总策划：戴清民
主　编：吴汉平

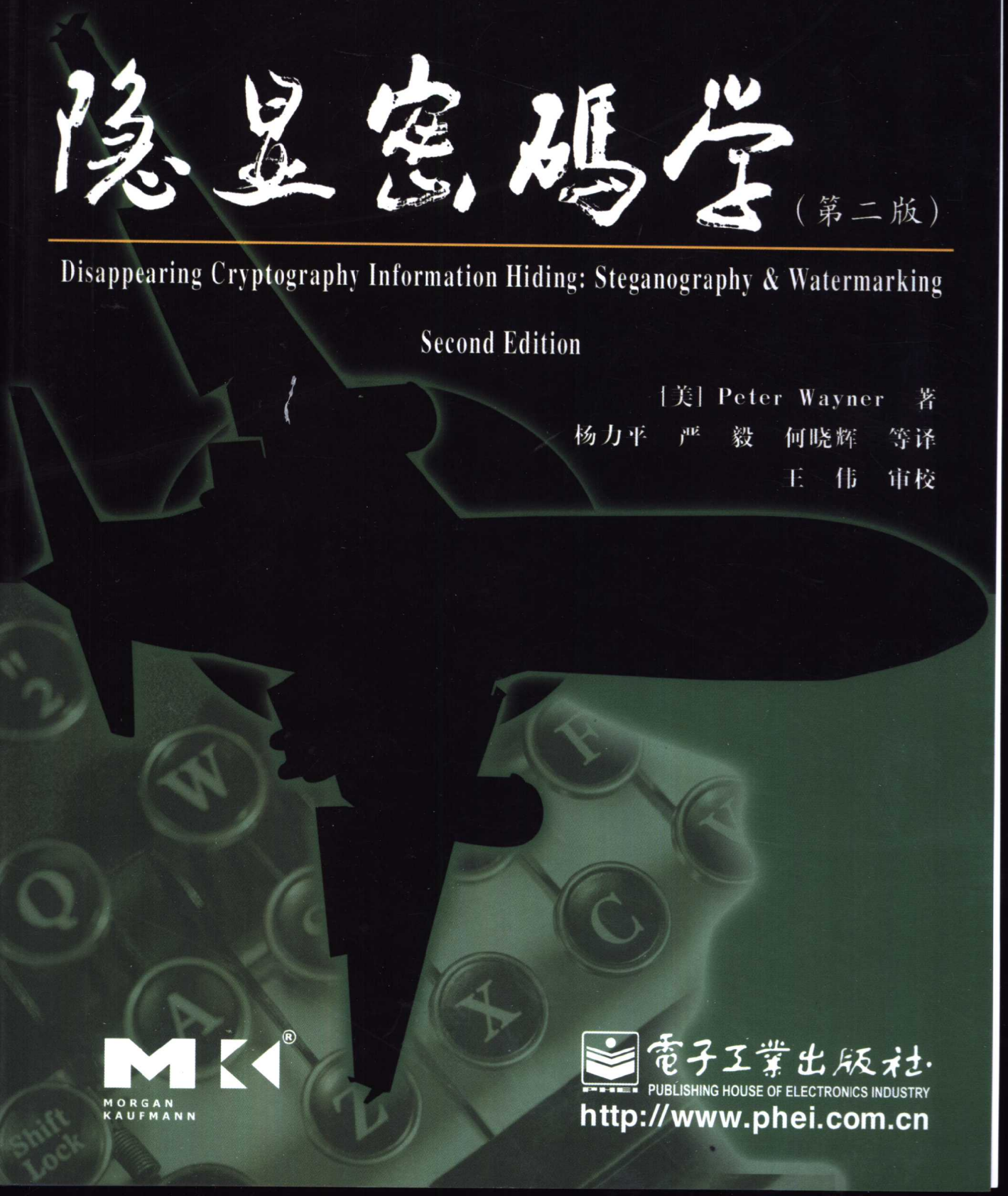
信息战名著翻译丛书

隐写密码学（第二版）

Disappearing Cryptography Information Hiding: Steganography & Watermarking

Second Edition

[美] Peter Wayner　著
杨力平　严　毅　何晓辉　等译
王　伟　审校



MORGAN KAUFMANN

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

总 策 划：戴清民
主　　编：吴汉平
书名题字：戴清民

信息战名著翻译丛书

Disappearing Cryptography

Information Hiding: Steganography & Watermarking Second Edition

隐显密码学

（第二版）

〔美〕 Peter Wayner 著

杨力平　严　毅　何晓辉　等译

王　伟　审校

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 提 要

本书描述了怎样利用字词、声音、图像以及怎样在数字数据中把它们隐藏起来，以便使它们看上去像其他的字词、声音、图像。当这些强大的技术被正确地使用时，就能使得跟踪信息发送者和信息接收者几乎成为不可能的事情。这些技术包括加密术（使数据变得不可理解），隐写术（把信息嵌入视频、音频和图形文件），水印术（把数据隐藏在图像和声音文件的噪声里），模仿（给数据穿上“衣服”使得它看上去像其他的数据），还有其他的一些技术。全书共17章，每一章都分为了几个小节，对那些想了解概念而又不想通读技术说明的人来说，每一小节都为他们提供了一个介绍和高水平的概要，并且为那些想自己编写程序的读者提供了更全面的细节。

本书适合信息安全管理人士、管理人员、军方及其他对信息安全感兴趣的读者阅读。

版权贸易合同登记号：01-2002-4132

图书在版编目（CIP）数据

隐显密码学（第二版）/（美）沃纳（Wayner, P.）著；杨力平等译. 一北京：电子工业出版社，2003.2
（信息战名著翻译丛书）

书名原文：Disappearing Cryptography Information Hiding: Steganography & Watermarking Second Edition

ISBN 7-5053-8395-7

Ⅰ. 隐… Ⅱ. ①沃… ②杨… Ⅲ. 密码—理论 Ⅳ. TN918.1

中国版本图书馆CIP数据核字（2002）第104583号

责任编辑：郝黎明
印　　刷：北京天竺颖华印刷厂
