 语音编码,采用逐帧按位取反的方式对 G. 729a 语音进行全面测试,以客观的语音质量感知评价(PESQ)标准为依据,总结 G. 729a 语音帧中可隐藏位(最低有效位).结合矩阵编码算法,提出一种基于 G. 729a 语音的信息隐藏方法,以大量的语音样本为载体,对该方法进行测试分析.实验结果表明:该方法能提供可观的隐藏容量(8 bit · 帧<sup>-1</sup>),较高的嵌入效率(2.678),很好的不可感知性(平均 PESQ 值为 3.593),且很好地符合了 VoIP 通信的实时性要求.

**关键词:** 信息隐藏; 隐写; IP 语音; G. 729a; 矩阵编码

**中图分类号:** TP 309.2; TP 309.7 **文献标志码:** A

信息隐藏是近年来提出的一种解决信息安全的新途径.它是利用人的听(视)觉对媒体信号的感觉冗余,将隐秘信息嵌入到数字媒体中.以隐藏嵌入过程的存在性,使除了执行嵌入行为的发送方和接收方之外的第三方不知道此过程<sup>[1]</sup>.随着研究的不断深入,信息隐藏已不仅仅局限于以图像、音频、文本等静态媒体为载体,对于能够提供可观隐藏空间的流媒体<sup>[2]</sup>也成为了学者们关注的热点.在基于流媒体的信息隐藏研究中,以 VoIP(voice over internet protocol)为代表的信息隐藏技术发展的最为迅速.其原因主要是 VoIP 所嵌入的隐秘信息具有强动态性,不易被攻击者篡改,且 VoIP 语音能够提供较好的隐蔽性和隐藏容量<sup>[2]</sup>.基于 VoIP 的信息隐藏有两种研究思路:一种是基于网络协议的信息隐藏方法<sup>[3]</sup>;另一种是基于载荷(语音流)方法<sup>[4-7]</sup>.为实现信息隐藏,通常首先分析编码器参数的冗余度以确定语音流中的可隐藏位.文献[8-9]分别分析了 G. 723.1 编码器在低比特率(5.3 kbit · s<sup>-1</sup>)和高比特率(6.3 kbit · s<sup>-1</sup>)下参数的冗余度.Liu 等<sup>[10]</sup>采用分段信噪比来评价语音质量,分析了 G. 729 语音帧的最低有效位(LSB),认为 G. 729 语音帧的冗余位主要分布在第二子帧的 LSP 参数上.除了对于编码参数中的最低有效位的探索,近年来研究者们也开始从语音编码原理的角度挖掘冗余性.Huang 等<sup>[11-12]</sup>先后提出利用不活跃语音帧来隐藏隐秘信息和将嵌入过程整合到编码过程中;类似地,刘程浩等<sup>[13]</sup>则将隐写过程整合到基音预测过程中.本文针对 VoIP 系统普遍采用 G. 729a 语音编码,以客观的语音质量感知评价标准为依据,分析和总结 G. 729a 语音帧中的可隐藏位;然后以此为基础,结合矩阵编码算法,提出了一种基于 G. 729a 语音的信息隐藏方法.

1 G. 729a 可隐藏位置分析

G. 729a 编解码器按照 8 kHz 采样,16 bit 量化且每帧长 10 ms,因而 G. 729a 语音帧被压缩成 80 位,每位都代表相应的参数向量.G. 729a 语音帧各比特所代表的意义,如表 1 所示.表 1 中:每个比特位

收稿日期: 2013-07-17

通信作者: 田晖(1982-),男,副教授,主要从事网络与多媒体信息安全、信息隐藏理论与技术的研究. E-mail: htian@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(61302094, 61370007, 61202468);福建省科技创新平台建设基金资助项目(2012H2002);福建省自然科学基金资助项目(2011J05151, 2013J01241);福建省中青年教师教育科研资助项目(JA13012);华侨大学青年教师科研提升资助项目(ZQN-PY115);中央高校基本科研业务费专项资金资助项目(JB-ZR1220);华侨大学高层次人才科研启动项目(11BS210)

表示为  $Px-y$  格式,  $P$  表示参数类型,  $x$  表示子帧号,  $y$  表示比特位置, 从低位 0 到高位. 在混合语音编码流中, 每比特对语音质量有不同程度的影响<sup>[10]</sup>, 将对语音失真影响不大的位称为冗余位, 即最低有效位. 通过对语音样本中各帧的每个比特逐位取反来测试 G. 729a 语音帧的最低有效位.

