



(12)发明专利申请

(10)申请公布号 CN 108962267 A
(43)申请公布日 2018.12.07

(21)申请号 201810742980.0

(22)申请日 2018.07.09

(71)申请人 成都信息工程大学

地址 610000 四川省成都市西南航空港经
济开发区学府路1段24号

(72)发明人 李孝杰 史沧红 吴锡 吕建成
王录涛 郭峰 伍贤宇 罗超

(74)专利代理机构 北京华仲龙腾专利代理事务
所(普通合伙) 11548

代理人 李静

(51)Int.Cl.

G10L 19/018(2013.01)

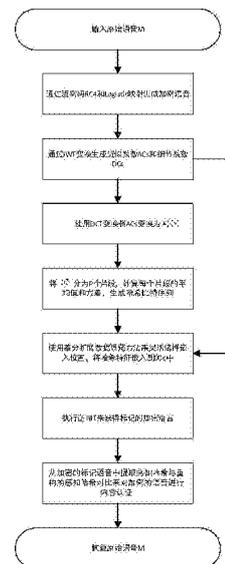
权利要求书3页 说明书9页 附图6页

(54)发明名称

一种基于哈希特征的加密语音内容认证方法

(57)摘要

本发明涉及一种基于哈希特征的加密语音内容认证方法,其包括:输入原始语音后,通过Logistic映射和流密码RC4对原始语音进行加密以生成加密语音,对加密语音进行分帧并对每帧加密语音执行整数小波变换和离散余弦变换,通过比较低频DCT系数的均值和方差来计算哈希特征,利用差分扩展将哈希特征作为水印嵌入到IWT的细节系数的高位比特中;然后对IWT近似系数和含哈希特征的细节系数执行逆IWT变换来获得含水印的加密语音,从含水印的加密语音中提取哈希特征与重构的哈希特征进行对比来对加密语音进行内容认证。本发明提高了云计算中的语音内容认证的鲁棒性,可以准确定位篡改语音帧,在实际应用中适用范围更广。



1. 一种基于哈希特征的加密语音内容认证方法,其特征在于,输入原始语音后,通过Logistic映射和流密码RC4对原始语音进行加密以生成加密语音,对加密语音进行分帧并对每帧执行整数小波变换IWT和离散余弦变换DCT,通过比较低频DCT系数的均值和方差来计算哈希特征,利用差分扩展将哈希特征作为水印嵌入到IWT的细节系数的高位比特中;然后对IWT近似系数和含哈希特征的细节系数执行逆IWT变换来获得含水印的加密语音,从含水印的加密语音中提取哈希特征与重构的哈希特征进行对比来对加密语音进行内容认证。

2. 如权利要求1所述的方法,其特征在于,包括以下步骤:

S1) 输入原始语音 $M = \{m_i, 1 \leq i \leq I\}$, 其中, $m_i \in (-32768, 32767)$;

S2) 通过流Logistic映射和密码RC4生成加密语音,该步骤包括:

S2.1) 转换一个样本值 m_i 成16位二进制 $\{v_{i,15}, v_{i,14}, \dots, v_{i,0}\}$,采用公式(1)计算,

$$v_{i,n} = \left\lfloor \frac{m_i}{2^n} \right\rfloor \bmod 2, n = 0, 1, \dots, 15 \quad (1)$$

$$\text{其中, } m_i = \sum_{n=0}^{15} v_{i,n} \cdot 2^n \quad (2)$$

S2.2) 计算加密语音样本 $V_{i,n}$,采用公式(3)计算:

$$V_{i,n} = v_{i,n} \oplus r_{i,n} \quad (3)$$

其中, $r_{i,n}$ 是以 K_{ENC} 为密钥的流密码RC4产生的二进制序列;

S2.3) 使用Logistic映射对 c_i 进行置乱来构造加扰结果, c_i 表示加密语音比特的十进制数,并且 c_i 采用公式(4)计算:

$$c_i = \sum_{n=0}^{15} V_{i,n} \cdot 2^n \quad (4)$$

S2.4) 设伪随机序列 $Y = \{y_q, 1 \leq q \leq Q\}$,其通过Logistic映射生成,Logistic映射用公式(5)来表示:

$$y_q = \rho \cdot y_{q-1} \cdot (1 - y_{q-1}), 3.5699 \leq \rho \leq 4 \quad (5)$$

设 K_{ENS} 为初始密钥,将伪随机序列 Y 按照升序排序从而得到升序序列 $y_{order(q)}$,采用公式(6)计算:

$$y_{order(q)} = \text{Sort}(y_q), q = 1, 2, \dots, Q \quad (6)$$

其中, $order(q)$ 是 q 的索引, $\text{Sort}(\cdot)$ 是排序函数;

S2.5) 使用索引 $order(q)$ 扰乱加密语音 C ,得到加扰加密语音 C' , $C' = \{c'_i, 1 \leq i \leq I\}$;

S3) 通过IWT变换生成近似系数 AC_s 和细节系数 DC_s ,其包括:

基于加扰加密语音 C' ,将 C' 分成 N 个非重叠帧,由 F 表示 $F = \{f_n | n = 1, 2, \dots, N\}$,设每帧包含 J 个样本,则 $N \cdot J = I$,其中 I 是原始语音样本的数量;

在加扰加密语音 C' 的每帧上执行 T 级IWT,将 f_n 定义为 f ,对每帧 f ,将 AC_s 定义为 $AC_s^T(b)$,将 DC_s 定义为 $DC_s^T(b)$,其中 $b = J/2, J/2^2, \dots, J/2^T, T = 1, 2, \dots$;

S4) 使用DCT变换将 $AC_s^T(b)$ 变换为 $D_s^T(b)$,其包括:

使用DCT变换将 $AC_s^T(b)$ 变换为以由 $D_s^T(b)$ 表示的特征,采用来自于 $D_s^T(d)$ 的 $2/3 \times J/2^T$ 个最低频DCT系数,定义为 D_s^{TD} ;

S5) 将 D_s^{TD} 分成 P 个片段,计算每个片段的平均值和方差,生成哈希比特序列,其包括:

S5.1) 将 D_f^{TD} 分成 P 个片段, 每个片段长度 $L = (2/3 \times J/2^1) / P$, 每个片段定义为 $D_f^{TD}(p, l)$, 其中 $p = 1, 2, \dots, P, l = 1, 2, \dots, L$; 采用公式 (7) 计算每个片段 $D_f^{TD}(p, l)$ 的平均值 $\bar{D}_f^{TD}(p)$, 然后采用公式 (8) 计算第 P 个片段的方差:

$$\bar{D}_f^{TD}(p) = \frac{1}{L} \sum_{l=1}^L D_f^{TD}(p, l) \quad (7)$$

$$O_f^{TD}(p) = \frac{1}{L} \sum_{l=1}^L [D_f^{TD}(p, l) - \bar{D}_f^{TD}(p)]^2 \quad (8)$$

其中, $O_f^{TD}(p)$ 是方差;

S5.2) 定义第 f 帧的哈希特征为 $H(f) = \{w_f^T(v) | v = 1, 2, \dots, P-1\}$, 定义

$$w_f^T(v) = \begin{cases} 1, & \text{if } \bar{D}_f^{TD}(p+1) \geq \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) \geq O_f^{TD}(p); \\ 0, & \text{if } \bar{D}_f^{TD}(p+1) \geq \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) < O_f^{TD}(p); \\ 0, & \text{if } \bar{D}_f^{TD}(p+1) < \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) \geq O_f^{TD}(p); \\ 1, & \text{if } \bar{D}_f^{TD}(p+1) < \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) < O_f^{TD}(p); \end{cases} \quad (9),$$

其中 $v \in [1, (P-1)]$ 是 f 的索引, $f \in [1, N]$, 最终的哈希特征定义为 $W(u) = \{u = 1, 2, \dots, (P-1) \times N\}$;

S5.3) 使用 Logistic 映射产生伪随机序列 D , 使用初始密钥 K_{ENW} 加密 $W(u)$; 定义 $D = \{d_r | d_r \in \{0, 1\}, r = 1, 2, \dots\}$, 其中 $d_r = \begin{cases} 0, & y_r < 1/2 \\ 1, & y_r \geq 1/2 \end{cases}$, y_r 是由 Logistic 映射生成的伪随机数, 加密后的哈希特征 $C[W(u)]$ 满足 $C[W(u)] = d_r \oplus W(u)$, 哈希特征的总长度为 $(P-1) \times N$;

S6) 用差分扩展数据隐藏方法来选择嵌入位置, 将哈希特征嵌入到 DCs 中, 其包括:

S6.1) 使用差分扩展数据隐藏方法来灵活选择嵌入位置, 对于第 f 组, 将 T 级 DCs $DC_f^T(b)$ 分为高位和低位比特; 高位比特定义为 $[H]_{2^x} = DC_f^T - (DC_f^T \bmod 2^x)$, 低位比特定义为 $[L]_{2^x} = DC_f^T - [H]_{2^x}$,

$$\text{其中 } DC_f^T = \begin{cases} [DC_f^T - (DC_f^T \bmod 2^x)] + (DC_f^T \bmod 2^x), & DC_f^T \geq 0; \\ [DC_f^T - (DC_f^T \bmod (-2^x))] + (DC_f^T \bmod (-2^x)), & \text{otherwise} \end{cases}, 2^x \text{ 是高位比特与低位比特之间的区分};$$