出版发行：电子工业出版社　http://www.phei.com.cn
　　北京市海淀区万寿路173信箱　邮编：100036
　　北京市海淀区翠微东里甲2号　邮编：100036
经　　销：各地新华书店
开　　本：787×1092　1/16　印张：18.5　字数：460千字
版　　次：2003年2月第1版　2003年2月第1次印刷
定　　价：32.00元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换，若书店售缺，请与本社发行部联系。联系电话：（010）68279077

对隐显密码学的批判和赞扬

在研究隐匿信息的领域里，隐显密码技术是机智的和有趣的。Peter Wayner先生在密码学这个领域就许多技术、应用和研究方向等提出了很直观的观点。本书以涉及知识面广博而使其成为真正的、独一无二的著作。本书是那些想开始研究隐藏信息的人的必读物。

Deepa Kundur博士，多伦多大学

Peter Wayner的《隐显密码学》非常透彻、精确并很有趣，它对语法、代码、研究等方面进行的确定显得特别有意义，而这些语法、代码、研究正是SDMI和RIAA这样的组织赢得“数字千年版权行动”的重要因素。Peter Wayner以易于接受的方式，介绍了隐藏信息的所有主要技术以及对其所进行的攻击。Peter Wayner的这本书很容易会成为标准密码参考书。

Mike Stay, staym@datawest.net

本书第二版为隐藏密码世界和隐藏真实信息科学提供了一个非常现代的、易于接受的介绍。本书囊括了围绕密码学和水印学竞争在内的各方面技术，这些竞争包括统计学家和密码学家之间的竞争，水印机和自由旋转数字复印的竞争。本书易于接受，知识范围广博。它含有经过整理的实例、源代码及有关文献的软件包及其参考附注的评论。

Adam Back博士，Cypherspace Internet Secutity创始人

致　　谢

本书已是第二版，所以，在这里更要感谢广大朋友的帮助。毫无疑问，我还欠所有“加密和编码”邮件列表工作的参与者一份感激，是他们最初的贡献才激励我写了第一本书，他们始终如一的工作热情使得这份邮件列表成为获取信息的最好来源。

一些最新的邮件列表都更多地把焦点放在了这个主题上。“水印”列表和“密码”列表都是以良好的信噪比系数提供高质量的讨论，其他像“RISKS”和“Dave Farber's Interest People”这些列表帮助提供了一些意想不到的途径。当然，像Slashdot，Kuro5hin和InfoAnarchy这些现代的列表网址是由固定的、适度的讨论来提供的，这些讨论可以帮助信号跳出噪声。对那些在高质量邮件里为我们提交足量的固定信息和深层次想法的社区成员来说，我们不可能根据名字一个一个地道谢。