设待测试样本经过 G. 729a 编码后有  $n$  个语音帧  $F, F=\{f_1, f_2, \cdots, f_n\}$  且  $f_i=\{b_{i,1}, b_{i,2}, \cdots, b_{i,80}\}$ . 其中:  $1 \leq i \leq n; b_{i,j} (1 \leq i \leq n, 1 \leq j \leq 80)$  表示第  $i$  个语音帧的第  $j$  位. 对  $F$  中的第  $j$  位取反且同时对  $b_{i,j} (1 \leq i \leq n)$  置反进行测试. 每次测试仅对 G. 729a 语音帧的一个特定位置进行取反, 并根据其对语音质量的影响来判断是否为最低有效位.

表 1 G. 729a 语音帧各比特的意义  
Tab. 1 Meanings of frame bits for G. 729a speech

位	码字	位	码字	位	码字	位	码字	位	码字	位	码字	位	码字
1	L0	11	L2-2	21	P1-5	<u>31</u>	<u>C1-9</u>	41	S1-3	51	GB1-0	<u>61</u>	<u>C2-8</u>
2	L1-6	12	L2-1	22	P1-4	<u>32</u>	<u>C1-8</u>	42	S1-2	52	P2-4	<u>62</u>	<u>C2-7</u>
3	L1-5	13	L2-0	23	P1-3	<u>33</u>	<u>C1-7</u>	43	S1-1	53	P2-3	<u>63</u>	<u>C2-6</u>
4	L1-4	<u>14</u>	<u>L3-4</u>	24	P1-2	<u>34</u>	<u>C1-6</u>	44	S1-0	54	P2-2	<u>64</u>	<u>C2-5</u>
5	L1-3	<u>15</u>	<u>L3-3</u>	25	P1-1	<u>35</u>	<u>C1-5</u>	<u>45</u>	<u>GA1-2</u>	55	P2-1	<u>65</u>	<u>C2-4</u>
6	L1-2	<u>16</u>	<u>L3-2</u>	26	P1-0	<u>36</u>	<u>C1-4</u>	46	GA1-1	<u>56</u>	<u>P2-0</u>	<u>66</u>	<u>C2-3</u>
7	L1-1	<u>17</u>	<u>L3-1</u>	27	P0	<u>37</u>	<u>C1-3</u>	47	GA1-0	<u>57</u>	<u>C2-12</u>	<u>67</u>	<u>C2-2</u>
8	L1-0	<u>18</u>	<u>L3-0</u>	<u>28</u>	<u>C1-12</u>	<u>38</u>	<u>C1-2</u>	48	GB1-3	58	C2-11	<u>68</u>	<u>C2-1</u>
9	L2-4	19	P1-7	29	C1-11	<u>39</u>	<u>C1-1</u>	<u>49</u>	<u>GB1-2</u>	<u>59</u>	<u>C2-10</u>	<u>69</u>	<u>C2-0</u>
10	L2-3	20	P1-6	<u>30</u>	<u>C1-10</u>	<u>40</u>	<u>C1-0</u>	<u>50</u>	<u>GB1-1</u>	<u>60</u>	<u>C2-9</u>	70	S2-3
												80	GB2-0

为了直观的评价语音质量的变化, 采用客观的语音质量评价标准 PESQ, 其较之完全通过人来打分的主观评价方式 MOS 技术, 具有成本低廉、客观、自动、可重复等优点, 近年来在通信语音质量评价领域得到了广泛应用. PESQ 的原理是通过比较参考信号和降级信号输出一个对语音质量失真的预测值, 其范围在  $-0.5 \sim 4.5$  之间, 值越高表示加噪语音信号的语音质量越好. 测试中将原始语音作为参考信号, 而将修改后的语音作为降级信号, 以评判各比特位被修改后对语音质量的影响.

为了测试最低有效位, 广泛收集了 2 800 个语音样本, 涵盖汉语男声、汉语女声、英语男声、英语女声 4 类. 对各比特位的平均 PESQ 值进行统计, 根据置反后平均 PESQ 值 (PESQ score) 的变化大小, 可得到 G. 729a 语音帧的 LSB (PESQ 值大于 3.5 的位), 如表 2 所示. 根据表 1 中的帧比特分布情况 (加下划线部分代表最低有效位), 不难看出 G. 729a 语音帧的最低有效位主要分布在自适应码书延迟、线性声谱对、固定码书索引以及码书增益等参数上.

