第2章 密码学基础

罗文坚

主要内容

2.1 密码学基础知识

2.2 古典替换密码

2.3 对称密钥密码

2.4 公开密钥密码

2.5 消息认证

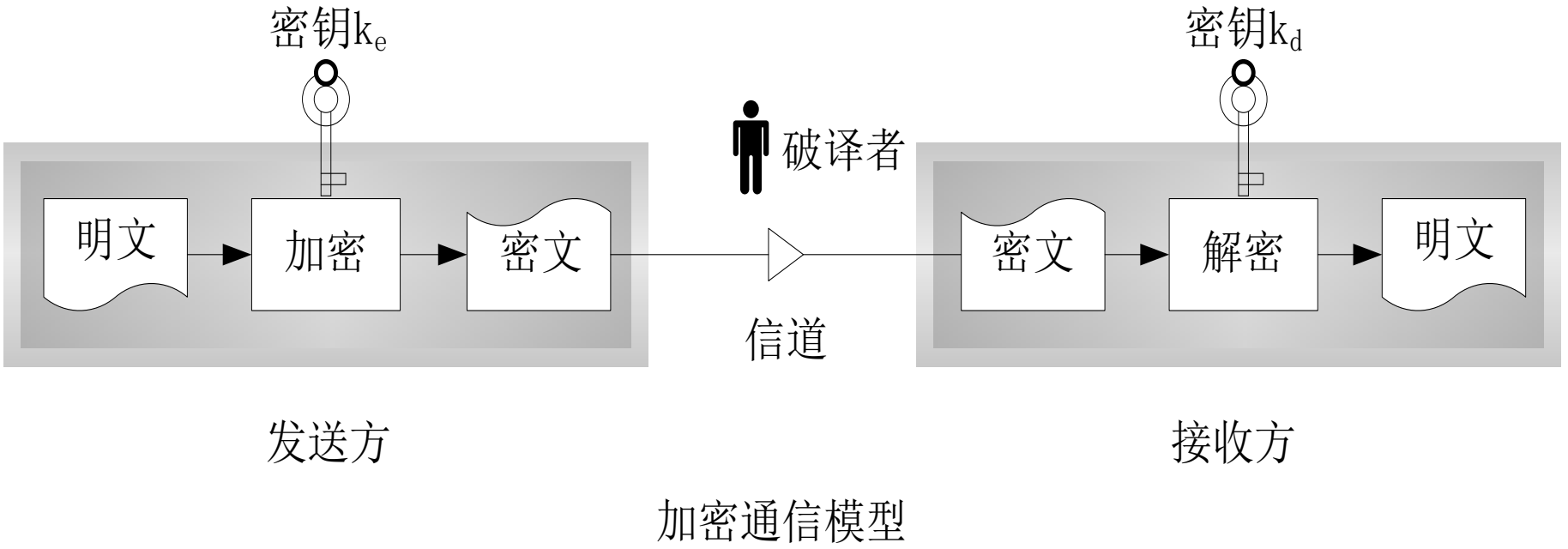
2.6 密码学新进展

引言

- 解决数据的机密性、完整性、不可否认性以及身份识别等问题均需要以密码为基础
  - 密码技术是保障信息安全的核心基础。
- 密码学(Cryptography)包括密码编码学和密码分析学两部分。
  - 将密码变化的客观规律应用于编制密码用来保守通信秘密的，称为密码编码学；
  - 研究密码变化客观规律中的固有缺陷，并应用于破译密码以获取通信情报的，称为密码分析学。
- 历史
  - 宋代的曾公亮、丁度等编撰《武经总要》
  - 1863年，卡西斯基所著的《密码和破译技术》
  - 1949年，香农发表了《秘密体制的通信理论》

密码体制

- 消息在密码学中被称为明文(Plain Text)。
- 伪装消息以隐藏它的内容的过程称为加密(Encrypt)。
- 被加密的消息称为密文(Cipher Text)。
- 把密文转变为明文的过程称为解密(Decrypt)。



加密通信模型

密码体制

- 完整密码体制要包括如下五个要素：
  - M是可能明文的有限集称为明文空间；
  - C是可能密文的有限集称为密文空间；
  - K是一切可能密钥构成的有限集称为密钥空间；
  - E为加密算法，对于任一密钥， 都能够有效地计算；
  - D为解密算法，对于任一密钥，都能够有效地计算。

