

MP3Stego 信息隐藏与检测方法研究^{*}

宋 华^{1,2}, 幸丘林¹, 李维奇¹, 戴一奇²

(1. 北京电子技术应用研究所, 北京 100091;

2. 清华大学计算机系, 北京 100084)

摘 要: 信息隐藏是一门正在迅速发展的学科, 广泛应用于秘密数据通信、知识产权保护等领域。从音频信息隐藏 MP3Stego 出发, 分析了各种 MP3Stego 的检测攻击技术, 然后通过对 MP3Stego 隐藏特性的研究, 提出了一种检测 MP3Stego 的新方法。这种方法通过计算 MP3 文件中 part23、stuffingBits 的统计量, 可以准确检测出经过 MP3Stego 处理的 MP3 文件。实验证明, 该方法简单有效, 可以用于实现 MP3Stego 信息隐藏的检测。

关键词: 信息隐藏; MP3Stego 检测方法

中图分类号: TP393 文献标识码: A 文章编号: 0529-6579 (2004) S2-0221-04

信息隐藏是一门正在迅速发展的学科, 广泛应用于秘密数据通信、知识产权保护等领域。参考尤新刚等人提出的定义^[1], 信息隐藏是利用人类感觉器官的不敏感 (感觉冗余), 以及多媒体数字信号本身存在的冗余 (数据特性冗余), 将信息隐藏于一个宿主信号 (掩护体) 中, 不被觉察到或不易被注意到, 而且不影响宿主信号的感觉效果和使用价值。相应的, 对信息隐藏的检测和攻击研究, 已经成为了一个重要的研究方向。

1 MP3Stego 信息隐藏

MP3Stego 是剑桥大学计算机实验室安全组开发的一个在 MP3 上隐藏信息的程序^[2], 可以在互联网上得到其源代码。由于 MP3 是目前互联网上最为流行的音乐格式, 而 MP3Stego 又简单易用, 因此成为很受欢迎的一个保护秘密信息的工具。

MP3Stego 作者认为, 通过 MP3Stego 进行信息隐藏是安全的, 人们无法有效地检测出 MP3 文件中是否隐藏有秘密信息。但是, 通过更多的分析和研究, 发现 MP3Stego 还是存在一些漏洞和缺陷。

1.1 MP3 音频编码简介

MP3 是 Fraunhofer-IIS 研究所的研究成果^[3]。MP3 是第一个实用的有损音频压缩编码。在 MP3 出现之前, 一般的音频编码即使以有损方式进行压缩能达到 4:1 的压缩比例已经非常不错了。但是, MP3 可以实现 12:1 的压缩比例, 这使得 MP3 迅速地流行起来。MP3 之所以能够达到如此高的压缩比例同时又能保持相当不错的音质是因为利用了听觉

音频编码技术, 也就是利用了人耳的特性, 削减音乐中人耳听不到的成分, 同时尽可能地维持原来的声音质量。

MP3 音频编码的编码过程见图 1, 声音信号首先通过滤波器组, 滤波器组包括子带滤波和 MDCT 变换, 将时域信号转换为频域信号。同时声音信号通过心理声学模型, 根据人耳的特性产生单独掩蔽阈值、总体掩蔽阈值和最小掩蔽阈值, 用于压缩声音信号。然后, 量化和编码模块根据这些掩蔽值, 对声音频域信号进行量化和编码。最后, 经过量化后的低比特率编码数据被格式化为 MP3 比特流。