信息隐藏专题研讨会的组织者通过举办关于这个主题的优秀的研讨会给这片领域带来了学术的严密性。创建、编辑、审查和出版书稿的规程在很多方法上提升了这一艺术形态。由Springer-Verlag出版的论文集对那些对这一领域发展有兴趣的人来说，是个很好的资源。

还有其他一些人从其他方面给了我很大的帮助。Peter Neumann仔细审查了本书初稿并提出了很好的改进建议。Bruce Schneier也非常热心，他从他的第一本书[Sch94]中给了我们文献目录的电子版本。我把它转变为Bibtex格式并在这本书里用它作一些参考。另外Anderson对文献目录的注解也给了我很大的帮助。

Scott Craver, Frank Hartung, Deepa Kundur, Mike Stay和三个匿名评论者查阅了第二版本。他们的评论帮助我修改了很多错误并提出了很多改善的建议。DavidMolnar和Greg Ruse也帮助鉴别了第一版的错误。

这本书最初是由AP PROFESSION（Harcourt-Brace公司的一家分公司）出版的。出版第一版的负责人包括：Chuck Glaser, Jeff Pepper, Mike Williams, Barbara Northcott, Don Deland, Tom Ryan, Josh Mills, Gael Tannwnbaum和Dave Hannon。

当然，这个新的版本如果没有得到Morgan Kaufmann的Tim Cox的支持的话，也不可能存在。感谢Tim、Stacie Pierce和Howard Severson对我的帮助以及鼓励。

译 者 序

在翻译本书的过程中，我们一直在想，翻译和出版这本《隐显密码学（第二版）》是很有意义的。书中介绍的用于隐藏信息的隐写术，是一种很独特的、重要而又有趣的信息隐藏技术，它与信息加密技术是不同的。它们两者虽然都能保护通信的秘密，但信息加密技术是明明白白地告诉别人：你传递了秘密的东西。而隐藏信息的隐写术也能传递秘密，并且别人不知道你曾经传递了秘密的东西，这是非常重要的。掌握这种技术，对于保护自己、保护好人，以及打击坏人都是很有作用的。本书的作者在书中也提到“9·11”世贸大厦受恐怖袭击的事件，在恐怖主义猖獗的今天，我们国家的有关部门和有关研究机构，以及信息安全爱好者都应当注重研究、掌握这门技术。

本书与原著一样有三种注解。一种是在一句话的后面标上“①”、“②”等，然后在该页的下端作出具体注解。第二种是在一句话的后面用方括号标出，如“[Blu82] ”、“[JJ98a]”等，这是关于参考书的注解，根据它读者可以到本书后部附录中找到具体的参考书目。第三种我们称其为“旁注”，在原著中，它是标在一段话的旁边，即书页的外侧空白处。本书中，将其放在这段话的后面，用楷体字标明是旁注。

由于本书所讲述的是一门比较新的技术，有的术语很不好译。同时，作者还经常以调侃的口吻说话，需要使之转换成适合书面的语言。经过我们的努力，并多方请教有关人士，终于比较准确地把全书翻译出来了。在此，对我们在北京的朋友张宏先生表示感谢，对译典通知讯网（http://www.dreye.com.cn/）、汉英论坛（http://www.overwaters.com/forum/contact.htm）以及论坛上的朋友们表示感谢。此外，读者对书中的阐述如有疑问，可以与我们联系（ylpcn@263.net）。

前　言

关于第二版的说明

自从本书的第一版出版后，密码学和隐显信息领域在近5年里发生了巨大的变化。来自科学界的兴趣增长了，专注于有关课题的专门会议也增多了，而且，出现了很多新观念、方法和技术，它们中的许多都包含在这本书里。

这种正在萌芽发展的兴趣并不只局限于科学实验室，企业界已经把目光投入到了这个领域，他们希望隐显信息的使用能够给影音和影像的创建者有一个控制结果的机会。这种隐显信息通常被称做水印。隐显有效负载可以包括有关创建者、版权持有者及购买者的信息，甚至包括有关谁被允许享用信息，以及他们隔多长时间被允许推动按钮的专门说明。

虽然私人公司也一直在致力于信息隐藏技术的研究，但有些时候，科学进步的推动并不与一些企业界的愿望要求相调和。科学家们想让有关隐写术运算法则优缺点的报道自由地交流，而一些企业家则害怕这些信息将会被用来攻击他们的系统，所以他们宁愿把这知识隐藏起来。