- 密码体系必须满足如下特性：
  1. 加密算法(E_k：M->C)和解密算法(D_k：C->M)满足：
     - $D_k(E_k(x))=x$，这里$x \epsilon M$；
  2. 破译者不能在有效的时间内破解出密钥k或明文x。

密码学的发展阶段

密码学的发展历程大致经历了三个阶段：

- 古代加密方法（手工阶段）
  - 例如，塞塔式密码，藏头诗等。
- 古典密码（机械阶段）
  - 文字替换，使用手工或机械变换的方式实现。
- 近代密码（计算机阶段）
  - 与计算机技术、电子通信技术紧密相关。
  - 摆脱了原先用铅笔和纸进行手工设计时易犯的错误。
  - 摆脱了用电子机械方式实现的密码机的高额费用。

密码的分类

- 依据密码体制的特点以及出现的时间分类：

1. 古典替换密码
   - ➢文字替换，使用手工或机械变换的方式实现。
   - ➢例如，单表代替密码、多表代替密码以及转轮密码。
2. 对称密钥密码
   - ➢加密过程和解密过程使用同一密钥来完成。
   - ➢又称为秘密密钥密码，或单密钥密码。
   - ➢分为分组密码和序列密码。
3. 公开密钥密码
   - ➢加密过程和解密过程使用两个不同的密钥来完成。
   - ➢又称为非对称密钥密码，双密钥密码。

密码的分类

- 依据处理数据的类型：
  1. 分组密码(block cipher)
     - 将定长的明文块转换成等长的密文，这一过程是在密钥控制下完成的。
     - 对于大部分分组密码，分组大小是64位；以后会增加。
     - 又称为分块密码或者块密码。
  2. 序列密码(stream cipher)
     - 加解密时一次处理明文中的一个或几个比特。
     - 又称为流密码。

- 非对称密码体制都是分组密码。

密码分析

- 密码分析也称为密码攻击。密码分析攻击主要包括：
  - 唯密文攻击：有一些消息的密文。
  - 已知明文攻击：有一些消息的密文以及对应的明文。
  - 选择明文攻击：不仅有一些消息的密文及对应的明文，而且可选择被加密的明文。
  - 自适应选择明文攻击：选择明文攻击的特殊情况。不仅能选择被加密的明文，而且也能基于以前加密的结果修正这个选择。
  - 选择密文攻击：能选择不同的被加密的密文，并可得到对应的解密的明文。例如，得到了一个防篡改的自动解密盒，但不知道密钥。
  - 选择密钥攻击：并不表示密码分析者能够选择密钥，只表示密码分析者具有不同密钥之间的关系的有关知识。不是很实际，但有时很有效。

主要内容

2.1 密码学基础知识

2.2 古典替换密码

2.3 对称密钥密码

2.4 公开密钥密码

2.5 消息认证

2.6 密码学新进展

简单代替密码

- 简单代替密码
  - 指将明文字母表M中的每个字母用密文字母表C中的相应字母来代替。
  - 例如：移位密码、乘数密码、仿射密码等。

- 移位密码
  - 具体算法是将字母表的字母右移k个位置，并对字母表长度作模运算。
  - 每一个字母具有两个属性，本身代表的含义，可计算的位置序列值。
  - 加密函数：$E_k(m) = (m + k) \bmod q$;
  - 解密函数：$D_k(c) = (c - k) \bmod q$;

移位密码举例

- 凯撒Caesar密码是一种移位密码。

- 本例采用英文字母表，则凯撒密码体系的数学表示为：
  - M=C={有序字母表}，q = 26，k = 3。
    - 其中q 为有序字母表的元素个数，即q = 26。
  - 使用凯撒密码对明文字符串逐位加密，结果如下：
    - 明文信息M = meet me after the toga party
    - 密文信息C = phhw ph diwho wkh wrjd sduwb