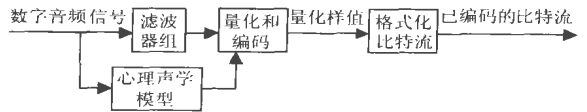


图 1 MP3 音频编码

Fig 1 The MP3 encoder

实现 MP3 音频编码的一个重要部分是两个嵌套的循环, 其中内循环是量化编码循环, Huffman 码表对较小的量化值分配较短的码长, 如果量化编码操作的结果产生的比特数超过了可用的最大比特数限制, 那么将增大量化步长, 得到更小的量化值。不断重复这一操作, 直到 Huffman 编码的结果比特数达到足够小。外循环是噪声控制循环, 根据屏蔽阈值对量化噪声进行控制。每个子带都有一个的比例因子, 如果发现在一个给定的子带中的量化噪声超过了屏蔽阈值, 那么该子带的比例因子就将

^{*} 收稿日期: 2004-09-11

被调整,以减少量化噪声。内循环很容易收敛,但是当两个循环嵌套到一起,有时候会造成无法收敛,这时候就需要有额外的检测条件来停止循环。

1.2 MP3Stego 信息隐藏

MP3Stego 是在 8 hz MP3 编码器的基础上实现的。命令“encode music.wav music.mp3”实现正常的 mp3 编码,即将 music.wav 压缩编码成 music.mp3。而命令

```
encode E hidden.txt P password music.wav
music.mp3
```

在实现 mp3 编码的同时,将秘密信息 hidden.txt 采用密码 password 加密后隐藏到 music.mp3 文件中。

数字音频的频域信号在量化和编码时,存在量化误差。这个量化误差是一个不确定值,例如采用不同的心理声学模型可以导致不同的量化误差,并且这个量化误差如何取值可以在量化编码程序中进行调整设定。采用不同的量化误差可以导致不同的量化结果,并在一定程度上影响最后的音质。MP3Stego 通过调节量化误差的大小,将量化和编码后的长度作为信息隐藏的方法,即长度为奇数代表信息 1,长度为偶数代表信息 0,从而将信息隐藏到最后的 MP3 比特流中。因此,MP3Stego 可以说是一种量化编码信息隐藏方法。

MP3Stego 的隐藏过程可以由如下流程说明:

```
static int inner_loop( ... )
{
    ...
    do {
        quantizerStepSize += 1.0;
        bits = quantize( );
        switch ( hiddenBit ) {
            case 2:
                embedRule = 0; break;
            case 0: case 1:
                embedRule = ((bits + part2length) % 2) != hiddenBit;
                break;
            default:
                ERROR("inner-loop: unexpected hidden bit.");
        } while( (bits > max_bits) || embedRule );
        return bits;
    }
}
```

MP3Stego 在实现量化和编码的内部循环函数 inner_loop() 中,将嵌入规则添加到量化编码的循环规则中。如果量化编码后 part23 的长度不符合嵌入要求(奇代表 1,偶代表 0),则重新调节量化误差,进行量化编码。如果量化编码后 part3 的长度不符合量化编码要求(长度超过最大值限定),

则需要调节量化误差,再进行量化编码。

2 已有的 MP3Stego 检测方法

对信息隐藏进行检测攻击,可以分为下面几类。第 1 类是最彻底的检测攻击,不仅可以发现秘密信息存在,还可以解码还原出秘密信息,可以对秘密信息实现修改、删除,能实现这 1 类的攻击是最有价值的。第 2 类检测攻击方法是只需要发现秘密信息存在,而不必还原出秘密信息或对秘密信息实现破坏,这一类检测攻击可用于发现秘密通信的存在。第 3 类检测攻击方法是只需要破坏秘密信息,而不必发现秘密信息存在或还原出秘密信息,这一类检测攻击可用于阻碍秘密通信,破坏知识产权知识保护等。

MP3Stego 的作者认为,MP3Stego 是一种比较安全的信息隐藏,可应用于秘密通信、知识产权保护等方面。对 MP3Stego 的检测攻击,只存在第 3 类攻击方法,即破坏秘密信息。通过将 MP3 文件解码,再重新压缩成 MP3 文件,秘密信息将被破坏。不过这种方法将严重影响 MP3 的音质。

