密码学导引

冯登国　裴定一　编著

科学出版社

密码学导引

冯登国　裴定一 编著

中国科学院应用研究与发展重大项目资助(项目编号:KY951-A1-102)
国家自然科学青年基金项目资助(项目编号:69703012)

科学出版社
1999

内 容 简 介

全书包括十一章和一个附录，系统地介绍了现代密码学的基本理论和技术，主要内容包括密码学的基本概念、信息理论基础、复杂性理论基础，流密码，分组密码的设计原则、工作模式和一些有代表性的分组密码算法及攻击分组密码的一些典型方法，公钥密码算法，各种数字签名方案和各种签名的应用环境，Hash 函数的分类、攻击方法和一些有代表性的 Hash 算法，时戳技术，大量的身份识别协议和基于身份的密码方案，一些密钥管理技术，电子货币以及电子选举协议、潜信道和健忘传输协议。附录介绍了本书中所用到的一些最基本的数学知识和美国国家标准技术研究所(NIST)最近公布的十五个 AES 候选算法。

本书可供从事信息安全专业的科技人员、硕士和博士研究生参考，也可供高等院校相关专业的师生阅读。

图书在版编目(CIP)数据

密码学导引/冯登国，裴定一编著. -北京：科学出版社，1999.4
ISBN 7-03-007295-2

Ⅰ.密… Ⅱ.①冯… ②裴… Ⅲ.密码-理论 Ⅳ.TN918.2

中国版本图书馆 CIP 数据核字(1999)第 02749 号

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717

新蕾印刷厂 印刷
新华书店北京发行所发行 各地新华书店经售

1999 年 4 月第 一 版 开本：787×1092 1/16
1999 年 4 月第一次印刷 印张：19
印数：1—3 000 字数 434 000

定价：30.00 元

(如有印装质量问题，我社负责调换〈环伟〉)

前　言

电子计算机和通信网络的广泛应用，一方面为人们的生活和工作带来了极大的方便，但另一方面也带来了许多亟待解决的问题，其中信息的安全性就是一个突出的问题。密码技术是保证信息的安全性的关键技术。信息的安全性主要有两个方面，即信息的保密性和认证性。保密的目的是防止对手破译系统中的机密信息。认证的目的有两个：一是验证信息的发送者是真正的，而不是冒充的；二是验证信息的完整性，在传送或存储过程中未被窜改、重放或延迟等。信息的认证性和信息的保密性是信息的安全性的两个不同方面，认证不能自动地提供保密性，而保密也不能自然地提供认证功能。在现有的大部分教材中侧重于信息的保密性的研究，本书将兼顾信息的保密性和认证性这两个方面的研究，并行处理。本书以密码算法和协议为主线，对当前的热门话题诸如电子货币，零知识协议，基于身份的密码算法和协议，以及密钥托管技术等作了比较详细的论述。

Shannon 在 1949 年发表了"保密通信的信息理论"（见第二章文献[1]），将密码学的研究纳入了科学的轨道，但该篇论文当时并没有引起人们的广泛重视，直到 70 年代中期，人类开始步入信息时代时才引起了普遍的重视。那时密码学研究出现了两件引人注目的事情：一件是 Diffie 和 Hellman 发表了"密码学的新方向"（见第一章文献[5]）一文，提出一种崭新的密码体制，冲破了长期以来一直沿用的私钥体制；另一件是美国国家标准局（NBS）公开征集，并于 1977 年正式公布实施的美国数据加密标准（DES）。这两个事件标志着现代密码学的诞生。

全书共分 11 章。第 1～3 章和附录介绍了密码学的基础知识，包括密码学的基本概念，密码学的信息理论基础，密码学的复杂性理论基础和本书中所用到的一些最基本的数学知识。第 4～12 章主要介绍了现有的有代表性的算法和协议，其中包括一些有代表性的私钥密码算法、一些有代表性的公钥密码算法、各种数字签名方案、一些典型的 Hash 算法、一些流行的识别协议、一些密钥分配和交换协议及密钥托管技术、电子货币以及一些特殊的密码协议。