乘数密码

- 将明文字母串逐位乘以密钥k并进行模运算。

- 数学表达式：$E_k(m)=k * m \bmod q$, $\gcd(k, q) = 1$。
  - gcd(k, q)=1表示k与q的最大公因子为1。

- 算法描述：
  - M=C=Z/(26)，明文空间和密文空间同为英文字母表空间，包含26个元素；q=26；
  - K={k∈整数集 | 0<k<26，gcd(k, 26)=1}，密钥为大于0小于26，与26互素的正整数；
  - $E_k(m) = k * m \bmod q$。
  - $D_k^{-1}(c)=k^{-1} * c \bmod q$，其中$k^{-1}$为k在模q下的乘法逆元。

乘数密码：密钥取值与乘法逆元

- 乘数密码的密钥k与26互素时，加密变换才是一一映射的。
  - k的选择有11种：3、5、7、9、11、15、17、19、21、23、25。
  - k取1时没有意义。

- k^{-1}为k在模q下的乘法逆元。
  - 其定义为$k^{-1} * k \bmod q = 1$，
  - 可采用扩展的欧几里德算法。欧几里德算法又称辗转相除法，用于计算两个整数a和b的最大公约数。

仿射密码

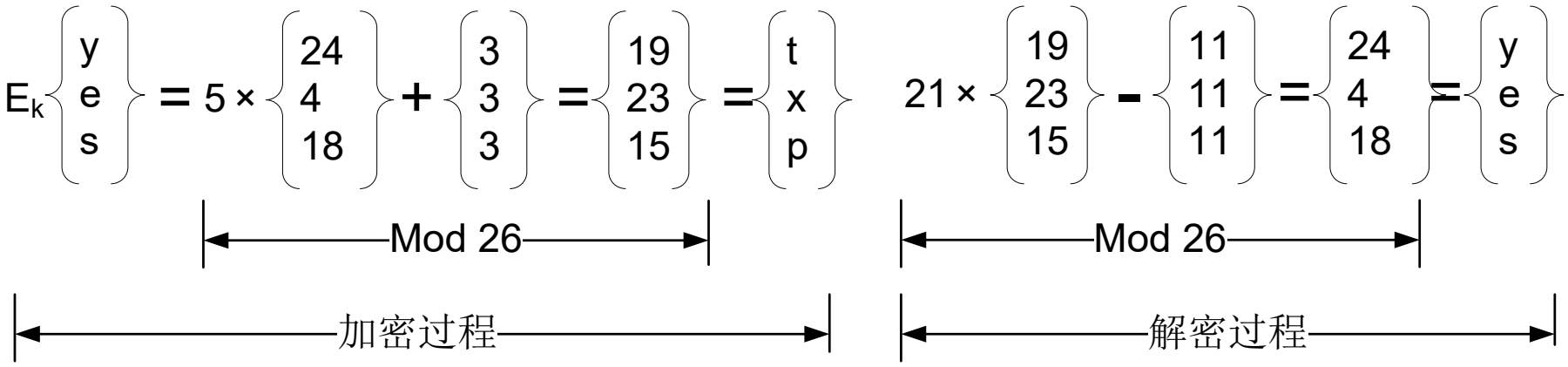
- 仿射密码可以看作是移位密码和乘数密码的结合。

- 密码体制描述如下：
  - M=C=Z/(26)； q=26；
  - K={ $k_1, k_2 \in Z \mid 0< k_1, k_2<26$， $gcd(k_1, 26)=1$}；
  - $E_k(m)=(k_1 m + k_2) \bmod q$；
  - $D_k(c)= k_1^{-1}(c - k_2) \bmod q$，其中$k_1^{-1}$为$k_1$在模q下的乘法逆元。