当记录行业开始聚焦于Scotta A. Craver、John P. Dan S. Wallach、Drew Dean和Edward W. Felten这些人所做的工作时，这种斗争就爆发成为一场公开的战役。这些程序员攻击了许多由SDMI（Secure Digital Music Initiative，安全数码音乐小组）分配的技术，这个组织是由音乐行业的一些成员发起主办的，主要致力于生产水印系统。这些攻击是SDMI在一个旨在测试运算法则力度的公开竞赛上招致的。不幸的是，SDMI的领导者想通过强制那些进入竞赛的人签署一份有关收集奖金的秘密誓约来制约他们。其实，SDMI是试图使想传播他们对公众详细审查结果的人保持沉默，并试图获得公众详细审查的所有政治优势。当这些程序员2001年4月在匹兹堡（美国宾夕法尼亚州西南部城市）的信息隐藏专题研究组上呈出他们的作品时，RIAA（Recording Industry Association of America，美国记录工业协会）发给他们一封信并建议公众讨论这些程序员是否应该被诉讼惩罚。程序员们则在信中加了些自己的方式内容，就是他们称RIAA和音乐业试图抑制他们的第一修改权。这些程序员后来于2001年8月在华盛顿的USENIX商讨会上呈现了他们的作品，很明显，战争的导火线依然存在。一方是希望信息公开共享（即使它可能产生不愉快的结果）的人，另一方是希望建立审查制度从而保持世界安定的人。

这种冲突看上去好像来自于对隐藏信息运算法则的脆弱的察觉。如果有人知道剧本播放机制，他们就可以通过重写或制造不规则的噪声来毁坏信息。记录业一直在担心有人会使用这种怎样破坏SDMI运算法则的知识来摧毁水印信息——当然，这种事很难做。在某些人眼里，惟一的解决办法就是通过禁止交流技术知识来加强安全保护。

这种看法和密码学所持有的观点完全不同。大多数行业认为，公众安全技术是创建安全运算法则的最佳方法。通过隐匿所实现的安全并不像设计良好的运算法则那样成功。结果，

公众安全技术便识别了密码运算法则的许多弱点，并且帮助研究者发展出能够久经考验的解决方法。

一些生产水印工具的公司感觉，除了保密以外，它们别无选择。水印工具还不足以安全到可以抵制攻击，所以，这些公司希望使用一些附加的安全技术来使得它们更加安全。

不幸的是，附加安全技术带来了额外的麻烦。隐藏信息很容易通过压缩、重新格式化和重新记录伪装信息等方法删除。在录音棚、音像商店和打印商店中使用的大多数普通工具都可以删除水印，对此你完全不用多虑。位就是位，信息就是信息，两个事物之间不可能有一成不变的固定的联系。

本书中关于版权持有者和科学家之间的斗争只是一个开始。隐秘的运算法则从来都没有被很久地使用，也没有任何原因来解释它们为什么要被使用下去。与此同时，希望读者能享受这本书带来的所有信息。你可以尽情地享受本书，无论你花多长时间读完它。

使用说明

本书的前几章主要讲的是构成基本的“欺骗包”的要素，例如私钥密码术、秘密共享和错误校正代码。后几章主要描述怎样用多种途径对隐藏信息使用这些技术。每一章都会给读者一个介绍，如果你想加以利用的话，所介绍的这些信息足够你能使用这些数据。

每一章的信息都是按照重点和难点的顺序安排的。对那些不想费力阅读技术细节又想理解概要的人来说，本书每一章的开头都有高水平的内容提要，对那些想通过掌握信息来创建自己的程序的人来说，本书对每套细节都有介绍。对更深层次的、较多数学的细节不感兴趣的人可以跳过每一章的最后一部分，但注意不要遗漏任何重要内容。对采用运算法则有兴趣的程序员则可以钻研一下最后一部分。

所有的章节（除第1章外）都是以讽喻的叙述来试图阐述章节的观点。你可能发现这些叙述有些很有趣，有些很愚蠢，但是希望你能在这些叙述的过程中获得更好的洞察力。在每一章的最后有一个章节小结，列出了这一章的主要重点。

本书的大部分都是以有趣的方式提供信息。但知识就是力量，有力量的人都想提高自己的控制力，所以本书的最后一章就是一些有关政治问题的评论。

目　录

第1章 概 述

1.1 简介

大家都知道，计算机里的信息看起来是被完美地定义和确定的，比如银行账目是1 432 442美元还是8.32美元，气温是73° F还是74° F，会议是下午4点开还是在4点半开。计算机仅仅能处理数字，而且数字还必须是准确的。