表 2 G. 729a 语音帧比特的可隐藏位  
Tab. 2 Available bits for steganography in G. 729a speech frame

位	PESQ	位	PESQ	位	PESQ	位	PESQ	位	PESQ	位	PESQ
18	4.267 2	15	4.073 8	79	3.734 3	32	3.708 4	40	3.594 0	36	3.546 3
17	4.205 1	56	3.915 3	35	3.727 7	61	3.700 1	34	3.583 5	65	3.537 0
16	4.200 1	49	3.869 6	38	3.722 4	28	3.682 9	69	3.583 3	39	3.533 7
14	4.192 3	78	3.861 3	50	3.722 3	57	3.677 8	63	3.576 0	68	3.527 6
31	4.176 5	45	3.797 5	64	3.719 2	37	3.603 7	30	3.554 0	33	3.521 2
60	4.171 5	74	3.792 1	67	3.709 6	66	3.594 5	59	3.549 1	62	3.507 9

2 基于矩阵编码的 G. 729a 的信息隐藏算法

为了减少对载体的修改从而提高嵌入的不可感知性, 利用矩阵编码指导隐秘信息的嵌入. 矩阵编码是最常用的一种提高感知透明性的隐写编码策略<sup>[14]</sup>, 它可在  $2^n-1$  比特载体信息中隐藏  $n$  bit 的秘密信息, 而只需修改不超过 1 bit 的载体信息, 因而可以有效提高嵌入效率并且降低比特改变率. 矩阵编码过程等价于根据一个哈希函数求出待隐藏信息在载体中的嵌入位置. 假设有长度为  $n$  bit 的隐秘信息分组  $S, S=\{s_1, s_2, s_3, \cdots, s_n\}$ , 要隐藏在长度为  $l$  bit 的载体分组  $C$  中,  $C=\{c_1, c_2, c_3, \cdots, c_l\}$ , 其中:  $n$  与  $l$  满

足  $l=2^n-1$ . 则隐秘信息的嵌入过程可表述为如下 3 个步骤.

**步骤 1** 根据载体  $C$  中各位的下标  $i(1\leq i\leq l)$ , 求得大小为  $n\times l$  的编码矩阵  $A$ , 即

$$A=\begin{bmatrix}a_{1,1}&a_{2,1}&\cdots&a_{l,1}\\a_{1,2}&a_{2,2}&\cdots&a_{l,2}\\\vdots&\vdots&\cdots&\vdots\\a_{1,n}&a_{2,n}&\cdots&a_{l,n}\end{bmatrix}.\tag{1}$$

式(1)中:  $i=a_{i,1}\times 2^0+a_{i,2}\times 2^1+\cdots+a_{i,n}\times 2^{n-1}, 1\leq i\leq l$ .

**步骤 2** 将载体  $C$  与  $A_m$  对应相乘并连续求异或, 得值  $x_m$ . 其中:  $A_m(1\leq m\leq n)$  表示编码矩阵  $A$  中的第  $m$  行;  $a_{m,j}(1\leq m\leq n, 1\leq j\leq l)$  表示编码矩阵  $A$  中的第  $m$  行第  $j$  列的值. 即

$$x_m=(c_1\times a_{m,1})\oplus (c_2\times a_{m,2})\oplus \cdots \oplus (c_j\times a_{m,l}).\tag{2}$$

**步骤 3** 比较  $s_m$  与  $x_m(1\leq m\leq n)$  是否相等. 若相等, 则令  $X_m$  为 0; 否则,  $X_m$  为 1(其中  $s_m$  为隐秘信息  $S$  中的第  $m$  位), 根据式(3)求嵌入位置  $X$ , 即

$$X=x_1\times 2^0+x_2\times 2^1+\cdots+x_n\times 2^{n-1}.\tag{3}$$

若求得的  $X$  为 0, 则不必修改(说明要嵌入的位与待嵌入的隐秘信息的值相同); 否则, 修改载体  $C$  中第  $X$  位的值.