- 密钥情况： k_1和k_2？

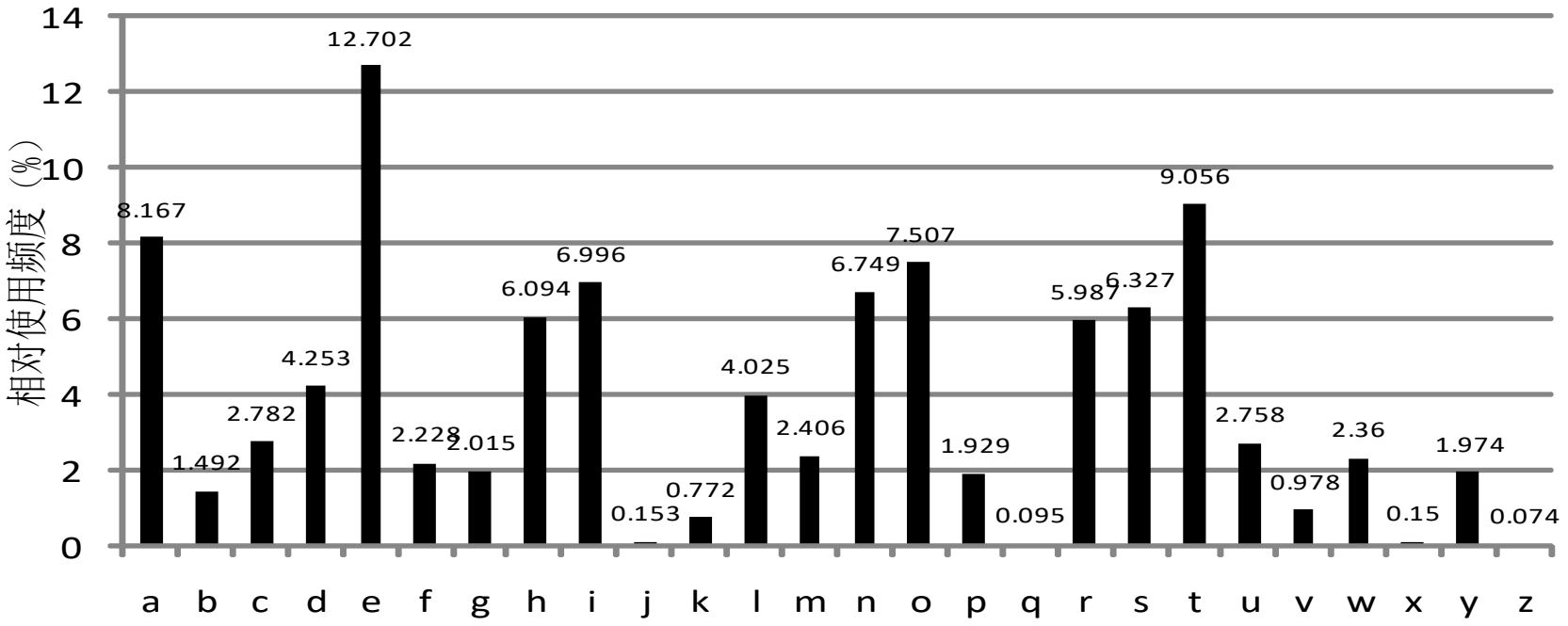
仿射密码举例

- 设k＝（5，3），注意到5^{-1} mod 26=21。
- 加密函数：
  - $E_k(x)=5x+3 \pmod{26}$，
- 解密函数：
  - $D_k(y)=21(y-3) \bmod 26 = 21y - 11 \pmod{26}$。
- 加密明文“yes”的加密与解密过程如下：



基于统计的密码分析

- 简单代替密码的加密是从明文字母到密文字母的一一映射。
- 攻击者统计密文中字母的使用频度，比较正常英文字母的使用频度，进行匹配分析。
- 如果密文信息足够长，很容易对单表代替密码进行破译。



多表代替密码

- 多表代替密码是以一系列代替表依次对明文消息的字母进行代替的加密方法。
- 多表代替密码使用从明文字母到密文字母的多个映射来隐藏单字母出现的频率分布。
  - 每个映射是简单代替密码中的一对一映射。

- 若映射系列是非周期的无限序列，则相应的密码称为非周期多表代替密码。
  - 非周期多表代替密码：对每个明文字母都采用不同的代替表（或密钥）进行加密，称作一次一密密码。
  - 这是一种理论上唯一不可破译的密码。但需要的密钥量和明文消息长度相同，难于广泛使用。

维吉尼亚Vigenère密码

- 经典的多表代换密码有：
  - Vigenère、Beaufort、Running Key、Vernam和轮转机等密码。

- 维吉尼亚Vigenère密码
  - 是以移位代替为基础的周期多表代替密码。
  - 加密时每一个密钥被用来加密一个明文字母，当所有密钥使用完后，密钥又重新循环使用。

- 维吉尼亚Vigenère密码算法如下：
  - $E_K(m)=C_1 C_2 \ldots C_n$，其中$C_i=(m_i + k_i) \bmod 26$；
  - 密钥K可以通过周期性反复使用以至无穷。