但生活并不是简单的。信号器和电子器件生产商假装认为数字数据是完美而不可改变的，就像是一个数字的琥珀结晶体一般。但是，自然界充满了噪声，数字只能近似地表现事情的发生。数字信息比自然界提供给它的信息要精确得多。

数字本身就是奇特的东西。所有的数字都能确定地运用算法、方程来进行运算，但同时也能用数值来误导和示错，用欺骗技巧来加密。统计学家称之为数字欺骗。汽车经销商和会计员可以在资产负债表中隐藏违规活动。加密可以只按一个按钮就使成批的数字看起来像是另一组数字。

语言本身常常超过了理性思想的把握。作者围绕着论题摇头晃脑和思考，并依赖细微差别、折射、暗示、隐喻和众多的修饰来表达信息。这些修饰工具没有一样是完美的，人们像是找一个方法辩论“是”那个字的定义。

本书描述了如何利用这些不确定性和不完全性来隐藏信息。本书介绍了如何把字符、声音、图像隐藏到数字数据当中，并使它们看起来像别的字符、声音、图像，把秘密转换成无害的噪声以便使秘密通过网络，并在位流的海洋之中消失。本书描述了如何使数据模拟别的数据来伪装其信号源，并隐藏数据的目的地；如何将会话浸没在噪声的洪流中，以至让别人无法知道会话的存在与否；如何把存在溶化成为虚无的状态，并把它们从虚无之中转换出来，使之再次存在。

传统的密码学以用数学的安全方法来锁住信息的秘密而取得成功。隐藏数据使之不被发现是一个相似的但又有其独特特性的过程，但也常常称之为密码学。这有许多历史性的例子，包括隔室隐藏和机械系统，例如微粒的、或者是瞬时脉冲传输，这将使信息很难被发现。还有别的技术，例如对单词的第一个字母进行编码来伪装内容，使之看起来像别的内容。所有的这些方法都被不断地使用。

数字信息提供了极好的机会，而不仅仅是隐藏信息，还可以为隐藏数据开发一般的理论框架结构。它使得从理论上描述一般的算法，以及描述一个没有密钥的人要想找到数据的难度成为可能。一些算法为自身的强度提供了很好的模型，而另一些并没有。

David Kahn的电码译员为这种技术提供了一个很好的发展历史[Kah67]。

隐藏信息的一些算法使用密钥来控制它们的工作。本书的一些算法用这样一个方法隐藏信息：它有可能再现这些信息而不需要知道密钥。这听起来像是密码学，尽管它是在伪装中隐藏信息。

把这些算法当做隐写术还是密码学？要想在它们之间划清一条界线是武断和危险的。许多很好的加密工具也能产生数据使之像是完美的随机数字，可以说这些工具试图用于欺骗，它像是产生随机噪声来隐藏信息。另一方面，许多密码加密算法是在学习了隐藏信息如何被发现的经验后，才被设计得不容易被直接破解。在一个阵营里面放置一个算法后，时常又忘记它为什么可以在另一个阵营中存在，所以，最好的解决方案就是把本书当做是处理数据的工具集，在提供每一种工具的同时都提供一些指示错误和一些安全性，用户可以结合几种不同的工具来完成自己的目的。

本书出版时的题目是“Disappearing Cryptography”（隐显密码学），这有个简单的原因，因为本书刚面世的时候还很少有人懂得“steganography”（隐写术）这个单词的意思。现在我们仍然保持这个题目，理由是题目并不是购买者靠封面判断书籍的一个方便工具。如果认为这些算法仅仅是隐藏信息的工具那就大错特错了：一些算法提供加密安全的同时，也提供有效的伪装；当一些算法表现出独立性时，另一些则在加密算法有了深层次的理论研究；当一些算法只是基础性的保护时，另一些则是不依靠密钥就极难破解。所以，试图把算法分类为纯粹的隐写术和密码学只是强人所难。

当我们的全部生活都可能是数字信息时，可以设想信息的形式、形状和外表将会有无穷大的量。

1.2 加密的原因

使用本书中的技术可能有好几种理由，其中一些原因是低俗下流的，甚至是麻烦的——我们所称的四骑士——贩毒者、恐怖分子、儿童色情作品作者和洗黑钱者将找到一个方法，就像和利用电话、汽车、飞机、麻醉药、割箱器、小刀、图书馆、摄像机以及其他公用设施一样，为了他们的非法利益而使用这些加密工具和技术。这就不需要解释，为什么这些人能隐藏在匿名和秘密的面纱后面犯下凶残的罪行。

