2004年上海大学博士学位论文 35

# 信息隐藏安全性研究
## ——理论研究以及在版权保护和隐蔽通信中的技术实现

作　者：张新鹏
专　业：通信与信息系统
导　师：王朔中

上海大学出版社
·上海·

2004 年上海大学博士学位论文

# 信息隐藏安全性研究
## ——理论研究以及在版权保护和隐蔽通信中的技术实现

作　者：张新鹏
专　业：通信与信息系统
导　师：王朔中

上海大学出版社
· 上 海 ·

Shanghai University Doctoral Dissertation (2004)

# Security of Information Hiding:
## Theoretical Study and Technical Implementation in Copyright Protection and Covert Communications

**Candidate:** Zhang Xin-peng

**Major:** Communication and Information System

**Supervisor:** Prof. Wang Shuo-zhong

**Shanghai University Press**

• Shanghai •

# 上 海 大 学

本论文经答辩委员会全体委员审查，确认符合上海大学博士学位论文质量要求。

**答辩委员会名单：**

| | | | |
|---|---|---|---|
| 主 任： | 余松煜 | 教 授，上海交大图象所 | 200030 |
| 委 员： | 张立明 | 教 授，复旦大学电子工程系 | 200030 |
| | 蒋昌骏 | 教 授，同济大学计算机系 | 200092 |
| | 张兆扬 | 教 授，上海大学通信学院 | 200072 |
| | 王治钢 | 研究员，上海航天局 809 研究所 | 200031 |
| | 顾亚平 | 研究员，中科院东海站 | 200010 |
| | 吴亚明 | 研究员，中科院微系统所 | 200050 |
| 导 师： | 王朔中 | 教 授，上海大学 | 200072 |

评阅人名单：

| | | |
|---|---|---|
| 梁庆林 | 教　授，北京大学电子学系 | 100871 |
| 杨永田 | 教　授，哈尔滨工程大学计算机学院 | 150001 |
| 余松煜 | 教　授，上海交大图像研究所院 | 200030 |

评议人名单：

| | | |
|---|---|---|
| 周思永 | 教　授，北京理工大学 | 100089 |
| 黄继武 | 教　授，中山大学信息学院 | 510275 |
| 宣国荣 | 教　授，同济大学 | 200092 |
| 严壮志 | 教　授，上海大学通信学院 | 200072 |
| 张兆扬 | 教　授，上海大学通信学院 | 200072 |
| 王治钢 | 研究员，上海航天局 809 研究所 | 200031 |

# 答辩委员会对论文的评语

张新鹏的博士学位论文研究信息技术的前沿课题，在二项国家自然科学基金资助下，对信息隐藏领域的两个主要分支——数字水印、密写与密写分析的一系列重要问题展开深入研究，取得了多项创新成果，对于信息安全和知识产权保护有重要意义和良好的应用前景.论文作者取得的主要成果可归纳如下：

(1) 数字水印研究.研究水印基础理论，对两种应用最广泛的信息隐藏算法进行分析，其结论对改进水印算法有指导意义.改进迭加方案，较大幅度地提高了水印性能.研究攻击条件下量化算法的信息容量，导出了受攻击后能保留的最大可靠信息量及其实现条件.分别提出能抵御恶意可逆性攻击和嵌入器攻击的水印协议和算法，有效地提高了安全性，并可整合形成一个对这两类攻击同时具有免疫力的多比特水印系统.论文进一步考虑了将其各项成果结合在一起，实现水印技术的安全性集成，有助于数字水印走向成熟并实现商用.

(2) 密写和密写分析研究.针对不同载体提出了一组安全性好、计算量小、信息容量大的密写方法，包括①面向非压缩载体的安全 LSB 密写，可抵抗任何利用直方图异常和位面特征非对称性的分析；②能抵抗奇异色分析的调色板图像安全密写；③能抵抗基于直方图异常或分块特性分析的安全 JPEG 图像密写方案.在密写分析方面，提出对 BPCS 和 PVD 密写的有效攻击.由此引入混合进制系统，提出了优于 BPCS 和 PVD 方法的安全密写法，消除了这两种密写中的安全隐患，并具有高度的灵活性和隐

蔽性.

本论文表明,张新鹏具有坚实宽广的基础理论和系统深入的专门知识,科研创新能力很强.论文论述正确，结构合理，叙述清晰，文字通畅.实验数据充足，结果可信.在答辩中表达清楚，回答问题正确.

鉴于张新鹏博士学位论文突出的创新性,并注意到张新鹏同学作为第一作者已发表论文 24 篇(含 2 篇录用)，在 SCI 源刊发表论文 5 篇,EI 源刊发表论文 7 篇.目前已检索到 SCI 收录 3 篇，EI 收录 7 篇，特推荐该论文为优秀博士学位论文.

