计算机安全保密技术

（国家自然科学基金研究课题）

张焕国　覃中平
丁玉龙　崔宝秋　编著

机械工业出版社

(京)新登字 054 号

本书较系统地介绍计算机安全保密的基本理论和实用技术，既简明扼要地介绍国内外的前沿研究成果，又详细介绍了广大计算机用户迫切需要的安全保密实用技术。全书共分十一章，分别介绍密码技术、数字签名、认证技术、计算机网络安全保密、操作系统安全保密、数据库安全保密、软件加密、硬件加密和计算机病毒防治技术。本书是作者在长期从事计算机安全保密技术科研和教学的基础上总结的成果。本书内容丰富实用，深入浅出，可供广大计算机用户，科研人员和管理人员作为技术参考书，又可供大专院校信息类专业用作教材或教学参考书。

图书在版编目(CIP)数据

计算机安全保密技术/张焕国等编著. —北京：机械工业出版社，1994.12

ISBN 7-111-04414-2

Ⅰ.计… Ⅱ.张… Ⅲ.①计算机管理-安全技术②电子计算机-密码术 Ⅳ.TP309

中国版本图书馆 CIP 数据核字(94)第 09965 号

出 版 人：马九荣(北京市百万庄南街 1 号 邮政编码 100037)

责任编辑：王中玉 汪小星 版式设计：王颖 责任校对：樊中英

封面设计：肖晴 责任印制：王国光

北京市密云县印刷厂印刷·新华书店北京发行所发行

1995 年 2 月第 1 版·1995 年 2 月第 1 次印刷

787mm×1092mm 1/16·15 印张·367 千字

0 001—3500 册

定价：18.80 元

前　言

在现代社会中，计算机在社会事务处理、军事决策和生产过程控制等方面的作用日益增大，社会对计算机的依赖也日益增大，如果计算机系统中的设备和数据遭到破坏，将会造成社会混乱和重大损失。因此，计算机安全保密问题已成为世人关注的社会问题，并成为计算机科学技术的一个重要课题。

目前，总的来说，信息的安全保密技术的研究落后于整个信息产业的发展，而且国内系统介绍信息安全保密技术的书籍不多。为了增进信息安全保密技术的交流，我们编著了这本抛砖引玉之作。

本书是作者在武汉大学计算机科学系多年教学和研究的基础上写成的。其研究工作得到国家自然科学基金的资助。作者试图从理论和实践相结合的角度较系统地介绍计算机安全保密的基本理论和实用技术。既介绍国内外的前沿研究成果，又介绍目前已经实用的具体技术；既注意介绍国外学者的研究成果，同时又用了较大篇幅介绍我国学者的研究成果。尽管作者有以上初衷，但因学术水平和篇幅所限，仍有许多信息安全保密技术和理论未能介绍，如零知识证明理论、网络协议安全性验证、防电磁辐射技术等。而且所述内容也会有不妥和错误之处。对此，作者恳请读者的理解和批评指正。

本书共分十一章。第二、三两章介绍密码技术，其中第二章主要介绍传统密码及其应用，第三章介绍公开密钥密码，主要介绍 RSA 和 FAPKC 公开钥密码体制。第四章介绍基于密码技术的数字签名技术。第五章介绍认证理论与技术。第六至八章介绍计算机系统的安全保密技术，其中第六章介绍计算机网络安全保密，第七章介绍计算机操作系统的安全保密，第八章介绍计算机数据库的安全保密。第九章介绍微机软件加密技术。第十章介绍硬件加密技术。第十一章介绍计算机病毒防治技术。每章后面列出一定数量的参考文献。

本书由张焕国、覃中平、丁玉龙、崔宝秋共同编写，由张焕国对全书进行统编。

作者衷心感谢陶仁骥教授和肖国镇教授长期以来给予作者的指导和关心。

目　录

第一章 概 论

信息、能源、材料是现代社会的三大支柱。电子计算机和数据通信网络的广泛应用是现代社会信息化的标志。在信息化社会中，人们用通信网络交换信息，用电子计算机管理和处理信息。这种信息交换和处理的自动化给人们的生活和工作带来极大的方便。但同时也带来许多亟待解决的问题，其中数据的安全保密便是一个突出的问题。这是因为，在信息化社会中国家的政治、军事、外交、经济及企业的财务、人事等重要数据都存储在计算机系统中，从而使计算机系统成为不法分子攻击的主要目标。而且计算机系统中的资源往往为多用户共享，不法分子可利用共享资源窃取和篡改他人的数据。其次是因为从计算机系统中非法窃取数据不会留下痕迹，而且并非能立即知道数据的安全是否已受到危害。其三是因为计算机和通信设备会发生故障，操作人员也会发生误操作，如不采取措施，则会造成数据泄密。