但是，这些工具和技术同样能够保护弱者。在保护本书的利益的同时，下面列出了一些可能好的用法：

1. 你能寻求深层次的个人问题的咨询，例如自杀的想法。
2. 你可以告知同事和朋友有关气味和个人卫生保健问题。
3. 你能安全地约会浪漫的伴侣。
4. 你可以扮演多种角色和以不同的身份开玩笑。
5. 你可以调查工作的可能性而不需要展示当前的工作和可能失去的工作。
6. 当你受到指责的时候，可以向当局匿名检举某人。
7. 你能向报纸杂志揭露不公平和不正当行为的数据。
8. 你可以参加好争吵的政治辩论而不伤害与恰好在辩论另一方的人的友谊。

9. 可以保护你的私人信息不被贩毒者、恐怖分子、儿童色情作品作者和洗黑钱者利用。

10. 警方可以通过秘密工具传递信息以便渗透到坏人之中。

还有许多别的原因，但作者十分惊讶政府官方并没有意识到这些自由在世界上的重要性。许多政府的功能只是幕后交易和权利游戏，匿名通信反映了政治水平的标准。作者常常相信政府摩擦会停止，当一些人想要信息受到约束的时候，信息就能够受到严格的控制，没有人会去做任何工作。他们只需花数小时的时间讨论：谁可以访问这些信息，谁不能访问这些信息。

总结中更详细地检查了这种技术的希望性和危险性。

举例来说，美国中央情报局曾被指责遗漏了前苏联即将解体的信息，他们对国家内讧时发展的苏维埃军队继续发布悲观的评估。一些人将此归咎于贪欲、权力和政治，作者认为应完全归咎于保护信息秘密的无能。情报首脑Bob不能从情报首脑Frad那儿共享秘密数据，因为每样东西都是独立划分的。当人们没有获得新的或可靠的信息，他们就只得仰赖他们的基本成见，这就会出现他们认为苏联是个新生帝国的情形。他们常常对一些问题进行隐蔽的分析，但这样将导致更低的效率。

信息的匿名散播是对社会这个吱吱作响的轮子的润滑剂。只要人们询问了它的有效性，并且认识到它的来源后，就不会乐意跟随在本文之后了，他们每个人都应该能够使用信息的功能了。当匿名信息正确到来时，匿名信息就仅仅是信息。它只是位流，不是子弹，也不是炸弹或者猛烈的攻击。共享信息通常能帮助社会寻求正当的利益。

秘密通信是安全的核心。警察和保安部门不仅仅是用来保护进度表、计划和商业事务的人。本书里的算法就像大门上和汽车上的锁，它将这个权力给每一个人，给每一个人这种权力去保护自己，防卫别人的犯罪和诋毁。这样，警察就不需要去每一个地方，因为人们能够自我保护。

因为这种种原因或者更多的原因，这些算法对保护个人以及私人数据是一个强有力的工具。

1.3 它是如何工作的

隐藏信息的方法有好多种。有一些提供了秘密行动，但并不是所有的这些方法都同样的强固。有些提供给用户一些初学层次的范例式的帮助，有些是采用了大量的自动化处理，还有一些为了安全还能结合特别需求提供多重的保护层次。它们在文件中全都采用一些随机位、一些不确定位和一些未指定位。下面是本书所使用的技术的概要。

- **使用噪声** 最简单的技术就是把信息和噪声放到一幅图片或是一组声音里面，一个数字文件是由一组用来表征光线和声音的强度在时间和空间上精确的点的数字组成。通常，这些数字都被额外精密地计算过，这样才能有效地避免被人所发觉。例如，一张图片上的一个点，可以由220个蓝色单位在总单位从0到255之间的区间内的变化所表征，当这个点由220个蓝色单位转换为219个蓝色单位时，用肉眼将不可能察觉。

如果这个过程被系统地运作，将可以在人眼的感觉极限范围下隐藏大量的信息。一张数字图像有2048×3072个像素，每个像素都包含有24位有关图像颜色的信息。多达756KB的数据可以被隐藏在每一个像素的每一种颜色里至少3位有意义的位之中。这可能比这本书的文字还要多。人类的眼睛不能够察觉这微妙的变化，但是电脑可以把它全部重新构造出来。

