语音信息隐藏中的 AERA 算法

白 剑[1,2], 景晓军[3], 杨 榆[1,2], 徐迎晖[1,2], 钮心忻[1,2], 杨义先[1,2]
(1. 北京邮电大学信息安全中心 126 信箱, 北京 100876; 2. 北京邮电大学网络与交换国家重点实验室, 北京 100876; 3. 北京邮电大学电信工程学院, 北京 100876)

摘 要: 本文提出了一种能够在全球移动通信系统(Global System for Mobile Communications, GSM)中使用的信息隐藏算法——基于分析合成的能量比调整(Analysis-By-Synthesis Energy Ratio Adjust, AERA)算法. 算法采用了分析合成(Analysis-By-Synthesis, ABS)技术, 在嵌入过程中根据输入明文语音实时的调整嵌入强度, 使得隐藏效果和解码效果都达到最佳值. 大量仿真试验结果表明算法对 GSM 中的规则脉冲激励长期预测编码(Regular Pulse Excited-Long Term Prediction Coding, RPE-LTP Coding)、自适应差分脉冲编码调制(Adaptive Differential Pulse Code Modulation, ADPCM)等语音压缩编码以及滤波操作有很强的鲁棒性, 携密语音的分段平均信噪比达到 37dB, 可以达到透明性要求. 算法简单易行并且是基于盲检测的, 具有很高的实用性.

关键词: 语音信息隐藏; 分析合成; 能量比; RPE-LTP 编码
中图分类号: TP391.41 文献标识码: A 文章编号: 0372-2112 (2005) 09-1541-04

AERA Algorithm on Speech Hiding System

BAI Jian[1,2], JING Xiao-jun[3], YANG Yu[1,2], XU Ying-hui[1,2], NIU Xin-xin[1,2], YANG Yi-xian[1,2]
(1. *Information Security Center, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*;
2. *State Key Laboratory of Switching and Networking, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*;
3. *Telecommunications Engineering School, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*)

Abstract: This paper presents a speech hiding algorithm Analysis-By-Synthesis Energy Ratio Adjust Algorithm (AERAA) that can be deployed on Global System for Mobile Communications (GSM). The algorithm is based on the technique of Analysis-By-Synthesis (ABS) and can adjust adaptively the strength of the security information component according to public speech. The hiding capacity of the algorithm is 50 bits per second. After a lot of experiments, it is proved that this algorithm has strong robustness to many attacks such as Regular Pulse Excited-Long Term Prediction coding (RPE-LTP coding) on GSM, Adaptive Differential Pulse Code Modulation (ADPCM) compression and most of filters. The average Signal Noise Ratio (SNR) of the middle speech is above 37db, therefore the imperceptibility can be ensured. Moreover the algorithm is based on blind-detection, it is simple and effective.

Key words: speech hiding; ABS; energy ratio; RPE-LTP coding

1 引言

信息隐藏技术是当前信息安全领域研究的热点问题. 作为信息隐藏技术的一个重要分支, 对语音信息隐藏技术的研究正日益升温. 当前语音隐藏的主要方法有以下几种: (1) 最低有效位(LSB)方法[1]: 该方法可以隐藏较多的数据, 但稳健性较差, 无法抵抗音频数据处理所带来的破坏; (2) 相位编码法[2]: 该算法通过修改音频数据的傅里叶系数的相位值将数据隐藏到音频数据中; (3) 频谱变换法[3]: 该算法借鉴扩频通信的原理, 将数据作为噪声隐藏到载体数据的频谱中, 该方法具有较高的健壮性; (4) 回声隐藏法[4]: 该算法通过改变语音回声的延迟来隐藏水印数据. 此外, 还有其他多种语音伪装算法, 但可以看作以上算法的改进.