的区分;

S6.2) 使用与划分 C' 相同的方法将加密哈希特征 $C[W(u)]$ 划分为 N 个组, 并用 $w_f^T(B_T)$ 表示每个组, 随机选择 $P-1$ 个 T 级 DCs 的并用 \hat{DC}_f^T 表示, $b - (P-1)$ 个未选中的 T 级 DCs 由 DC_f^T 表示, 嵌入方法有溢出, 系数 \hat{DC}_f^T 使用公式 (10) 进行预处理,

$$DC_f^{\hat{T}}(v) = \lfloor DC_f^T(v) / 2 \rfloor, \quad v \in \{1, 2, \dots, (P-1) \times N\} \quad (10)$$

然后将 $(P-1)$ 个哈希比特嵌入到系数为 \hat{DC}_f^T 的 IWT 中, 使得 \hat{DC}_f^T 为嵌入后的细节系数, 采用公式 (11) 计算:

$$DC_f^{\tilde{T}} = \begin{cases} 2[H]_{2^x} + 2^x w_f^T(B_T) + [L]_{2^x}, & DC_f^{\hat{T}} \geq 0; \\ 2[H]_{2^x} - 2^x w_f^T(B_T) + [L]_{2^x}, & \text{otherwise} \end{cases} \quad (11)$$

并用 $DC_f^{\tilde{T}}$ 和 DC_f^T 替换 $DC_f^T(b)$;

S7) 重复步骤S4到S6直至哈希特征的嵌入完成, 然后对 AC_f^T 、 DC_f^T 和 \tilde{DC}_f^T 执行逆IWT来获得含水印的加密语音 \tilde{C} ;

S8) 从含水印的加密语音中提取哈希特征与重构的哈希特征来对加密的语音进行内容认证, 如果这两个特征认证距离小于某个阈值, 则认证成功; 若特征认证距离大于这个阈值, 则认证失败; 该步骤包括:

S8.1) 获得含水印的加密语音 \tilde{C} 的IWT系数 $\tilde{AC}_f(b)$ 和 $\tilde{DC}_f(b)$, 其中 $b=J/2, J/2^2, \dots, J/2^T$, $T=1, 2, \dots$;

S8.2) 定义重构的哈希特征为 $W'(u) = \{u=1, 2, \dots, (P-1) \times N\}$, 第 f 帧的哈希特征为 $H'(f) = \{w_f^T(v) | v=1, 2, \dots, P-1\}$;

S8.3) 给定 T 级DCs系数 $\tilde{DC}_f(b)$, 高位比特定义为 $[\tilde{H}]_{2^x} = DC_f^T - (DC_f^T \bmod 2^x)$, 其中 DC_f^T 采用公式(12)计算:

$$DC_f^T = \begin{cases} [DC_f^T - (DC_f^T \bmod 2^x)] + (DC_f^T \bmod 2^x), & DC_f^T \geq 0; \\ [DC_f^T - (DC_f^T \bmod (-2^x))] + (DC_f^T \bmod (-2^x)), & otherwise \end{cases} \quad (12),$$

并且 $DC_f^T(v)$ 采用公式(13)计算:

$$DC_f^T(v) = 2 \times DC_f^T(v), \quad v=1, 2, \dots, (P-1) \times N \quad (13)$$

使用密钥 K_{ENW} , 哈希特征提取满足公式(14):

$$\tilde{w}_f^T(B_f) = \frac{[\tilde{H}]_{2^x}}{2^x} \bmod 2, \quad T=1, 2, \dots \quad (14)$$

S8.4) 对每帧, 使用公式(14)提取对应的哈希比特; 定义总的哈希特征为 $\tilde{W}(u) = \{u=1, 2, \dots, (P-1) \times N\}$, 第 f 帧的哈希位为 $\tilde{H}(f) = \{\tilde{w}_f^T(v) | v=1, 2, \dots, P-1\}$;

S8.5) 对第 f 帧, 逐帧比较提取的哈希位 $\tilde{H}(f)$ 和重构的哈希位 $H'(f)$; 对于第 f 帧, 如果 $\tilde{H}(f)$ 和 $H'(f)$ 不相等比特数大于4, 则表明第 f 帧被篡改了; 如果 $\tilde{H}(f)$ 和 $H'(f)$ 不相等比特数小于等于4, 则第 f 帧是完好的;

S9) 恢复原始语音, 其包括:

S9.1) 使用公式(15)恢复系数 $DC_f^T(b)$:

$$DC_f^T = \begin{cases} \frac{[\tilde{H}]_{2^x} - 2^x \tilde{w}_f^T(B_f)}{2} + [\tilde{L}]_{2^x}, & DC_f^T \geq 0; \\ \frac{[\tilde{H}]_{2^x} + 2^x \tilde{w}_f^T(B_f)}{2} + [\tilde{L}]_{2^x}, & otherwise \end{cases} \quad (15)$$

并使用 $DC_f^T(b)$ 替换 $\tilde{DC}_f(b)$;

S9.2) 在系数 AC_f^T 和 DC_f^T 上执行逆IWT变换, 从而得到加密语音 C' ;

S9.3) 使用密钥 K_{ENS} 和 K_{ENC} 以正确解密加密语音 C' , 从而得到恢复的原始语音 M 。

一种基于哈希特征的加密语音内容认证方法

技术领域

[0001] 本发明属于加密语音内容认证领域,尤其涉及一种基于哈希特征的加密语音内容认证方法。

背景技术

[0002] 语音信号是一种重要的多媒体信号,其可以应用于例如军事指挥,司法部门记录证据和在线音频指令等场景。由于语音文件需要很大的存储空间,很多语音文件都存储在云中。然而这是不安全的,因为任何人都可以下载、读取和篡改语音文件的内容。因此,云计算中保护语音文件内容并判断语音文件的完整性非常重要。

[0003] 加密技术是语音内容保护最有效的方法之一,因为它可以将原始数据转换为不可理解的数据。为了信息安全和隐私保护,数据通常在上传和传输到云端之前进行加密。数字水印是信息安全中的一项重要技术,可以保护信息的完整性和真实性。为了提高安全性并保护用户的隐私,许多研究将加密和数字水印技术相结合。其他研究通过使用异或来加密图像并将额外数据嵌入到加密的图像中。也有研究使用Paillier同态加密来加密原始图像,并使用同态性质将数据嵌入到加密图像中。大多数数字水印方法可以在加密图像中找到。然而,这些技术还没有经过系统的研究,也未应用于加密语音。而且,用于加密语音的内容认证方案很少。

[0004] 此外,语音加密领域还存在一些限制,例如加密语音内容是随机的,原始语音的特征消失了,从原始语音中提取特征的大多数常规方法不能够直接应用于语音加密领域。

发明内容

[0005] 针对现有技术之不足,本发明提出了一种基于哈希特征的加密语音内容认证方法,其包括:输入原始语音后,通过Logistic映射和流密码RC4对原始语音进行加密以生成加密语音,对加密语音分帧并对每帧执行整数小波变换IWT和离散余弦变换DCT,通过比较低频DCT系数的均值和方差来计算哈希特征,利用差分扩展将哈希特征作为水印嵌入到IWT的细节系数的高位比特中;然后对近似系数和含水印的细节系数执行逆IWT来获得含水印的加密语音,从含水印的加密语音中提取哈希特征与重构的哈希特征对比来对加密语音进行内容认证。

[0006] 根据一个优选实施方式,本发明的加密语音内容认证方法包括以下步骤:

[0007] S1) 输入原始语音 $M = \{m_i, 1 \leq i \leq I\}$,其中, $m_i \in (-32768, 32767)$;

[0008] S2) 通过Logistic映射和流密码RC4生成加密语音,该步骤包括:

[0009] S2.1) 转换一个样本值 m_i 成16位二进制 $\{v_{i,15}, v_{i,14}, \dots, v_{i,0}\}$,采用公式(1)计算,

$$[0010] \quad v_{i,n} = \left\lfloor \frac{m_i}{2^n} \right\rfloor \bmod 2, n = 0, 1, \dots, 15 \quad (1)$$

$$[0011] \quad \text{其中, } m_i = \sum_{n=0}^{15} v_{i,n} \cdot 2^n \quad (2)$$

[0012] S2.2) 计算加密语音样本 $V_{i,n}$,采用公式(3)计算:

$$[0013] \quad V_{i,n} = v_{i,n} \oplus r_{i,n} \quad (3)$$

[0014] 其中, $r_{i,n}$ 是以 K_{ENC} 为密钥的流密码RC4产生的二进制序列;