本书是作者在长期从事科研和教学的实践基础上编写的。为了适应不同层次和不同专业的读者，我们对内容作了精心的安排，汲取了国内外现有文献中的精华。各章末附有注记和文献，介绍与该课题有关的资料和发展现状，以便感兴趣的读者进一步研究。

作　者

1998 年 9 月 10 日

目　录

第1章 引 论

密码学是一门古老而又年青的科学，它用于保护军事和外交通信可追溯到几千年前。在当今的信息时代，大量的敏感信息如病历、法庭记录、资金转移、私人财产等常常通过公共通信设施或计算机网络来进行交换，而这些信息的秘密性和真实性是人们迫切需要的。因此，现代密码学的应用已不再局限于军事、政治和外交，其商用价值和社会价值也已得到了充分肯定。

密码学的发展历史大致可划分为三个阶段：

第一个阶段为从古代到1949年。这一时期可看作是科学密码学的前夜时期，这段时期的密码技术可以说是一种艺术，而不是一种科学，密码学专家常常是凭直觉和信念来进行密码设计和分析，而不是推理证明。

第二个阶段为从1949年到1975年。1949年Shannon发表的“保密系统的信息理论”一文为私钥密码系统建立了理论基础，从此密码学成为一门科学，但密码学直到今天仍具有艺术性，是具有艺术性的一门科学。这段时期密码学理论的研究工作进展不大，公开的密码学文献很少。1967年Kahn出版了一本专著《破译者》(The Codebreakers)[1]，该书没有任何新的技术思想，只记述了一段值得注意的完整经历，包括政府仍然认为是秘密的一些事情。它的意义在于它不仅记述了1967年之前密码学发展的历史，而且使许多不知道密码学的人了解了密码学。70年代初期，IBM发表了Feistel和他的同事们在这个学科方面的几篇技术报告[2,3,4]。

第三个阶段为1976年至今。1976年Diffie和Hellman的“密码学的新方向”[5]一文导致了密码学上的一场革命。他们首次证明了在发送端和接收端无密钥传输的保密通信是可能的，从而开创了公钥密码学的新纪元。

本章主要介绍密码学的一些基本概念和一些有代表性的古典密码体制及其密码分析。

1.1 密码学的基本概念

密码学(cryptology)是研究密码系统或通信安全的一门科学。它主要包括两个分支，即密码编码学(cryptography)和密码分析学(cryptanalytics)。密码编码学的主要目的是寻求保证消息保密性或认证性的方法，密码分析学的主要目的是研究加密消息的破译或消息的伪造。

采用密码技术可以隐蔽和保护需要保密的消息，使未授权者不能提取信息。被隐蔽的消息称作明文(plaintext)，隐蔽后的消息称作密文(ciphertext)或密报(cryptogram)。将明文变换成密文的过程称作加密(encryption)，其逆过程，即由密文恢复出原明文的过程称作解密(decryption)。对明文进行加密操作的人员称作加密员或密码员(cryptographer)。密码员对明文进行加密时所采用的一组规则称作加密算法(encryption algorithm)，传送

消息的预定对象称作接收者(receiver),他对密文进行解密时所采用的一组规则称作解密算法(decryption algorithm)。加密和解密算法的操作通常都是在一组密钥(key)控制下进行的,分别称为加密密钥(encryption key)和解密密钥(decryption key)。

根据密钥的特点,Simmons[6]将密码体制分为对称和非对称密码体制(symmetric 和 asymmetric cryptosystem)两种。对称密码体制又称单钥(one-key)或私钥(private key)或传统(classical)密码体制,非对称密码体制又称双钥(two-key)或公钥(public key)密码体制。在本书中,我们采用私钥和公钥密码体制这两个术语。在私钥密码体制中,加密密钥和解密密钥是一样的或彼此之间容易相互确定。按加密方式又可将私钥密码体制分为流密码(stream cipher)和分组密码(block cipher)两种。在流密码中,将明文消息按字符逐位地加密。在分组密码中,将明文消息分组(每组含有多个字符),逐组地进行加密。在公钥密码体制中,加密密钥和解密密钥不同,从一个难于推出另一个,可将加密和解密能力分开。现有的大多数公钥密码属于分组密码,只有概率加密体制属于流密码。

