

音频隐写的传统方法

原创

Meiyan_Rabbit 于 2022-03-03 10:47:18 发布 1535 收藏

分类专栏：[音频隐写](#) 文章标签：[其他](#) [音频](#) [音视频](#)

版权声明：本文为博主原创文章，遵循 [CC 4.0 BY-SA](#) 版权协议，转载请附上原文出处链接和本声明。

本文链接：<https://blog.csdn.net/whmemeda/article/details/123187342>

版权



[音频隐写](#) 专栏收录该内容

2 篇文章 0 订阅

订阅专栏

音频隐写的传统方法

一、简介

评估任何隐写技术的主要指标有：

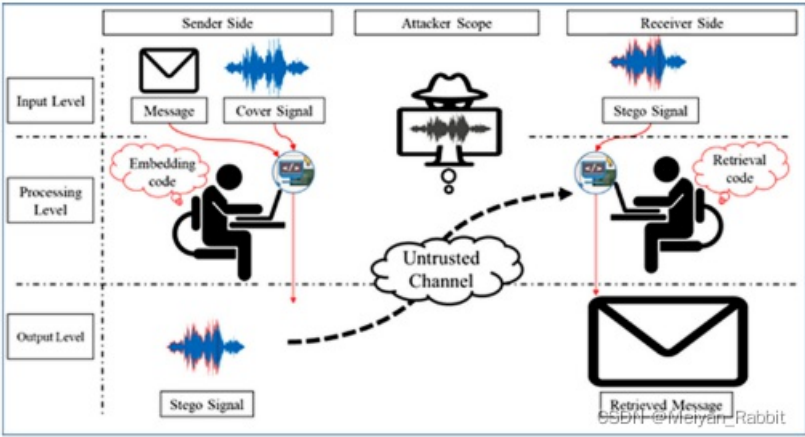
- 1. 感知透明度
- 2. 隐藏容量
- 3. 鲁棒性



各指标含义
感知透明度：表示隐藏信息后，音频的质量影响程度
隐藏容量：隐藏秘密信息量的大小
鲁棒性：表示隐藏后文件对攻击的抵抗力

二、音频隐写框图

音频隐写框图如下：



音频隐写三个主要的组成部分为：

- 1. DOE
- 2. MODI(Method of data injection)
- 3. 载波类型

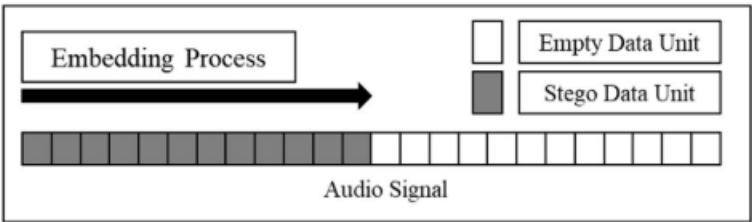
其中

- 1. DOE：域，包括时域、频域等
- 2. MODI：嵌入流程
- 3. 载波类型：包括WAV、MP3和AMR等类型

三、隐写方法

1、Linear or sequential embedding(线性或顺序嵌入)

