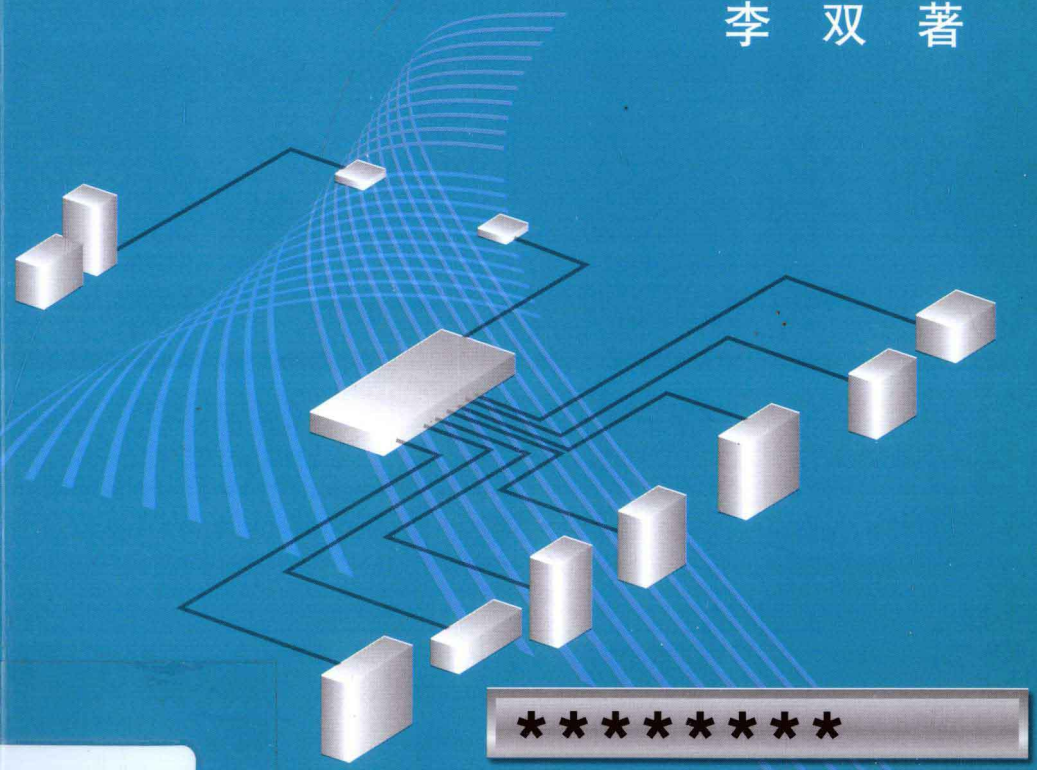
访问控制与加密

李 双 著



机械工业出版社
CHINA MACHINE PRESS

访问控制与加密

李 双 著

机械工业出版社

本书内容包括为三部分。第一部分为第1章，介绍密码学的相关概念和知识。第二部分为第2章～第5章，介绍了访问控制技术，包括自主访问控制、强制访问控制、基于角色的访问控制、基于对象的访问控制和基于任务的访问控制。特别对基于对象的访问控制展开讨论，提出了一种扩展的基于角色的访问控制模型，并谈论了实现细节。第三部分对应书中第6章和第7章，讲述可搜索加密方案的背景、相关知识和现状，并提出了基于属性的可搜索加密方案。

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

访问控制与加密/李双著. —北京：机械工业出版社，2012.8

ISBN 978-7-111-39128-9

Ⅰ. ①访… Ⅱ. ①李… Ⅲ. ①电子计算机－安全技术 Ⅳ. ①TP309

中国版本图书馆CIP数据核字（2012）第154242号

机械工业出版社（北京市百万庄大街22号 邮政编码100037）

策划编辑：徐明煜 责任编辑：徐明煜 顾 谦

版式设计：霍永明 责任校对：张 媛

封面设计：姚 毅 责任印制：张 楠

北京振兴源印务有限公司印刷

2012年8月第1版第1次印刷

148mm×210mm · 5.5印张 · 149千字

0001—2500册

标准书号：ISBN 978-7-111-39128-9

定价：20.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务 网络服务

社服务中心：（010）88361066 教材网：http://www.cmpedu.com

销售一部：（010）68326294 机工官网：http://www.cmpbook.com

销售二部：（010）88379649 机工官博：http://weibo.com/cmp1952

读者购书热线：（010）88379203 **封面无防伪标均为盗版**

前　言

目前，许多敏感的信息和技术大都是通过计算机来控制和管理的，随着计算机技术，特别是网络技术的发展，大型网络应用系统或数据管理系统所面临的一个难题就是复杂的数据资源的管理，如何确保它们的安全成为当今计算机安全领域研究的热点。ISO（国际标准化组织）（ISO 7498-2）中定义了5个层次型安全服务：身份认证服务、访问控制服务、数据保密服务、数据完整性服务、不可否认服务。访问控制和数据保密是其中的重要部分。