在消息传输和处理系统中,除了意定的接收者外,还有非授权者,他们通过各种办法如搭线窃听、电磁窃听、声音窃听等来窃取机密信息,称其为截收者(eavesdropper)。他们虽然不知道系统所用的密钥,但通过分析可能从截获的密文推断出原来的明文,这一过程称作密码分析(cryptanalysis)。从事这一工作的人称作密码分析员或密码分析者(cryptanalyst)。对一个密码系统采取截获密文进行分析的这类攻击称作被动攻击(passive attack)。密码系统还可能遭受的另一类攻击是主动攻击(active attack),非法入侵者(tamper)主动向系统窜扰,采用删除、更改、增填、重放、伪造等手段向系统注入假消息,以达到损人利己的目的。所谓一个密码是可破的(breakable),是指如果通过密文能够迅速地确定明文或密钥,或通过明文-密文对能够迅速地确定密钥。通常假定密码分析者或敌手(opponent)知道所使用的密码系统,这个假设称作 Kerckhoff 假设。当然,如果密码分析者或敌手不知道所使用的密码系统,那么破译密码是更难的,但是我们不应该把密码系统的安全性建立在敌手不知道所使用的密码系统这个前提之下。因此,在设计一个密码系统时,我们的目的是在 Kerckhoff 假设下达到安全性。

破译或攻击(break 或 attack)密码的方法有穷举破译法(exhaustive attack method)和分析法两种。穷举法又称作强力法(brute force method)或完全试凑法(complete trial-and-error method),这种方法是对截获的密文依次用各种可能的密钥试译,直到得到有意义的明文,或在密钥不变的情况下,对所有可能的明文加密直到得到与截获密文一致为止。只要有足够的计算时间和存储空间,原则上穷举法总是可以成功的,但在实际中,时间和存储空间都受到约束,因此,这种方法往往是不可行的。分析破译法又分确定性分析法和统计分析法两类。确定性分析法是利用一个或几个已知量用数学关系式表示出所求未知量(如密钥等)。统计分析法是利用明文的已知统计规律进行破译的方法,密码分析者对截收的密文进行统计分析,总结出其间的统计规律,并与明文的统计规律进行对照比较,从中提取出明文和密文之间的对应或变换信息。密码分析之所以能够成功地破译密码,最根本的原因是明文中有多余度。

根据密码分析者破译时已具备的前提条件,通常人们将攻击类型分为四种,即唯密文攻击(ciphertext-only attack)、已知明文攻击(known plaintext attack)、选择明文攻击(chosen plaintext attack)和选择密文攻击(chosen ciphertext attack)。

(1)唯密文攻击:密码分析者有一个或更多的用同一个密钥加密的密文,通过对这些截获的密文进行分析得出明文或密钥。

(2)已知明文攻击:除待解的密文外,密码分析者有一些明文和用同一个密钥加密这些明文所对应的密文。

(3)选择明文攻击:密码分析者可得到所需要的任何明文所对应的密文,这些密文与待解的密文是用同一个密钥加密得来的。

(4)选择密文攻击:密码分析者可得到所需要的任何密文所对应的明文(这些明文可能是不大明了的),解密这些密文所使用的密钥与解密待解的密文的密钥是一样的。

上述四种攻击类型的强度按序递增,唯密文攻击是最弱的一种攻击,选择密文攻击是最强的一种攻击。选择密文攻击主要用于分析公钥密码体制。如果一个密码系统能够抵抗选择明文攻击,那么它当然能够抵抗唯密文攻击和已知明文攻击。