梁敬弘等^[4]指出,由于 MP3Stego 的秘密信息隐藏在 part23 长度的奇偶性中,因此通过改写 part23 长度的最低位,即改变奇偶性,可以将秘密信息破坏,从而实现对 MP3Stego 的攻击。这种方法也属于第 3 类攻击,并且也将影响 MP3 的音质。因为 part23 长度代表的是频域编码的有效位,改写 part23 长度的最低位,也就是改写了频域编码有效位的长度,使得本来有效的高频信息位被简单舍去,从而造成 MP3 的音质的下降。

Andreas Westfeld^[5]指出,MP3Stego 存在第 2 类检测方法,通过统计分析块长度,即 part23 的长度,可以检测出 MP3Stego 是否隐藏有秘密信息。在实现量化和编码的内部循环函数 inner_loop() 中,正常情况下当 part23 长度小于可用最大值的要求后就可以结束内部循环,但在信息隐藏方式下,part23 长度还需要满足嵌入要求,因此可能需要进一步扩大量化步长,得到更小的量化值,part23 长度也随之变小。由于 MP3 文件的帧长度和帧速率是恒定的,下一帧可用的最大比特数也就变大。因此,下一帧 part23 长度也就会变大。虽然 part23 长度的均值保持不变,但是在信息隐藏方式下,part23 长度的方差会发生很大的变化(图 2)^[5]。

用于检测 MP3Stego 的方差计算公式为

$$s^2 = \frac{\sum x^2 - 1/n \left(\sum x \right)^2}{n - 1}$$

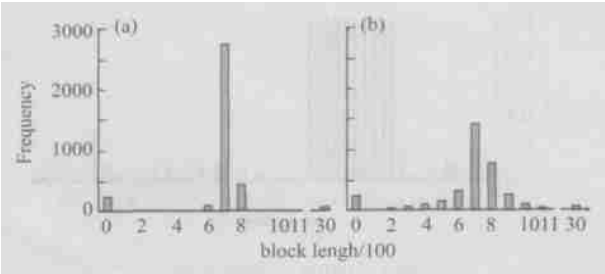


图2 不含秘密信息 (a) 与含最大量秘密信息 (b) 的块长度直方图

Fig 2 Histogram of block length without steganography (a) and with the maximum of embedded data (b)

其中, n 代表参与计算的 part23 长度的个数, $\sum x$ 代表 part23 长度的总和, $\sum x^2$ 代表 part23 长度平方的总和。通过计算方差公式, 可以检测出包含 0.01% 秘密信息的 MP3 文件。MP3Stego 实现的秘密信息隐藏量一般在 0.05% ~ 0.1% 左右, 因此方差公式可以很准确的检测出 MP3Stego 信息隐藏。

3 一种新的 MP3Stego 检测方法

在实现 MP3 编码的时候, 格式化比特流的帧长度是以字节为单位的, 而在实现量化编码时, 编码后产生的 Main Data 是比特流, 是以 bit 为单位的, 因此在将数据打包成帧的时候, 可能会产生 1 ~ 7 的附加比特位, 来满足字节为单位的要求。另外, 不同的帧有不同的压缩率, 为了满足恒定帧速率的要求, 有时候也会产生附加比特位来填充帧中空余的比特位。这些附加比特位被保存在 stuffingBits 或 ancillary 的位置, 如图 3 所示。

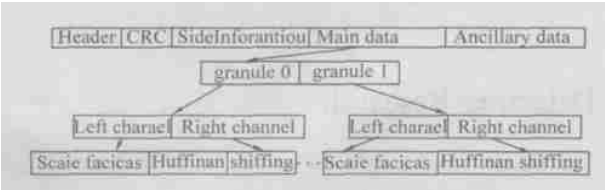


图3 MP3 帧结构

Fig 3 The MP3 frame structure

将数据打包成帧, 添加附加比特位程序如下:

```
void ResvFrameEnd (L3~side~info~t *l3~side, int mean~bits)
{
    if (stuffingBits) {
        #ifdef MP3STEGO
            if (stuffingBits % 2) {
                gi->part2~3~length += stuffingBits - 1;
                stuffingBits = 1;
                l3~side->resvDrain = stuffingBits;
            }
            else gi->part2~3~length += stuffingBits;
        #else
            // ...
        #endif
    }
}
```

```
gi->part2~3~length += stuffingBits;
#endif
} ...
}
```