互联网的高速发展，越来越多的企业和个人已经把大量数据交给第三方服务器存储，保护私有数据的机密性和隐私成为急需解决的问题。加密技术是保护数据机密性和隐私的一种较为有效的手段，然而对加密后的数据进行检索却是一项非常困难的工作。特别是在非可信环境的情况下，如何对加密数据进行高效地查询引起了人们普遍的关注。因此，可搜索加密（关键词可搜索公钥加密方案，简称“可搜索加密”）成为了近几年的研究热点之一。

就本书的内容而言，涉及密码学基础知识、访问控制和可搜索加密三个方面。密码学基础知识是本书中的第1章，主要介绍密码学简介、私钥密码体制、公钥密码体制、Hash函数和密码分析相关知识。

关于访问控制技术对应本书的第2章～第5章。主要讨论基于角色的访问控制。基于角色的访问控制是由美国国家标准与技术研究局（NIST）的Ferraiolo等人在20世纪90年代提出来的。2001年8月，RBAC的技术提出者和NIST提出了一个建议标准，建立了NIST RBAC2001技术的参考模型，定义了各种模型构件，并给出了一套系统与管理功能规范。本书的第3章介绍NIST RBAC2001技术的参考模型及相关细节。

尽管 NIST 对 RBAC2001 定义了建议标准，但描述简单，只规定了最基本的概念需求。因此 RBAC 还是一个开放的课题，还有许多方面值得研究。为使 RBAC 适用于更广泛的范围，我们在第 4 章提出一种可代理的、具有条件约束的权限清晰的 ERBAC 模型。此 ERBAC 模型不仅继承了传统 RBAC 的所有优势，而且可以不作任何修改地嵌入到工作流系统中，可在工作流系统中实现动态的访问控制。关于 RBAC 的实现，书中第 5 章主要对访问控制中的安全策略作了详细讨论，提出了将 ERBAC 模型与 RBMLS（基于角色的多级安全）模型结合起来，实现在自主中渗透着强制，在强制中蕴涵着自主的访问策略。其中 RBMLS 模型是支持等级分类级别不同信息的单向或双向流动，且机密性与完整性兼顾的多级安全政策模型。在 RBAC 的具体实施中，文章针对网络终端不固定的用户的网络访问，具体讨论了基于口令的实施方案。

可搜索加密技术的讨论对应本书第 6 章和第 7 章。2004 年，Boneh 等人利用匿名的基于身份加密方案构造了一个公钥可搜索加密方案（Public Key Encryption with Keyword Search，PEKS），该方案是针对在加密的邮件系统中邮件网关搜索带特定关键词邮件的应用场景而提出的，解决了特定环境下对加密数据进行检索的这一困难工作。随后有很多科学家投入到可搜索加密的研究中，将 PEKS 扩展到支持多个关键词查询、支持多种检索操作、对检索关键词的门限有时间限制等。第 6 章首先介绍可搜索加密技术，第 7 章给出了基于属性的可搜索方案（ATT-PEKS）的定义和构造方法，此方案与 PBEKS 不同之处在于基于属性的可搜索方案是适应群组的公钥加密搜索方案，扩大了信息的共享性，实现了信息的多方查询，节省了第三方信息存储的空间，并且进行了一致性和安全性证明。

可搜索加密方案的研究是目前密码学领域重要的课题，有很多问题需要研究，如在加密检索中如何能够保护用户身份信息、访问模式、检索标准等敏感信息，这是一个重要的课题；同时相关的研究也不仅仅局限于方案本身，甚至会引发研究者们对一些深层次的数学问题的研究和发现。并且随着新应用的出现，必将会诞生更多、更新、更好的方案。