基于矩阵编码信息隐藏的提取算法是根据上述哈希函数即编码矩阵  $A$  求得, 可视作嵌入算法的一个逆过程. 只需求出秘密信息的位  $s_j$ , 即

$$s_j=c'_1\times a_{j,1}+c'_2\times a_{j,2}+\cdots+c'_l\times a_{j,l}.\tag{4}$$

式(4)中:  $c'_i$  为载密信息  $C'=\{c'_1,c'_2,c'_3,\cdots,c'_l\}$  中第  $i$  位的值;  $a_{j,i}(1\leq j\leq n, 1\leq i\leq l)$  表示编码矩阵  $A$  中的第  $j$  行第  $i$  列的值;  $s_j$  为提取的隐秘信息  $S=\{s_1,s_2,s_3,\cdots,s_n\}$  中第  $j$  位的值.

在信息隐藏过程中<sup>[14]</sup>, 通常定义嵌入率为隐藏的隐秘信息比特数与载体比特数的比值; 比特改变率为载体中被改变的比特数与载体比特数的比值; 嵌入效率为隐藏的隐秘信息比特数与载体中被改变的比特数的比值. 在应用矩阵编码的隐藏过程中, 嵌入率( $\alpha$ )、比特改变率( $\beta$ )和嵌入效率( $\lambda$ )分别表示为

$$\alpha=\frac{n}{l}=\frac{n}{2^n-1},\tag{6}$$

$$\beta=\frac{0\times\frac{1}{2^n}+\sum_{i=1}^{2^n-1}\frac{1}{2^n}}{2^n-1}=\frac{1}{2^n},\tag{7}$$

$$\lambda=\frac{\alpha}{\beta}=\frac{\frac{n}{2^n-1}}{\frac{1}{2^n}}=\frac{n\cdot 2^n}{2^n-1}.\tag{8}$$

不同的隐秘信息分组长度  $n$  下的矩阵编码隐藏方法的隐藏性能, 如表 3 所示. 众所周知, 传统的 LSB 替换方法中, 嵌入率和比特改变率分别为 1 和 0.5, 也就是说 LSB 替换方法的嵌入效率为 2. 由此可知: 采用矩阵编码可以获得较之传统 LSB 替换更高的嵌入效率. 同时, 矩阵编码的嵌入效率随隐秘信息分组长度  $n$  的增大而增大, 而嵌入率却随之减少. 因此, 在实际应用中, 需在嵌入效率和嵌入率之间进行合理折中. 为兼顾嵌入效率和隐藏容量, 根据上节的分析结果选取每帧中隐藏效果最佳的 12 个最低有效位, 并按 3 bit 长度进行分组, 然后在每组上分别执行基于矩阵编码的隐藏算法, 即每组可隐藏 2 bit 隐秘信息, 从而每帧共可隐藏 8 bit 隐秘信息.

表 3 矩阵编码在不同隐秘信息分组长度  $n$  下的隐藏性能

Tab. 3 Steganographic performance of the matrix encoding with different lengths of the secret messages $n$									
$n$	$l=2^n-1$	$\alpha$	$\beta$	$\lambda$	$n$	$l=2^n-1$	$\alpha$	$\beta$	$\lambda$
2	3	0.666 7	0.250 0	2.666 7	6	63	0.095 2	0.015 6	6.095 2
3	7	0.428 6	0.125 0	3.428 6	7	127	0.055 1	0.007 8	7.055 1
4	15	0.266 7	0.062 5	4.266 7	8	255	0.031 4	0.003 9	8.031 4
5	31	0.161 3	0.031 3	5.161 3	9	511	0.017 6	0.002 0	9.017 6

### 3 测试与分析

为了测试文中算法的有效性,先选择英文对话“Question11. W: Did you hear that Anna needs to stay in bed for 4 weeks? M: Yeah. She injured her spine in a fall and a doctor told her to lie flat on her back for a month so it can mend. Q: What can we learn from the conversation?”(记为  $C_1$ )和中文短句“人之初,性本善,性相近,习相远”(记为  $C_2$ )作为载体,以华侨大学英文简介(共 4 810 字节)作为隐秘信息进行测试. 经过 G. 729a 编码之后, $C_1$  共有 2 010 帧,可隐藏 2 010 字节隐秘信息; $C_2$  共有 710 帧,可隐藏 710 字节隐秘信息(由于载体长度限制,因此不能将整段隐秘信息完全嵌入). 嵌入隐秘信息前后的声谱,如图 1 所示. 由图 1 可知:载体语音在嵌入隐秘信息前后的声谱图基本一致,说明隐藏过程未对载体语音频谱造成明显的改变.