- **向外传播信息** 许多复杂的装置能够向外传送含有像素和瞬时声音的文件。无论是人类还是计算机的这种散布，它保护了信息，而且也不容易察觉。许多技术都属于这种类别，工程师首先用这种技术来使无线通信场地减少干扰、抑制干扰和添加一些秘密。现在，要把它们改编成为数字通信并不困难。向外传播信息常常要增加结构的回复力以防止随机的和恶意的威胁。向外传播数据通常这样描述：并不是所有的位都需要重编原始数据，如果一些部分被损坏了，信息仍然能够穿过。许多这样的传播技术把信息隐藏在图片或者声音文件的噪声里面，但这也能被传递其他形式数据所使用。
- **采用统计描绘** 数据常常隐藏在图形中，而计算机则试着通过查看图形来决定数据。例如，在英文文本里，使用字母“p”的时候远远多于使用字母“q”。如果数据能够还原，采用英语语言的统计描述，这样，计算机系统就能够注意到是“p”还是“q”。

很多技术都和产生加密安全随机数字的过程密切相关——那就是说，一个随机数字通量是不能被预测的。一些算法使用数字通量来选择具体位置，而另一些则用隐藏信息把这些随机数值混合在一起，并用信息代替其中一些随机数值。

- **采用结构描绘** 模拟文件类型的统计量仅仅是一个开始。许多复杂的解决方案依赖建立在基础数据的复杂模型中以便更好地摹仿。例如，信息可以隐藏为看起来像棒球比赛的成绩单一样。这些位被隐藏在用来选择的名词、动词和文本的其他部分之间，数据通过文本的分类，以及把单词和位选择配对而得以恢复。这种技术能产生惊人的结果，尽管信息的容量常常看上去是多环的和无方向的，这对于“笨”一点的人和被编程用来扫描特殊字样的电脑来说，已经完全能够解决。
- **随机替代** 许多软件程序使用随机数字发生器来增加场景、声音和游戏的真实感。怪物在数学上的定义看上去更像是随机数字发生器在一张平滑的皮肤上加上斑点、瑕疵、痣、疤痕和沟槽。信息能够隐藏在这些随机数字当中，斑点和疤痕的位置都可以携带信息。
- **次序变换** 一张食品单或许仅仅就只是一张食品单，但那些项目的次序可能携带着大量的惊人的信息。
- **分离信息** 首先声明，这并不是数据需要通过数据包传播的原因。分离信息技术的做法是把数据分成许多数据包通过不同的路径到达目的地。复杂算法能够通过分离信息使n个部分的k个子集被重组为完整的信息。
- **隐藏信息源** 一些算法允许人们广播式地发布信息而不公开他们的身份，即不公开信息的来源。这和隐藏他们的信息本身不同，在某些情况下这是很有价值的。

上述这些不同的技术可以通过不同的方法结合起来应用。首先，信息可以隐藏在序列当中；其次，再隐藏到某些文件的噪声当中；然后，隐藏数据信息源进行广播。

1.4 隐写术的用法

在产品和协议中隐藏数据有多种使用方法。隐藏稍微不同的信息或者结合使用不同的算法可以创建不同用法的不同工具。这里有一些最有趣的应用：

- **改进数据结构** 许多程序员都知道，标准数据结构已经陈旧过时，在新的时期计划外的信息就要既能添加到格式当中又不损坏旧的软件。隐写术就是一个解决方案。举个例子，你添加有关相片的额外信息到相片本身里，这个信息将跟随相片一起传播，但并不干扰那些旧的、不知道信息存在的软件。一些研究者建议了另一个用法：你可以嵌入一段从在后台做X射线数字化的放射线学家的注释。这个文件仍然能在标准工具下使用，节约了医院更新设备的费用。
- **强力水印** 有些数字内容，比如书籍、电影和音像文件的创建者，可能需要添加隐藏信息到文件中，用来描述他们对文件设定的限制。这些信息简单的可能像 “This file copyright 2002 by Big Fun”，复杂的可能像“This file can only be played twice before 12/31/2002 unless you purchase three cases of soda and submit their bottle tops for rebate, in which case you get four song plays for every bottle top.”