## 答辩委员会表决结果

经答辩委员会表决,全票同意通过张新鹏同学的博士学位论文答辩，建议授予工学博士学位.

答辩委员会主席：余松煜

2004 年 3 月 6 日

# 摘 要

信息隐藏是信息安全领域中的一项新兴技术，主要有数字水印和数字密写两大分支，分别用于多媒体数据的版权保护和隐蔽通信. 安全性，即抵抗各种敌对攻击的能力，是信息隐藏技术的核心. 无论是数字水印还是密写，都是在与各种攻击技术的对抗中获得进步，而攻击技术也随着数字水印和密写的进步不断更新. 矛盾的双方在对抗中互相促进、同步得到发展. 本文围绕信息隐藏中的安全性问题，将攻击、防守统一考虑，对取得的一系列研究成果进行总结.

(1) 信息隐藏的理论基础进行了研究. 分析和比较了迭加与量化这两种应用最广泛的信息隐藏算法的性能，得出了迭加法适用于高频、量化法适用于低频的结论. 改进了迭加方案，使每个嵌入到数字媒体中的序列都负载多比特信息，大大提高信息隐藏性能，有效缓解了稳健性、隐蔽性和嵌入量之间的矛盾. 讨论了攻击条件下量化隐藏的信息容量，得到结论：当隐藏者保密变换方式，并以特定的统一量化步长、嵌入特定的比特数，可以在攻击后保留最大的可靠信息量，此时的信息量与载体数据量和密写强度成正比，与攻击强度成反比.

(2) 开发能够抵御恶意攻击的水印算法和协议，有效地提高数字水印安全性. 一方面提出了可抵抗嵌入器攻击的新水印算法.这个算法中，即便在密钥相同的条件下水印信号也互相独立选取，嵌入器攻击不但不能去除合法水印，反而会在数字媒体中留下攻击的痕迹. 另一方面，基于对可逆性攻击的深入分析提出

了安全水印协议该协议指出攻击者可实施一种更强有力的基于算法构造的可逆性攻击，突破现有的为抵抗可逆性攻击而设置的限制，构成对版权保护的更大威胁，新的水印协议通过建立水印算法与嵌入信号之间的联系，使得攻击者无法混淆版权归属. 新协议与新算法可以整合在一个水印系统中，形成一个可嵌入多比特、并对可逆性攻击和嵌入器攻击具有免疫力的水印系统.

(3) 站在密写方的立场提出了一系列应用于不同类型载体的密写方法，实际上这些密写新方法也是建立在对已有密写技术进行分析，即攻击的基础上提出的. 首先，提出了面向非压缩载体的更为安全的 LSB 密写方案. 与简单 LSB 方案相比既不减少嵌入量，又不增加失真度，而且保证了 $F_1$、$F_{-1}$ 之间的对称性、维持了原始直方图不变，因而可抵抗所有利用直方图异常和 $F_1$、$F_{-1}$ 非对称性的分析. 其次，指出了调色板图像中的 OPA 密写并不安全，可利用奇异颜色检测出秘密信息的存在性，为了对付这种分析，提出了新的调色板密写方法，既消除了奇异颜色又考虑了原始图像局部特性，具有更优良的隐蔽性和安全性. 第三，提出了一种安全的 JPEG 图像密写方案，将秘密信息嵌入非 0 非直流 DCT 系数时，既不改变原始图像的 DCT 系数直方图，也不增加载体图像的分块效应，因此可以有效抵抗基于直方图失真或分块特性的密写分析.

(4) 对基于视觉特性的密写进行了分析并提出安全的密写方案. 通过分析复杂度直方图和象素灰度差值直方图中的异常现象可以有效攻击 BPCS 密写和 PVD 密写，还可以准确估计嵌入量. 而新密写方案能够更细致地分辨每个象素的隐藏能力，并且通过引入混合进制系统，不必在每个象素一定嵌入整数比特，具有更好的灵活性和隐蔽性，而且攻击者在不知密钥的情况下难以

检测秘密信息是否存在.

本文在数字水印方面的工作可以结合在一起，尝试了数字水印技术的安全性集成，将有助于数字水印走向成熟并最终实现全面商用. 本文设计的密写方案具有安全性好、计算量小、负载率高的优点，将有助于密写技术在未来的信息战中发挥重要作用.

**关键词** 信息隐藏, 安全性, 攻击, 数字水印, 密写, 密写分析

## Abstract

As a new branch of information security, information hiding technology including digital watermarking and steganography aims to protect intellectual property rights of multimedia contents and to send secret messages under the cover of a carrier signal. The capability of resisting malicious attacks, *i.e.*, security, is a key requirement in information hiding. While numerous progresses are being made in watermarking and steganographic techniques against various attacks, the opposite side of the combat is also striving to advance their attack strategies in order to defeat the effort of information hider. This thesis summarizes the results obtained in the author's investigation into various aspects of information hiding both in digital watermarking and steganography/steganalysis. The study has been carried out in view of the information warfare with the data hider on one side, and the attacker/steganalyst on the other.