目前，计算机系统的安全保密问题已成为世人注目的社会问题。

世界主要工业化国家中每年因利用计算机犯罪所造成的直接经济损失令人吃惊，仅美国每年因利用计算机犯罪所造成的直接经济损失就高达几百亿美元，远远超过普通盗窃金额的十倍以上。我国因计算机应用不普及，故利用计算机犯罪的总数较小，但据悉仅最近两年内我国已发生利用计算机盗窃金钱的案件已达上千起，其增势之迅猛令人震惊！

除金融信息外，政治、军事等重要数据也是不法分子攻击的重点。德国几名青年曾打入五角大楼和北约的计算机数据库。美国通用动力公司的一名软件设计员设计的逻辑炸弹破坏了太空导弹工程数据库，致使电脑数据库的数据无法恢复，造成了无法弥补的损失。英国、法国和韩国也均有类似事件发生。

过去被认为是科学幻想的计算机病毒，现已活生生地出现在人们的面前，对计算机安全构成极大的威胁。1988 年 11 月 3 日美国康乃尔大学一年级研究生罗特·莫里斯编制的称为蠕虫的计算机病毒通过美国全国范围的计算机网络 INTERNET 大面积传播，致使 6000 多个基于 UNIX 的工作站和小型机被传染，直接经济损失达 6000 万美元以上。1988 年 8 月日本最大的微机网络 PC-VAN 也曾大面积被病毒传染。至于个人微机被病毒传染的病例就更普遍了。在我国个人微机被病毒传染也很普遍，特别是大中学校公用实验室的微机几乎均被病毒感染过。流行的病毒除从国外传入者外，还有许多国内不法分子自己编制的国产病毒。由于人们对计算机病毒不甚了解，加上新闻机构的渲染，使人们对之有一种恐惧感，大有谈毒色变之势。

面对如此严重危害计算机数据安全的种种威胁，必须采取措施确保计算机数据的安全保密。本书旨在介绍确保计算机数据安全保密的一些基本技术。

确保数据安全保密就是要保护数据免受未授权的泄露、篡改和毁坏，主要包含数据秘密性(secrecy)和数据真实性(authenticity)两个侧面，此外还涉及数据的完整性(integrity)。

要确保计算机数据的安全保密，必须综合采取各种措施才能奏效。除法律、行政、教育等措施外，最重要的是要采用技术保护措施。

确保计算机数据安全保密的技术措施可分为访问控制技术和密码技术两大类。

访问控制要对访问的申请、批准、执行、撤消全过程进行控制，以确保只有合法用户的合法

访问才给以批准，而且被批准的访问只能执行授权的操作。

访问控制的第一道设防是用户身份的认证，以识别区分合法用户和非法用户，从而阻止非法用户访问系统。目前常用的办法是口令认证，然而口令的不安全性已被普遍认识。理想的用户身份认证应是基于用户生理特征的身份认证，诸如指纹、眼网膜纹等。然而由于生理特征识别技术比较复杂、成本高，故应用尚不能普及。安全的认证过程应当是用户和系统平等地互相认证的过程。智能卡是一种理想的用户身份持证。它装有微电脑，具有数据处理能力，因而可以采用复杂的认证协议、实现人机的对等相互识别，在国外已得到广泛应用。

用户被批准访问系统后，仍然要对访问的执行施加控制。这主要包括：授权，决定哪个主体有资格访问哪个客体；确定访问权限，决定本次访问是否具有读、写、执行、删除、附加及转移等权力；实施访问权限。在这里授权策略和控制机制是重要的，授权策略确保授权的安全性，而控制机制具体实施授权的安全策略。除了对直接的访问进行控制外，还应对信息的流动和推理攻击施加控制。如果一次访问将产生信息和权力的流动，则应注意这种流动是否可能造成泄密。推理攻击是指用户通过多次合法访问的结果，进行推理计算得出他无权访问的秘密信息。推理攻击是危害统计数据库安全的主要威胁。

审计跟踪是访问控制的一个重要方面。它对用户使用何种系统资源、使用时间、执行的操作等问题进行完整的记录，以备非法事件发生后能够有效地追查。它的存在对不法分子构成心理威慑，是对系统安全实施有效监控的一种重要手段。

密码技术是一门古老的技术，大概自人类社会出现战争便产生了密码。由于密码长期仅用于军事、公安、外交等要害部门，其研究本身也只限于秘密进行，所以密码技术被蒙上神秘的面纱。然而今天它已广泛用于计算机和通信系统中。

密码技术的基本思想是伪装信息，伪装就是对数据施行一组可逆的数学变换。伪装前的数据称为明文，伪装后成为密文。伪装的过程称为加密，去掉伪装恢复明文的过程称为解密，加解密要在密钥的控制下进行。将数据以密文形式存储在计算机文件中或在网络中传输，且只给合法用户分配密钥。这样，既使密文被非法窃取，因其没有密钥而不能得到明文，从而达到保密的目的。同样也不能伪造合理的密文，因而篡改数据必然被发现，从而达到确保数据真实性的目的。