一个密码通信系统可用图 1.1.1 表示,它由以下几个部分组成:明文消息空间 P;密文消息空间 C;密钥空间 K_1 和 K_2,在私钥体制下 $K_1=K_2=K$,此时密钥 K 需经安全的密钥信道由发方传送给收方;加密变换 $E_{k_1}:P\to C,k_1\in K_1$,由加密器完成;解密变换 $D_{k_2}:C\to P,k_2\in K_2$,由解密器实现。对每一个密钥 $k_1\in K_1$(k_1 确定一个加密变换 E_{k_1}),有一个匹配的密钥 $k_2\in K_2$(k_2 确定一个解密变换 D_{k_2})使得对一切 $m\in P$,有 $D_{k_2}(E_{k_1}(m))=m$。

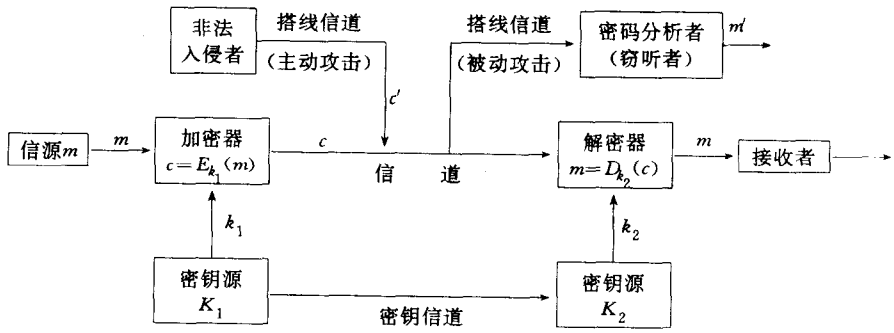

图1.1.1 密码系统模型

对于给定的明文消息 $m\in P$ 和密钥 $k_1\in K_1$,加密变换将明文 m 变换为密文 c:

$$c=f(m,k_1)=E_{k_1}(m),\qquad m\in P,\qquad k_1\in K_1 \tag{1.1.1}$$

接收者利用通过安全信道传送来的密钥 k_1(私钥体制下)或用本地密钥生成器产生的解密密钥 $k_2\in K_2$(公钥体制下)控制解密操作 D,对收到的密文进行变换恢复明文消息 m:

$$m=D_{k_2}(c),m\in P,k_2\in K_2 \tag{1.1.2}$$

而密码分析者,则用选定的变换函数 h,对截获的密文 c 进行变换,得到的明文是明文空间中的某个元素 m'

$$m'=h(c) \tag{1.1.3}$$

一般地,$m'\neq m$。

令 $\varepsilon=\{E_{k_1}:P\to C\mid k_1\in K_1\}$,$D=\{D_{k_2}:C\to P\mid k_2\in K_2\}$,则称六元组$(P,C,K_1,K_2,\varepsilon,D)$为一保密系统(secrecy system)。

为了保护信息的机密性,抵抗密码分析,保密系统应当满足下述要求:

(1)系统即使达不到理论上是不可破的,即 $p_r(m'=m)=0$,也应当是实际上不可破

的。也就是说，从截获的密文或某些已知明文-密文对，要确定密钥或任意明文在计算上是不可行的。

(2)系统的保密性不依赖于对加密体制或算法的保密(Kerckhoff 假设)，而依赖于密钥。

(3)加密和解密算法适用于所有密钥空间中的元素。

(4)系统既易于实现又便于使用。

防止消息被窜改、删除、重放和伪造的一种有效的方法是使发送的消息具有被验证的能力，使接收者或第三者能够识别和确认消息的真伪，实现这类功能的密码系统称作认证系统(authentication system)。消息的认证性和消息的保密性不同，保密性是使截获者在不知道密钥的条件下不能解读密文的内容，而认证性是使任何不知道密钥的人不能构造出一个密报，使意定的接收者解密成一个可理解的消息(合法的消息)。认证理论和技术是最近 20 年来随着计算机通信的普遍应用而迅速发展起来的一个重要的密码学研究领域。如传统的手写签名正在被更迅速、更经济和更安全的数字签名(digital signature)所取代。