[0015] S2.3) 使用Logistic映射对 c_i 进行置乱来构造加扰结果, c_i 表示加密语音比特的十进制数,并且 c_i 采用公式(4)计算:

$$[0016] \quad c_i = \sum_{n=0}^{15} V_{i,n} \cdot 2^n \quad (4)$$

[0017] S2.4) 设伪随机序列 $Y = \{y_q, 1 \leq q \leq Q\}$,其通过Logistic映射计算得到,Logistic映射用公式(5)来表示:

$$[0018] \quad y_q = \rho \cdot y_{q-1} \cdot (1 - y_{q-1}), 3.5699 \leq \rho \leq 4 \quad (5)$$

[0019] 设 K_{ENS} 为初始密钥,将伪随机序列 Y 按照升序排序从而得到升序序列 $y_{order(q)}$,采用公式(6)计算:

$$[0020] \quad y_{order(q)} = \text{Sort}(y_q), q = 1, 2, \dots, Q \quad (6),$$

[0021] 其中, $order(q)$ 是 q 的索引, $\text{Sort}(\cdot)$ 是排序函数;

[0022] S2.5) 使用索引 $order(q)$ 置乱加密语音 C ,得到置乱的加密语音 C' , $C' = \{c'_i, 1 \leq i \leq I\}$;

[0023] S3) 对 C' 进行认证,其包括:

[0024] 基于置乱的加密语音 C' ,将 C' 分成 N 个非重叠帧,由 F 表示 $F = \{f_n | n = 1, 2, \dots, N\}$,设每帧包含 J 个样本,则 $N \cdot J = I$,其中 I 是原始语音样本的数量;

[0025] 在加扰加密语音 C' 的每帧上执行 T 级IWT,将 f_n 定义为 f ,对每帧 f ,将 AC_s 定义为 $AC_f^T(b)$,将 DC_s 定义为 $DC_f^T(b)$,其中 $b = J/2, J/2^2, \dots, J/2^T, T = 1, 2, \dots$;

[0026] S4) 使用DCT变换将 $AC_f^T(b)$ 变换为 $D_f^T(b)$,其包括:

[0027] 使用DCT变换将 $AC_f^T(b)$ 变换为以由 $D_f^T(b)$ 表示的特征,采用来自于 $D_f^T(b)$ 的 $2/3 \times J/2^T$ 个最低频DCT系数,定义为 D_f^{TD} ;

[0028] S5) 将 D_f^{TD} 分成 P 个片段,计算每个片段的均值和方差,生成哈希比特序列,其包括:

[0029] S5.1) 将 D_f^{TD} 分成 P 个片段,每个片段长度 $L = (2/3 \times J/2^T) / P$,每个片段定义为 $D_f^{TD}(p,l)$,其中 $p = 1, 2, \dots, P, l = 1, 2, \dots, L$;采用公式(7)计算每个片段 $D_f^{TD}(p,l)$ 的平均值 $\bar{D}_f^{TD}(p)$,然后采用公式(8)计算第 P 个片段的方差:

$$[0030] \quad \bar{D}_f^{TD}(p) = \frac{1}{L} \sum_{l=1}^L D_f^{TD}(p,l) \quad (7)$$

$$[0031] \quad O_f^{TD}(p) = \frac{1}{L} \sum_{l=1}^L [D_f^{TD}(p,l) - \bar{D}_f^{TD}(p)]^2 \quad (8)$$

[0032] 其中, $O_f^{TD}(p)$ 是方差;

[0033] S5.2) 定义第 f 帧的哈希比特为 $H(f) = \{w_f^v(v) | v = 1, 2, \dots, P-1\}$,定义

$$[0034] \quad w_f^T(v) = \begin{cases} 1, & \text{if } \bar{D}_f^{TD}(p+1) \geq \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) \geq O_f^{TD}(p); \\ 0, & \text{if } \bar{D}_f^{TD}(p+1) \geq \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) < O_f^{TD}(p); \\ 0, & \text{if } \bar{D}_f^{TD}(p+1) < \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) \geq O_f^{TD}(p); \\ 1, & \text{if } \bar{D}_f^{TD}(p+1) < \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) < O_f^{TD}(p); \end{cases} \quad (9),$$

[0035] 其中 $v \in [1, (P-1)]$ 是 f 的索引, $f \in [1, N]$, 最终的哈希特征定义为 $W(u) = \{u=1, 2, \dots, (P-1) \times N\}$;

[0036] S5.3) 使用 Logistic 映射产生伪随机序列 D , 使用初始密钥 K_{ENW} 加密 $W(u)$; 定义 $D = \{d_r | d_r \in \{0, 1\}, r=1, 2, \dots\}$, 其中 $d_r = \begin{cases} 0, & y_r < 1/2 \\ 1, & y_r \geq 1/2 \end{cases}$, y_r 是由 Logistic 映射生成的伪随机数, 加密后的哈希特征 $C[W(u)]$ 满足 $C[W(u)] = d_r \oplus W(u)$, 哈希特征的总长度为 $(P-1) \times N$;

[0037] S6) 用差分扩展数据隐藏方法来选择嵌入位置, 将哈希特征嵌入到 DCs 中, 其包括:

[0038] S6.1) 使用差分扩展数据隐藏方法来灵活选择嵌入位置, 对于第 f 组, 将 T 级 DCs $DC_f^T(b)$ 分为高位比特和低位比特; 高位比特定义为 $[H]_{2^x} = DC_f^T - (DC_f^T \bmod 2^x)$, 低位比特定义为 $[L]_{2^x} = DC_f^T - [H]_{2^x}$,

[0039] 其中 $DC_f^T = \begin{cases} [DC_f^T - (DC_f^T \bmod 2^x)] + (DC_f^T \bmod 2^x), & DC_f^T \geq 0; \\ [DC_f^T - (DC_f^T \bmod (-2^x))] + (DC_f^T \bmod (-2^x)), & otherwise \end{cases}$, 2^x 是高阶与低阶之间的区分;

[0040] S6.2) 使用与划分 C' 相同的方法将加密哈希特征 $C[W(u)]$ 划分为 N 个组, 并用 $w_f^T(B_r)$ 表示每个组, 随机选择 $P-1$ 个 T 级 DCs 的并用 \hat{DC}_f^T 表示, $b - (P-1)$ 个未选中的 T 级 DCs 由 \tilde{DC}_f^T 表示, 嵌入方法有溢出, 系数 \hat{DC}_f^T 使用公式 (10) 进行预处理,

$$[0041] \quad \hat{DC}_f^T(v) = \lfloor \hat{DC}_f^T(v) / 2 \rfloor, \quad v \in \{1, 2, \dots, (P-1) \times N\} \quad (10)$$

[0042] 然后将 $P-1$ 个哈希比特嵌入到系数为 \hat{DC}_f^T 中, 使得 \tilde{DC}_f^T 为嵌入后的 DCs, 采用公式 (11) 计算:

$$[0043] \quad \tilde{DC}_f^T = \begin{cases} 2[H]_{2^x} + 2^x w_f^T(B_r) + [L]_{2^x}, & \hat{DC}_f^T \geq 0; \\ 2[H]_{2^x} - 2^x w_f^T(B_r) + [L]_{2^x}, & otherwise \end{cases} \quad (11)$$

[0044] 并用 \tilde{DC}_f^T 和 \hat{DC}_f^T 替换 $DC_f^T(b)$;

[0045] S7) 重复步骤 S4 到 S6 直至哈希特征的嵌入完成, 然后对 \hat{AC}_f^T 、 \tilde{DC}_f^T 和 \hat{DC}_f^T 上执行逆 IWT 来获得含水印的加密语音 \tilde{C} ;

[0046] S8) 从含水印的加密语音中提取哈希特征与重构的哈希特征来对加密的语音进行内容认证, 如果特征认证距离小于某个阈值, 则认证成功; 若特征认证距离大于这个阈值, 则认证失败; 该步骤包括:

[0047] S8.1) 获得标记的加密语音 \tilde{C} 的系数 $\tilde{AC}_f(b)$ 和 $\tilde{DC}_f(b)$, 其中 $b = J/2, J/2^2, \dots, J/2^T$, $T = 1, 2, \dots$;

[0048] S8.2) 定义重构的哈希特征为 $W'(u) = \{u=1, 2, \dots, (P-1) \times N\}$, 第 f 帧的哈希位为

$H'(f) = \{w_f^T(v) | v=1, 2, \dots, P-1\}$;

[0049] S8.3) 给定T级DCs系数 $\tilde{DC}_f(b)$, 高位比特定义为 $[\tilde{H}]_{2^x} = DC_f^T - (DC_f^T \bmod 2^x)$, 其中 DC_f^T 采用公式(12)计算:

$$[0050] \quad DC_f^T = \begin{cases} [DC_f^T - (DC_f^T \bmod 2^x)] + (DC_f^T \bmod 2^x), & DC_f^T \geq 0; \\ [DC_f^T - (DC_f^T \bmod (-2^x))] + (DC_f^T \bmod (-2^x)), & otherwise \end{cases} \quad (12),$$

[0051] 并且 $DC_f^T(v)$ 采用公式(13)计算:

$$[0052] \quad DC_f^T(v) = 2 \times DC_f^T(v), \quad v=1, 2, \dots, (P-1) \times N \quad (13)$$

[0053] 使用密钥 K_{ENW} , 哈希特征满足公式(14):

$$[0054] \quad \tilde{w}_f^T(B_T) = \frac{[\tilde{H}]_{2^x}}{2^x} \bmod 2, \quad T=1, 2, \dots \quad (14)$$

