〔美〕D.E.R.丹宁 著

密码学与数据安全

王育民 肖国镇 译 肖国镇 校

国防工业出版社

密码学与数据安全

〔美〕 D. E. R. 丹宁 著
王育民 肖国镇 译
肖国镇 校

国防工业出版社

(京)新登字 106 号

内 容 简 介

数据安全性是防止计算机和通信系统中的数据被非授权泄露或修改的科学和研究方法。本书的目的是介绍数据安全性的数学原理以及阐明这些原理如何用于操作系统、数据库系统和计算机网。其内容包括:引论;加密算法;密码技术;接入控制;信息流控制;推理控制。

本书曾作为普都(Purdue)大学研究生一学期的计算机科学课程,适宜计算机专业,通信专业的研究生、大学高年级学生做为教材或自学参考书。

CRYPTOGRAPHY AND DATA SECURITY

Dorothy Elizabeth Robling Denning

Addison-Wesley Publishing Company 1982

*

密 码 学 与 数 据 安 全

[美]D. E. R. 丹宁 著

王育民 肖国镇 译

肖国镇 校

*

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 100044)

新华书店经售

国防工业出版社印刷厂印装

*

850×1168 1/32 印张 14 7/8 391 千字

1991 年 11 月第一版 1991 年 11 月第一次印刷 印数:0 001- 1000 册

ISBN 7-118-00434-0/TN · 87 定价:12.10 元

前　言

电子计算机已经从四十年代稀有的实验性企业发展成为八十年代丰富多彩的数据处理系统。当我们转向依赖这些系统去处理和存储数据时，我们也就开始担心有关它们保护有价值数据的能力了。

数据安全性是防止计算机和通信系统中的数据被非授权泄露或修改的科学和研究方法。本书的目的是介绍数据安全性的数学原理，以及阐明这些原理如何用于操作系统、数据库系统和计算机网。本书适用于想了解这些原理的入门知识的学生和专业人员。书中列出了许多参考文献，想进一步研究特殊专题的读者可以参考。

数据安全问题自 1975 年以来一直在迅速地发展。我们已经看到了在密码学上令人鼓舞的进展：公钥加密、数字签字、数据加密标准、密钥安全保卫体制和密钥分配协议。我们已经发展了证实程序不泄露机密数据，或以较低的安全许可级将机密数据传给用户的技术。我们已经找到了为保护统计数据库中数据的新的控制方法，以及攻击这些数据库的新方法。我们对于安全性在理论和实际上的限制已有了更好的理解。

因为这一领域发展得非常迅速，所以很难写一本既紧凑又现时的书。甚至当手稿还未写好，在这个领域中又有了新的进展。虽然我尽力体现一些这方面的进展，但仍未如我所愿将它们写入本书。在很多情况下，我只能列入一些参考文献。

有些领域还未成熟，而我也没能将它们处理得令我满意。这些领域之一是操作系统的证实，另一个是将密码控制纳入操作系统和数据库系统。我希望在本书以后的版本中能将这些论题写得好一些。

数据安全性学科从数学和计算机科学中吸取得极多。我假定读者已有一些程序、数据结构、操作系统、数据库系统、计算机结构、概率论和线性代数的背景知识。因为我发现大多数计算机科学的学生在信息论和数论方面的知识不多，所以我写了理解本书所需的这些论题的有关内容。由于复杂性理论是相当新的领域，我也简要地加以介绍。

本书曾用于普都(Purdue)大学研究生一学期的计算机科学课程。对学生指定习题、程序设计和一项学期(课程)设计。本书适用于作为研究生或大学高年级学生的教本和自学。每章末都有一些习题，大多数习题安排得使读者能自己识别出正确的答案。我有意地不列出解答。自然，其中也有一些难题。

下面简要综述各章内容：

第一章，**引论**，介绍密码学、数据安全、信息论、复杂性理论和数论的基本概念。

第二章，**加密算法**，描述经典的和现代的加密算法，其中包括数据加密标准(DES)和公钥算法。

第三章，**密码技术**，研究将密码控制纳入计算机系统的各种有关技术，其中包括密钥管理。

第四章，**接入控制**，研究控制主体(例如用户或程序)接入目标(例如文件或记录)的各种机构的基本原理。这类机构调节针对目标的直接接入，而不涉及包含在这些目标中的信息所产生的问题。

第五章，**信息流控制**，描述调节信息传播的控制。为了防止程序泄露机密数据，或防止将分类的数据传播给只有较低安全许可的用户，这类控制是必须的。

第六章，**推理控制**，描述防止将机密数据作为有关个人小组的统计量来公布的控制。

下面为谢辞(略)

目　录

第一章 引 论

§1.1 密 码 学

密码学是研究密写的一门科学。**密码**是把**明**文(或**平**文)变换成**密**文(有时叫**密报**)的一种密写方法。把明文变成密文的过程叫**加密**;其逆过程,即把密文变换成明文的过程叫**解密**。加密与解密均由一个密钥或一组密钥来控制(见图 1.1)。