限于本人水平，书中难免有疏漏和错误之处，恳请读者批评指正。

李　双

2012 年 6 月于北京

目　录

第 1 章　密码学基础

1.1　密码学简介

密码学是研究密码系统或通信安全的一门科学。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消息。被隐蔽的消息称作明文（Plaintext），隐蔽后的消息称作密文（Ciphertext）。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）。对明文进行加密时所采用的一组规则称作加密算法（Encryption Algorithm），对密文进行解密时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）。加密是建立在将明文变换成一种不可读或不可识别的密文的加密算法基础上，解密是将不可读或不可识别的密文恢复成明文的逆过程（见图 1.1.1）。

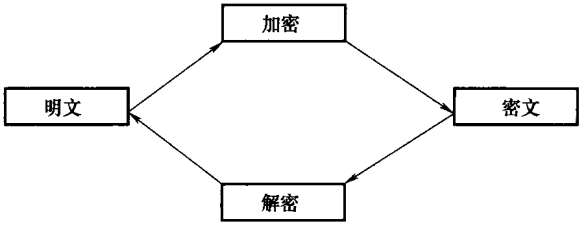


图 1.1.1　加密和解密过程

密码体制（Cryptosystem）是一些算法和相关的隐藏信息和显露信息的集合。根据密钥的特点，Simmons[1]将密码体制分为对称密码体制（Symmetric Cryptosystem）和非对称密码体制（Asymmetric Cryptosystem）两种。对称密码体制又称私钥（Private Key）密码体

制和公钥（Public Key）密码体制。在私钥密码体制中，加密密钥和解密密钥是一样的或是彼此之间容易相互确定。在公钥加密体制中，加密密钥和解密密钥不同，从一个难于推出另一个，可将加密能力和解密能力分开。按加密方式又可将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。现有的大多数公钥密码体制属于分组密码。

与密码体制相对应的是密码分析（Cryptanalysis）。密码分析是分析一个密码体制的过程，或者是检验密码体制的完整性，或者是为了攻破它。密码分析的过程虽然不知道系统所用的密码，但通过分析可能从截获的密文推断出原来的明文。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），攻击者（Attacker）主动扰乱系统，采用删除、更改、增添、重放、伪造等手段向系统注入假消息。攻击一个密码体制的过程通常称为破译（Cracking）。所谓一个密码体制是可破的（Breakable），是指通过密文能够迅速地确定明文或密钥，或通过明文—密文对能够迅速确定密钥。通常假设密码分析者（Cryptanalyst）或敌手（Attacker）知道所使用的密码系统，这个假设称作 Kerckhoff 假设。

一个密码通信系统可用图 1.1.2 表示，它由以下部分组成：明文消息空间 P；密文消息空间 K；密钥源 K_1 和 K_2，在私钥体制下 $K_1=K_2=K$，此时密钥 K 需经安全的秘密通道由发送方传送给接收方；加密变换 E_{k_1}：$P\rightarrow C$，$k_1\in K_1$，由加密器实现；解密变换 D_{k_2}：$C\rightarrow P$，$k_2\in K_2$，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}），使得对所有 $m\in P$，有 $D_{k_2}(E_{k_1}(m))=m$。

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：

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

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

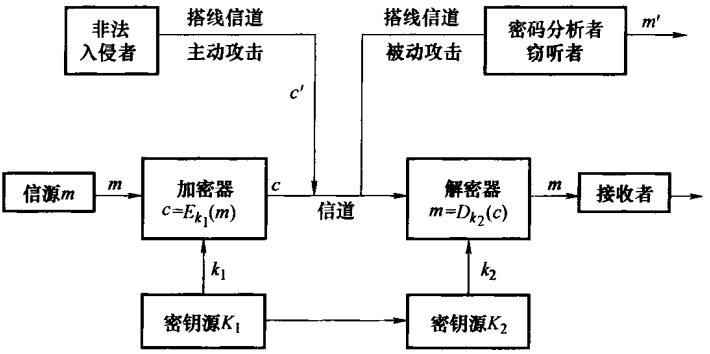


图 1.1.2　密码系统模型

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，对接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 $E = \{E_{k_1}: P \rightarrow C \mid k_1 \in K_1\}$，$D = \{D_{k_2}: C \rightarrow P \mid k_2 \in K_2\}$，则称六元组$(P, C, K_1, K_2, E, D)$为一保密系统(Secrecy System)。