8 Hz mp3 在正常 MP3 编码的时候, 首先考虑将 1 ~ 7 的 stuffingBits 添加到 part23 长度上去, 使得 part23 长度为整字节, 然后再考虑在 ancillary 的位置添加附加比特位。

在 MP3Stego 实现信息隐藏时, 由于 stuffingBits 是计算累加到 part23 的长度上去, 因此为了保证 part23 长度为奇数代表信息 1, 长度为偶数代表信息 0, 附加的比特位长度必须是偶数。由于附加比特位有一半的情况是奇数, 这时就会剩余 1 bit 的附加位, 这个 1 比特的附加位被 MP3Stego 添加到 ancillary 的位置。

这种情况和正常情况不相符合。通过统计 part23 长度、part23 中 stuffingBits 的长度、和 ancillary 的长度, 可以比较出正常 MP3 编码和在 MP3Stego 实现信息隐藏时 MP3 编码的差别。表 1 给出在将一个 WAVE 文件编码成 MP3 文件时的统计量。

表1 MP3Stego、8 hz mp3 和 Lame 编码时 part23 长度、stuffingBits 和 ancillary 模 8 后的统计量

Tab 1 The statistics of part23 length, stuffingBits, and ancillary of the MP3Stego, the 8 hz~mp3 and the Lame encoder								
编码	%8	0	1	2	3	4	5	6 7
MP3 - Stego	part23 长度	396	0	0	0	0	0	396
	stuffingBits	985	0	189	0	224	0	186 0
	ancillary	396	396	0	0	0	0	0
8 hz~mp3	part23 长度	792	0	0	0	0	0	0
	stuffingBits	894	91	96	96	102	116	95 94
	ancillary	792	0	0	0	0	0	0
Lame	part23 长度	106	101	90	110	96	101	103 86
	stuffingBits	1585	0	0	0	0	0	0
	ancillary	106	86	103	101	96	110	90 101

可以看出, MP3Stego 并没有将所有的 part23 长度调整为整字节, 而是 50% 为整字节, 50% 模 8 为 7; StuffingBits 模 8 在奇数分布为 0, 在 2、4、6 均匀分布; ancillary 的 50% 为 0, 50% 为 1。而 8 hz~mp3 在编码时将 part23 长度调整为整字节, StuffingBits 在 1 ~ 7 均匀分布, ancillary 为 0。两者有明显区别 (表 1)。

MP3 有很多种编码器, 为了统计各种编码器的不同编码特性, 还需要计算别的编码器对音频进行编码后所产生的统计量。下面给出另一种典型的 MP3 编码器 Lame 的统计量 (表 1)。

可以看出, Lame 编码器没有调整 part23 的长度, 而是将附加比特位添加到 ancillary 的位置, 所以所有的 stuffingBits 都为 0, 同一帧中 part23 的长度和 ancillary 的长度相加后模 8 为 0。

通过比较这些不同编码器统计量的不同, 最后我们给出一个用于检测 MP3Stego 的统计公式

$$R = \frac{\text{abs}(\sum_{i=0}^7 (\frac{1}{8}L - \frac{1}{8}L)) + \frac{4 * \text{stereo}}{3}}{\sum_{i=0}^7 \frac{1}{8}L} .$$

$$\frac{\text{abs}(\sum_{i=0}^7 (\frac{1}{8}N - \frac{1}{8}N)) + \text{abs}(\sum_{i=0}^7 (\frac{1}{8}N - \frac{3}{8}N)) + \text{abs}(\sum_{i=0}^7 (\frac{1}{8}N - \frac{5}{8}N))}{\sum_{i=0}^7 \frac{1}{8}N}$$

