**杭州电子科技大学**

**毕业设计（论文）开题报告**

|  |  |
| --- | --- |
| **题 目** | **基于数字签名的数字音频内容认证方法的分析与实现** |
| **学 院** | **网络空间安全学院** |
| **专 业** | **网络工程** |
| **姓 名** | **廖越强** |
| **班 级** | **18272412** |
| **学 号** | **18041618** |
| **指导教师** | **王慧** |

**一、综述本课题国内外研究动态，说明选题的依据和意义**

互联网的高速发展，使得多媒体数字音频的创作、存储与传输都变得及其便利，但同时也带来了一些问题。通过互联网渠道下载的音频文件的不能够确定其内容是否遭到恶意篡改，可信度会受到怀疑。其根本原因在于音频文件的编辑修改可以在极低的成本下进行，通过使用一些剪辑软件对音频片段的删减或者是顺序调换，可以轻易地篡改文件所表达的语义。为了满足音频等多媒体数据的信息管理与保护，数字音频认证算法的研究是非常有必要的。

音频内容的认证技术实际上就是对其数据的完整性和真实性进行保护的技术手段，可以确保接受方得到的音频数据在传输过程中没有遭到攻击者的恶意编辑篡改。这项技术可应用的领域非常广阔，例如警局口供录音、法庭辩护、国家机密安全、商业机密、音乐、军事等领域对音频数据的认证都有着非常高的需求，研究出实用的认证算法可以带来很大的社会效益。

目前音频文件的认证方式主要分为两类，一类是鲁棒签名算法，另一类是数字水印法。而认证的标准也分为硬认证和软认证两种，硬认证拒绝对内容的任何修改，仅接受无损的压缩或格式转换，而软认证则相对宽松，允许某些方面的编辑修改，这些处理被称为可容许操作，而拒绝其他的恶意操作。本次研究主要是针对数字签名方式的软认证算法的设计与分析。

数字签名作为一种文件内容的认证的有效手段，在密码学中已经有非常成熟的研究与实践应用，其过程简单地描述就是在签名系统中通过使用Hash函数得到数字摘要，加密后得到数字签名后绑定到原始文件中。数字签名可以保证文件的完整性，且其本身具有不可抵赖性，校验真伪时无需获取原始文件，仅需要使用签名和待认证文件即可进行认证检验。其应用研究在图像等领域已经有大量成果公布，但在音频领域的完整性认证上，相关研究相比之下就比较少了。音频信号领域和图像领域的认证重点有所不同，人们对于图像的感知是视觉上的，而对于音频的感知主要是听觉层面，对音频的认证应该侧重于其内容含义上的一致而不是比特流的完整性，因此不能直接以传统的哈希函数直接处理音频数据。如何根据音频信号在听觉上的特征来生成认证所需的数字摘要将是本次研究的重点。

把提取数字摘要的哈希函数应用在音频领域的研究开始于本世纪初，国外如纽约Ploy科技大学、Delft大学、PHILIPS研究院对此已有较多优秀的研究成果。

2008年哈尔滨工业大小信息对抗技术研究所在《感知哈希综述》一文中以人类为感知模型，明确了感知哈希的定义和性质，同时也指出未来感知哈希在音频领域的应用将会有更多研究出现。

Radhakrishna等在A model of co-saliency based audio attention一文中提出了一种根据音频内容的某些特征来进行认证的思路，其原理是两个听觉相似的音频的掩蔽曲线也是几乎稳定的高度相似。实验证明，这种基于内容的哈希值可以区分MP3格式转换等音频处理和恶意篡改。

Haitsma等则设计了另一种以能量差作为鲁棒特征来计算哈希值，通过对比两音频的哈希值来进行内容鉴别。实验结果表明该方法对于保持内容的操作的处理效果也是非常好的，具有极低的误报率。