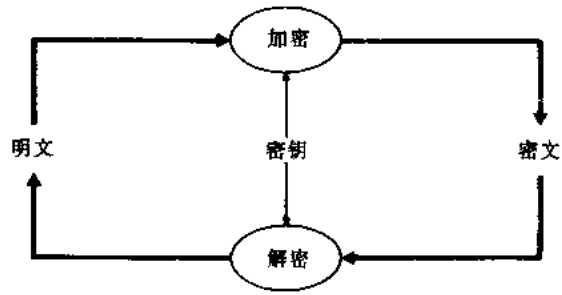

图 1.1 密写

有两种基本类型的密码:**移位**与**代换**。移位密码是把数据中的位或字符重排。例如,对于"栅栏密码",是把明文消息字母仿照栅栏的模式写出,然后再按行移动。如下所示:

DISCONCERTED COMPOSER

↓

D O R C O

I C N E T D O P S R

S C E M E

↓

DORCOICNETDOPSRSCEME

这种密码的密钥由栅栏的深度给出，在此例中为3。

代换密码是用代换的方式代替位、字符或字符组。一种简单类型的**代换密码**是将英文字母向前移动 K 位(移过 Z 之后再循环返至 A)。K 便是这种密码的密钥。这种密码通常叫**凯撒密码**，这是因为朱利叶斯·凯撒曾经就 $K=3$ 的情形使用过这种密码。凯撒方法如下所示：

IMPATIENT WAITER

↓

LPSDWLHQW ZDLWHU

代码是代换密码的一种特殊类型，它使用"代码本"作为密钥。明文字或词组与它的密文代换被一起编入代码本，如下所示：

字	代码
BAKER	1701
FRETTING	5603
GAITARIST	4008
LOAFING	3790
⋮	⋮

LOAFING BAKER

↓

3790 1701

代码的术语有时被用于任何类型的密码。

在计算机应用中，移位常常与代换结合起来使用。例如，数据加密标准(DES)便是使用移位与代换的组合按 64 位分组加密(见第二章)。

密码分析学是研究破译密码的一门科学。所谓一个密码是**可破的**，是指如果通过密文能够决定明文或密钥，或通过明文-密文对能够决定密钥。有三种基本破译方法：单纯密文破译、已知明文破译及选择明文破译。

所谓**单纯密文破译**，密码分析者必须仅由截获的密文来决定密钥，不过在这里加密方法，明文语言，密文所涉及的主题以及某些可能出现的词可以是已知的。例如，描述一件财宝被埋藏的位置的一条消息，(在英语中)大概会含有诸如 BURIED，TREASURE，NORTH，TURN，RIGHT，MILES 等词。

所谓**已知明文破译**，是指密码分析者知道某些明文-密文对。作为一个例子，假设由一个用户终端传给计算机的一条加了密的消息被密码分析者截获，他知道这条消息从一个标准的报头诸如"LOGIN"开始。作为另外一个例子，密码分析者可能知道关于物理学的一份包含密文的特定记录的部门区段；事实上，密码分析者可能知道在数据库中的每一份记录的部门区段。在某些情况下，对于某些可能出现的词的知识会十分接近已知明文破译。加了密的程序特别易受攻击，因为在程序语言中会有规律地出现一些诸如 **begin**，**end**，**var**，**procedure**，**if**，**then** 等关键词。甚至即使密码分析者不知道这些加了密的关键词的精确位置，他同样可以对它们做合理的猜测。在今天，一类密码能够考虑被采纳，仅当它们在假定密码分析者具有任意多数量的明文-密文对之下能够经得起已知明文的攻击。

所谓**选择明文破译**，是指密码分析者能够获得选择的明文所相应的密文。对于密码分析者而言，这是最有利的情形。数据库系统可能最易受这种类型的攻击，如果用户能够将某些要素插入数

据库，然后再观察所贮存的密文的变化。拜尔与梅茨格称这种问题为**诈骗存储问题**。

公开密钥体制(定义见§1.3)中引入了第四种破译：**选择密文破译**。虽然原文是不大明了的，但密码分析者可用它推断密钥。

一种密码叫做**无条件安全**的，如果不论截获多少密文，在密文中仍然没有足够的信息用以唯一地决定明文。在§1.4中我们将给出无条件安全密码的正式定义。除一个例外，所有的密码都是可破的，只要给定无限的资源。因此，对于那种密码，破译它在计算上是不可行的，我们更有兴趣。一种密码叫做**计算上安全**的，或强的，如果以可用的资源不可能破译它。

包括密码学与密码分析学知识的科学分支，叫**保密学**。

§1.2 数据安全

经典密码学提供了在消息可能被窃听和截取的信道上对信息的保密手段。发信者选取一种密码和加密钥，然后直接把它传送给收信者，或者间接地在一个低速但安全的信道上传送它(典型的是通过可靠的信使)。消息与回执均在不安全信道上用密文传输(见图1.2)。经典的加密方案在第二章中介绍。