一个安全的认证系统应当满足下述的基本要求：

(1)意定的接收者能够检验和证实消息的合法性和真实性。

(2)消息的发送者对所发送的消息不能抵赖。

(3)除了合法的消息发送者外，其他人不能伪造合法的消息。而且在已知合法密文 c 和相应消息 m 下，要确定加密密钥或系统地伪造合法密文在计算上是不可行的。

(4)当通信双方(或多方)发生争执时，可由称作仲裁者(arbitrator)的第三方解决争执。

这里值得一提的是，密码学中的术语“系统或体制”(system)、“方案”(scheme)和“算法”(algorithm)本质上是一回事，本书中按作者的习惯交替使用了这些术语。

1.2 古典密码学

本节简要介绍几种古典密码体制及对这些体制的一些破译方法，用来说明设计和分析密码的基本方法。虽然这些密码大都比较简单而且容易破译，但研究这些密码的设计原理和分析方法对于理解、设计和分析现代密码是十分有益的。

1.2.1 古典密码体制

1.2.1.1 代换密码和置换密码

令 A 表示含 N 个“字母”或“字符”的明文字母表，例如，可以是普通的英文字母 $A\sim Z$，也可以是数字、空格、标点符号或任何可以表示明文消息的符号。因此可以将 A 抽象地表示为一个整数集 $Z_N=\{0,1,\cdots,N-1\}$。在加密时通常将明文消息划分成长为 L 的消息单元，称为明文组，以 m 表示，如 $m=(m_0,m_1,\cdots,m_{L-1})$，$m_l\in Z_N$，$0\leqslant l\leqslant L-1$。$m$ 也称作 L-报文，它是定义在 Z_N^L 上的随机变量，$Z_N^L=Z_N\times Z_N\times\cdots\times Z_N$($L$ 个)$=\{m=(m_0,m_1,\cdots,m_{L-1})\mid m_l\in Z_N,0\leqslant l\leqslant L-1\}$。$L=1$ 为单字母报(1-gram)，$L=2$ 为双字母报(digrams)，$L=3$ 为三字母报(trigrams)。明文空间 $P=Z_N^L$。

令 A' 表示含 N' 个“字母”或“字符”的密文字母表，抽象地可用整数集 $Z_{N'}=\{0,1,\cdots,N'-1\}$ 来表示。密文单元或组为 $c=(c_0,c_1,\cdots,c_{L'-1})$（$L'$ 个），$c_{l'}\in Z_{N'}$，$0\leqslant l'\leqslant L-1$。$c$ 是定义在 $Z_{N'}^{L'}$ 上的随机变量。密文空间 $C=Z_{N'}^{L'}$。

一般地，明文和密文由同一字母表构成，即 $A'=A$。

加密变换是从明文空间到密文空间的映射 $f:P\to C$。加密变换通常是在密钥控制下变化的，因此，一般记为

$$c=f(m,k)=E_k(m)\qquad k\in K\qquad m\in P\qquad c\in C \tag{1.2.1}$$

K 为密钥空间。

假定 f 是一个单射，对固定的 $k\in K$，令 $C_k=\{c=f(m,k)=E_k(m)\mid m\in P\}\subseteq C$，因此对给定的密文组 $c\in C_k$，有且仅有一个对应的明文组 m，也就是说，对于此函数 f，存在逆映射 $f^{-1}:C_k\to P$，使

$$f^{-1}(c)=f^{-1}\{f(m)\}=m\qquad m\in P\qquad c\in C_k \tag{1.2.2}$$

即 f^{-1} 为解密变换。

一个密码系统就是在 f 作用下由 Z_N^L 到 $Z_{N'}^{L'}$ 的映射，在这种意义上，称此种密码为代换密码(substitution cipher)，如图 1.2.1 所示。$L=1$ 时，称作单字母代换，也称作流密码(stream cipher)。$L>1$ 时，称作多码代换，也称作分组密码(block cipher)。