[0055] S8.4) 对每帧, 使用公式(14)提取对应的哈希位; 定义总的哈希特征为 $\tilde{w}(u) = \{u=1, 2, \dots, (P-1) \times N\}$, 第f帧的哈希位为 $\tilde{H}(f) = \{\tilde{w}_f^T(v) | v=1, 2, \dots, P-1\}$;

[0056] S8.5) 对第f帧, 逐帧比较提取的哈希位 $\tilde{H}(f)$ 和重构的哈希位 $H'(f)$; 对于第f帧, 如果 $\tilde{H}(f)$ 和 $H'(f)$ 不相等比特数大于4, 则表明第f帧被篡改了; 如果 $\tilde{H}(f)$ 和 $H'(f)$ 不相等比特数小于等于4, 则第f帧是完好的;

[0057] S9) 恢复原始语音, 其包括:

[0058] S9.1) 使用公式(15)恢复系数 $DC_f^T(b)$:

$$[0059] \quad DC_f^T = \begin{cases} \frac{[\tilde{H}]_{2^x} - 2^x \tilde{w}_f^T(B_T)}{2} + [\tilde{L}]_{2^x}, & DC_f^T \geq 0; \\ \frac{[\tilde{H}]_{2^x} + 2^x \tilde{w}_f^T(B_T)}{2} + [\tilde{L}]_{2^x}, & otherwise \end{cases} \quad (15)$$

[0060] 并使用 $DC_f^T(b)$ 替换 $\tilde{DC}_f(b)$;

[0061] S9.2) 对系数 AC_f^T 和 DC_f^T 执行逆IWT变换, 从而得到加密语音 C' ;

[0062] S9.3) 使用密钥 K_{ENS} 和 K_{ENC} 解密加密语音 C' , 从而得到恢复的原始语音 M 。

[0063] 本发明具有以下有益效果:

[0064] 本发明提出了一种基于哈希特征的加密语音内容认证方法, 通过对加密语音执行整数小波变换IWT和离散余弦变换DCT, 可以计算鲁棒的哈希特征并嵌入到高位比特中来完成语音内容认证。本发明具有较高的安全性, 可以准确定位篡改语音帧, 并且对一些常见的信号处理操作具有很强的鲁棒性。此外, 本发明提高了云计算中的语音内容认证的鲁棒性, 在一些常见信号处理操作情况下, 能对篡改语音帧进行精确定位, 在实际应用中适用范围更广。

附图说明

[0065] 图1示出了本发明加密语音内容认证方法的流程图;

[0066] 图2示出了原始和加密语音的直方图;

[0067] 图3示出了原始和加密语音的语谱;

- [0068] 图4示出了嵌入在不同位置的水印的SNR和SNRseg值；
 [0069] 图5示出了原始语音信号、具有散列特征的加密语音及解密语音的波形图；
 [0070] 图6示出了一个加密语音帧的假拒绝概率；
 [0071] 图7示出了经过插入攻击的加密语音的检测位置结果；
 [0072] 图8示出了经过替换攻击的加密语音的检测位置结果；
 [0073] 图9示出了经过删除攻击的加密语音的检测位置结果。

具体实施方式

[0074] 为使本发明的目的、技术方案和优点更加清楚明了，下面结合具体实施方式并参照附图，对本发明进一步详细说明。应该理解，这些描述只是示例性的，而并非要限制本发明的范围。此外，在以下说明中，省略了对公知结构和技术的描述，以避免不必要地混淆本发明的概念。

[0075] 如图1所示，本发明的基于哈希特征的加密语音内容认证方法包括以下步骤：

[0076] S1) 输入原始语音 $M = \{m_i, 1 \leq i \leq I\}$ ，其中， $m_i \in (-32768, 32767)$ ；

[0077] S2) 通过流密码RC4和Logistic映射生成加密语音，该步骤包括：

[0078] S2.1) 转换一个样本值 m_i 成16位二进制 $\{v_{i,15}, v_{i,14}, \dots, v_{i,0}\}$ ，采用公式 (1) 计算，

$$[0079] \quad v_{i,n} = \left\lfloor \frac{m_i}{2^n} \right\rfloor \bmod 2, n = 0, 1, \dots, 15 \quad (1)$$

$$[0080] \quad \text{其中, } m_i = \sum_{n=0}^{15} v_{i,n} \cdot 2^n \quad (2)$$

[0081] S2.2) 计算加密语音样本 $V_{i,n}$ ，采用公式 (3) 计算：

$$[0082] \quad V_{i,n} = v_{i,n} \oplus r_{i,n} \quad (3)$$

[0083] 其中， $r_{i,n}$ 是以 K_{ENC} 为密钥的流密码RC4产生的二进制序列。如果原始语音样本足够多，那么加密样本的可能性就更大。这使得加密结果足够安全。

[0084] S2.3) 设 c_i 表示加密语音比特的十进制数，并使用Logistic映射对 c_i 进行置乱来构造加扰结果， c_i 采用公式 (4) 计算：

$$[0085] \quad c_i = \sum_{n=0}^{15} V_{i,n} \cdot 2^n \quad (4)$$

[0086] S2.4) 设伪随机序列 $Y = \{y_q, 1 \leq q \leq Q\}$ 通过Logistic映射计算可得，Logistic映射用公式 (5) 表示：

$$[0087] \quad y_q = \rho \cdot y_{q-1} \cdot (1 - y_{q-1}), 3.5699 \leq \rho \leq 4 \quad (5)$$

[0088] 设 K_{ENS} 为初始值，也就是密钥，将伪随机序列 Y 按照升序排序，得到升序序列 $y_{\text{order}(q)}$ ，采用公式 (6) 计算：

$$[0089] \quad y_{\text{order}(q)} = \text{Sort}(y_q), q = 1, 2, \dots, Q \quad (6),$$

[0090] 其中， $\text{order}(q)$ 是 q 的索引， $\text{Sort}(\cdot)$ 是排序函数，

[0091] S2.5) 使用索引 $\text{order}(q)$ 扰乱加密语音 C ，得到加扰结果 C' ， $C' = \{c'_i, 1 \leq i \leq I\}$ 。

[0092] S3) 基于加扰加密语音 C' ，将 C' 分成 N 个非重叠帧，由 F 表示 $F = \{f_n | n = 1, 2, \dots, N\}$ 。设每帧包含 J 个样本，则 $N \cdot J = I$ ，其中 I 是原始语音样本的数量。

[0093] 在 C' 的每帧上执行 T 级 IWT，将 f_n 定义为 f ，对每帧 f ，将 ACs 定义为 $AC'_T(b)$ ，将 DCs 定

义为 $DC_f^T(b)$,其中 $b=J/2, J/2^2, \dots, J/2^T, T=1, 2, \dots$;

[0094] S4) 使用DCT变换将 $AC_f^T(b)$ 变换为 $D_f^T(b)$,其包括:

[0095] 使用DCT变换,将 $AC_f^T(b)$ 变换为以由 $D_f^T(b)$ 表示的特征,采用来自于 $D_f^T(d)$ 的 $2/3 \times J/2^T$ 个最低频DCT系数,定义为 D_f^{TD} 。

[0096] S5) 将 D_f^{TD} 分成P个片段,计算每个片段的平均值和方差,生成哈希比特序列,其包括:

[0097] S5.1) 将 D_f^{TD} 分成P个片段,每个片段长度 $L = (2/3 \times J/2^T) / P$,每个片段定义为 $D_f^{TD}(p,l)$,其中 $p=1, 2, \dots, P, l=1, 2, \dots, L$ 。计算每个片段 $D_f^{TD}(p,l)$ 的平均值得到 $\bar{D}_f^{TD}(p)$,采用公式(7)计算;然后计算第P个片段的方差,采用公式(8)计算:

$$[0098] \quad \bar{D}_f^{TD}(p) = \frac{1}{L} \sum_{l=1}^L D_f^{TD}(p,l) \quad (7)$$

$$[0099] \quad O_f^{TD}(p) = \frac{1}{L} \sum_{l=1}^L [D_f^{TD}(p,l) - \bar{D}_f^{TD}(p)]^2 \quad (8)$$

[0100] 其中, $O_f^{TD}(p)$ 是方差。

[0101] S5.2) 定义第f帧的哈希比特为 $H(f) = \{w_f^T(v) | v=1, 2, \dots, P-1\}$,定义

$$[0102] \quad w_f^T(v) = \begin{cases} 1, & \text{if } \bar{D}_f^{TD}(p+1) \geq \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) \geq O_f^{TD}(p); \\ 0, & \text{if } \bar{D}_f^{TD}(p+1) \geq \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) < O_f^{TD}(p); \\ 0, & \text{if } \bar{D}_f^{TD}(p+1) < \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) \geq O_f^{TD}(p); \\ 1, & \text{if } \bar{D}_f^{TD}(p+1) < \bar{D}_f^{TD}(p) \text{ and } O_f^{TD}(p+1) < O_f^{TD}(p); \end{cases} \quad (9)$$