在实际话音通信系统中存在各种编码、压缩、AD 重采样、噪声等干扰因素. 为了在实际语音通信系统中应用语音信息隐藏技术, 要求隐藏算法能够适应实际通信的应用环境. 如为了在 PSTN 网络中应用语音信息隐藏技术, 要求隐藏算法能够抵御 AD 重采样、噪声、A 律、U 律压缩、ADPCM 编码等攻击. 为了在 GSM 移动通信系统中使用语音信息隐藏技术, 则要求语音信息隐藏算法具有抵抗 RPE-LTP 语音编码、噪声攻击等性能. 由于 RPE-LTP 算法采用了脉冲激励线性预测的算法进行语音编码, 对携密语音的波形进行了重构, 因此现有的很多语音隐藏算法尤其是时域波形隐藏算法在经过 GSM 中的 RPE-LTP 编码解码后不能正确解码. 为了获得一种能够和现

收稿日期: 2004-07-01; 修回日期: 2005-03-06
基金项目: 国家重点基础研究发展规划项目(No. TG1999035804); 教育部优秀青年教师资助计划项目; "十五"国家密码发展基金

有大多数 GSM 手机兼容的语音隐藏移动通信系统, 要求设计一种能够在手机 GSM 语音编码芯片之前完成信息隐藏的算法. 并且要求该算法能够抵抗 GSM 中的 RPE-LTP 编码解码.

本文提出的 AERA(ABS Energy Ratio Adjust) 算法利用 GSM 中的 RPE-LTP 编解码以后输出语音和原始语音的相邻段能量比基本保持一致的特性来进行信息隐藏. 同时采用了 ABS[5,6] (Analysis-By-Synthesis) 技术, 在隐藏算法中根据输入明文语音实时的调整相邻段能量比, 使得隐藏效果和解码效果都达到最佳值. 大量仿真试验结果表明该算法能够抵抗 GSM 中的 RPE-LTP 编码解码, 其明文语音质量下降不多. 每秒明文能够嵌入 50bit 的密文信息. 是一种简单有效的 GSM 移动通信系统语音隐藏算法.

2 GSM 编码算法特性研究

经过对大量语音信号的 RPE-LTP 编码前后特性分析研究发现, 一段语音信号经过编解码以后, 能量改变的幅度不大. 对大量语音(采样率 8kHz, 以 80 个采样点为一个分段) 经过 RPE-LTP 编解码前后的每段能量比进行了分析和统计(见表 1), 结果表明编码前后的能量比集中在 0.9~ 1.5, 尤其是 0.9 ~ 1.3 之间.

表 1 RPE-LTP 编码解码前后能量比统计表

编码前后能量比	落在该能量比的语音分段数量	占总数量的百分比
0.9- 1.1	9486	31%
1.1- 1.3	13715	45%
1.3- 1.5	4775	15%
1.5- 2.0	1725	5%
其他	1362	4%

由以上分析结果可知, 可以利用相邻语音段能量比来进行信息隐藏. 如果将相邻段能量比定位越大, 越能保证正确解码但是隐藏效果越差. 反之亦然. 可以在隐藏算法中采用 ABS 技术, 对每一段输入的明文进行 RPE-LTP 预编码, 分析出编码以后的大致能量比范围, 然后实时的调整输入信号能量比. 这样可以既获得好的隐藏效果, 又能保证解码效果.

3 信息嵌入算法

信息嵌入算法的流程框图如图 1 所示:

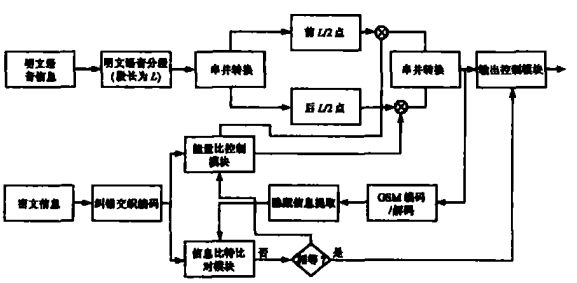

图 1 信息嵌入算法流程框图

具体实现流程如下:

(1) 将需要隐藏的信息 S 使用 Gray 码进行纠错交织编码, 形成 q 比特的的隐藏信息 X:

$$X=\{x(i),0<i<q\},\ x(i)\in\{0,1\} \tag{1}$$

(2) 将明文语音 P 进行分段, 设 $P=\{p(j),0<j<M\}$ 为 M 个样点的公开语音, 分段后表示为:

$$p(k)=p(k\times L+j),0<k<K \text{ 且 } 0<j<L \tag{2}$$

式中 $p(k)$ 表示第 k 段语音, K 表示明文语音的总段数, L 表示每段样点数. 若要保证隐藏信息 X 能够完整嵌入, 必须满足不等式: $q\leqslant L$ 或者 $q\times L\leqslant M$. 本文取 $L=160$. 将 p 分段分别计算前 $L/2$ 个点的能量和后 $L/2$ 个点的能量:

$$E_1=\sum_{j=0}^{L/2-1}(p(k\times L+j))^2 \tag{3}$$

$$E_2=\sum_{j=L/2}^{L}(p(k\times L+j))^2 \tag{4}$$

(3) 在对应段嵌入信息. 由于透明性和鲁棒性是一对矛盾, 嵌入深度 β 增加, 鲁棒性随之提高, 但势必导致携密语音 p' 质量的下降, 因此这里的 β 值将依据 ABS 算法进行自适应调整:

将 $x(i)$ 嵌入到 p 中的具体过程如下:

(1) 确定一个比较小的 β 初始值, 文中选择 β 初始值为 1.1.

(2) 计算前 $L/2$ 段的放大增益:

$$\beta 1=\begin{cases}\sqrt{\beta\times E_2/E_1}, & x(i)=1 \text{ 且 } E_1/E_2<\beta\\ 1, & \text{其他}\end{cases} \tag{5}$$

其次计算后 $L/2$ 段的放大增益:

$$\beta 2=\begin{cases}\sqrt{\beta\times E_1/E_2}, & x(i)=0 \text{ 且 } E_2/E_1<\beta\\ 1, & \text{其他}\end{cases} \tag{6}$$

(3) 最后根据 β1、β2 的值将 $x(i)$ 嵌入到明文语音 p 中

$$p'(k\times L+j)=\begin{cases}p(k\times L+j)\times\beta 1, & 0\leqslant j<L/2\\ p(k\times L+j)\times\beta 2, & L/2\leqslant j\leqslant L\end{cases} \tag{7}$$

(4) 获得携密的明文语音 p'. 再对 p' 进行 RPE-LTP 编码, 获得编码语音 p'_c.

(5) 将 p'_c 进行 RPE-LTP 解码, 获得 p'_d.

(6) 利用信息提取算法对 p'_d 进行提取, 获得提取密文信息 $x(i)'$.

(7) 如果 $x(i)=x(i)'$, 那么 p' 即为需要的携密语音. 将 p' 储存或者传输. 如果 $x(i)$ 不等于 $x(i)'$, 那么提高 β 的值(文中定 $\beta=\beta+1$), 并且跳转到(2) 步执行.

利用该算法将一段长为 4049 比特的密文信息隐藏到一段明文语音中, 表 2 显示嵌入中使用的 β 值和该 β 值使用次数.

表 2 嵌入算法中 β 值使用次数列表

使用的 β 值	β 值使用次数	β 值所占百分比
1.1	3196	78.27%
1.2	533	13.16%
1.3	213	5.26%
1.4	96	2.37%
1.5	38	0.94%

4 信息提取算法

信息提取算法的流程框图如图 2 所示:

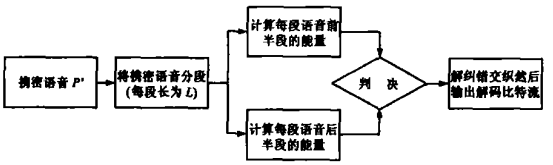

图 2　信息提取算法流程框图

在信息提取时, 首先将接收到的携密的明文语音 P' 进行分段处理, 每段段长为 L 个采样点. 分段后表示为:

$$p'(k)=p'(k\times L+j),\ 0<k<K \text{ 且 } 0<j<L \tag{8}$$

然后计算前 $L/2$ 点能量 E_1' 和后 $L/2$ 点能量 E_2':

$$E'_1=\sum_{j=0}^{L/2-1}(p'(k\times L+j))^2 \tag{9}$$

$$E'_2=\sum_{j=L/2}^{L}(p'(k\times L+j))^2 \tag{10}$$

然后进行判决获得通过信道的密文信息 X'. 判决规则如下:

$$X'(i)=\begin{cases}1, & E'_1>E'_2\\ 0, & E'_2\geqslant E'_1\end{cases} \tag{11}$$

将 X' 经过解交织解纠错编码即可得到恢复出的的密文信息.