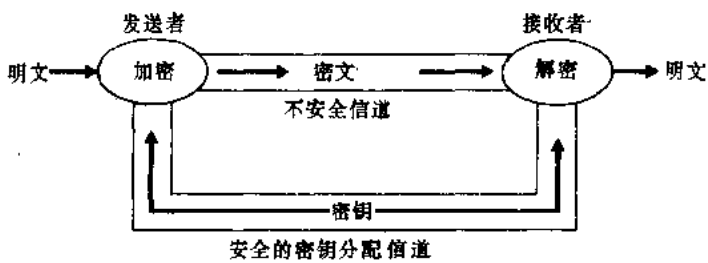

图1.2 经典信息信道

现代密码学可使在高速电子线路上传输的数据或存储在计算机系统中的数据受到保护。这里有两个基本的主题：**保密性**(或**专有性**)，以防止数据被非特许的泄露，以及**确证性**(或**完整性**)，以防止对数据被非特许的修改。

在电子线路上传输信息易受被动窃密(它威胁保密性)和主动窃密(它威胁确证性)的攻击(见图 1.3)。**被动窃密**(窃听)是指对消息的截收，一般无法检测出是否被截收了。虽然它通常用于揭露消息内容，但在计算机网络中，它同样也用于通过该网络来监测信号流用以判断谁与谁在通信。保护消息内容不被暴露的手段由第二章所述的加密变换以及第三章所述的加密技术来提供。防止信号流分析的手段可由控制最终加密结果来提供，这在第三章中加以叙述。

主动窃密(**窜改**)是指蓄意对消息流加以修改。这可以达到任意改变一条消息或通过重放早先消息数据来代替现有消息中数据的目的(例如，用早先的金额记录字段"CREDIT JONES'S ACCOUNT WITH $ 5000"来代替现有的金额交易字段"CREDIT SMITH'S ACCOUNT WITH $ 10")。这还可以达到注入假消息，注入原先消息的重放(例如重复信用记录)或删除消息(例如阻止一笔交易"DEDUCT $ 1000 FROM SMITH'S ACCOUNT")的目的。加密通过使对手无法制造可译成有意义明文的密文来防止消息被窜改以及防止假消息的注入。但是，应当注意，虽然这种技术可查出消息是否被窜改，却无法阻止这种窜改。

单纯的加密技术不能防止重放，因为对手能够简单地重放以前的密文。为防止这种问题发生的密码技术在第三章中叙述。虽然加密技术不能阻止消息删除，在第三章中所述的密码技术却可查出在一条消息流中组或字符的删除。整个消息的删除可通过要求消息应答的通信协议来查出。

在计算机系统中的数据易受类似威胁的攻击(见图 1.4)。对于保密性的威胁包括浏览、泄露与推理。**浏览**是指通过主存储器及

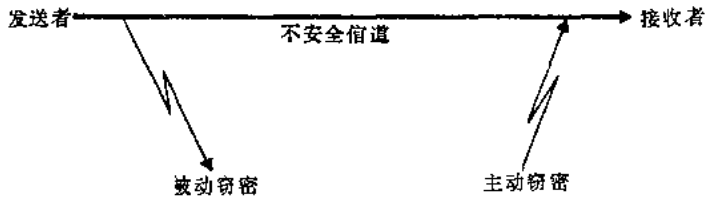

图 1.3 对于安全通信的威胁

辅助存储器来搜索信息(例如机密数据或专用软件程序)。这类似于在通信信道中的窃听,但有以下两点重要区别。第一,存储在计算机系统中的信息具有较长的寿命;在这种意义上说,浏览要比窃听造成更为严重的威胁。第二,在电子线路上传输的信息即使当对系统的访问被拒绝时也易遭窃听(抽头)的攻击。仅当用户能够访问该系统且访问非特许的存储区域时,浏览才是可能的。在第四章所讲的接入控制便可防止这一点。

密码学可通过把信息变得难于理解的办法来防止对手浏览。密码学可以补充接入控制之不足,并且对于保护磁带及磁盘上的数据特别有用,因为它们一旦被窃取就不可能再受到该系统的保护。然而,密码学不能保护数据在按明文形式处理时不被泄露。为此目的,就需要接入控制,并且这些控制必须包含在使用中清除存贮器的程序以保证机密数据不致无意地被暴露。如果对接入不加控制,加密的数据也易遭**密文搜索**的攻击(例如发现职员通过搜索具有同一密文薪金的记录来制造等同的薪金);对于这一问题的密码学解答将在第三章中叙述。

泄露是指在合法的接入数据的过程中数据被传输给非特许的用户。例如,在编制专用软件程序的过程中,编制者可能会泄露这种程序。一种所得税程序可能会泄露用户的机密信息。一个文件编辑者可能会把机要军事数据在没有安全清除的情况下泄露给用户。密码学与接入控制之不足可由第五章所讨论的信息流控制加