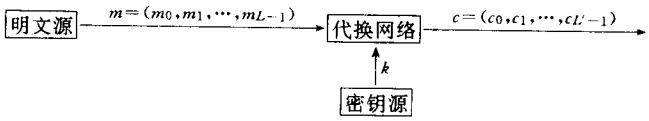

图 1.2.1　代换密码框图

一般地，选择相同的明文和密文字母表。此时，若 $L=L'$，则映射 f 可构造成一对一的映射，密码无数据扩展。若 $L<L'$，则有数据扩展，可将映射 f 设计成一对多的映射，即明文组可能找到多于一个密文组来代换，这称之为多名(或同音)代换密码(homophonic substitution cipher)。若 $L>L'$，则明文数据将被压缩，此时映射 f 不可能构造成可逆映射，从而从密文有时也就无法完全恢复出原明文消息，因此保密通信中必须要求 $L\leqslant L'$。但 $L>L'$ 的映射可以用在认证系统中。

在 $A=A'$，$N=N'$ 和 $L=1$ 时，若对所有明文字母，都用一种固定的代换进行加密，则称这种密码为单表代换(monoalphabetic substitute)。若用一个以上的代换表进行加密，这就称作是多表代换(polyalphabetic substitute)。这是古典密码中的两种重要体制，曾被广泛地使用过。

在代换密码中有一种特殊的代换密码，即代换并没有改变明文字母，而只改变了它们的位置，密码学史上把这种代换密码称作置换密码(permutation cipher)，又称换位密码(transposition cipher)。下面是置换密码的一个详细描述。

设 $P=C=Z_N^L$，K 表示所有的 $\{0,1,\cdots,L-1\}$ 上的置换构成的集合。对每一个给定的密钥 $k=\pi$（一个置换），置换密码的加密变换定义为 $E_\pi(m_0,m_1,\cdots,m_{L-1})=(m_{\pi(0)},m_{\pi(1)},\cdots,m_{\pi(L-1)})=(c_0,c_1,\cdots,c_{L-1})$。因而解密变换为 $D_{\pi^{-1}}(c_0,c_1,c_{L-1})=(c_{\pi^{-1}(0)},c_{\pi^{-1}(1)},\cdots,c_{\pi^{-1}(L-1)})$，这里 π^{-1} 是 π 的逆置换。

1.2.1.2 单表代换密码

单表代换密码是对明文的所有字母都用一个固定的明文字母表到密文字母表的映射，即 $f:Z_N \to Z_N$。令明文 $m=m_0m_1\cdots$，则相应的密文为 $c=E(m)=c_0c_1\cdots=f(m_0)f(m_1)\cdots$。若明文字母表为 $A=Z_N=\{0,1,\cdots,N-1\}$，则相应的明文字母表为 $A'=\{f(0),f(1),\cdots,f(N-1)\}$，$A'$ 是 A 的某种置换。本小节主要来介绍一类简单的单表代换密码——仿射密码。

仿射密码(affine cipher)是一种特殊的代换密码。明文空间 P 和密文空间 C 均为 Z_N，密钥空间为 $K=\{k=(k_1,k_0)\mid \gcd(k_1,N)=1,k_0,k_1\in Z_N\}$。对给定的密钥 $k=(k_1,k_0)\in K$，其加密变换为

$$E_k(i)=(ik_1+k_0)\bmod N \tag{1.2.3}$$

解密变换为

$$D_k(j)=k_1^{-1}(j-k_0)\bmod N \tag{1.2.4}$$

其中 $\gcd(k_1,N)$ 表示 k_1 与 N 的最大公因子，$x\bmod N$ 表示 x 除以 N 所得的余数，密钥空间 K 的大小为 $N\varphi(N)$。关于剩余类环 Z_N 和欧拉函数 $\varphi(\cdot)$ 的定义和性质参见附录。

当 $k_0=0$ 时，称为乘法密码(multiplicative cipher)，又称采样密码(decimation cipher)。当 $k_0=1$ 时，称为移位密码(shift cipher)，又称加法密码(additive cipher)。

我们将通过建立英文字母和模 26 的剩余之间的对应关系来使用移位密码加密普通的英文消息。

例 1.2.1 假定移位密码的密钥为 $k_0=3$，明文消息为 we will meet at midnight。我们首先利用表 1.2.1 将该明文转化成一列整数：22 4 22 8 11 11 12 4 4 19 0 19 12 8 3 13 8 6 7 19。其次加 3 到每个值，并进行模 26 运算得：7 15 7 19 22 22 23 15 15 4 11 4 23 19 14 24 19 17 18 4。最后，再利用表 1.2.1 将这一列整数转化为字母，从而获得密文：HPHTWWXPPELEXTOYTRSE。解密过程与加密过程类似，不同的只是进行模 26 减 3，而不是模 26 加 3。