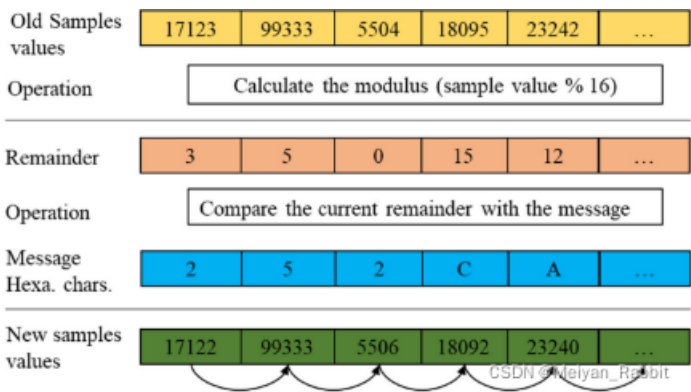
顺序嵌入基于顺序访问原始音频文件，然后执行MODI，其执行过程如下图所示：



- (1) 优点：实现了最高的嵌入容量和最高的嵌入效率
- (2) 缺点：对于任何攻击，鲁棒性都非常低

因此，针对这种方法，数据加密，数字签名等其他支持技术的实现是该方法的主要重点。

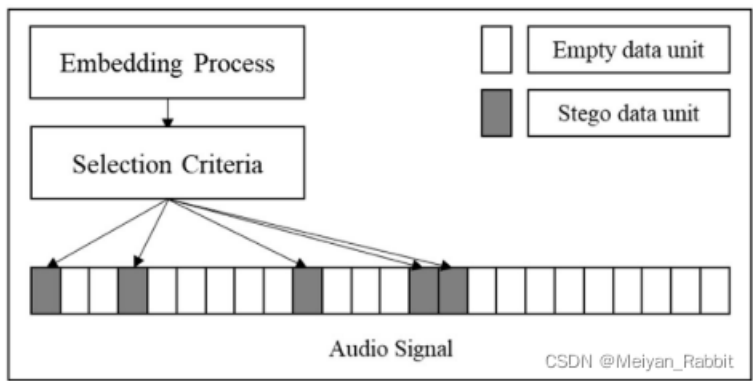
下图是实现该方法的一个案例：
(案例来源文献: B. Datta, S. Tat, Robust high capacity audio steganography using modulo operator, in: Computer, Communication, Control and Information Technology (C3IT), 2015 Third International Conference on. IEEE, 2015.)



- ①首先对将初始的数据进行模16，取得余数，即为上图的Remainder；
- ②将Remainder与秘密信息Message进行相减，并取绝对值，得到新的余数；
- ③组后将旧数据和新的余数相加，即得到隐写之后的音频数据。

2、Selective-based embedding(选择性嵌入)

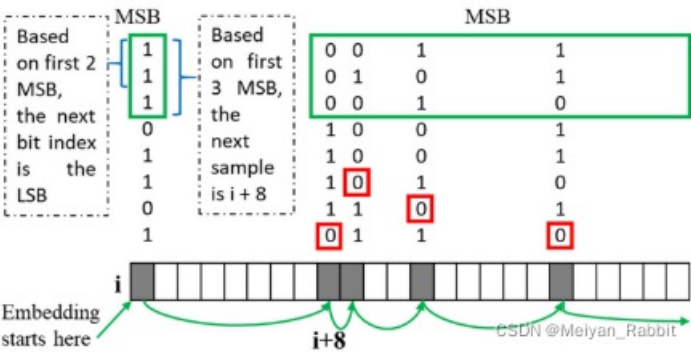
选择性嵌入与顺序嵌入方法相反，选择性嵌入根据定义的选择标准选择嵌入位置，如共享密钥，随机序列或者最高有效位嵌入。下图是嵌入图示。



- (1)优点：
- 选择嵌入过程可以在数据的访问域上执行，与MODI无关，最常见的域是时域；
 - 主要特点是将随机性和音频隐写术结合起来以提高安全性；
 - 其次，这种方法将消息数据随机分散在原始信号上，从而降低了通过统计隐写分析方法检测到的可能性，并减少了消息检索攻击的脆弱性，从而得到较高的SNR和PSNR值。
- (2)缺点：容量和嵌入效率会降低；
- 其次，由于发送方嵌入时候基于链式形式进行嵌入，因此接收方检索时也是基于链式执行，当一位错误时，秘密信息便会提取失败；这类方法大多对消息破坏攻击的鲁棒性较低。