主要内容

2.1 密码学基础知识

2.2 古典替换密码

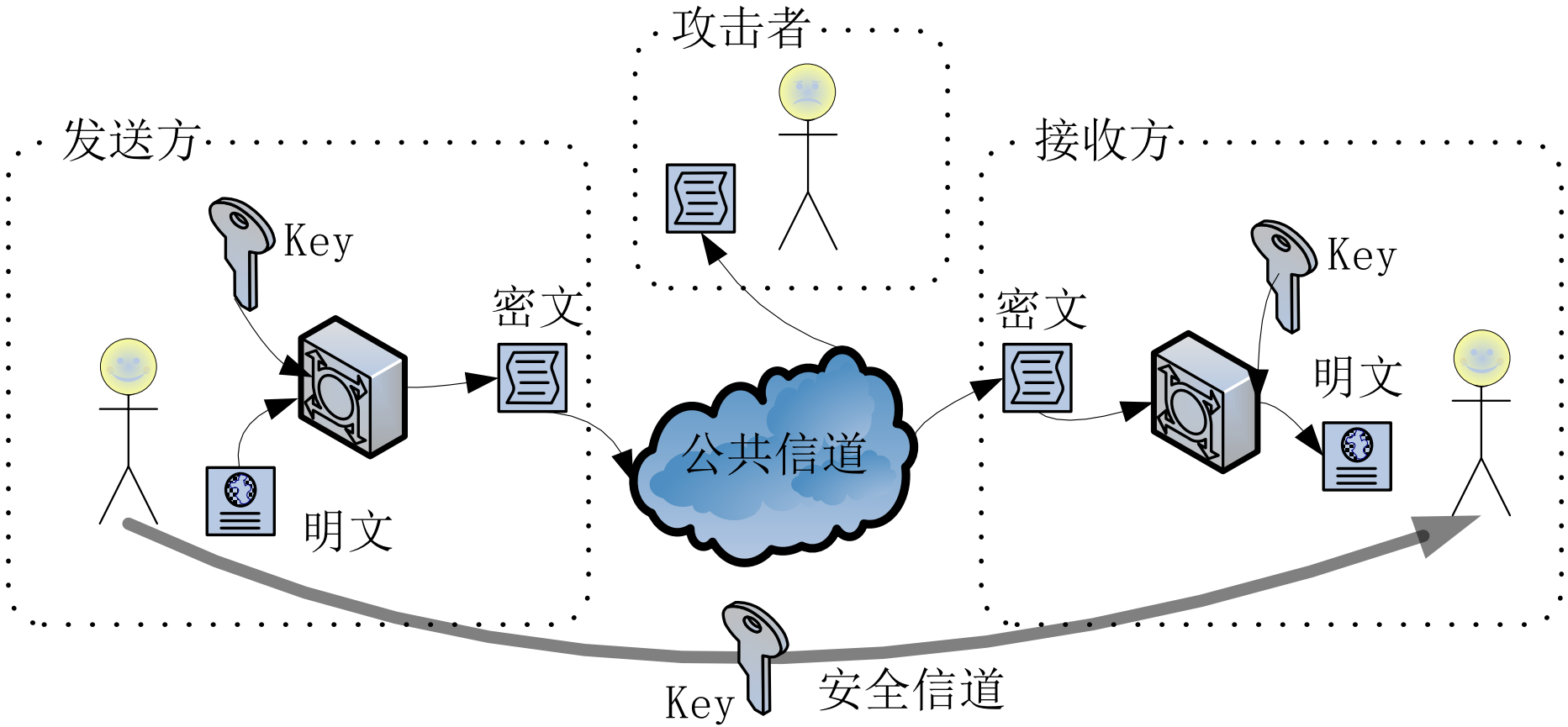
2.3 对称密钥密码

2.4 公开密钥密码

2.5 消息认证

2.6 密码学新进展

对称密钥密码

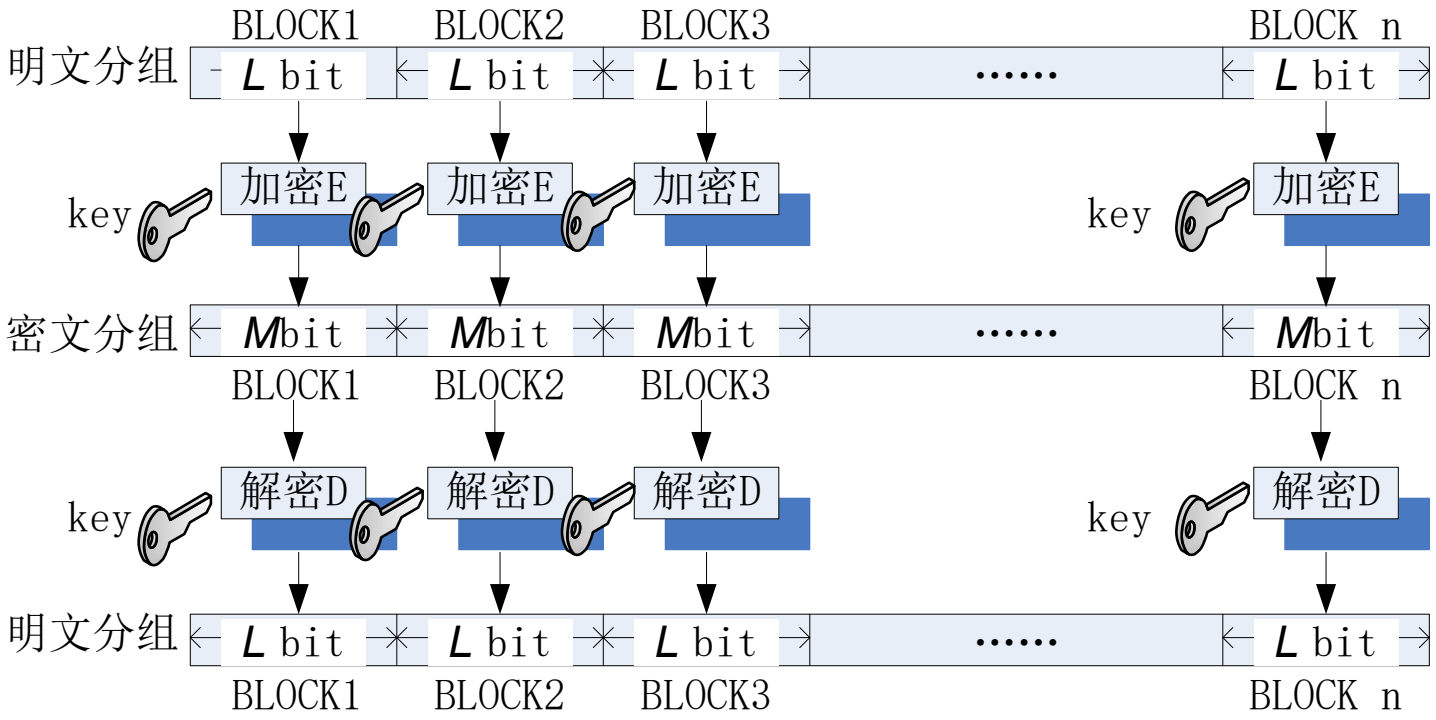


对称密钥密码的模型

- 对称密钥密码的通讯安全性取决于密钥的机密性，与算法本身无关，算法是公开的。

对称密钥密码加密模式

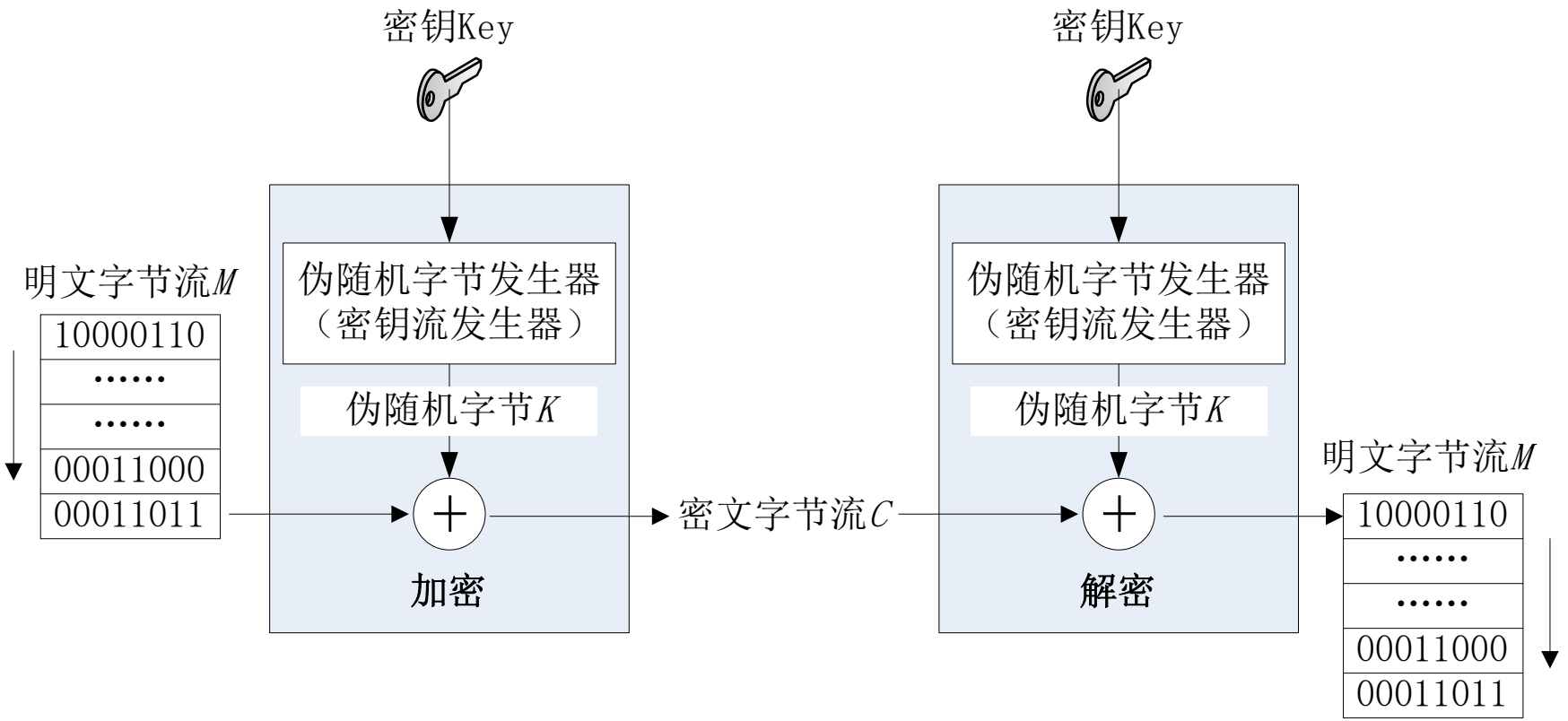
- 对称密码加密系统从工作方式上可分为：
  - 分组密码、序列密码