表 1.2.1 英文字母和模 26 的剩余之间的对应关系

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

1.2.1.3 多表代换密码

多表代换密码是以一系列(两个以上)代换表依次对明文消息的字母进行代换的加密方法。令明文字母表为 Z_N，$f=(f_1,f_2,\cdots)$ 为代换序列，明文字母序列 $m=m_1m_2\cdots$，则相应的密文字母序列为 $c=E_k(m)=f(m)=f_1(m_1)f_2(m_2)\cdots$。若 f 是非周期的无限序列，则相应的密码称为非周期多表代换密码。这类密码，对每个明文字母都采用不同的代换表(或密钥)进行加密，称作一次一密密码(one-time pad cipher)，这是一种在理论上唯一不可破的密码(参见第 2 章)。这种密码可以完全隐蔽明文的特点，但由于需要的密钥量和明文消息长度相同而难于广泛使用。为了减少密钥量，在实际应用中多采用周期多表代换密码，即代换表个数有限，重复地使用，此时代换表序列为 $f=(f_1,f_2,\cdots,f_d,f_1,f_2,\cdots,f_d,\cdots)$，

相应于明文字母 m 的密文为 $c=E_k(m)=f(m)=f_1(m_1)f_2(m_2)\cdots f_d(m_d)f_1(m_{d+1})f_2(m_{d+2})\cdots f_d(m_{2d})\cdots$。当 $d=1$ 时就退化为单表代换，因此可以说多表代换密码是单表代换密码的一种推广。

有名的多表代换密码有 Vigenère、Beaufort、Running-Key、Vernam 和转轮机(rotor machine)等密码。这里只介绍 Vigenère 多表代换密码，其它的参见文献[7,8,9,10]等。

Vigenère 密码是由法国密码学家 Blaise de Vigenère 于 1858 年提出的一种密码，它是一种以移位代换(当然也可以用一般的字母代换表)为基础的周期代换密码。d 个代换表 $f=(f_1,f_2,\cdots,f_d)$ 由 d 个字母序列给定的密钥 $k=(k_1,k_2,\cdots,k_d)\in Z_N^d$ 决定，其中 $k_i(i=1,2,\cdots,d)$ 确定明文的第 $i+td$ 个字母(t 为正整数)的移位次数，即加密公式为

$$c_{i+td}=E_{k_i}(m_{i+td})=(m_{i+td}+k_i)\bmod N \tag{1.2.5}$$

从而解密公式为

$$\begin{aligned} m_{i+td}&=D_{k_i}(c_{i+td})=E_{N-k_i}(c_{i+td})\\ &=(N-k_i+m_{i+td}+k_i)\bmod N=m_{i+td} \end{aligned} \tag{1.2.6}$$

称 k 为用户密钥(user key)或密钥字 (key word)。密钥量为 N^d，当 N 与 d 较大时，密钥量是很大的。将用户密钥 k 周期地延伸就给出了整个明文加密所需的工作密钥(working key)。

例 1.2.2 假定我们仍使用表 1.2.1，$d=6$，$k=$cipher，明文串是 this cryptosystem is not secure。

首先将 k 及明文串转化为数字串：$k=(2,8,15,7,4,17)$，$m=(19,7,8,18,2,17,24,15,19,14,18,24,18,19,4,12,8,18,13,14,19,18,4,2,20,17,4)$。

其次模 26“加”密钥字 $k=(2,8,15,7,4,17)$得：

19	7	8	18	2	17		24	15	19	14	18	24																	
2	8	15	7	4	17		2	8	15	7	4	17																	
21	15	23	25	6	8		0	23	8	21	22	15																	
18	19	4	12	8	18		13	14	19	18	4	2																	
2	8	15	7	4	17		2	8	15	7	4	17																	
21	1	19	19	12	9		15	22	8	25	8	19																	
20	17	4																											
2	8	15																											
22	25	19																											