下面是基于该方法的案例：

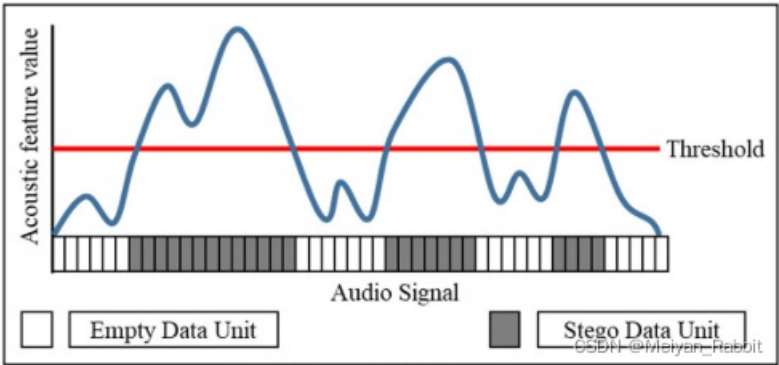
(文献来源：M. Asad, J. Gilani, A. Khalid, Three layered model for audio steganography,in: Proc. - 2012 Int. Conf. Emerg. Technol., ICET 2012, 2012, pp. 270–275,http://dx.doi.org/10.1109/ICET.2012.6375438.)



3、Frequency masking and amplitude thresholding（频率掩蔽和幅度阈值）

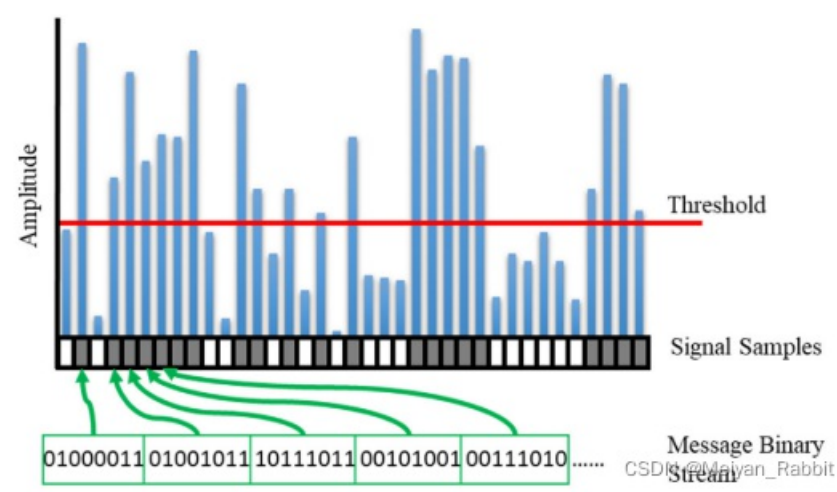
该方法的主要思想是应用于声音信号分量的声学条件，围绕调节或定制数据单元中的嵌入而构建的（主要原理是人类听觉系统无法识别高于或低于特定频率范围的声音信号）。在时域和小波域的方法中，这个方法被称为幅度阈值，在基于频率的方法中，这个方法称为频率掩蔽。

下图是该方法的示例图：



- (1) 优点：该方法的主要优点是可以欺骗人的耳朵，因而减少被人主动发现的情况
- (2) 缺点：没有明确的针对某种方法鲁棒性，且SNR和PSNR对音频质量的表示是可变的，并且没有显著的质量改进的方案。

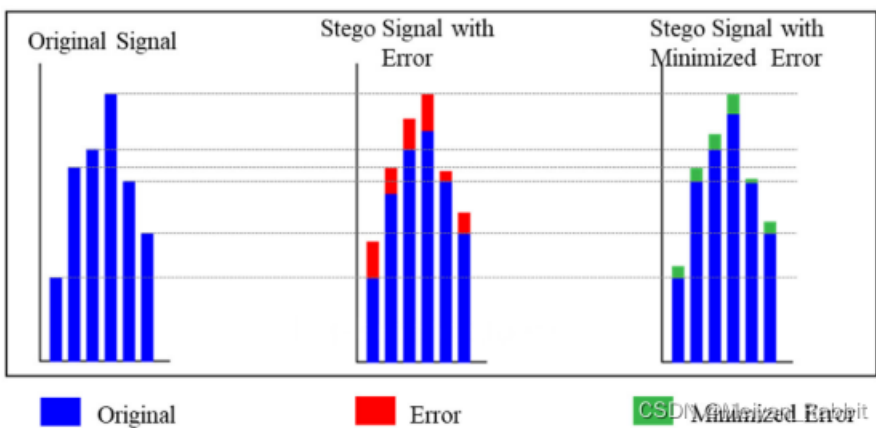
实现该方法的案例：
(文章来源于： M.A. Ahmed, M.L.M. Kiah, B.B. Zaidan, A.A. Zaidan, A novel embedding method to increase capacity and robustness of low-bit encoding audio steganography technique using noise gate software logic algorithm, J.Appl. Sci. 10 (1) (2010) 59–64.)



