多载体安全隐写技术的研究与应用

DUOZAITI ANQUAN YINXIE JISHU DE YANJIU YU YINGYONG

陈够喜 著

国防工业出版社
National Defense Industry Press

多载体安全隐写技术的研究与应用

陈够喜 著

国防工业出版社
·北京·

图书在版编目(CIP)数据

多载体安全隐写技术的研究与应用 / 陈够喜著. —
北京：国防工业出版社，2012.3
ISBN 978-7-118-07940-1

Ⅰ. ①多… Ⅱ. ①陈… Ⅲ. ①信息安全—安全技术—研究
Ⅳ. ①G203

中国版本图书馆 CIP 数据核字(2012)第 023008 号

国防工业出版社出版发行
(北京市海淀区紫竹院南路 23 号 邮政编码 100048)
国防工业出版社印刷厂印刷
新华书店经售
*
开本 850×1168 1/32 **印张** 4 **字数** 96 千字
2012 年 3 月第 1 版第 1 次印刷 **印数** 1—2500 册 **定价** 29.00 元

前　言

随着互联网技术的不断发展和推广，电子政务、电子商务、航天和军工等行业得到了前所未有的提升和发展。伴随互联网技术的蓬勃发展，各种文字、图像、音频和多媒体数据等及时传播到世界各地，人们进行着一种高效和成本低廉的信息传递工作，政府、企业、个人和各种组织等普遍利用这一个开放互联的公共网络信息平台，构建适合自己的信息渠道，传递大量的公开或秘密的信息。同时，对这些平台或网络上的数据信息的安全性保护也引起世界各地用户的普遍关注。

信息隐藏是信息安全领域的一个重要分支，主要通过公开的媒体或公共网络对隐秘信息进行隐蔽传递或存储，采用人眼不易察觉的秘密携带方式实现安全高效的信息传递，掩盖“正在进行通信”的事实。隐写技术是信息隐藏的重要组成部分，是隐蔽通信的重要实现方式。本书在分析单载体隐写的隐写容量、安全性和鲁棒性等局限性基础上，结合前人的研究成果，构建了多载体批量隐写模型，提出了二值图像和灰度图像等批量隐写方案，设计和实现了多载体隐写系统。

本书的主要创新点如下：

(1) 针对单载体隐写容量和安全性的局限性，提出了多载体批量隐写模型。从信息论和假设检验角度证明了在一定约束条件下多载体隐写模型的安全性，并形式化描述了载体的最大有效载荷与载体数量的相关性。实验结果和分析表明了该模型

的有效性。

(2) 基于二值图像分块嵌入思想，提出了基于二值图像的批量隐写算法。结合多载体隐写模型和二值图像的特点，设计了大尺寸批量隐写实现算法，并分析了批量二值图像的隐写编码效率和失真度量。

(3) 面向灰度图像灰度值变化的特点，提出了灰度图像的批量隐写算法，基于矩阵编码技术，实现了高效的批量隐写。结合 Bernstein 多项式，提出了面向分存的批量隐写算法。构建了基于比特流的灰度图像隐写模型，实现了各种数字图像的归一化隐写容量。

总之，通过分析和研究单载体各种隐写技术的局限性，构建多载体批量隐写模型，可以为信息隐藏技术今后的理论研究和发展奠定一定的基础，同时对于研究隐写分析、数字图像认证和网络协议构建也大有帮助。

全文共分为 5 章。第 1 章为绪论，阐述本书的研究基础和研究意义，并介绍目前信息隐藏国内外的研究现状。第 2 章为多载体隐写的基本理论。首先对单载体隐写模型进行了综述，分析了单载体隐写在隐写容量和安全性方面的局限性；然后，提出了多载体隐写的模型，并从信息论和假设检验两个方面讨论了单载体隐写与多载体隐写系统的安全性，形式化地描述了完备多载体隐写系统和不完备多载体隐写系统，并证明了在一定的约束条件下不完备多载体隐写系统的安全性。第 3 章为基于二值图像的批量隐写，研究和分析了目前二值图像隐写的特点，提出了二值图像的批量隐写模型；针对大尺寸图像的批量隐写，实现了隐秘信息的嵌入和提取，并进行了实验验证和分析；最后，研究了批量二值图像的隐写编码和视觉失真度量。第 4 章为基于灰度图像的批量隐写，针对灰度图像的特点，分析了 LSB