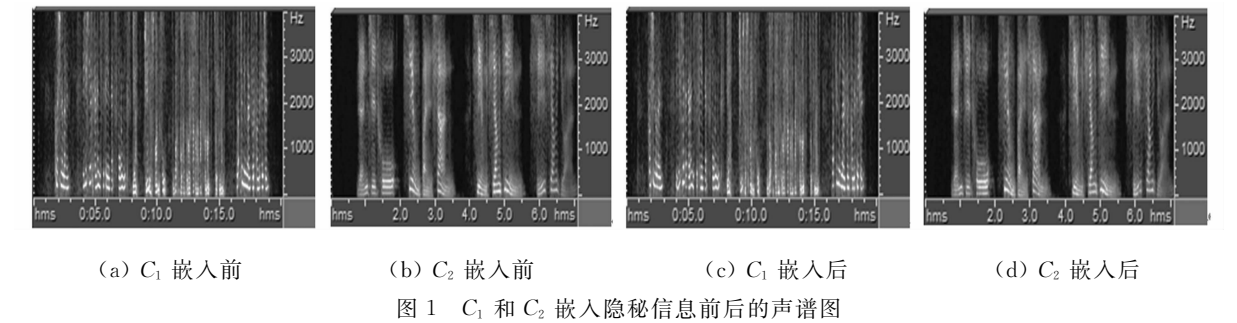


图 1  $C_1$  和  $C_2$  嵌入隐秘信息前后的声谱图

Fig. 1 Spectrogram of  $C_1$  and  $C_2$  before and after embedding secret message

为进一步测试文中算法的隐藏效果,选取 ITU P. 501 标准附录 B 中提供的多种音频库中的语音作为载体样本进行 PESQ 值测试. 该样本库覆盖了世界上的主要 10 种语言,如中文、英语、法语、德语、意大利语、日语、波兰语、西班牙语等,每种语言又包含了男声和女声,具有良好的多样性,近年来在语音测试中被广泛采用.

在测试过程中,首先,将载体样本转换为 8 kHz 采样,16 bit 量化的 PCM 格式语音;然后,将其经 G. 729a 编码器编码、嵌入隐秘信息;最后,分别解码已嵌入隐秘信息及未执行任何操作的语音样本为 PCM 格式语音. 将未执行任何操作的 PCM 语音样本作为 PESQ 的原始语音,而经嵌入后的 PCM 语音样本作为 PESQ 的降级语音,进行 PESQ 测试.

各类语音样本嵌入隐秘信息后的 PESQ 值,如图 2 所示. 由图 2 数据可知:载密语音的语音质量与原始载体语音的语种存在一定的关系,如载密法语语音的平均 PESQ 值为 3.7 左右,而载密西班牙语语音的平均 PESQ 值为 3.4 左右,但就整体而言,载密语音的平均 PESQ 值均维持在 3.4 以上,说明文中方法对语音质量影响较小,能够提供良好的感知透明性.

表 4 进一步给出了 10 个语种的语音样本在嵌入上述隐秘信息之后的嵌入效率,及各样本嵌入效率的波动性(即方差)的统计结果. 由表 4 可知:文中算法的平均嵌入效率为 2.678;此外,对同一语种而言嵌入效率波动性较小,说明本算法能够提供较为稳定的隐藏性能.

表 4 不同语种的嵌入效率均值及其方差

Tab. 4 Mean embedding efficiencies and their variance for different languages

语种	均值	方差	语种	均值	方差
中文	2.671	0.000 58	德语	2.700	0.000 57
英国英语	2.667	0.000 37	意大利语	2.690	0.000 05
美国英语	2.678	0.000 26	日语	2.669	0.001 01
芬兰语	2.664	0.000 72	波兰语	2.676	0.000 17
法语	2.670	0.000 34	西班牙语	2.693	0.000 41

除了采用客观评价标准评价本方法的隐藏效果外,还进行了 ABX 测试<sup>[15]</sup>,从主观上评价执行嵌入操作的语音样本与原始语音样本之间的相似度. ABX 测试,是让受测者同时听三段同样内容的样本语音 A、B 和 X,并让受测者判断 X 与 A 和 B 中的哪个一致. 在测试过程中,若样本 X 被判错的比例越高,则说明人耳越难分辨 A 和 B. 文中选取了 11 个受测者,对 ITU P. 501 标准附录 B 中提供的 40 组样本

分别进行了 ABX 测试,其中,A 和 B 中的一者为原始语音样本,另一者为嵌入隐秘信息后的语音样本;X 则为原始语音样本和嵌入隐秘信息后的语音样本之一.对不同语种的语音样本进行 ABX 测试的平均判错率( $\eta$ ),如图 3 所示.