最后将所得的密文数字串利用表 1.2.1 转化成密文字母串即

VPXZGIAXIVWPUBTTMJPWIZITWZT

解密过程与加密过程类似，不同的只是进行模 26 减，而不是模 26 加。

1.2.1.4 多字母代换密码

前面介绍的仿射密码和 Vigenère 密码都是以单个字母作为代换对象的。如果每次对 $L>1$ 个字母进行代换就是多字母代换密码(polygram substitution cipher)。多字母代换的优点是容易将字母的自然频度隐蔽或均匀化而有利于抵抗统计分析。这种密码主要有 Playfair 密码、Hill 密码等。这里主要介绍 Hill 密码,其它的参见文献[7,8,9,10]等。

令明文字母表为 Z_N,若采用 L 个字母为单元进行代换,则多码代换是映射 $f:Z_N^L \to Z_N^L$。若映射 f 是线性的,则称 f 是线性变换,可用一个 Z_N 上的 $L\times L$ 阶矩阵 T 表示。若 T 是满秩的,则变换为一对一映射,且存在逆变换 T^{-1},使 $TT^{-1}=T^{-1}T=I$(I 为 $L\times L$ 阶单位矩阵)。将 L 个字母的数字表示为 Z_N 上的一个向量 $m=(m_1,m_2,\cdots,m_L)$,则 Hill 密码的加密变换为

$$c=(c_1,c_2,\cdots,c_L)=mT \bmod N \tag{1.2.7}$$

解密变换为

$$m=cT^{-1} \bmod N \tag{1.2.8}$$

类似于单字母仿射代换密码,可构造多字母仿射代换密码。令 $b=(b_1,b_2,\cdots,b_L)\in Z_N^L$,$T$ 为 Z_N 上的 $L\times L$ 阶满秩矩阵,则可通过下述仿射变换对明文组 $m=(m_1,m_2,\cdots,m_L)$加密得密文 $c=(c_1,\cdots,c_L)$,即

$$c=(mT+b) \bmod N \tag{1.2.9}$$

解密变换为

$$m=(c-b)T^{-1} \bmod N \tag{1.2.10}$$

当 $T=I$ 时,就是上一小节介绍的 Vigenère 密码。

例 1.2.3 假定 $L=2$,明文空间和密文空间均为 $P=C=Z_{26}$,密钥

$$K=\begin{pmatrix}11 & 8\\ 3 & 7\end{pmatrix},\quad K^{-1}=\begin{pmatrix}7 & 18\\ 23 & 11\end{pmatrix}$$

现在加密明文 july,由表 1.2.1 知,ju 对应于数字组(9,20),ly 对应于数字组(11,24)。计算$(9,20)K \bmod 26=(3,4)$,$(11,24)K \bmod 26=(11,22)$。再由表 1.2.1 知,(3,4)对应于 DE,(11,22)对应于 LW,故密文为 DELW。

解密过程:由表 1.2.1 知,DE 对应于(3,4),LW 对应于(11,22)。计算$(3,4)K^{-1} \bmod 26=(9,20)$,$(11,22)K^{-1} \bmod 26=(11,24)$。再由表 1.2.1 知,(9,20)对应于 ju,(11,24)对应于 ly,故明文为 july。

1.2.2 古典密码体制分析

简单的单表代换密码,如移位密码极易破译。仅统计标出最高频度字母再与明文字母表字母对应决定出移位量,就差不多可以得到正确解了。其它如乘法密码、一般的仿射密码要复杂些,但多考虑几个密文字母统计表与明文字母统计表的匹配关系也不难解出。另外单表代换密码如移位密码也很容易用穷举密钥搜索来破译,因为密钥量仅为 N。可见,一个密码系统是安全的一个必要条件是密钥空间必须足够的大,使得穷举密钥搜索破译是不可行的,但这不是一个密码系统安全的充分条件。

多表代换密码的破译要比单表代换密码的破译难得多,因为在单表代换下,字母的频