其中, $\frac{1}{8}L$ 表示 part23 长度模 8 后等于 i 的统计量, $\frac{1}{8}N$ 表示 stuffingBits 模 8 后等于 i 的统计量。阈值可以设置为 0.5, 当 R 大于 0.5 的时候表示检测到 MP3Stego, 当 R 小于 0.5 的时候表示没有检测到 MP3Stego。

4 检测结果

为了检验上一节得出的检测公式, 在网上随机下载了 90 个 MP3 文件, 再混入 10 个采用 MP3Stego 信息隐藏的 MP3 文件, 然后对这 100 个 MP3 文件进行辨别。辨别的结果如图 4 所示。

从图 4 可以清楚看出, 采用 MP3Stego 信息隐藏的 MP3 文件, 其统计值 R 大约等于 1, 网上的大部分 MP3 文件, 其统计值稍大于 0, 有一些 MP3 文件在 0.1 ~ 0.3 附近。通过该辨别公式, 可以清楚辨别出采用了 MP3Stego 信息隐藏的 MP3 文件。

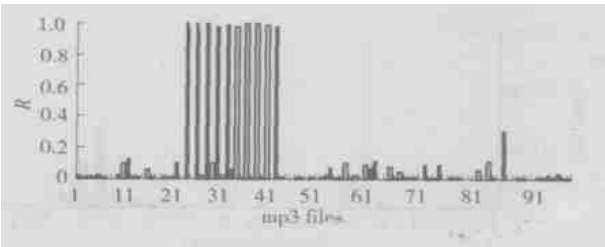


图 4 MP3Stego 检测结果

Fig 4 The result of the MP3Stego detection

5 结 论

通过对 MP3Stego 信息隐藏特性的仔细分析, 本文发现了一种检测 MP3Stego 信息隐藏的新方法。这种方法通过计算 MP3 文件中 part23、stuffingBits 的统计量, 可以准确检测出经过 MP3Stego 处理的 MP3 文件。实验证明, 该方法简单有效, 可以用于实现 MP3Stego 信息隐藏的检测。

参考文献:

[1] 尤新刚, 周琳娜, 郭云彪. 信息隐藏学科的主要分支及术语 // 全国第三界信息隐藏学术研讨会 Cj . 2001.

[2] <http://www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/>

[3] KARLHEINZ B, GERHARD S. ISO-MPEG-1 Audio: A generic standard for coding of high quantity digital audio [J]. J Audio Eng Soc, 1994, 42 (10): 780—792.

[4] 梁敬弘, 王道顺, 黄连生, 等. 一种对 MP3Stego 的攻击研究 [J]. 计算机辅助设计与图形学学报, 2003, 15 (8): 954—960.

[5] ANDREAS W. Detecting low embedding rates [C]. Information Hiding 5st Int Workshop, Noordwijkerhout, 2002; 329—342.

MP3Stego Hiding and Detecting Research

SONG Hua^{1,2}, XING Qiu lin¹, LI Wei qi¹, DAI Yi qi²

(1. Beijing Institute of Electronic Technology Application, Beijing 100091, China;

2; Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: Information hiding is the art and science of communicating in such a way that the presence of a message cannot be detected. It can be used for secret communicating and copyright marking and receives more attention from the research community and from industry today. This paper introduced the audio information hiding techniques first, with an emphasis on MP3Stego, then introduced the existing MP3Stego attacking and detecting techniques. After analyzing the properties of MP3Stego, a new detecting method was found. This method can detect mp3 files produced by MP3Stego correctly by calculating the statistic of part23 and stuffingBits in mp3 files. The experiments proved that this method is effective and can be used to detect MP3Stego.

Key words: information hiding; MP3Stego detecting methods