隐写、K 隐写、变换域隐写和分存隐写技术，提出了灰度图像的批量隐写算法；基于 Bernstein 多项式，提出了面向分存的批量隐写算法，将各种图像的隐写容量进行归一化处理，提出了基于比特流的灰度图像隐写算法；最后对上述算法进行了实验和结果分析，验证算法的有效性。第 5 章为总结和展望。

本书得到了山西省科技攻关项目(20090322004)资助和国防工业出版社王京涛老师的大力帮助。

由于作者水平所限，书中可能存在诸多纰漏和不足，敬请各位同行专家和广大读者批评指正。

目　录

第1章 绪 论

1.1 引 言

1.1.1 基本概念

信息隐藏技术是20世纪90年代以来在国外兴起的一门前沿技术，属于信息安全领域的研究范畴。信息隐藏技术也称信息伪装术，集信息论、计算机通信、计算机图形与图像、计算机网络攻防与密码分析等多学科知识，是国际学术界的研究热点之一。简而言之，信息隐藏(Information Hiding)是利用公开的载体在时间或空间的冗余特性，将隐秘信息嵌入到公开的载体之中，从而形成了携密载体，实现利用公共网络进行隐秘信息的隐蔽通信或存储[1-3]。这些公开载体包括网页、文字、图形图像、音频和视频、软件等相关数字多媒体信息。隐藏后的隐秘信息对于非授权者一般而言是不可见或触摸不到的，从而达到对版权、密钥和秘密信息的保护或传递。可见，信息隐藏是利用人类感觉器官的不敏感性，以及多媒体数字信号本身的冗余特性，将隐秘信息隐藏在宿主信号之中，不被人感觉或注意到，而且不影响宿主信号的使用效果[4,5]。

1.1.2 信息隐藏的分类

信息隐藏按照不同的标准可以进行不同的分类。图1-1是根据信息隐藏应用背景进行的分类。

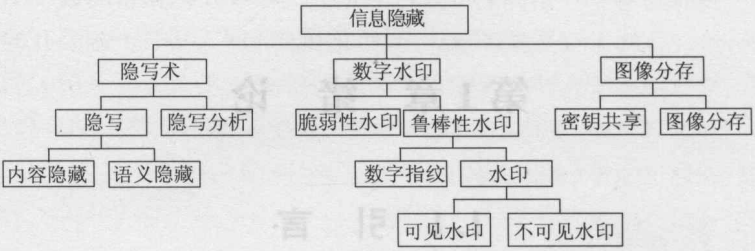

图 1-1　信息隐藏的分类

信息隐藏技术主要包括隐写术(Steganography)、数字水印(Digital Watermarking)和图像分存(Images Sharing)三大类型。隐写术和数字水印的主要区别:前者的目的是隐藏秘密信息,而后者的目的是通过隐藏少量的秘密信息进行版权保护[6]。图像分存是为了将通过置乱、加密和分拆等处理后的秘密信息,嵌入到不同的载体之中,不因其中的一个携密载体损毁而无法提取秘密信息。有时也将两种技术混合使用,以提高攻击与隐写分析的难度。

隐写术分为隐写(Steganography)与隐写分析(Steganalysis)两个方面,是防守与攻击的一对矛盾体[7]。它主要包括空域与时域隐写、变换域隐写、文本隐写和语义隐写等。基于数字图像的典型隐写技术体现在空域与时域之中。在空域中主要的算法有:LSB 隐写、MLSB 隐写、$\pm K$ 与随机调制隐写、调色板图像隐写以及二值图像隐写等,而在变换域或 DCT(Discrete Cosine Transform,DCT 变换)域中主要算法有:Jsteg 隐写、F5 隐写、OutGuess 隐写和 MB 隐写等[8,9]。隐写分析主要分为特定算法隐写分析与通用盲检测两类。