然而对于传统密码，通信双方必须约定使用相同密钥，而密钥的分配只能通过其它保密途径（如派信使等）。在计算机网络中，设共有 n 个用户，任意两个用户均要进行保密通信，故需要 $n(n-1)/2$ 种不同的密钥，当 n 较大时这一数值是很大的。另一方面，为了安全性要求通信密钥经常更换。如此大量的密钥要经常地产生、分配与更换，其困难性和危险性是可想而知的。而且有时甚至不可能事先约定密钥。如商业、企业间想通过网络洽谈生意而又要保守商业秘密，在许多情况下不可能事先预约密钥。因此，传统密码由于其在密钥管理上的困难性而不适于计算机网络应用。

1976 年美国学者 Diffie 和 Hellman 提出公开密钥密码的概念。公开密钥密码从根本上克服了传统密码在密钥分配上的困难，特别适合计算机网络应用。然而公开密钥密码算法的研究并非易事，而且尚无有效的数学方法证明其安全性。国外众多的公开密钥密码算法中，研究得比较充分的首推 RSA 算法。RSA 密码的安全性建立在大合数因子分解的困难性之上。我国学者陶仁骥、陈世华提出的一种基于有限状态自动机可逆性理论的公开密钥密码算法（简称为 FAPKC）具有一些显著的优点，得到国际的关注。由于其运算只涉及逻辑运算，因而软硬件实

现都较容易，而且加解密速度快，密钥量适中，密钥产生容易。FAPKC 的安全性建立在构造非线性弱可逆有限自动机的弱逆及矩阵多项式因式分解的困难性之上。

目前，公开密钥密码已从理论研究进入实际应用。美国、日本、德国、英国、比利时等国都研究成功 RSA 硬件设备。基于 RSA 的商用计算机网络已经应用。美国于 1991 年 8 月提出一种基于公开密钥密码的数字签名建议。我国一些单位也进行了 RSA 的软件实现。作者及其课题组用软硬件相结合的方法实现了 FAPKC，并在微机 NOVEL 网环境上开发成功一个完整的微机网络信息保护系统。

第二章　传统密码

本章讨论密码学的基本概念、古典的和近代的传统密码及其在计算机系统中的应用。

第一节　密码学的基本概念

密码学(cryptology)是一门古老的科学。大概自人类社会出现战争便产生了密码，以后逐渐形成一门独立的学科。在密码学形成和发展的历程中，科学技术的发展和战争的刺激都起了积极的推动作用。电子计算机一出现便被用于密码破译。电子计算机对密码学的发展产生了巨大的影响和推动。1949 年商农发表了题为《保密系统的通信理论》的著名论文，把密码学置于坚实的数学基础之上。1977 年美国联邦政府正式颁布数据加密标准(DES)，这是密码史上的一个创举。1976 年 W・Diffie 和 M・Hellman 提出公开密钥密码，这是密码学发展的又一个里程碑。大规模集成电路的迅猛发展为密码学的进一步发展提供了强有力的物质基础，计算机的广泛应用为密码学的进一步发展提出了新的客观需要。特别是传统密码不适应计算机网络的应用，在这种情况下产生了公开密钥密码。公开密钥密码从根本上克服了传统密码在密钥分配管理方面存在的弱点，因而特别适合计算机网络和分布式计算机系统的应用。此外，除了计算机通信的数据传输保密之外，计算机的操作系统和数据库的安全保密问题也很突出，由此产生了计算机密码学。

密码学的研究方式由过去的单纯秘密进行转向公开和秘密两条战线同时进行。自古以来，密码主要用于军事、政治、外交等要害部门，因而密码学的研究本身也是秘密地进行的。密码学的知识和经验主要掌握在军事、政治、外交等保密机关，不便公开发表，这是过去密码学的书籍一向很少的原因。然而由于微电子学、计算机科学的发展使得计算机和通信网络的应用进入了人们的日常生活和工作领域。出现了电子转帐、电子邮政、办公室自动化等必须确保数据安全的系统，使得民间和商业界对数据安全保密的需要大大增加。于是，在民间产生了一批不从属于保密机关的密码学者，他们可以毫无顾忌地发表文章，互相竞争，公开地进行密码学研究。事实证明，正是这种公开的研究和秘密的研究相结合的局面促成了今天密码学的空前繁荣。

研究编制密码的科学称为密码编制学(cryptography)，研究破译密码的科学称为密码分析学(cryptanalysis)，而密码编制学和密码分析学共同组成密码学。