5　性能分析

为了测试该算法的性能, 进行了各类仿真实验. 实验的数据中均采用 8kHz 采样, 16bits 量化的语音, 语音 S 共 323840 个样点. 语音段长 L 为 160 个采样点, 嵌入深度 β 初始值 1.1. 分别在无攻击和多种攻击情况下分析本文算法的性能.

5.1　性能指标

本文衡量音质的主要性能指标有归一化相关系数、分段平均信噪比[7].

其中归一化相关系数 ρ 的定义为: 对于原始序列 W 和变化后的序列 W', 有:

$$\rho(W,W')=\frac{\sum_i w(i)w'(i)}{\sqrt{\sum_i w(i)^2}\sqrt{\sum_i w'(i)^2}} \tag{12}$$

分段平均信噪比 SNR 用于衡量中间语音 P' 的客观音质, 定义为各段语音信噪比的平均值. 当 $p(i)=p'(i)$ 时, 本段 $p(i)$ 和 $p'(i)$ 信噪比的分母为 0, 此时令信噪比为 100. 由此可得分段平均信噪比 SNR 定义为:

$$\mathrm{SNR}=\frac{1}{K}\sum_{i=0}^{K-1}\mathrm{SNR} \tag{13}$$

其中

$$\mathrm{SNR}_i=\begin{cases}100, & p(i)=p'(i)\\ \dfrac{\sum_{j=0}^{L-1}p^2(j\times L-j)}{\sum_{j=0}^{L-1}[p(j\times L+i)-p'(j\times L+i)]^2}, & p(i)\neq p'(i)\end{cases} \tag{14}$$

5.2　携密语音的质量

携密语音 P' 质量的好坏反映了算法的透明性指标, 是衡量算法好坏的重要标志. 表 3 给出了本文算法在无攻击情况下的分段平均信噪比 SNR(dB) 和原始明文和携密明文归一化相关系数指标.

表 3　无攻击情况下携密语音 P' 的音质

算法	性能指标	
	分段平均信噪比 SNR(dB)	$\rho(P,P')$
固定能量比调整隐藏算法($\beta=1.5$)	34.7	0.984
AERA 算法	37.14	0.9896

从主观听觉测试结果表明, 本文所提算法获得的携密语音仅仅比原始明文略多一点噪声, 并不影响整体听觉效果. 图 3、4 给出了原始语音 P 和携密语音 P' 之间波形和频谱对比图.

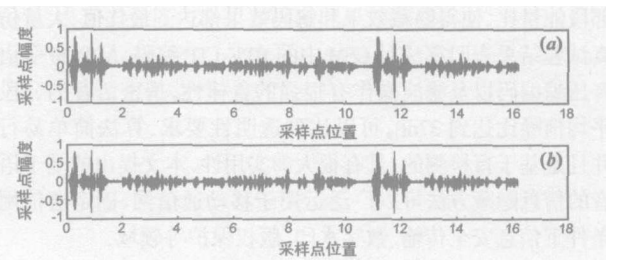

图 3 (a)携密语音 P'; (b) 原始语音 S

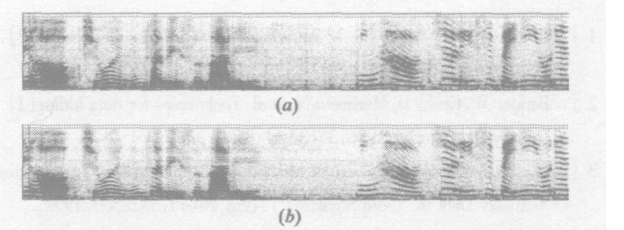

图 4 (a)携密语音 P' 的语谱图; (b) 原始语音 S 的语谱图

5.3　抗语音压缩测试

抗压缩能力是衡量算法鲁棒性的重要指标, 也是通信网络环境对实用信息隐藏技术的要求. 为了测试本文所提算法抵抗语音压缩的性能, 我们将中间语音 P' 分别使用 RPE-LTP 编码、G.721 32kb/s ADPCM 以及 G.723 24/40kb/s 的 ADPCM 编码进行压缩后再进行信息提取. 表 3 显示了经过不同语音编码算法处理以后信息提取的误码率以及经过纠错以后的误码率.