具体方法： If the sample is equal or above the appointed threshold,8 LSBs will be substituted with message bits.

4、 Error minimization-based embedding(基于误差最小化的嵌入)

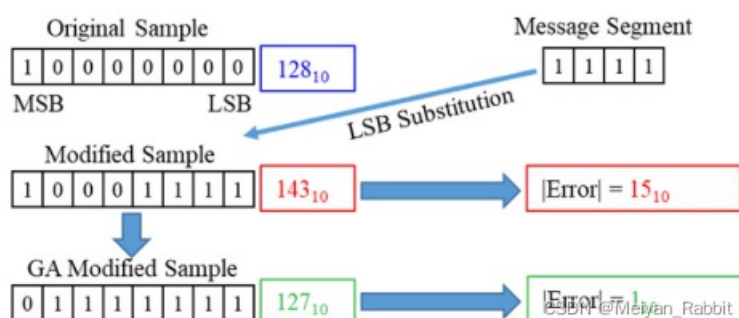
这种方法通过减少结果误差从而减少嵌入的影响，从而减少整体的信号失真，主要方法有：MES,STCs和GA。



其中：
Matrix Embedding Strategy(MES)和Syndrome-Trilles Codes(STC)算法在嵌入之前执行
而Genetic Algorithm(GA)算法在嵌入之后执行
(1) 优点： 该方法可以提高感知透明度，同时针对统计隐写分析具有较好的鲁棒性
(2) 缺点： 该方法是在时域实现，因此对信号破坏的鲁棒性低

实现该方法的案例：

(文章出处: M. Zamani, A.B.A. Manaf, Genetic algorithm for fragile audio watermark-ing, Telecommun. Syst. 59 (3) (2015) 291–304, <http://dx.doi.org/10.1007/s11235-014-9936-x>.)

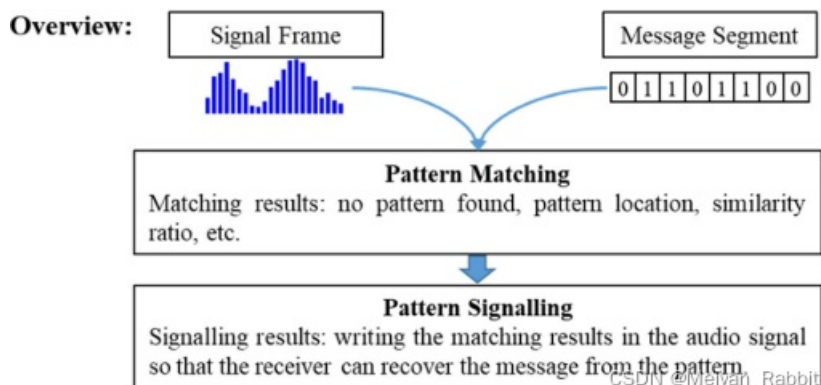


上述方法就是执行GA算法的流程：

首先使用LSB方法将秘密信息嵌入到载体信息中，然后执行GA算法

5、Pattern-matching-based embedding(基于模式匹配的嵌入)