密码技术的基本思想是伪装信息，使未授权者不能理解它的真实含义。所谓伪装就是对信息进行一组可逆的数学变换。伪装前的原始信息称为明文(plaintext)，伪装后的信息称为密文(ciphertext)，伪装的过程称为加密(encryption)。加密在加密密钥(key)的控制下进行。用于对数据加密的一组数学变换称为加密算法。发信者将明文数据加密成密文，然后将密文数据送入数据通信网络或存入计算机文件。授权的收信者接收到密文后，施行与加密变换相逆的变换，去掉密文的伪装恢复出明文，这一过程称为解密(decryption)。解密在解密密钥的控制下进行。用于解密的一组数学变换称为解密算法。因为数据以密文的形式存储在计算机文件中，或在数据通信网络中传输，因此即使数据被未授权者非法窃取或因系统故障和操作人员误操作而造

成数据泄露，未授权者也不能理解它的真正含义，从而达到数据保密的目的。同样，未授权者也不能伪造合理的密文，因而不能篡改数据，从而达到确保数据真实性的目的。

一个密码系统，通常简称为密码体制(crytosystem)，由五个部分组成：

1)明文空间 $\boldsymbol{M}$，它是全体明文的集合。

2)密文空间 $\boldsymbol{C}$，它是全体密文的集合。

3)密钥空间 $\boldsymbol{K}$，它是全体密钥的集合。其中每一个密钥 K 均由加密密钥 K_e 和解密密钥 K_d 组成，即 $K=(K_e,K_d)$。

4)加密算法 $\boldsymbol{E}$，它是一族由 $\boldsymbol{M}$ 到 $\boldsymbol{C}$ 的加密变换。

5)解密算法 $\boldsymbol{D}$，它是一族由 $\boldsymbol{C}$ 到 $\boldsymbol{M}$ 的解密变换。

对于每一确定的密钥 $K=(K_e,K_d)$，加密算法将确定一个具体的加密变换，解密算法将确定一个具体的解密变换，而且解密变换是加密变换的逆过程。对于明文空间 $\boldsymbol{M}$ 中的每一个明文 M，加密算法在加密密钥 K_e 的控制下将 M 加密成密文 C

$$C=E(M,K_e) \tag{2-1}$$

而解密算法在解密密钥 K_d 的控制下从密文 C 中解出同一个明文 M

$$M=D(C,K_d)=D(E(M,K_e),K_d) \tag{2-2}$$

如果一个密码体制的 $K_e=K_d$，或由其中一个很容易推出另一个，则称为单密钥密码体制或对称密码体制或传统密码体制。否则，称为双密钥密码体制或非对称密码体制。进而，如果在计算上 K_d 不能由 K_e 推出，这样将 K_e 公开也不会损害 K_d 的安全，于是便可以将 K_e 公开。这种密码体制称为公开密钥密码体制。

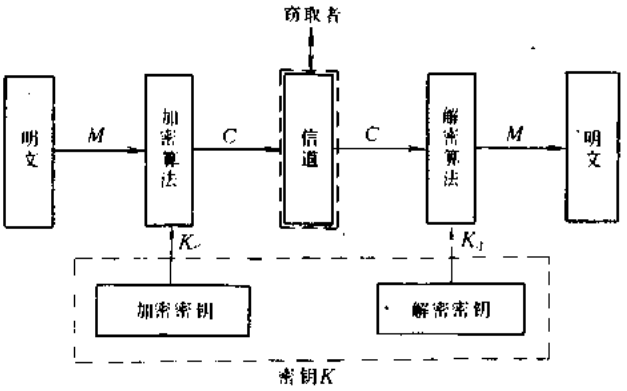

图 2-1 密码体制

根据对明文的划分与密钥的使用方法不同可将密码体制分为分组密码和序列密码体制。

设 M 为明文，分组密码将 M 划分为一系列明文块 $M_1,M_2,\cdots,M_n$，通常每块包含若干字符，并且对每一块 M_i 都用同一个密钥 K_e 进行加密。即

$$C=(C_1,C_2,\cdots,C_n)$$

其中

$$C_i=E(M_i,K_e) \qquad i=1,2,\cdots,n \tag{2-3}$$

而序列密码将 M 划分为一系列的字符或位(bit)$m_1,m_2,\cdots,m_n$，并且对于这每一个 m_i 用密钥序列 $K_e=(k_{e1},k_{e2},\cdots,k_{en})$ 的第 i 个分量 k_{ei} 来加密，即

$$C=(c_1,c_2,\cdots,c_n)$$

其中

$$c_i=E(m_i,k_{ei}) \qquad i=1,2,\cdots,n \tag{2-4}$$

分组密码一次加密一个明文块，而序列密码一次加密一个字符或一个位(bit)。两种密码在计算机系统中都有广泛应用。

如果能够根据密文确定出明文或密钥，或者能够根据明文-密文对确定出密钥，则我们说这个密码是可破译的。否则，我们说这个密码是不可破译的。

密码分析者攻击密码的方法主要有以下三种。

1)穷举攻击　所谓穷举攻击就是指密码分析者用试遍所有密钥的方法来破译密码。穷举攻击所花费的时间等于尝试次数乘以一次解密(加密)所需的时间。显然可以通过增大密钥量或加大解密(加密)算法的复杂性来对抗穷举攻击。当密钥量增大时，尝试的次数必然增大。当解密(加密)算法的复杂性增大时，完成一次解密(加密)所需的时间增大。从而使穷举攻击在实际上不能实现。