在过去的几十年里，本质是数学的密码算法发展得非常先进。密码学并不像许多人认为的那样仅用于军事和外交通信。现实中，密码学有许多商业用途：保护公司机密信息、保护电话通话、电子商务、网银的安全使用。密码学完全是用于保护个人和组织的保密信息。例如，密码学用于防止伪造的火车票和火车票泄漏乘客个人信息。具体做法是每张火车票上都印有乘客身份证号的密文，检票时同时使用火车票和身份证，读取设备会首先解密火车票的身份证号密文，得到明文与身份证上的信息进行比对。除非伪造者已经分析出了所使用的密码体制，否则他们将不可能伪造出有效火车票。同时，因为火车票上印的是乘客身份证号的密文，即使乘客不小心丢失，也不致泄漏身份证号如此重要的个人信息。

1.2 私钥密码体制

私钥密码体制根据对明文消息加密方式的不同可分为两大类，即流密码和分组密码。流密码将消息分成连续的符号 $x=x_1, x_2, \cdots$，用密钥流 $k=k_1, k_2, \cdots$第 i 个元素 k_i 对 x_i 加密，即 $E_k(x)=E_{k_1}(x_1)E_{k_2}(x_2)\cdots$。如果流密码经过 d 个符号后重复，则称该流密码是周期的，否则称之为非周期的。一次一密密码是非周期的。在分组密码中，明文被分成若干组，每组含有 m 个符号 $x=(x_1, x_2, \cdots, x_m)$，每一组明文在密钥 $k=(k_1, k_2, \cdots, k_m)$下变换成对应的密文 $y=(y_1, y_2, \cdots, y_m)$，记为 $y=E_k(x)$，而且每组明文用同一个密钥 k。

流密码与分组密码的区别在于记忆性(见图 1.2.1)。

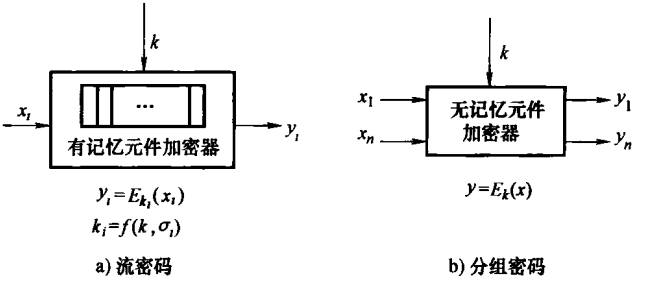


图 1.2.1 流密码与分组密码的加密方式

在流密码中，密钥流元素 k_i 的产生由第 i 时刻流密码的内部状态 σ_i 和种子密钥 k 决定，可写为 $k_i=f(k, \sigma_i)$。加(解)密变换 E_{k_i} (D_{k_i})是时变的，其时变性由加(解)密器中的记忆元件来保证。而分组密码的加(解)密变换 $E_{k_i}(D_{k_i})$不是时变的，加(解)密器中不存在记忆元件。

1.2.1 流密码

在流密码中，加密器中存储的状态随时间而变化，可以用状态转移函数来描述，记为 f_σ。根据状态转移函数 f_σ 是否依赖于输入的明文(字符或 bit)，可将流密码分为两类，即同步流密码(Synchro-

nous Steam Cipher)和自同步流密码(Self - synchronous Steam Cipher)。

在同步流密码中，转移状态函数 f_σ 与输入的明文无关。此时，密钥流 $\{k_i = f(k, \sigma_i)\}_{i=1}^{\infty}$ 与明文无关，而 i 时刻输出的密文 $c_i = E_{k_i}(m_i)$ 也不依赖于 i 时刻之前的明文。因而可将同步流密码的加密器划分成密钥流生成器[或滚动密钥生成器(Running Key Generator)]，或伪随机序列生成器(Pseudorandom Sequence Generator)和加密变换器(单纯利用滚动密钥 k_i 对输入的明文符号 x_i 进行加密的变换器)两部分(见图1.2.2)。

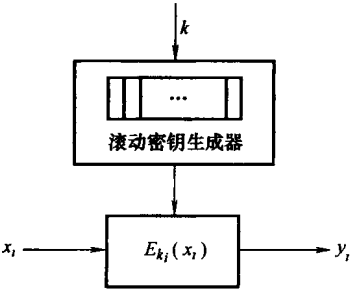


图1.2.2 分解后的流密码加密器

在同步流密码中，只要发送方和接收方有相同的种子密钥 k 和内部状态，就能产生出相同的密钥流。此时，我们说发送方和接收方的密钥生成器是同步的。一旦两者不同步，解密过程立即失败，密码系统此时要能提供某种辅助手段以重建同步。