从全部样本整体测试结果来看,平均判错率( $\eta$ )为 48.18%,接近随机猜想的概率 50%.说明人耳很难准确的区分原始语音和嵌入隐秘信息后的语音样本,同时说明本方法能够提供良好的嵌入透明性.此外,从不同语种的测试结果来看,不同语种的 ABX 测试判错率略有不同,其规律与图 2 中 PESQ 值的测试结果基本对应,即 PESQ 值较高者,ABX 平均判错率也较高.这一现象可能表明语种(或语音内容)对语音质量和隐藏效果有一定程度上的影响,后续将对这一问题做更深入的研究.

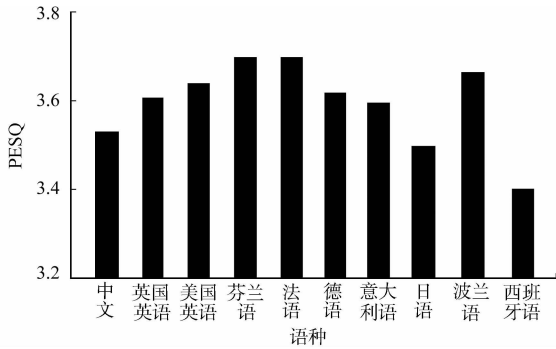


图 2 不同语种的 PESQ 值

Fig. 2 PESQ score for different languages

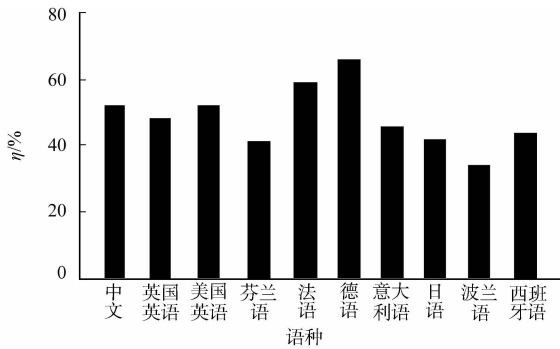


图 3 对不同语种的 ABX 测试的错误率

Fig. 3 Error rate of ABX test for different languages

此外,针对 G. 729a 语音流提出的隐藏方法可应用的另一个重要前提是满足语音通信的实时性要求.为此,在一台计算机(主要配置参数为 Intel(R) Core(TM)2 Duo CPU E7500 @2.93 GHz 2.94 GHz;2.00 GB(RAM);Windows 7 旗舰版 Service Pack 1)上,对所提出方法嵌入和提取时间进行测试.实验结果表明:本方法执行过程中,每帧上隐秘信息的嵌入时间平均为 6.868  $\mu$ s,每帧上提取隐秘信息的平均时间为 7.142  $\mu$ s,这远小于 G. 729a 编码器所允许每帧的编码时间 15 ms(注:G. 729a 编码器以 10 ms 为一帧进行编码,并允许 5 ms 的预留).由此说明:本方法能够很好地满足实时语音通信的要求.

### 4 结论

对基于 G. 729a 语音流的信息隐藏方法进行了初步研究.以客观的语音质量感知评价标准为依据,通过逐帧按位取反对 G. 729a 语音帧的最低有效位进行了全面测试和分析;根据分析结果,结合矩阵编码算法,提出了一种基于 G. 729a 语音的信息隐藏方法.最后,选取大量语音样本,对提出的方法进行了测试和分析.结果表明:该方法能提供可观的隐藏容量,较高的嵌入效率和较好的不可感知性.

值得指出的是,本方法从本质上来说属于最低有效位修改的隐藏方法,因而针对载体的大范围篡改,尤其是针对最低有效位的主动攻击会在一定程度上造成隐秘信息的丢失.当然,这种攻击方式对于正常的语音通信也是致命的,也就是说正常的语音通信都将难以为继.因此,在研究隐蔽通信时,常常不考虑这种极端的情形.对于非人为的载体影响(如噪声干扰、数据包丢失等)可能造成的隐秘信息丢失,在隐蔽通信系统中可采用差错重传技术<sup>[7]</sup>予以解决.