2)统计分析攻击　所谓统计分析攻击是指密码分析者通过分析密文和明文的统计规律来破译密码。统计分析攻击在历史上为破译密码作出过极大的贡献。许多古典密码都可以通过分析密文字母和字母组的频率而破译。对抗统计分析攻击的方法是设法使明文的统计特性不带入密文。这样，密文不带有明文的痕迹，从而使统计分析攻击成为不可能。

3)数学分析攻击　所谓数学分析攻击是指密码分析者针对加密算法的数学依据通过数学求解的方法来破译密码。为了对抗这种数学分析攻击，应选用具有坚实数学基础和足够复杂的加密算法。

此外，根据密码分析者可利用的数据来分类，可将破译密码的类型分为以下三种：。

1)仅知密文攻击(ciphertext-only attack)所谓仅知密文攻击是指密码分析者仅根据截获的密文来破译密码。

2)已知明文攻击(known-plaintext attack)所谓已知明文攻击是指密码分析者根据已经知道的某些明文-密文对来破译密码。例如，密码分析者可能知道从用户终端送到计算机的密文数据以一个标准词“LOGIN”开头。又例如，加密成密文的计算机程序特别容易受到这种攻击。这是因为诸如“BEGIN”、“END”、“IF”、“THEN”，“ELSE”等词有规律地在密文中出现，密码分析者可以合理地猜测它们。近代密码学认为，一个密码仅当它能够经得起已知明文攻击时才是可取的。

3)选择明文攻击(chose-plaintext attack)　所谓选择明文攻击是指密码分析者能够选择明文并获得相应的密文。这是对密码分析者最有利的情况。计算机文件系统和数据库特别容易受到这种攻击，因为用户可随意选择明文，并得到相应的密文文件和密文数据库。例如，IBM-PC 微机 BASICA 程序加密的情况便是如此。

一个密码，如果无论密码分析者截获了多少密文和用什么方法进行攻击都不能被攻破，则称为是绝对不可破译的。绝对不可破译的密码在理论上是存在的。但是，如果能够利用足够的资源，那么任何实际的密码都是可破译的。因此，对我们更有实际意义的是在计算上不可破译的(computationlly unbreakable)密码。如果一个密码不能被密码分析者根据可利用的资源所破译，则称为在计算上是不可破译的。

第二节　古典密码

虽然用近代密码学的观点来看，许多古典密码是很不安全的，或者说是极易破译的。但是我们不应忘记古典密码在历史上所发挥的巨大作用。本节介绍几种著名的古典密码。

一、置换密码

把明文中的字母重新排列，字母本身不变，但其位置改变了，这样编成的密码称为置换密码。

最简单的置换密码是把明文中的字母顺序倒过来，然后截成固定长度的字母组作为密文。

例如：

明文：明晨5点发动反攻

MING CHEN WU DIAN FA DONG FAN GONG

密文：**GNOGN AFGNO DAFNA IDUWN EHCGN IM**

另一种置换密码是把明文按某一顺序排成一个矩阵，然后按另一顺序选出矩阵中的字母以形成密文，最后把密文截成固定长度的字母组。

例如：

明文：**MING CHEN WU DIAN FA DONG FAN GONG**

矩阵：**MINGCH**　　选出顺序：按列

ENWUDI

ANFADO

NGFANG

ONG

密文：**MEANO INNGN NWFFG GUAA CDDN HIOG**

由此可以看出，改变矩阵的大小和选出顺序可以得到不同形式的密码。其中有一种巧妙的方法：首先选用一个词语作为密钥，然后按字母的字典顺序给密钥中各字母一个编号（重复字按先左后右顺序编号或去掉重复字母）。于是得到一组与密钥词语对应的数字序列。最后按此数字序列中的数字顺序选出密文。