同步流密码有两种基本的工作模式。一种是输出—分组反馈模式[2] (OMF)，另一种是 Diffie 和 Hellman[3] 给出的计数模式。同步流密码的一个优点是无错误传播，一个传输错误只影响一个符号，不会影响一组符号。但这也是一个缺点，因为敌手篡改一个符号比篡改一组符号容易。通过附加非线性纠错码可以克服这个缺陷。

在自同步流密码中，状态转移函数 f_σ 与输入的明文有关。此时密钥流 $k_i = f(k, \sigma_i)$ 与明文符号有关，而 i 时刻的密文 c_i 不仅仅依赖于明文符号 x_i。这种思想可追溯到16世纪 Vigenere 所发明的自身密码[4]。Vigenere 自身密码的密钥流通过在初始 k_1 后附加每个密文符号来产生，即 $k_i = y_{i-1}(i > 1)$。显然，这种密码是很不安全的，但是用所加密的消息来产生非重复密钥流的思想对密码学确是一大贡献。

密码反馈模式(CFB)是自同步流密码的一种最常用的工作模式。每个密文符号 y_i 在生成之后，立即送到移位寄存器 R 的一端(另一端的符号被丢掉)。在每次迭代中，R 的值作为分组加密算法 E 的输

入，而输出组的最低位符号用作下一个密文符号。

对于密码反馈模式，传输错误影响反馈圈。如果一个密文符号在传输中出错或丢失，等到该错误移出寄存器才能同步。因此一个错误至多影响 n 个符号，这里 n 为每组符号的个数。

“一次一密”密码是当今流密码体制的原型。“一次一密”要求用户在安全信道中传送长度不小于明文消息符号数量的密钥符号，这样做不切实际。上面介绍的流密码体制都不是完善保密系统，但是它们应当是条件安全的或是计算上不可破译的，为此由密钥 k 扩展成密钥流序列 $\{k_i\}_{i=1}^{\infty}$ 应当满足一定的要求，诸如极大的周期、良好的统计特性、能对抗已知的若干种攻击方法(如线性逼近、分别征服攻击等)。

从公开发表的文献来看，目前的绝大多数有关流密码的研究成果都是同步流密码方面的，由于自同步流密码系统一般需要密文反馈，因而使得分析工作复杂化。但自同步流密码具有抵抗密文搜索攻击和认证功能等优点。一个流密码是否具有很高的密码强度主要取决于密钥流生成器的设计。为了设计安全的密钥流生成器，必须在生成器中使用非线性变换，这就给生成器的理论分析工作带来了很大的困难。

1.2.2 分组密码

分组密码是将明文消息编码表示后的数字序列 x_1，x_2，…划分成长为 m 的组 $x=(x_1, x_2, \cdots, x_m)$，各组(长为 m 的矢量)分别在密钥 $k=(k_1, k_2, \cdots, k_m)$ 的控制下变换成等长的输出数字序列 $y=(y_1, y_2, \cdots, y_n)$(长为 n 的矢量)，分组密码的模型如图 1.2.3 所示。它与流密码的不同之处在于输出的每一位数字不是只与相应时刻输入明文数字有关，而是与一组长为 m 的明文数字有关。分组密码有其自身的优点：首先分组密码容易被标准化，因为在今天的数字网络通信中，信息通常被成块地处理和传输；其次，使用分组密码容易实现同步，因为一个密文组的传输错误不会影响其他组，丢失一个明文组不会对其后组的解密正确性产生影响。分组密码的主要缺点表现在两方面：一是分组加密不能隐蔽数据模式，即相同的

密文组蕴含着相同的明文组；二是分组加密不能抵抗组的重放、嵌入和删除等攻击。但分组密码的上述缺陷可以通过在加密处理中引入少量的记忆来克服。

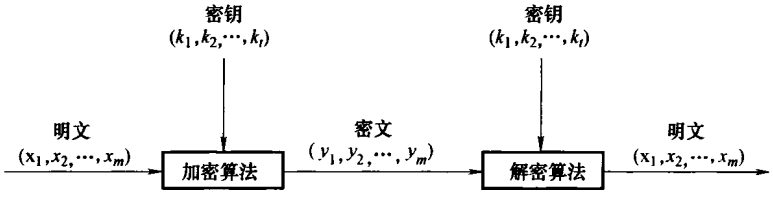


图 1.2.3　分组密码模型

若 $n>m$，则它为有数据扩展的分组密码，若 $n<m$，则它为有数据压缩的分组密码，若 $n=m$，则它为无数据扩展和压缩的分组密码。通常研究的是二元的情形，即明文 x 和密文 y 均为二元数字序列，它们的每个分量 x_i，$y_i\in\{0,1\}$。设 F_2 是一个二元域，F_2^s 表示 F_2 上的 s 维矢量空间。假定明文空间和密文空间均为 F_2^m，密钥空间为 S_K，S_K 是 F_2^t 的一个子集合。m 是明文组和密文组以 bit 形式出现的长度，称为分组长度；t 是密钥以 bit 形式出现的长度，称为密钥长度。