未来将针对文中方法在实际隐蔽通信系统中应用时的鲁棒性和可靠性问题进一步做深层次的研究和探讨.此外,如何进一步提高嵌入效率,各最低有效位的隐藏操作之间是否存在联系或相互影响等,也是文中将进一步深入的重要问题.

### 参考文献:

[1] GU Chun-ying,GAO Xiao-li. Research on information hiding technology[C]// Proceedings of the 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet). Yichang:IEEE Press,2012:2035-2037.

[2] MAZURCZYK W. VoIP steganography and its detection: A survey[J]. ACM Computing Surveys,2013,2(46):20.

- [3] FRACZEK W, MAZURCZYK W, SZCZYPIORSKI K. Hiding information in a stream control transmission protocol [J]. Computer Communications, 2012, 2(35): 159-169.
- [4] XU Er-chi, LIU Bo, XU Li-yang. Adaptive VoIP steganography for information hiding within network audio streams [C]// Proceedings of the 14th International Conference on Network-Based Information Systems (NBIS). Tirana: IEEE Press, 2011: 612-617.
- [5] MIAO Rui, HUANG Yong-feng. An approach of covert communication based on the adaptive steganography scheme on voice over IP [C]// Proceedings of the 46th IEEE International Conference on Communications (ICC). Kyoto: IEEE Press, 2011: 1-5.
- [6] LIU Jin, ZHOU Ke, TIAN Hui. Least-significant-digit steganography in low bitrate speech [C]// Proceedings of the 47th IEEE International Conference on Communications (ICC). Ottawa: IEEE Press, 2012: 1109-1112.
- [7] TIAN Hui, JIANG Hong, ZHOU Ke, et al. Adaptive partial-matching steganography for voice over IP using triple m sequences [J]. Computer Communications, 2011, 34(18): 2236-2247.
- [8] XU Ting-ting, YANG Zhen. Simple and effective speech steganography in G. 723. 1 low-rate codes [C]// Proceedings of the 1st International Conference on Wireless Communications & Signal Processing. Nanjing: IEEE Press, 2009: 1-4.
- [9] 杨俊, 柏森, 朱桂斌, 等. G. 723. 1 高速率语音流的信息隐藏算法 [J]. 计算机工程, 2011, 37(22): 114-115, 118.
- [10] LIU Li-hua, LI Ming-yu, LI Qiong, et al. Perceptually transparent information hiding in G. 729 bitstream [C]// Proceedings of the 4th International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Harbin: IEEE Press, 2008: 406-409.
- [11] HUANG Yong-feng, TANG Shan-yu. Steganography in inactive frames of VoIP streams encoded by source codec [J]. IEEE Transactions on Information Forensics and Security, 2011, 6(2): 296-306.
- [12] HUANG Yong-feng, LIU Cheng-hao, TANG Shan-yu. Steganography integration into a low-bit rate speech codec [J]. IEEE Transactions on Information Forensics and Security, 2012, 7(6): 1865-1875.
- [13] 刘程浩, 柏森, 黄永峰, 等. 一种基于基音预测的信息隐藏算法 [J]. 计算机工程, 2013, 39(2): 137-140.
- [14] TIAN Hui, ZHOU Ke, FENG Dan. Dynamic matrix encoding strategy for voice-over-IP steganography [J]. Journal of Central South University of Technology, 2010, 17(6): 1285-1292.
- [15] 贾俊, 王朔中. 基于 LabVIEW 平台的 ABX 测试系统 [J]. 声学技术, 2005, 24(1): 46-49.

## An Information Hiding Method Based on G. 729a Speech

GUO Shu-ting, TIAN Hui

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

**Abstract:** For G. 729a speech codec widely used in VoIP systems, a one-by-one bit flipping test for G. 729a speech is carried out, and the available bits for steganography (least significant bits) are correspondingly obtained based on perceptual evaluation of speech quality. To combine with matrix encoding, an information hiding approach based on G. 729a speech is proposed. The approach is tested and evaluated with a large number of speech samples as the carriers. Experimental results demonstrate that the proposed approach can provide a relatively high hiding capacity (8 bits/frame), higher embedding efficiency (2.678), good imperceptibility (the average PESQ score is 3.593) and nice real-time performance that meets the requirement of VoIP systems well.

**Keywords:** information hiding; steganography; voice over internet protocol; G. 729a; matrix encoding

(责任编辑: 黄晓楠      英文审校: 吴逢铁)