例如：

明文：**MING CHEN WU DIAN FA DONG FAN GONG**

密钥：玉兰花

YU LAN HUA

数字序列：653142

矩阵：**MINGCH**

ENWUDI

ANFADO

NGFANG

ONG

密文：**GUAA HIOG NWFFG CDDN INNGN MEANO**

这种置换密码的密钥是矩阵的大小及选出的顺序，而所选用的密钥词语仅仅是使密钥便于记忆罢了。

置换密码比较简单，它经不起已知明文攻击。但是，把它与其它密码相结合，可以得到十分有效的密码。

二、代替密码

首先构造一个或多个密文字母表，然后用密文字母表中的字母或字母组来代替明文字母

表中的字母或字母组。各字母或字母组的相对位置不变，但其本身改变了。如此编成的密码称为代替密码。

按代替过程所使用的密文字母表的个数可将代替密码分为单表代替密码、多表代替密码和多名代替密码。

1. 单表代替密码　单表代替密码又称简单代替密码，它使用一个密文字母表，并且用密文字母表中的一个字母来代替明文字母表中的一个字母。

设 $\boldsymbol{A}=\{a_0,a_1,\cdots,a_{n-1}\}$ 为含 n 个字母的明文字母表，$\boldsymbol{B}=\{b_0,b_1,\cdots,b_{n-1}\}$ 是含 n 个字母的密文字母表。定义一个由 $\boldsymbol{A}$ 到 $\boldsymbol{B}$ 的一一映射

$f:\boldsymbol{A}\to\boldsymbol{B}$

$f(a_i)=b_i$

设明文 $M=(m_0,m_1,\cdots,m_{l-1})$，则相应的密文 $C=(f(m_0),f(m_1),\cdots,f(m_{l-1}))$。可见，简单代替密码的密钥就是映射 f 或密文字母表 $\boldsymbol{B}$。

下面介绍几种典型的简单代替密码。

1)加法密码　加法密码的映射函数为

$$f(a_i)=a_j$$
$$j\equiv i+k \mod n \tag{2-5}$$

其中，$a_i\in\boldsymbol{A}$，k 是满足 $0<k<n$ 的正整数。

最著名的加法密码是古罗马 Caesar 大帝使用过的一种密码。Caesar 密码取 $k=3$，因此其密文字母表就是把明文字母表循环移位 3 位后得到的字母表。例如：

A：A　B　C　D … X　Y　Z

B：D　E　F　G … A　B　C

明文：**MING　CHEN　WU　DIAN　FA　DONG　FAN　GONG**

密文：**PLQJ　FKHQ　ZX　GLDQ　ID　GRQJ　IDQ　JRQJ**

2)乘法密码　乘法密码的映射函数为

$$f(a_i)=a_j$$
$$j\equiv ik \mod n \tag{2-6}$$

其中，要求 k 与 n 互素。这是因为仅当 $(k,n)=1$ 时，才存在两个整数 x,y，使得 $xk+yn=1$，才有 $xk\equiv 1 \mod n$，才有 $i\equiv xj \mod n$，密码才能正确解密。

例如，当用英文字母表作明文字母表而取 $k=13$ 时，便会出现

$$f(\mathrm{A})=f(\mathrm{C})=f(\mathrm{E})=\cdots=f(\mathrm{Y})=\mathrm{A}$$
$$f(\mathrm{B})=f(\mathrm{D})=f(\mathrm{F})=\cdots=f(\mathrm{Z})=\mathrm{N}$$

又若选 $k=5$，便得到如下的密文字母表：

A：A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

B：A F K P U Z E J O T Y D I N S X C H M R W B G L Q V

3)仿射密码　乘法密码与加法密码相结合便构成仿射密码。仿射密码的映射函数为

$$f(a_i)=a_j$$
$$j\equiv ik_1+k_0 \mod n \tag{2-7}$$

其中，要求 $(k_1,n)=1$，$0<k_0<n$。

仿此可构造更复杂的多项式密码：

$$f(a_i)=a_j$$
$$j\equiv i^t k_t+i^{t-1}k_{t-1}+\cdots+ik_1+k_0 \mod n \tag{2-8}$$

其中，要求$(k_t,n)=1$，$0<k_0<n$。

4)密钥词语代替密码　用一词组或短语作密钥(在词语密钥控制下编密码的方法在置换密码中曾得到应用，这一方法同样可以用到代替密码中)。首先随机地选择一个词组或短语作密钥，去掉密钥中的重复字母，把结果作为矩阵的第一行。其次在明文字母表中去掉矩阵第一行中的各字母，并将剩余字母依次写入矩阵的各行。最后按某一顺序从矩阵中取出字母构成密文字母表。例如：

密钥：**HONG　YE**

矩阵：**H O N G Y E**　　选出顺序：按列

A B C D F I

J K L M P Q

R S T U V W

X Z

A：A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

B：H A J R X O B K S Z N C L T G D M U Y F P V E I Q W

明文：**MING CHEN WU DIAN FA DONG FAN GONG**

密文：**LSTBJ KXTEP RSHTO HRGTB OHTBG TB**

2. 多表代替密码　简单代替密码很容易被破译，其原因在于只用一个密文字母表加密，从而使得明文中的一个字母只有用唯一的一个密文字母来代替。提高代替密码强度的一种方法是采用多个密文字母表，使明文中的每一个字母都有多种可能的代替。