由Ingemar J. Cox，Matthew L. Miller和Jeffrey A. Bloom编写的*Digital Watermarking*是对水印以及对子域（subfield）独特挑战的一个很好的介绍[CMB01]。

一些水印意味着文件经过大量失真后仍能找见。较理想的是水印应该在被修剪、选择、转换和压缩后仍然可以被发觉。惟一真正摧毁水印的途径就是改变文件，使文件不再被识别。

另有一些水印则故意做得尽可能的脆弱，其用途是如果某人想篡改文件，水印则会消失。当文件有可能被篡改的时候，将强力水印和脆弱水印结合应用不失是一个好方法。

- **文件跟踪工具** 隐藏信息可以识别文件的合法用户。如果文件被泄露或散布给未经授权的人，它可以跟踪并将有关情况返回到合法的用户那里。对于作品生产商和使用分级信息的政府机构来说，在每个文件添加单独的标签是一个很有吸引力的主意。
- **文件识别** 包含在文档里的隐藏信息同样包含有数字签名以确保真实性。一个常规软件程序只是显示（或演示）文档，但如果某人需要一些保证，就可以对文件里包含的经过认证标志的数字签名进行核实。
- **秘密通信** 在通信变得不安全后，隐藏信息和保留匿名的软件在政治环境也同样有用，这常常是两个人不能交换信息的时候，因为他们的敌人在监听。许多政府部门仍然浏览网络，进行电子会谈，这正是监视的好机会。在这种情况下，隐藏频道就提供一个政治上脆弱的机会，以逃避控制网络的强权[Sha01]。

隐藏信息的许多使用并不像隐写术和加密学那样分类。任何想要处理旧的数据格式和旧的软件的人都知道，程序员不会总是提供所有文件理想的数据结构。许多基础工具和本书里的隐写工具并没有什么太大的不同。智能一些的程序能包装一些以前从未用过的信息来扩展数据格式，这种程序能比人们对隐写技术所设想的用途产生更多的应用。也许在某一天的某一个地方，拯救一个孩子的生命还要感谢智能数据的操作和隐写术呢！

1.5 隐写术受到的攻击

隐写算法为信息提供保密和安全，可是保密和安全的等级是很难测量的。当一个数据混合进背景之后，什么时候才能有效地消失？评价某个隐写算法的强度的一个方法，就是尝试进行不同的攻击，然后评定算法所抵挡住的攻击。这种方法要达到完美还有很远的距离，但却是最实用的。即使你不断地尝试，也没有任何方法可以预测到所有的攻击。

攻击隐写算法和攻击加密算法非常类似，应用到很多相同的技术。当然，隐写算法允诺了一些保密和安全，更容易受到更多额外的攻击。

下面列出了一些可能受到的攻击：

- **针对文件** 攻击者访问文件并判断里面是否有隐藏信息。这是最弱的攻击方式，但这也正是隐写术起码应抵挡住的攻击的最小极限。许多这种形式的基础攻击依靠对数字图像或声音的统计分析显示文件里面信息的存在。这种类型的攻击对科学来说更多的是技巧，因为人们隐藏信息通过调整统计可以尝试着计算出攻击。
- **文件和原始拷贝** 在某些情况下，攻击者可能需要一份编码信息的拷贝和一份原始的、未编码信息的拷贝。显然，侦测出某些隐藏信息是微不足道的操作，如果两个文件不一样，那么里面肯定含有隐藏信息。真正的问题在于攻击者将对这些数据进行怎样的处理。攻击者可能会销毁隐藏信息，可以通过置换原始信息来达到目的。攻击者也可能尝试分离出信息，或者用他自己的信息来置换。
- **多重编码文件** 攻击者用n种不同的信息获得文件n种不同的拷贝，它们其中的一个可能是或可能不是原始未改变的文件。这种情形可能发生在当公司在把不同的跟踪信息插入到每个文件的时候，攻击者就可以聚集大量不同的版本。如果音乐公司出售有私人化水印的数字音乐文件，那么几个有合法拷贝的歌迷就可以在一起比较彼此的文件。一些攻击者可能会试着破坏跟踪信息，而另一些则试着置换为自己的版本信息。在这种情况下最简单的攻击就是把文件混合在一块，平均化文件的每一个要素，或者创建一个由各个文件提取出的不同部分的混和物。
- **访问文件和算法** 一个理想的隐写加密算法可以经受得起推敲，即使攻击者知道了它的算法。很明显，简单的隐写信息算法不能抵挡这种攻击，任何知道算法的人都能用它提取信息。但是如果你能保持加密算法秘密的一些部分，并且使用密钥来解密信息，就能够抵御这些攻击。本书里的许多算法就是使用密码保护随机数字发生器，从而控制信息是如何隐藏进文件中的。这些随机数字流的作用就像是一把密钥，如果你不知道它，就无法产生随机数字流，也就不能恢复混合隐藏后的信息。