数字水印[10-12]的基本思想是在数字图像、音频和视频等数字多媒体载体之中嵌入秘密信息,其目的是保护该数字产品的版权即产品的真实性或相关信息。嵌入的秘密信息可以使公司标示、产品商标、用户信息或产品的相关信息。根据实际应用需要,分为脆弱性水印和鲁棒性水印。脆弱性水印主要用于保护法律意义上的证据,

而鲁棒性水印用于产品的版权保护。

图像分存[13,14]研究的主要目的是把一幅或几幅秘密的数字图像分解成几幅无意义或者杂乱无章的图像，并嵌入到若干幅公开的数字媒体中进行存储或传输。它与隐写术和数字水印技术最大的不同在于它可以避免由于少数图像信息的丢失而造成严重的事故，同时在通信中个别图像信息的泄露不会引起整个图像信息的丢失。

1.1.3 研究背景

随着 Intranet 网络和计算机技术的广泛应用，电子商务、电子政务、军事、政治以及重要经济信息保护的需求不断提高，信息安全已成为国家的战略安全的一个重要组成部分，成为关系国计民生的重要课题。信息安全的目标是信息的机密性、完整性、抗否认性和可用性。信息安全是一门交叉学科，包括安全性理论研究、应用技术研究和安全管理研究等。安全性理论涉及身份认证、访问控制、审计追踪和安全协议等内容，应用技术研究包括入侵检测、渗透攻击、漏洞扫描、物理安全、网络安全、数据安全、边界安全和用户安全等方面[15]，而信息隐藏技术是信息安全的一个重要分支。

计算机网络的快速普及促进数字媒体业的高速发展，对数字媒体或相关产品的版权保护的需求也越来越高[16-18]。例如：基于网络传输的数字产品的非法下载和使用引起的版权纠纷；基于电子商务活动中的重要信息或数据的篡改；重要的学历和档案、文字材料等的存储和查询等。可见，对数字媒体产品在不影响其使用的情况下进行有效保护和鉴别（包括版权保护、身份认证和票据防伪等）的需求已迫在眉睫。

对于其他领域，隐蔽通信技术的研究已成为各国军事和政治领域的重要议题。10 年前举世震惊的“9·11”恐怖袭击事件结束后，美国前空军情报局作战主任 Marc Enger 说：“美国特工发现，本·拉登的组织曾用信息隐藏技术（将信息隐藏在平面图像里，比如文本信息隐藏在图像文件里）和色情网站进行通信”[19]。由此可

见，研究信息隐藏技术对于打击和监控恐怖分子和“三股势力”的活动，保护国家和平民百姓的安全具有重大意义。而且在重要领域，敏感信息和文件的存储与传输同样具有极高的实用价值[20]。

虽然信息隐藏技术在各国学者的推动下取得许多成果，但是该技术目前仍处于发展阶段。该学科还没有形成自己完整和成熟的理论，因此，从信息对抗的角度出发，尽快构建信息隐藏和检测技术的架构和理论，实现在互联网、局域网及无线传感器网等多种网络环境中信息传递的安全性，以及保证网络运行安全具有重要意义。

1.1.4 信息隐藏与密码学

信息隐藏和密码学是一对相互联系且有区别的重要技术，特别是隐写技术与密码学有着千丝万缕的联系，但它们的实现目的和采用的手段均不同。信息隐藏技术特别是隐写术是利用人类的感觉器官的冗余特性实现将秘密信息嵌入到公开的载体之中或利用公共通道进行秘密信息的隐蔽通信，目的是隐藏秘密信息的存在性。而密码学研究的目的是将秘密信息采用技术手段变为毫无意义的乱码，不能正常理解其含义，但秘密信息就在其中[21-23]。隐写术通常利用数字媒体的冗余特性，让它们携带秘密信息，攻击者无法得到这些携密载体或者尽管得到了携密载体也无法获知秘密信息的存在，对于提取嵌入其中的秘密信息就更难了。