基于秘密信息中的一个二进制序列和载体信息中的二进制序列进行比较，从而得到匹配项（嵌入是由两个二进制序列之间的匹配级别指导），在该方法中，DOE主要有时域和DWT；基于模式匹配的嵌入和基于误差最小化的嵌入类似，都是为了减少嵌入的错误。



该方法包括两个步骤，matching operation和signaling operation:

在matching operation中，识别和定位相似或匹配的项，在signaling operation中，对这些匹配的项进行编码

(1) 优点：鲁棒性好

(2) 缺点：时间和空间的高度复杂性；由于该方法的主要目的是减少覆盖的数量，因此嵌入效率低；如果消息是文本类型，对算法的要求非常高（即必须找到准确的二进制类型），从而这种方法对消息类型非常敏感。

实现方法的案例

* (文章出处: M. Srivastava, M.Q. Rafiq, A novel approach to secure communication using audio steganography, Adv. Mater. Res. 403–408 (2011) 963–969)

- A mismatch at the 1st MSB, 1 bit is embedded and the LSB is made **00**.
- A mismatch at the 2nd MSB, 2 bits are embedded and the LSB is made **01**.
- A mismatch at the 3rd MSB, 3 bits are embedded and the LSB is made **10**.

Message Binary Stream: 0 1 0 0 0 1 1 0 1 0 1 0

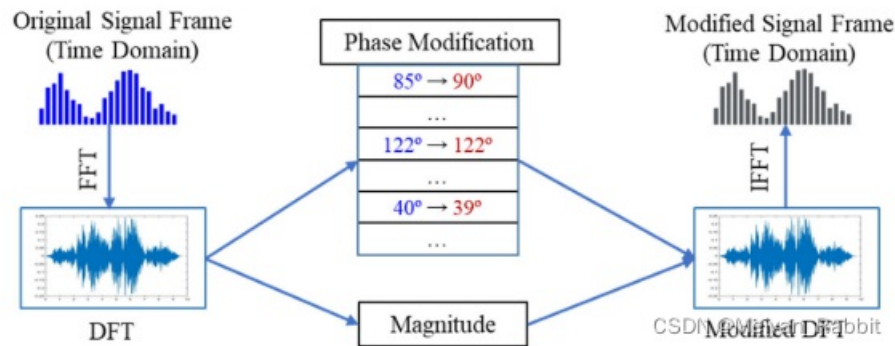
Original Sample	Message bits	Matching Result	Signalling Result
0 1 1 0 1 0 0 1	0 1 0	Mismatch at 3 rd MSB	0 1 1 0 1 0 1 0
0 1 0 0 1 1 1 1	0 0 1	Mismatch at 2 nd MSB	0 1 0 0 1 1 1 0
0 0 1 0 0 0 1 1	1 1 0	Mismatch at 1 st MSB	0 0 1 0 0 0 0 0
1 1 1 0 0 1 0 0	1 0 1	Mismatch at 2 nd MSB	1 1 1 0 0 1 0 1
0 1 1 1 1 0 1 1	1 0 1	Mismatch at 1 st MSB	0 1 1 1 1 0 0 0
0 1 1 0 1 1 1 1	0 1 0	Mismatch at 3 rd MSB	0 1 1 0 1 1 1 0

6、Phase coding（相位编码）

顾名思义，相位编码方法基于修改频率分量的相位值：

- (1) 优点：鲁棒性好（针对 **signal manipulation** 和 **message retrieval**），其感知透明度基于对原始载体的改变程度以及嵌入率
- (2) 缺点：基于这种形式进行嵌入的方法有一定的相似度，因此将选择以样本或毫秒位单位的帧长度作为特征，以来突出 **MODI** 在以及这种形式下的不同方法的内部差异