[0103] 其中 $v \in [1, (P-1)]$ 是f的索引, $f \in [1, N]$ 。因此,最终的哈希比特序列具有 $(P-1) \times N$ 个二进制比特。这是加密语音片段基于内容的特征。最终的哈希特征定义为 $W(u) = \{u=1, 2, \dots, (P-1) \times N\}$ 。

[0104] S5.3) 使用Logistic映射产生伪随机序列D,使用初始密钥 K_{ENW} 加密 $W(u)$ 。定义 $D = \{d_r | d_r \in \{0, 1\}, r=1, 2, \dots\}$,其中 $d_r = \begin{cases} 0, & y_r < 1/2 \\ 1, & y_r \geq 1/2 \end{cases}$, y_r 是由Logistic映射生成的伪随机数。加密后的哈希特征 $C[W(u)]$ 满足 $C[W(u)] = d_r \oplus W(u)$,哈希特征的总长度为 $(P-1) \times N$ 。

[0105] S6) 用差分扩展数据隐藏方法来选择嵌入位置,将哈希特征嵌入到DCs中,其包括:

[0106] S6.1) 使用差分扩展数据隐藏方法来灵活选择嵌入位置。在此方案中,嵌入位置位于高位比特中,其比最低有效位(LSB)更鲁棒。对于第f组,将T级DCs $DC_f^T(b)$ 分为高位比特和低位比特。高位比特定义为 $[H]_{2^x} = DC_f^T - (DC_f^T \bmod 2^x)$,低位比特定义为 $[L]_{2^x} = DC_f^T - [H]_{2^x}$,

$$[0107] \quad \text{其中 } DC_f^T = \begin{cases} \left[\left[DC_f^T - (DC_f^T \bmod 2^x) \right] + (DC_f^T \bmod 2^x), & DC_f^T \geq 0; \\ \left[DC_f^T - (DC_f^T \bmod (-2^x)) \right] + (DC_f^T \bmod (-2^x)), & \text{otherwise} \end{cases}, 2^x \text{ 是高阶与低阶之间的区}$$

分。

[0108] S6.2) 使用与划分 C' 相同的方法将加密哈希特征 $C[W(u)]$ 划分为N个组,并用 $w_f^T(B_r)$ 表示每个组。随机选择P-1个DCs并用 DC_f^T 表示, $b - (P-1)$ 个未选中的DCs由 DC_f^T 表示。嵌入方

法有溢出,系数 DC_f^T 使用公式(10)进行预处理,

$$[0109] \quad DC_f^T(v) = \lfloor DC_f^T(v)/2 \rfloor, \quad v \in \{1, 2, \dots, (P-1) \times N\} \quad (10)$$

[0110] 然后将 $P-1$ 个哈希比特嵌入到系数 DC_f^T 中,使得 DC_f^T 为嵌入后的DCs,采用公式(11)计算:

$$[0111] \quad DC_f^T = \begin{cases} 2[H]_{2^x} + 2^x w_f^T(B_T) + [L]_{2^x}, & DC_f^T \geq 0; \\ 2[H]_{2^x} - 2^x w_f^T(B_T) + [L]_{2^x}, & otherwise \end{cases} \quad (11)$$

[0112] 并用 DC_f^T 和 DC_f^T 替换 $DC_f^T(b)$ 。

[0113] S7) 重复步骤S4到S6直至哈希特征的嵌入完成,然后在 AC_f^T 、 DC_f^T 和 DC_f^T 上执行逆IWT来获得标记的加密语音 \tilde{C} 。

[0114] S8) 从标记的加密语音中提取的哈希特征与重构的哈希特征进行对比来对加密的语音进行内容认证,如果特征认证距离小于某个阈值,认证成功,反之,认证失败。设 $\tilde{C} = \{\tilde{c}_i | i=1, 2, \dots, I\}$ 表示正在检测的加密标记语音,认证步骤如下:

[0115] S8.1) 获得加密标记语音 \tilde{C} 的系数 $\tilde{AC}_f(b)$ 和 $\tilde{DC}_f(b)$,其中 $b = J/2, J/2^2, \dots, J/2^T, T = 1, 2, \dots$;

[0116] S8.2) 定义重构的哈希特征为 $W'(u) = \{u=1, 2, \dots, (P-1) \times N\}$,第 f 帧的哈希比特为 $H'(f) = \{w_f^T(v) | v=1, 2, \dots, P-1\}$;

[0117] S8.3) 给定 T 级DCs系数 $\tilde{DC}_f(b)$,高为比特定义为 $[\tilde{H}]_{2^x} = DC_f^T - (DC_f^T \bmod 2^x)$,其中 DC_f^T 采用公式(12)计算:

$$[0118] \quad DC_f^T = \begin{cases} [DC_f^T - (DC_f^T \bmod 2^x)] + (DC_f^T \bmod 2^x), & DC_f^T \geq 0; \\ [DC_f^T - (DC_f^T \bmod (-2^x))] + (DC_f^T \bmod (-2^x)), & otherwise \end{cases} \quad (12),$$

[0119] 并且 $DC_f^T(v)$ 采用公式(13)计算:

$$[0120] \quad DC_f^T(v) = 2 \times DC_f^T(v), \quad v=1, 2, \dots, (P-1) \times N \quad (13)$$

[0121] 使用密钥 K_{ENW} ,哈希特征满足公式(14):

$$[0122] \quad \tilde{w}_f^T(B_T) = \frac{[\tilde{H}]_{2^x}}{2^x} \bmod 2, \quad T=1, 2, \dots \quad (14)$$

[0123] S8.4) 对每帧,使用公式(14)提取对应的哈希比特。定义总的哈希特征为 $\tilde{W}(u) = \{u=1, 2, \dots, (P-1) \times N\}$,第 f 帧的哈希比特为 $\tilde{H}(f) = \{\tilde{w}_f^T(v) | v=1, 2, \dots, P-1\}$ 。

[0124] S8.5) 对第 f 帧,逐帧比较提取的哈希特征 $\tilde{H}(f)$ 和重构的哈希特征 $H'(f)$ 。对于对第 f 帧,如果 $\tilde{H}(f)$ 和 $H'(f)$ 不相等的比特大于4,则表明第 f 帧被篡改了,否则,第 f 帧是完好的。

[0125] S9) 恢复原始语音,其包括:

[0126] S9.1) 使用公式(15)恢复系数 $DC_f^T(b)$;

$$[0127] \quad DC_f^T = \begin{cases} \frac{[\tilde{H}]_{2^x} - 2^x \tilde{w}_f^T(B_T)}{2} + [\tilde{L}]_{2^x}, & DC_f^T \geq 0; \\ \frac{[\tilde{H}]_{2^x} + 2^x \tilde{w}_f^T(B_T)}{2} + [\tilde{L}]_{2^x}, & otherwise \end{cases} \quad (15)$$

[0128] 并使用 $DC_f^T(b)$ 替换 $\tilde{DC}_f(b)$;

[0129] S9.2) 为了得到加密语音 C' , 对系数 AC_f^T 和 DC_f^T 执行逆IWT变换;

[0130] S9.3) 使用密钥 K_{ENS} 和 K_{ENC} , 解密加密语音 C' , 得到恢复的原始语音 M 。

[0131] 针对现有云存储中加密算法存在的不足, 本发明基于流密码RC4和Logistic映射对原始语音进行加密, 对加密语音执行整数小波变换 (IWT) 和离散余弦变换 (DCT), 通过比较低频DCT系数的均值和方差来计算哈希特征, 利用差分扩展将哈希特征嵌入到IWT的细节系数的高位比特中。另外本发明提高了云存储中的语音内容认证的鲁棒性, 篡改的加密语音在一些常见信号处理操作下仍能精确的被定位, 在实际应用中适用范围更广。

[0132] 图2示出了对原始语音和加密语音进行直方图分析而得到的直方图。从图2中的(c)图可以看出, 加密语音近似于白噪声, 并且两个直方图没有明显的函数分布关系, 证明本发明的方案具有优越的混淆和扩散特性。

[0133] 图3是原始和加密语音的频谱图, 一个好的数据隐藏方案应该具有良好的不可理解性。

[0134] 为了评估所提出的方案的不可理解性, 应用信噪比 (SNR) 和分段信噪比 (SNRseg) 来测量具有哈希特征的解密语音的质量。

[0135] 图4示给出了嵌入在不同比特位置的水印的SNR和SNRseg值。如图4所示, 水印嵌入的位置越高, SNR和SNRseg值越小。

[0136] 图5(a)是原始语音信号的波形、图5(b)是含有哈希特征的加密语音、图5(c)是含有哈希特征的解密语音的波形。具有哈希特征的加密语音也可以直接解密, 含有哈希特征的解密语音仍具有高质量 ($X=3$), 其中 A 为 34.9498。可以看出, 图5(a)和图5(c)之间没有明显的差异, 这意味着嵌入的哈希特征对原始语音没有显著影响。

[0137] 图6是一个加密语音帧的错误拒绝概率, 当FRP较小时, 认证的性能更好。从图6可以看出, 当哈希比特数大于9时, FRP趋于零。在所提出的方案中, 每个单帧有32个还行比特, 足以满足FRP的要求。