图 1-2 为加密的基本原理框图，图 1-3 为加密与隐写的基本原理对比示意图。在图 1-2 中，加密算法又分为单钥和双钥算法，将明文转变为密文，所以，密钥 1 和密钥 2 可以相同(单钥)，也可以不同(双钥)。现代隐写术常与秘密学紧密结合，增加攻击难度，实现对秘密信息的更安全保护。图 1-4 为现代隐写技术的基本原理，其中信息的加密处理已成为隐写前的预处理部分。

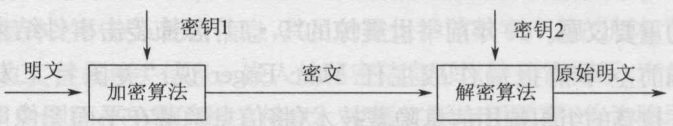

图 1-2 加密的基本原理框图

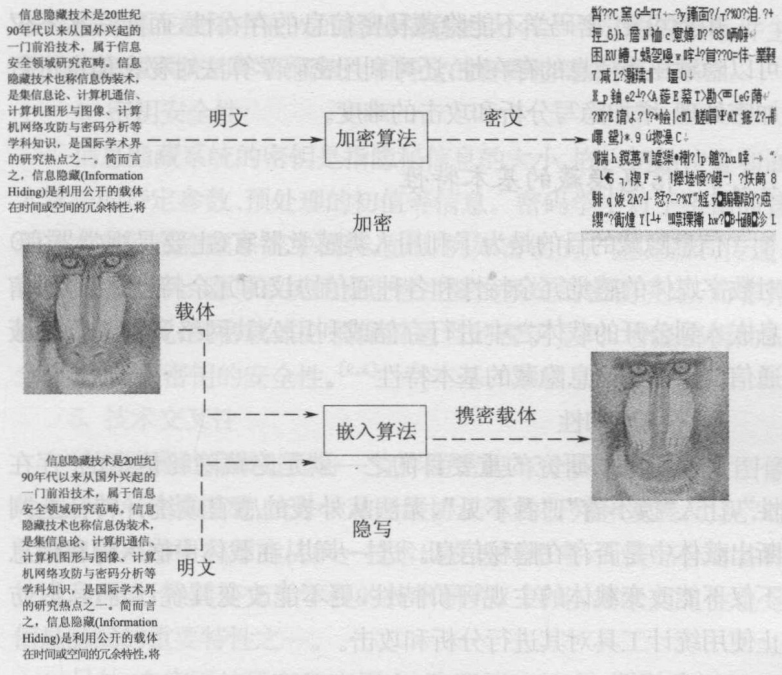

图 1-3　加密与隐写比较

图 1-4　现代隐写术的基本原理

由此可见，密码学不能隐藏秘密信息的存在性，而隐写术不仅可以隐藏秘密信息的存在性，还可利用密码学算法对秘密信息进行加密处理，增加隐写分析和攻击的难度。

1.1.5 信息隐藏的基本特性

信息隐藏的目的是为了利用人类感觉器官（主要是视觉器官）对数字媒体的感觉冗余特性和各种通信协议的冗余特性，将隐秘信息嵌入到公开的载体之中进行存储或利用公共网络资源进行隐蔽通信。以下是信息隐藏的基本特性[3,9]。

1. 不可感知性

信息隐藏的研究的重要目的之一就是隐藏隐秘信息的“存在性”，让人“摸不着”、“看不见”，无法从外表的感官或统计特征上判断出载体中是否存在隐秘信息。进一步讲，在载体中嵌入隐秘信息不仅不能改变载体的主观评价特性，更不能改变其统计特性，以防止使用统计工具对其进行分析和攻击。

2. 鲁棒性

秘密信息在传输中会受到各种主动或被动攻击，可能造成携密载体的部分破坏。鲁棒性就是用来度量隐写系统抗击各种有意或无意攻击携密载体的能力的程度，它是信息隐藏安全性的重要指标之一。通常，在传输秘密信息之前都要进行预处理操作，主要是进行置乱、压缩编码或其他处理，目的是降低原始秘密信息的大小，增加原始秘密信息的安全性，例如基于混沌映射的加密。根据实际需求，确定隐写系统的鲁棒性，不能一味追求高鲁棒性，使得隐写系统的算法复杂度过高，影响秘密信息的提取速度，或者在受到一定的攻击时将无法提取，即隐写系统的自恢复性降低。选择适当的鲁棒性和自恢复性需要综合考虑实际的需求环境情况。