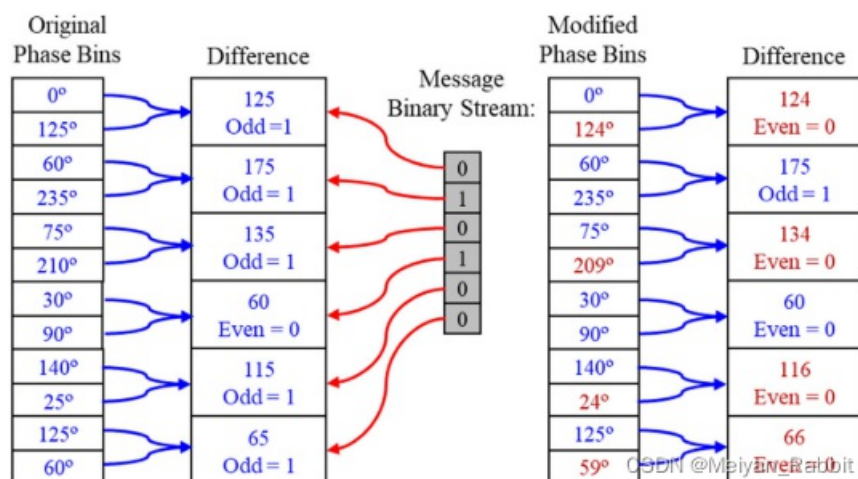
Overview:



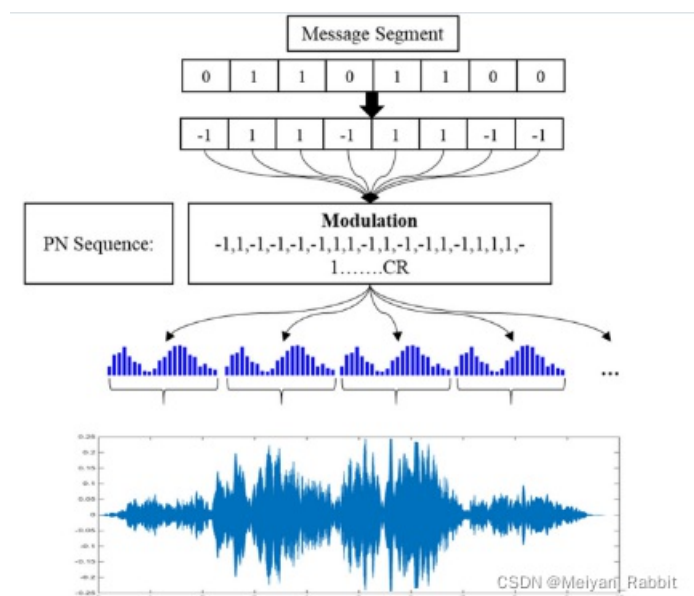
过程：

- ①对时域的音频帧进行FFT变换，得到其频域信息
- ②对频域信息，分离其幅度和相位分量
- ③对相位分量执行嵌入算法
- ④将修改后的相位分量和幅度进行结合
- ⑤执行DFT，得到嵌入秘密信息的音频帧

(文章来源: N. Parab, M. Nathan, K.T. Talele, *Audio steganography using differential phase encoding*, *Technol. Syst. Manag.* (2011) 146–151.)



SS最初是作为一种用于增强健壮性和确保消息传递的传输方法，出于同样的原因，SS被用于音频隐写，下图是基于SS进行音频隐写的框图：



- ①首先将消息从二进制更改为(1,-1)格式
- ②将消息的每个bit乘以伪随机序列(1,-1)，进而将每个bit传播n次
- ③当每个消息位的传播次数增加时，对应的鲁棒性就会增加，嵌入的容量也会降低

以上是关于音频隐写的传统方法，以上的几种传统方法再性能方面都有其优缺点。实施和想法的多样性也会产生不同的结果。

因为后续有想法在音频隐写方面做继续的探索，基于深度学习，或者基于语音编解码格式进行探索性研究，但归根结底，其音频的传统的隐写也应有所了解，于是阅读了近两年的一篇综述*（Digital audio steganography: Systematic review, classification, and analysis of the current state of the art）。

并对其中的部分方法代码搬砖复现，后续会逐渐发布。

侵删