随着音频信息服务的普及深入，其相应的内容信息管理和保护需求也越来越受到人们关注。为了在认证中消除对感知内容上的冗余，国内外学者开始针对性地投入到对音频感知冗余的技术研究，将人类听觉特性与数字摘要相关联起来。数字签名作为数字摘要的一种应用，能够保证数据的完整性和不可抵赖性，因此将数字签名应用到音频的内容认证这一研究是具有一定的研究意义的。

**二、研究的基本内容，拟解决的主要问题：**

与其他数据不同，音频数据可以用不同的格式来表达同样的内容。比如一段WAV格式的音频，转化为MP3压缩格式仍可以表达出非常接近的听觉效果，在高比特率情况下几乎无法靠听觉区分出差别，可以认为用户得到的是相同的信息。因此这里研究的重点是音频内容上的认证，而不是想传统认证一样简单地保护数据比特流。

研究的最终目的是要实现基于音频内容上的分析认证。这里将对音频的编辑修改分为两类，一是保持内容的合法操作，二是恶意篡改内容的非法操作。

（1）合法操作：对于不涉及音频内容上的修改编辑，将归类为合法操作。比如音量调节、噪声去除、高比特率有损压缩、滤波、音频格式转码等音频信号处理。

（2）非法操作：对于改变了原始音频内容语义的处理，如局部的删除、替换、插入，或者是重排时间轴上的音频片段等处理，归类为非法操作。

**拟解决问题：**

（1）选取合适的音频信号特征。音频的内容认证实际上是听觉感知层面的认证，这要求所选取的特征应该和人耳的分辨特性有相关性，因此如何选取音频信号的特征指标是本次研究的关键点。

（2）生成数字签名。根据第一步中得到的特征数据集生成数字签名。预计得到的签名也是数据集

（3）保证数字签名的安全性、真实有效性。为了避免攻击者得到原始音频后伪造签名，应当让签名具有密钥依赖性，保证原始作品的签名的真实有效性。

（4）设计音频内容认证时，数字签名的匹配算法。数字签名的实际应用中，通常只能得到原作品的签名和待测作品，而无法得到原作品文件。因此认证的关键在于如何匹配待测作品的签名与原始签名的相似性，根据二者的相似程度来判断是否认证成果。

**三、研究步骤、方法及措施：**

（1）**音频特征的选取**

通过阅读信号处理相关论文以及查询对应知识点了解到，音频信号有时域和频域两个维度，而能同时兼顾两者的信号特征提取方法主要有小波变换和短时傅里叶变换。可以考虑采用以若干样点数m做为帧长度进行分帧，对于每一帧数据都进行特征向量提取，得到的是该帧的特征数据矩阵。

（2）**根据特征矩阵，设计数字签名的生成算法**

对步骤一中每一帧得到的数据矩阵，可以求出如逼近系数绝对值的平均值（E）和标准差（σ）等数据指标，再考虑把这些指标转化为二进制的形式，最后再使用Hash运算，最终得到一个Hash集合（H），这个Hash集合也就是初步得到的用于认证匹配的集合（S）。

需要考虑签名的安全性。数字签名需要有抵抗破解的能力，即使签名的生成算法公开暴露，攻击者也无法根据签名生成算法伪造出正确的签名。这里为签名添加密钥性依赖，主要有下述三个思路可以考虑：

1. 直接加密原始音频
2. 加密上文得到的音频特征矩阵或是求得的数据指标。
3. 加密生成的数字签名

本文考虑采用第三种方式。从处理数据量的大小来考虑，上文步骤二得到的Hash集合的数据体量预计不会太大，因此初步考虑使用非对称加密算法，使用私钥对生成的Hash结果进行加密，最终生成的加密后的数字签名密文，以保证签名的真实有效性。

（4）**认证算法的设计与分析**

由于认证算法要具有鲁棒性，需要考虑一些对音频文件的合法操作，因此算法的认证结果不适合直接输出明确的是/否二值，而应当是根据认证场景来设置一个阈值（NC），当待认证音频与原始音频的相似程度超过阈值时才判定为认证成功。