[0138] 图7是经过插入攻击的加密语音的检测结果, 插入攻击涉及从另一个语音信号的采样插入到加密语音。对于我们的插入攻击实验, 从第5121到第5140位和从第10241到第10255位的采样点被插入了另一语音信号的采样值, 如图7(a)所示。图7(b)中的结果显示只有两个语音帧被篡改, 这是因为只被插入了两帧语音采样值。可以看出, 本发明所提出的方案可以准确检测插入攻击。

[0139] 图8是经过替换攻击的加密语音的检测结果, 替换攻击涉及用加密的语音信号本身采样替换加密的语音。对于替换攻击实验, 从第5121到第5140位和从第10241到第10255位的采样点被替换并示于图8(a)中。图8(b)中的结果显示了第21帧和第41帧中的篡改。因此, 可以得出结论, 本发明方法能够准确地检测和定位篡改。

[0140] 图9是经过删除攻击的加密语音的检测结果, 删除攻击涉及删除部分加密语音。对于我们的删除攻击实验, 从第5121到第5140位以及从第10241到第10255位的采样被删除。

定位结果如图9 (b) 所示,从这些结果中,可以看出第21帧和第41帧已被篡改。因此,可以得出结论,本发明方法能够准确地检测和定位篡改。

[0141] 在实验中,使用归一化互相关系数 (NC) 和误码率 (BER) 来测试所提出的方案的鲁棒性,在这个框架中,具有散列特征的加密语音经受一些常见的信号处理操作。在[11]中,通过量化把数据嵌入在到离散小波变换DWT和DCT混合域中。表1显示了各种攻击下的含水印的加密语音的NC和BER值。通过比较NC和BER值,可以验证本发明所提出的方案对增加噪声和重量化具有鲁棒性。其中,Proposed scheme表示本发明所提出的方案,Ref. [11]表示对比方案。

[0142] 表1

CSP	Proposed scheme	Ref.[11]	Proposed scheme	Ref.[11]
	Speech signal type1		Speech signal type2	
	NC		BER	
Not attack	1/1	1/1	1/1	1/1
Add noise (65dB)	0.9062/0.9962	0.5032	0.0955/0.0038	0.4917
Add noise (52dB)	0.9036/0.9958	0.5114	0.0969/0.0042	0.4867
Low-pass filtering (3.5kHz)	0.4877/0.5042	0.5008	0.5094/0.4938	0.4958
Low-pass filtering (3.95kHz)	0.5060/0.5001	0.5037	0.4951/0.5065	0.5005
Re-quantification (16bit-12bit-16bit)	0.6550/0.9891	0.4990	0.3424/0.0106	0.4962
Re-quantification (16bit-8bit-16bit)	0.5055/0.8419	0.5053	0.5027/0.1583	0.4976

[0144] 此外,还将本发明的方案与语音哈希算法进行了比较。表2展示出了在各种信号处理操作下含有哈希特征的加密语音的BER值。通过比较BER值,可以验证本发明提出的方案对于增加噪声和缩放更加鲁棒。其中,Proposed表示本发明所提出的方案,Ref. [15]表示对比方案。

[0145] 表2

CSP	Proposed	Ref.[15]
	Speech signal type I	
	BER	
Not attack	1	1
Add noise	0.0001	0.2450
Low-pass filtering	0.0393	—
Scaling 2%	0.0001	0.3830
Scaling -2%	0.0001	0.4380
Re-quantification	0.0001	—
Re-sampling	—	0.0320

[0147] 需要注意的是,上述具体实施例是示例性的,本领域技术人员可以在本发明公开内容的启发下想出各种解决方案,而这些解决方案也都属于本发明的公开范围并落入本发明的保护范围之内。本领域技术人员应该明白,本发明说明书及其附图均为说明性而并非构成对权利要求的限制。本发明的保护范围由权利要求及其等同物限定。