构造d个密文字母表

$$\boldsymbol{B}_j=(b_{j0},b_{j1},\cdots,b_{jm-1}) \qquad j=0,1,\cdots,d-1$$

字义d个映射

$$f_j:\boldsymbol{A}\to\boldsymbol{B}_j$$
$$f_j(a_i)=b_{ji} \tag{2-9}$$

设明文$M=(m_0,m_1,\cdots,m_{d-1},m_d,\cdots)$，则相应的密文$C=(f_0(m_0),f_1(m_1),\cdots,f_{d-1}(m_{d-1}),f_0(m_d),\cdots)$。由于加密过程中用到多个密文字母表，故称为多表代替密码。多表代替密码的密钥就是这d个映射或密文字母表。

最有名的多表代替密码要算16世纪法国密码学者 Vigenere 使用过的 Vigenere 密码。

Vigenere 密码使用26个密文字母表，像加法密码一样，它们是依次把明文字母表循环移位0,1,…,25位的结果。选用一个词组或短语作密钥，以密钥字母控制使用哪一个密文字母表。

把26个密文字母表排列在一起称为 Vigenere 方阵。表2-1给出了 Vigenere 方阵。

表2-1 Vigenere 方阵

		明文字母																							
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
	C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
	D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
	E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
	F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
	G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
	H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
	I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
	J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
	K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
密	L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
钥	M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
字	N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
母	O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
	W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
	Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Vigenere 密码的代替规则是，用明文字母在 Vigenere 方阵中的列与相应的密钥字母在 Vigenere 方阵中的行的交点处的字母来代替该明文字母。例如，若明文字母为 P，密钥字母为 Y，则用字母 N 来代替明文字母 P。

利用 Vigenere 密码加密的过程如下：

1)把明文写在上面一行；

2)把密钥写在下面一行。通常选用一个词语或某书中的一段话作为密钥；

3)对明文的每一个字母，根据密钥和 Vigenere 方阵进行代替。

例如，

明文：**MING CHEN WU DIAN FA DONG FAN GONG**

密钥：**XING CHUI PING YE KUO YUE YONG DA JIANG LIU**

密文：**JQAME OYVLC QOYRP URMHK DOAMR NP**

Vigenere 密码的解密过程如下：

1)把密钥写在上面一行；

2)把密文写在下面一行；

3)首先在密钥字母所对应的行中找到密文字母。该密文字母所在的列向上所对应的明文字母便是该密文字母的明文字母。

3. 多名代替密码　为了抵抗频率分析，希望密文中不残留明文字母频率的痕迹。一种明显的方法是设法将密文字母的频率分布拉平。这便是多名代替密码的出发点。

设明文字母表 $A=\{a_0,a_1,\cdots,a_{n-1}\}$，对于每一个明文字母 a_i，作一个与之对应的字符子集合 B_i，且使 B_i 中的字符个数正比于 a_i 在明文中的相对频率，称 B_i 为 a_i 的多名字符集。以集合 $\boldsymbol{B}=\{B_i|i=0,1,\cdots,n-1\}$ 作为密文字母表。定义映射函数

$$f:\boldsymbol{A}\rightarrow\boldsymbol{B}$$

$$f(a_i)=b_{ij}\qquad \text{而 } b_{ij}\in B_i \tag{2-10}$$

即，函数 f 将明文字母 a_i 映射到它的一个多名字符 b_{ij}。

设明文 $M=(m_0,m_1,\cdots,m_{n-1})$，则相应的密文 $c=(f(m_0),f(m_1),\cdots,f(m_{n-1}))=(c_0,c_1,\cdots,c_{n-1})$，其中 c_i 是根据映射函数从其多名字符集中随机地选取的一个多名字符。

例如，我们可以用0到99这100个整数构成26个子集合，分别作为26个英语字母的多名字符集合，而且保证每个多名字符集合的整数的个数正比于相应英语字母的相对频率以及不同的多名字符集合之间没有相同的整数。表2-2给出了一种多名代替密文字母表。

表2-2　多名密文字母表

A	B												
A	3	16	29	94	31	47	68	52					
B	87	71											
C	80	26	7										
D	11	40	62	93									
E	2	15	28	37	54	41	60	73	89	90	21	57	76
F	9	70											
G	34	82											
H	99	43	51	24	0	17							
I	4	19	27	81	33	46	66						
J													
K	39												
L	1	45	14	96									
M	10	88											
N	13	5	20	36	69	50	49						
O	6	18	95	74	59	48	23	30					
P	91	53											
Q													
R	92	79	42	58	12	38							
S	35	44	61	56	85	77							
T	8	25	63	55	72	64	86	97	98				
U	32	65	83										
V	78												
W	67	75											
X													
Y	22	84											
Z													