表 4　本文算法抗压缩性能测试结果

压缩算法	性能指标	
	直接提取信息的误码率	经过纠错以后提取信息误码率
GSM RPE-LTP 编码语音编码	1.7%	0
24kb/s ADPCM	1.5%	0
32kb/s ADPCM	0.45%	0
40kb/s ADPCM	0	0

5.4　抗低通滤波性能

将携密语音 P' 经过截止频率分别为 500~3500Hz 的低通滤波器, 滤波后提取隐藏信息. 试验结果表明即使 P' 在 2kHz

低通滤波的攻击下,经过纠错编码以后仍能够完全解出隐藏信息.

5.5 抗中值滤波性能

中值滤波是语音处理中常用的操作. 为了测试本文所提算法抵抗中值滤波的性能,我们将携密语音 P' 经过 3 点中值滤波后进行信息提取. 结果表明直接提取后信息的误码率为 2.66%,而经过 Gray 码纠错以后,隐藏信息能够完全恢复.

6 结论

本文基于一般语音编码对相邻段语音能量比改变不大这一特性,提出了一种能够在 GSM 移动通信网络中使用的信息隐藏算法——AERA (ABS Energy Ratio Adjust) 算法. 算法采用了 ABS 技术,在隐藏算法中根据输入明文语音实时的调整相邻段能量比,使得隐藏效果和解码效果都达到最佳值. 大量仿真试验结果表明算法对 GSM 中的 RPE-LTP 编码、ADPCM 等语音压缩编码以及滤波操作有很强的鲁棒性,携密语音的分段平均信噪比达到 37dB,可以达到透明性要求. 算法简单易行并且是基于盲检测的,具有很大的实用性. 本文提出的基于语音的信息隐藏方法可以广泛运用于移动通信网、固定通信网条件下信息安全传输、数字水印、版权保护等领域.

参考文献:

[1] Cooperman M, Moskowitz s. Steganographic Method and Device[P]. USA: Patent: 5687236, 11, 1997.

[2] Bender W, Gruhl D, Morimoto N, et al. Techniques for data hiding[J]. IBM System Journal, 1996, 35(3&4): 313- 336.

[3] Lee C, Moallemi K, Warren R. Method and Apparatus for Transporting Auxiliary Data in Audio Signals[P]. USA Patent: 5822360, 1998.

[4] Gruhl D, Bender W, Lu A. Echo hiding[A]. 1994 First International Workshop on Information Hiding, Cambridge, England[C]. Springer: Lecture Notes in Computer Science, 1996. 295- 315.

[5] Zhi-jun Wu, Wei Yang, Yi-xian Yang. ABS-based speech information hiding approach[J]. Electronics Letters, 2003, 39(22): 1617- 1619.

[6] 吴志军,钮心忻,杨义先. 语音隐藏的研究及实现[J]. 通信学报, 2002, 23(8): 99- 104.
Zhi-jun Wu, Xin-xin Niu, Yi-xian Yang. Research and Implementation for Speech Information Hiding[J]. Journal of China Institute of Communications, 2002, 23(8): 99- 104. (Chinese source)

[7] 陈亮,张雄伟. 基于语音参数模型的语音隐藏算法[J]. 计算机学报, 2003, 26(8): 974- 104.
Liang Chen, Xiong-wei Zhang. Speech Hiding Algorithm Based on Speech Parameter Mode[J]. 26(8): 974- 981. (Chinese source)

作者简介:

白 剑 男, 1979 年出生于江西, 现为北京邮电大学信息安全中心博士研究生, 研究领域包括: 语音信息隐藏, 信息安全技术, 软件无线电, 数字信号处理等. E-mail: Baijbupt@263.net.

景晓军 男, 北京市人, 副教授, 1995 年获通信与信息系统专业硕士学位, 1999 年获通信与信息系统专业博士学位, 2000 年至 2002 年在北京邮电大学从事博士后研究工作, 在国内外学术刊物和会议上发表学术论文 30 多篇, 合作出版著作二部, 负责在研项目 5 项, 主要研究方向为信息融合、模式识别、图像处理.

杨义先 男, 四川绵阳人, 博士, 北京邮电大学信息安全中心教授, 博士生导师, 长江学者特聘教授, 第九届全国政协委员, 主要研究领域为编码密码学、信息安全与网络安全、电子商务、移动通信的安全等.