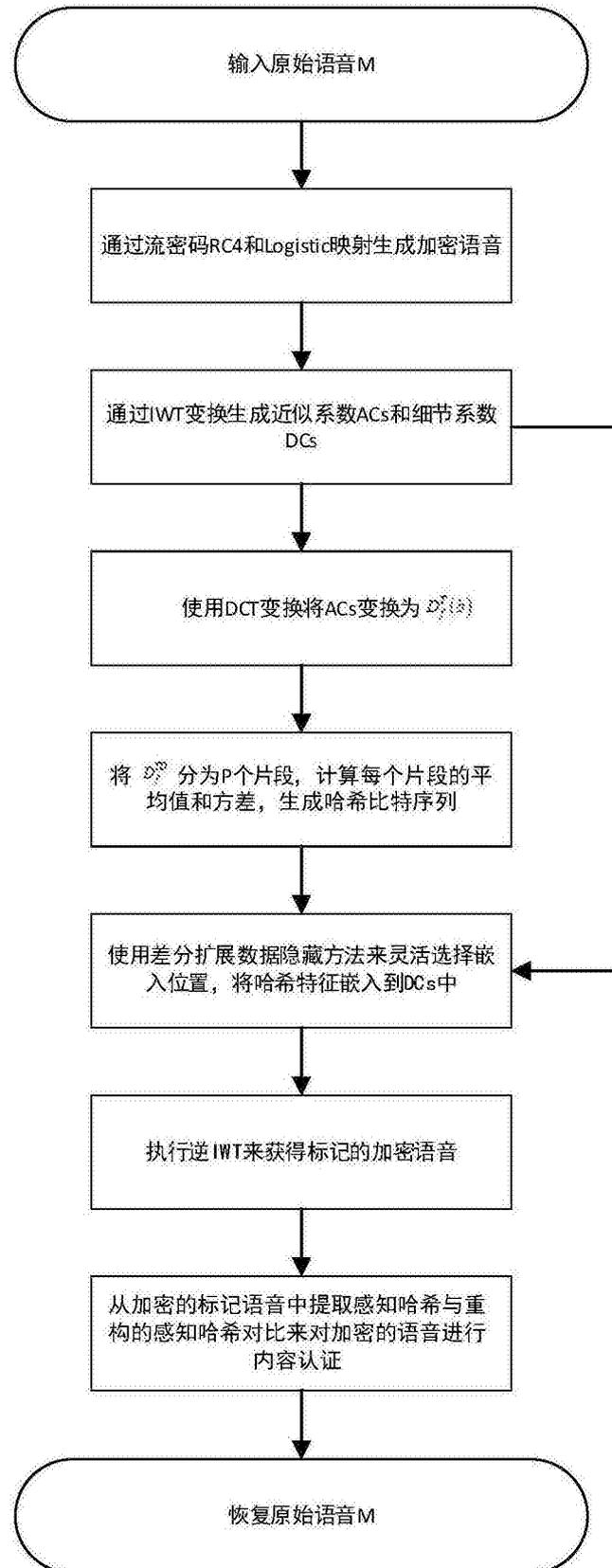


图1

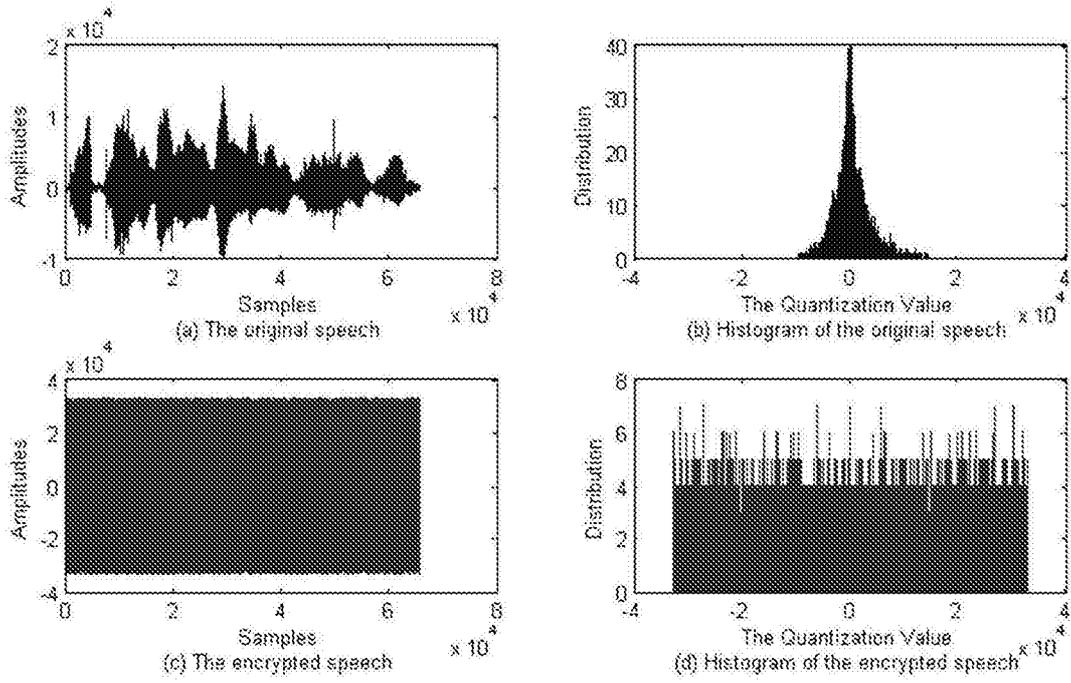


图2

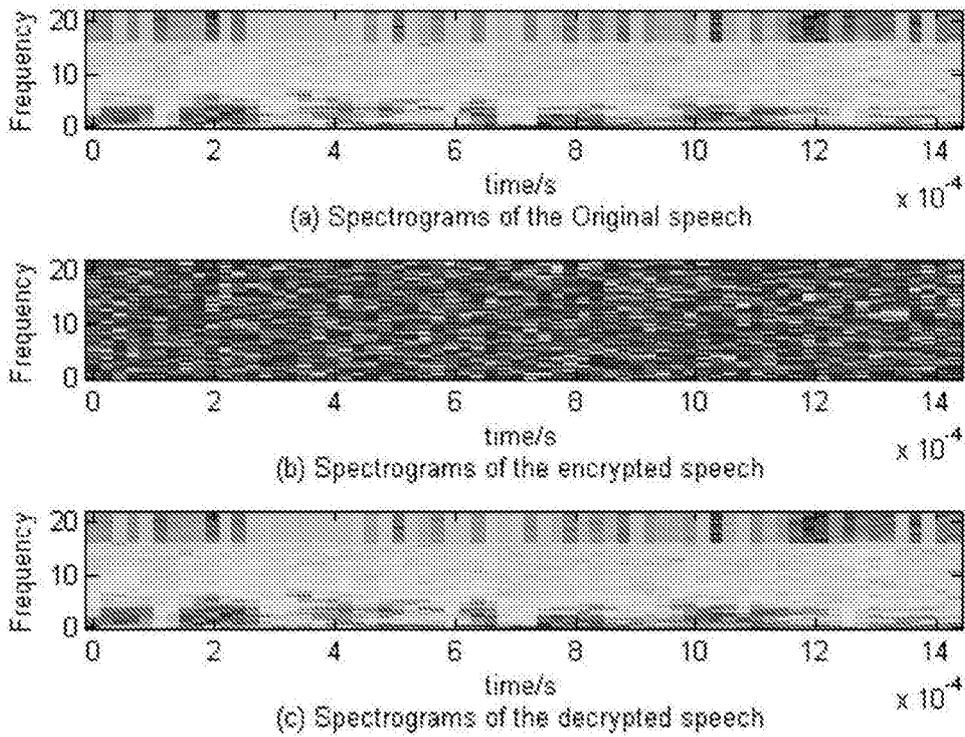


图3

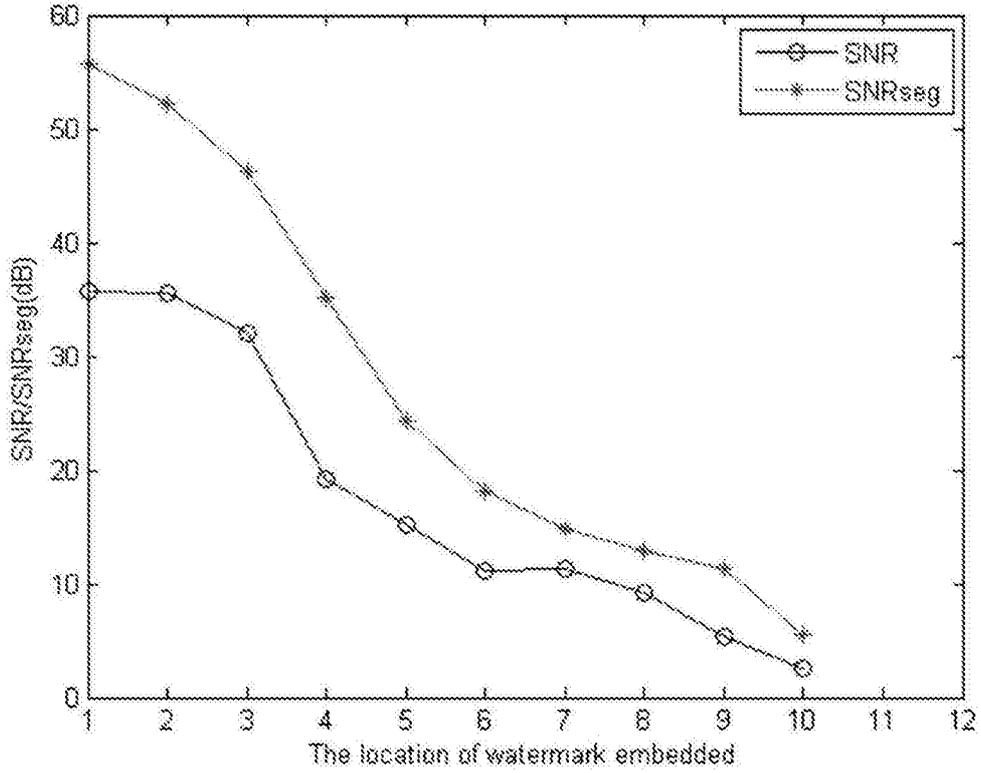


图4

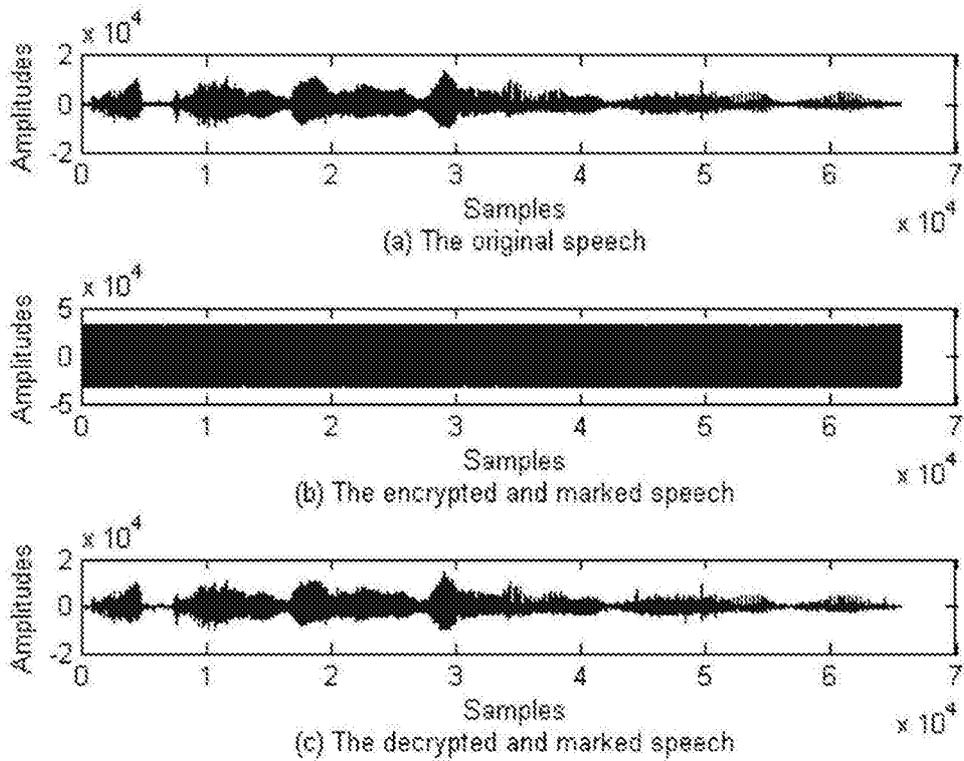