3. 隐写容量

在满足一定的约束条件下，信息隐藏系统或隐写系统最大的携带隐秘信息的比特数称为最大隐写容量，也称为最大有效载荷。隐

写容量就是在载体中嵌入的隐秘信息的大小。一定的约束条件主要是指隐写系统的不可感知性和特定的鲁棒性等条件。

4. 密钥安全性

信息隐藏系统的密钥是指隐秘信息的大小、嵌入初始位置和嵌入算法的特定参数、预处理的初值等信息。密码学中对于密钥的管理技术在信息隐藏系统中同样适用，例如密钥的产生、发放、传递、保存和管理等。隐写系统的安全性主要由系统的密钥决定，而非算法本身。因此，携密载体必须保留适当的密钥空间，增强密钥的安全强度，确保密钥的安全性。

5. 技术交叉性

信息隐藏的发展过程将一直伴随密码与分析学、信息论、图像处理、DNA 计算模型、统计学和网络协议等的研究和发展，可见，对于信息隐藏的研究不能从单一技术出发进行思考，而需综合对各种技术和理论，特别是交叉学科的技术研究。因此，技术交叉性也是信息隐藏的重要特性之一。

另外，在实际的研究和应用中，隐写算法的实时性、复杂度、载体库的选择以及网络配置等也需要综合考虑和分析，才能构造出一种实用、高效和安全的隐写系统。

1.2 信息隐藏的研究意义

从信息安全保护应用角度来说，信息隐藏技术应用主要体现于下述五个方面[24-29]：

(1) 版权保护。随着互联网和电子商务技术的快速发展，基于互联网的数字多媒体信息高速膨胀，从网上完全可以下载和复制数字多媒体产品，因而对其版权的保护尤其重要。基于网络的数字产品包括特供新闻、数字电影、数字图书馆和 MP3 等。这类数字产品的版权保护采用传统的方法无法得以实现，必须利用信息隐藏技术在数字媒体中嵌入鲁棒性数字水印才能够有效解决。当用户通过

合法渠道购买数字产品时，可将版权单位的类似 LOGO 和序列号等相关版权信息嵌入到该数字产品中。如果出现版权纠纷，则可通过提取其中的版权信息来确认真伪，也可对侵犯版权的非法复制者进行进一步的追查和处理，从而保护产品的版权持有人的合法利益。

(2) 身份认证。在互联网上进行通信时，采用信息隐藏技术将通信时的身份特征嵌入到发送和接收到的信息中，防止任何一方的抵赖行为，包括否认发出信息或接收到对方的信息。通过身份认证，确保通信系统的合法和高效运行。

(3) 保密通信。基于信息隐藏技术，将军事、政治、商务和金融等领域的重要数据或文件采用公开载体或公共通道进行隐蔽通信，确保敏感信息的安全存储和传输。目前，在这些领域将有更加广泛的应用前景。

(4) 电子商务中的网页保护与票据防伪。随着电子商务的不断发展，涉及电子商务的网页及其展示的信息必须加以保护，防止非法用户或入侵者篡改和链接，侵犯商家和合法用户的利益。对于使用较广的票据特别是电子票据采用嵌入水印技术进行有效保护和防伪。

(5) 图像取证。数字图像是一把双刃剑，采用各种软件工具对其进行方便快捷的编辑和修改的同时，如何保证原始数字图像的完整性和真实性的问题随之而生。采用数字取证和隐写分析技术，可有效对数字图像的来源进行追踪和认证。

从理论研究角度来说，信息隐藏技术是一门多学科交叉和融合的新技术，对于开展新的研究方向和技术应用，将具有十分广泛和深远的科学意义。

本书研究的意义主要表现在下述五个方面：

(1) 探索多载体隐写的最大/最小有效载荷与载体数的关系。针对不同的载体和人类感觉的最小变化率，研究该载体的最大/最小有效载荷与载体本身的形式化逻辑关系。分析在不同的约束条