例如：

明文：**DATA SECURITY**

密文：40 94 8 16 61 37 7 32 92 19 98 67

由于多名字符集中字符的个数正比于相应明文字母的相对频率，而且每个具体多名字符

的选取又完全是随机的，所以多名代替密码的字符频率分布将是平坦的，这大大地增强了密码的强度。

三、代数密码

美国电话电报公司的 Gilbert Vernam 在1917年为电报通信设计了一种非常方便的密码，后来被称为 Vernam 密码。Vernam 密码在近代计算机和通信系统中得到广泛应用。

Vernam 密码的明文，密钥和密文均用二元数字序列表示。设明文 $M=(m_0,m_1,\cdots,m_{n-1})$，密钥 $k=(k_0,k_1,\cdots,k_{n-1})$，密文 $c=(c_0,c_1,\cdots,c_{n-1})$，则

$$c_i=m_i\oplus k_i,\qquad i=0,1,\cdots,n-1 \tag{2-11}$$

这说明要编制 Vernam 密码，只需要先把明文和密钥表示成二元序列，再把它们按位模2相加便可。根据式(2-11)，有

$$m_i=c_i\oplus k_i \tag{2-12}$$

这说明要解密 Vernam 密码；只需要把密文和密钥的二元序列按位模2相加便可。可见，Vernam 密码的加密和解密非常简单，而且特别适合计算机和通信系统的应用。

例如：

明文：**DATA**

1000100 1000001 1010100 1000001

密钥：**LAMB**

1001100 1000001 1001101 1000010

密文：0001000 0000000 0011001 0000011

Vernam 密码属于序列密码。它的一个突出优点是加密变换和解密变换相同，都为模2相加。这使得加密和解密的软硬件实现极为简单，加密和解密可共用同一软件模块或硬件电路，工作量减少一半。在数学上，如果一个变换的正变换与逆变换相同，$f=f^{-1}$，则称其为对和变换。模2相加运算⊕即为一种对合运算。因此，在密码设计中都希望将其加密运算设计成对和运算。著名的 DES 和 FEAL 密码，其加密运算均为对和运算。

对于 Vernam 密码，如果同一密钥重复使用或密钥本身包含重复，则在已知明文攻击面前就非常脆弱了。这是因为，

$$k_i=c_i\oplus m_i$$

只要知道了某些明文-密文对，便可迅速确定出密钥。据此，为了增强 Vernam 密码的强度，应避免密钥重复。一种极端情况是，

1)密钥是真正的随机序列；

2)密钥和明文一样长；

3)一个密钥只用一次。

如果能够作到这些，则密码就绝对不可破译了。这便是著名的一次一密密码(one time pad)。然而在实际上一次一密是行不通的。首先，一次一密要求密钥是真正的随机序列，这在实际上是不可能完全作到的。其次，一次一密要求密钥与明文等长而且一个密钥只使用一次。这意味着必须保存大量的、很长的密钥，并且能够通过安全的途径将每次所使用的密钥告诉收信者。这在实际上也是困难的。可见，一次一密在实际的密钥管理与分配方面是极脆弱的。

虽然一次一密在实际上是行不通的，但它在理论上的成功却给我们展示出一个令人向往的目标。密码学者认为，如果能够用某种实际方式来模仿一次一密，那么将会得到一种保密性

极好的实用密码。无疑这是设计密码体制的一种途径。

第三节　古典密码的统计分析

任何自然语言都有许多固有的统计特性。如果明文语言的这种统计特性在密文中有所反应，则密码分析者便可通过分析明文和密文的统计规律而密码破译。许多著名的古典密码均可用统计分析的方法破译。

一、语言的统计特性

随便阅读一篇英文文献，立刻就会发现，其中字母 E 出现的次数比其它字母都多。如果进行认真统计，并且所统计的文献的篇幅足够长，便可发现各个字母出现的相对频率十分稳定。而且，只要文献不特别专门化，对不同的文献进行统计所得的频率大体相同。表2-3给出了英文字母的频率，同时显示出英文字母频率的分布模式。

表2-3　英文字母频率分布

字母	频率	
A	8.167	* * * * * * * * * * * * * * * * *
B	1.492	* * *
C	2.782	* * * * * *
D	4.253	* * * * * * * *
E	12.702	* *
F	2.228	* * * *
G	2.015	* * * *
H	6.094	* * * * * * * * * * * *
I	6.966	* * * * * * * * * * * * * *
J	0.153	
K	0.772	* *
L	4.025	* * * * * * * *
M	2.406	* * * * *
N	6.749	* * * * * * * * * * * * * *
O	7.507	* * * * * * * * * * * * * * *
P	1.929	* * * *
Q	0.095	
R	5.987	* * * * * * * * * * * *
S	6.327	* * * * * * * * * * * * * *
T	9.056	* * * * * * * * * * * * * * * * * * * *
U	2.758	* * * * *
V	0.978	* *
W	2.360	* * * * *
X	0.150	
Y	1.974	* * * *
Z	0.074	

注：* 代表频率0.5

进一步，根据各字母频率的大小可将英文字母分为几组。表2-4示出这一分组情况。