图5

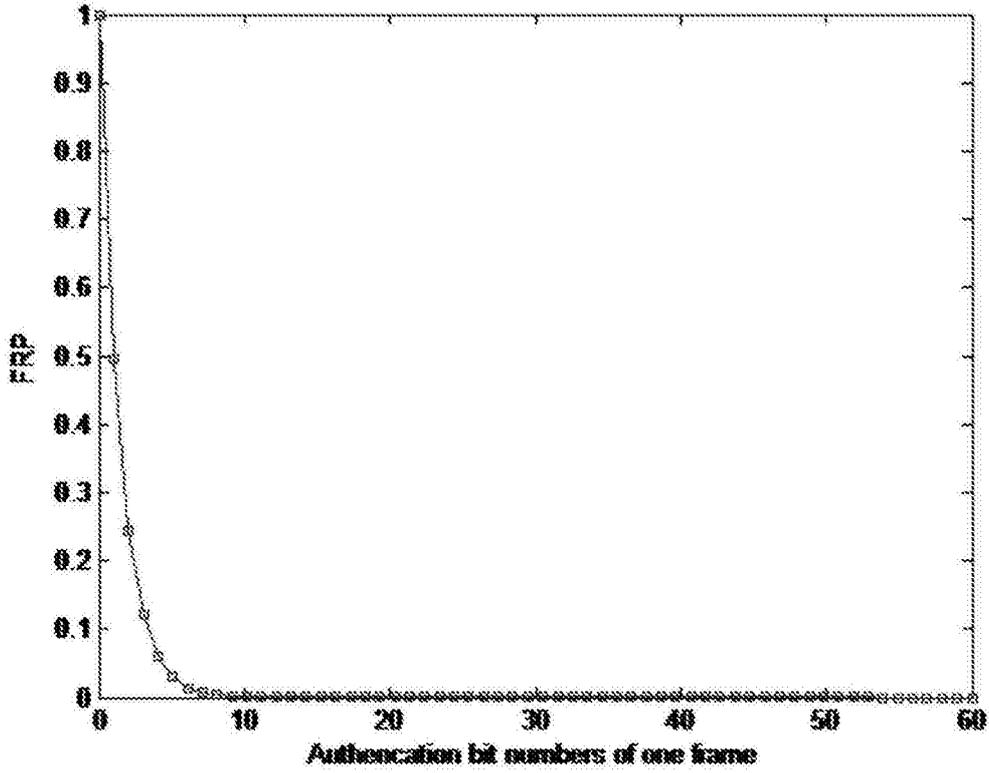


图6

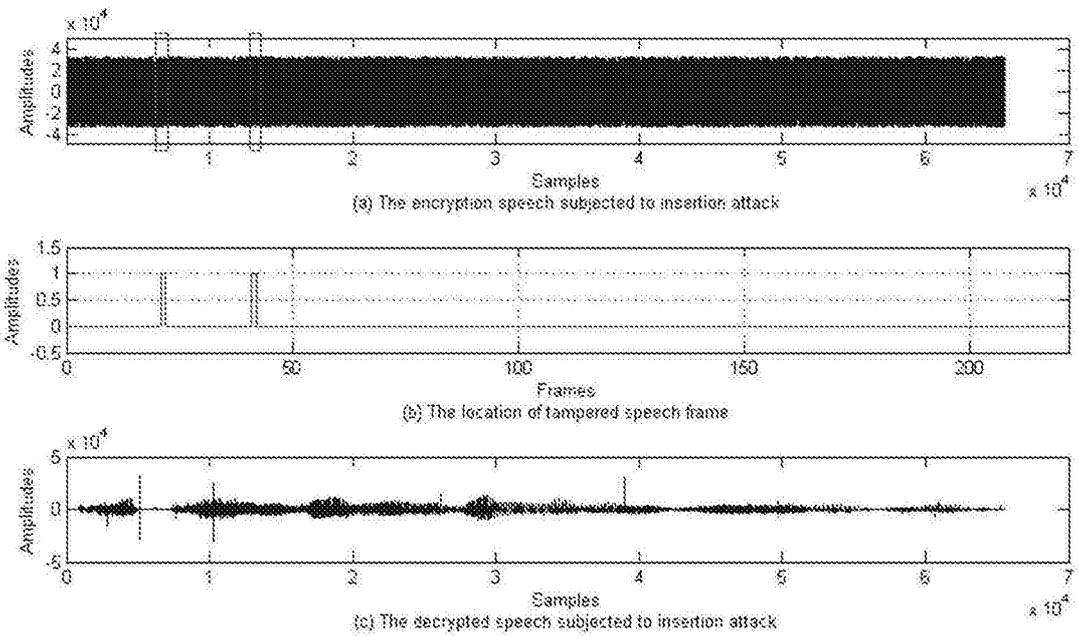


图7

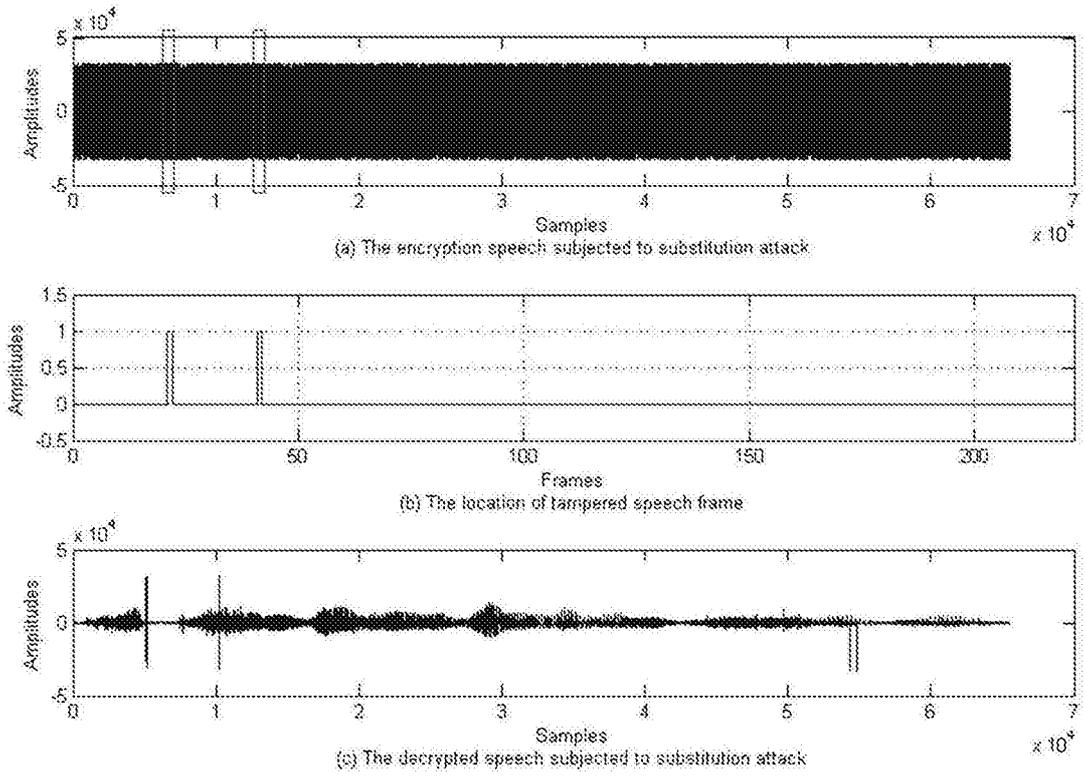


图8

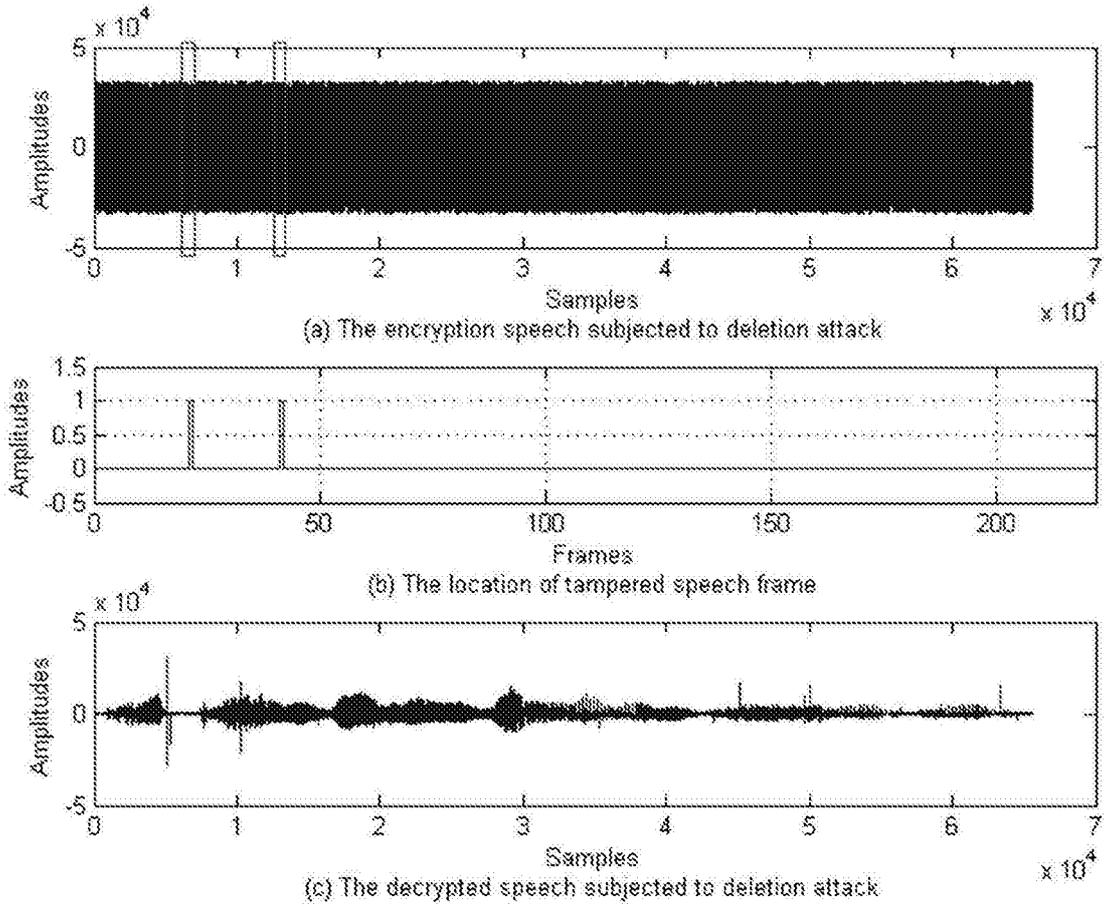


图9