实际认证中，通常可以得到原始音频的签名和待认证音频。可以先根据上述签名算法来生成待认证音频的数字签名，再将生成的签名与原始音频的签名进行相似度匹配。如果是两段相似度非常高的音频，那么根据上述算法生成的签名也会有很高的重合率，可以考虑使用归一化相关系数[-1, 1]来判断二者的相似度。因此认证的结果为落在[-1, 1]区间的一个数值，该数值越高，表明两段音频信号的相似程度越高，可将结果与前文提到的预设阈值（NC）进行比较，如果大于NC，则可判定为认证成果，反之则为认证失败，认为待认证作品已经遭到严重篡改或破坏。

**拓展**：

实际上除了生成数字签名、签名的匹配算法外，还有一个值得研究的点是如何传输、存储生成的数字签名。当然形式可以有很多种，可以与原音频分开存储，也可以考虑使用零水印的方式进行嵌入。

**四、研究工作进度：**

|  |  |  |
| --- | --- | --- |
| **序号** | **时间** | **内容** |
| **1** | 2021.7.1-2021.7.25 | **查阅相关资料，理解选题** |
| **2** | 2021.7.26-2021.7.31 | **任务书** |
| **3** | 2021.9.22-2021.9.30 | **开题报告会** |
| **4** | 2021.10.1-2021.10.15 | **设计算法框架** |
| **5** | 2021.10.15-2021.10.30 | **编写程序代码** |
| **6** | 2021.10.31-2021.11.15 | **做实验，收集实现数据，并分析其结果** |
| **7** | 2021.11.16-2021.11.30 | **优化、改进和修正算法，并准备论文撰写** |
| **8** | 2021.12.1-2021.12.16 | **撰写毕业论文** |
| **9** | 2021.12.17-2021.12.22 | **论文评审及查重** |
| **10** | 2021.12.27-2022.1.5 | **答辩报告会** |

**五、主要参考文献：**（所列出的参考文献不得少于10篇，其中外文文献不得少于2篇，发表在期刊上的学术论文不得少于4篇。）

1. Zhu L., Liu D., Yu L., Xie Y. and Wang M., Content integrity and non‐repudiation preserving audio‐hiding scheme based on robust digital signature[J]. Security and Communication Networks, 2013, 6(11): 1331-1343.
2. Yang X., Wu X. and Zhang M. , Audio Digital Signature Algorithm with Tamper Detection[C], 2009 Fifth International Conference on Information Assurance and Security, 2009: 15-18.
3. Cano P. and Batlle E., A review of audio fingerprinting[J], 2005 Journal of VLSI Signal Processing, 2005, 41(3): 271-284.
4. 汪竹蓉. 基于音乐内容分析的音频认证算法研究[D]. 复旦大学, 2011.
5. 李伟, 汪竹蓉, 李晓强, 刘亚多. 数字音频认证研究综述[J]. 计算机科学, 2009, 36(10): 21-24.
6. 王向阳, 祁薇. 用于版权保护与内容认证的半脆弱音频水印算法[J]. 自动化学报, 2007(09): 936-940.
7. 钱清, 王宏霞, 刘正辉. 基于基音周期的语音内容认证算法[J]. 铁道学报, 2014, 36(09): 60-67.
8. 吴琼, 李国辉, 涂丹, 孙韶杰. 面向真实性鉴别的数字图像盲取证技术综述[J]. 自动化学报, 2008, 34(12): 1458-1466.
9. 牛夏牧, 焦玉华. 感知哈希综述[J]. 电子学报, 2008(07): 1405-1411.
10. 李伟, 李晓强, 陈芳, 王淞昕. 数字音频指纹技术综述[J]. 小型微型计算机系统. 2008(11): 2124-2130.

**六、开题答辩小组评审意见：**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **考核点** | **背景及意义阐述情况** | **研究内容与任务书的匹配程度** | **研究方案合理性** | **进度安排情况** | **答辩情况** | **总分** |
| **满分** | **20** | **30** | **30** | **10** | **10** | **100** |
| **评分** |  |  |  |  |  |  |

开题答辩小组负责人签字：

年 月 日