The research achievements on information hiding security are classified into the following four categories.

The first is a study on a number of basic theoretical problems in information hiding. Based on a comparison of the performance of two popular techniques, theoretical analysis and numerical experiments show that the method of addition is suitable for watermark embedding into high frequency components, whereas the quantization technique is more appropriate for embedding into low

frequency components. An improved approach based on sequence addition is proposed that effectively exploits the information carrying capability of each binary sequence by mapping a number of bits to a single sequence picked up from an orthogonal set, leading to significant performance improvements. On the other hand, embedding capacity of the quantization method in the presence of attack and the optimal embedding strategy are studied. It is shown that the maximum attainable payload is determined by the size of cover media and the ratio between distortion levels caused by the data-hider and the active attacker respectively.

The second category is the development of watermarking algorithm or protocol capable of withstanding hostile attacks. It is demonstrated that, by randomly selecting watermark signals, which are mutually independent even derived from a single key, the inserter-based attack can no longer remove the legitimate mark. In addition, with this new method, a trace is left if the product was maliciously tampered. It is shown that, even with a noninvertible watermarking algorithm or an asymmetric watermarking protocol, it is still possible to affect an invertibility attack, which relies on a forged watermarking algorithm, a counterfeit mark, and a fake key. As a solution, a secure watermarking protocol is presented, which establishes correlation between the watermarking algorithm and the embedded mark. The proposed protocol and algorithm can be integrated into a watermarking system.

The third is to propose several secure steganographic schemes used in cover media with different formats. For an uncompressed

cover image, a novel steganographic scheme is described which avoids asymmetry inherent in conventional LSB embedding techniques and minimizes the histogram abnormality so that both histogram-based and asymmetry-based attacks are disabled. It is also shown that some peculiar colors can cause vulnerability of the optimal parity assignment (OPA) steganography in palette images. Based on this analysis, an advanced steganography is developed that, while keeping the advantage of low distortion of the OPA, avoids the above-mentioned peculiar colors and makes use of the local properties in the host image to improve the security of embedded information. In another data hiding technique for JPEG cover image, when AC DCT coefficients with non-zero values are used to carry secret bits, the DCT coefficient histogram and blockiness property are reserved. In this way, the histogram-based and blockiness-based steganalyses are effectively defeated.

The last category is steganographic and steganalytic techniques based on the human vision system. It is shown that, by analyzing the histogram of block complexity or pixel-value difference, the presence of secret message embedded by bit-plane complexity segmentation (BPCS) or pixel-value differencing (PVD) is detectable. In view of this vulnerability, a new steganographic scheme is proposed, in which the capacity for information hiding at each pixel is determined by the gray level variation of its immediate neighboring pixels, thus provides a better security and imperceptibility than BPCS and PVD.

In summary, the achievements on digital watermarking in the

paper, which are respectively immune to different attacks, and their integration, as a valuable attempt for acheiving universal security, will be helpful to consummate the digital watermarking technique. On the other hand, all the proposed steganographic schemes in this thesis possess good security, low computation complexity and high payload, so they will be useful to perform covert communication in the information warfare.

**Key words** information hiding, security, attack, digital watermarking, steganography, steganalysis

# 目　录

# 第一章 绪 论

随着信息时代的到来，数字多媒体产品在人们工作、生活中的作用越来越重要. 由于数字产品非常易于复制和修改，版权保护问题就越来越突出地表现出来. 另一方面，冷战结束后世界格局和国际关系日益复杂，各种势力之间的斗争愈演愈烈，“9.11”事件后的反恐呼声不断高涨，公共传媒中的信息安全也越来越得到人们的关注. 正是在这样的时代背景下，信息安全领域中的一项新兴技术——信息隐藏应运而生，并得到了迅速发展.

## 1.1 信息隐藏概述

信息隐藏(information hiding)是在不对载体信号产生过分影响的条件下将额外的信息嵌入数字媒体中，以实现版权保护、隐蔽通信等功能[1]. 信息隐藏可分为数字水印(digital watermarking)、密写(steganography)、匿名等分支，其中数字水印和密写是最为重要的两项，分别用于版权保护和隐蔽通信. 信息隐藏、数字水印、密写与信息安全的关系见图 1.1.1. 本文的研究工作主要涉及数字水印和密写（密写分析）.

数字水印技术是将标志产品作者、所有者、发行者、使用者、出品时间等信息按一定的算法嵌入载体信息中. 嵌入的水印不能过分影响载体信息的商用价值，并且可以从含水印的载体数据中检测或提取出来. 数字水印的目的是为了保护载体信息的版