一个私钥分组密码可以定义如下：

定义 1.2.1　一个私钥分组密码是一种满足下列条件的映射 E：$F_2^m\times S_K\to F_2^m$，对每个 $k\in S_K$，$E(\cdot,k)$ 是从 F_2^m 到 F_2^m 的一个置换。将一个分组密码简记为 $Y=E(X,K)$。

通常称 $E(\cdot,k)$ 为密钥 k 下的加密函数，称 $E(\cdot,k)$ 的逆为密钥 k 下的解密函数，记为 $D(\cdot,k)$。分组密码的真正的密码规模被定义为 $l=\log_2|S_K|$ bit。因而，密钥长度等于真正的密码规模的当且仅当 $S_K=F_2^t$。

我们知道，F_2^m 上的置换共有 $2^m!$ 个。由定义 1.2.1 可知，一个分组密码是 F_2^m 上的全体置换所构成集合的一个子集合。可见，设计分组密码的问题在于找到一种算法，能在密钥控制下从一个足够大且足够“好”的置换子集合中简单而迅速地选出一个置换，用来对当前输入的明文数字组进行加密变换。一个好的分组密码应该是既难

破译又容易实现的，加密函数 $E(\cdot, k)$ 和解密函数 $D(\cdot, k)$ 都必须是很容易计算的，但是要从方程 $y=E(x, k)$ 和 $x=D(y, k)$ 中求出密钥 k 应该是一个困难问题。

下面介绍关于分组密码的工作模式。

直接使用分组密码算法的工作模式称为电码本(ECB)模式，前面我们已指出这种工作模式的缺陷，为了克服这些缺陷，我们不得不改变工作模式。所谓一个分组密码的“工作模式”就是以这个分组密码为基础用不同的方式构造一个分组密码系统。目前已提出许多种分组密码的工作模式，比如密码分组链接(CBC)模式、密码反馈(CFB)模式、输出反馈(OFB)模式、级连模式(CM，又称多重加密)、计算器模式、分组链接(BC)模式、扩散密码分组(PCBC)链接模式等。其中 CBC 模式、CFB 模式、OFB 模式、CM 是四种最主要和最基本的模式。

1. CBC 模式

如图 1.2.4 所示，在 CBC 模式下，每个明文组 x_i 加密之前，先与反馈至输入端的前一组密文 y_{i-1} 按位模 2 求和后，再送至加密器中加密，即 $y_i = E_k(x_i \oplus y_{i-1})$，其中 $y_0 = \mathrm{IV}$ 是一个初始矢量，无需保密，但需随消息更换，首发双方必须选用同一个 IV。

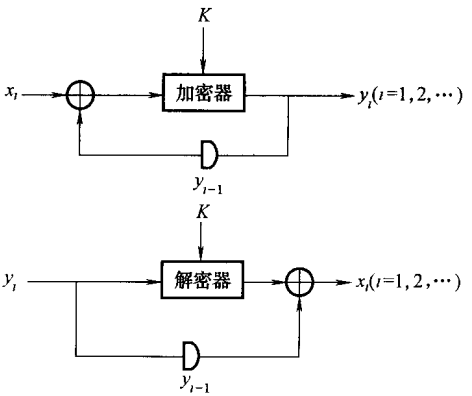


图 1.2.4 CBC 模式

显然各密文组 y_i 不仅与当前明文组 x_i 有关，而且通过反馈作用还与以前的明文组 x_1，x_2，…，x_{i-1} 有关。解密过程为 $x_i = D_k(y_i) \oplus y_{i-1}$。

CBC 模式的优点：一是能够隐蔽明文的数据模式，$x_{i+n} = x_i$ 未必蕴含着 $y_{i+n} = y_i$；二是在某种程度上能防止数据篡改，诸如组的重放、嵌入和删除等。CBC 模式的缺点是会出现错误传播，密文中任意一位发生变化会涉及后边一些组。CBC 模式的错误传播不大，一