- 分组密码原理：
  - 明文消息分成若干固定长度的组，进行加密；解密亦然。



分组密码工作原理示意图

序列密码（流密码）

- 通过伪随机数发生器产生性能优良的伪随机序列（密钥流），用该序列加密明文消息流，得到密文序列；解密亦然。



序列密码工作原理示意图

数据加密标准DES

- 1973年，美国国家标准局NBS公开征集国家密码标准方案，要求：
  - ① 算法必须提供高度的安全性；
  - ② 算法必须有详细的说明，并易于理解；
  - ③ 算法的安全性取决于密钥，不依赖于算法；
  - ④ 算法适用于所有用户；
  - ⑤ 算法适用于不同应用场合；
  - ⑥ 算法必须高效、经济；
  - ⑦ 算法必须能被证实有效；
  - ⑧ 算法必须是可出口的。
- 1974年，NBS开始第二次征集时，IBM公司提交了算法LUCIFER。1977年，LUCIFER被美国国家标准局NBS作为“数据加密标准FIPS PUB 46”发布，简称为DES。

S-DES加密算法

- S-DES： Simplified DES，简化的DES。
- S-DES是由美国圣达卡拉大学的Edward Schaeffer教授提出的，主要用于教学，其设计思想和性质与DES一致，有关函数变换相对简化，具体参数要小得多。

- 输入为一个8位的二进制明文组和一个10位的二进制密钥，输出为8位二进制密文组。

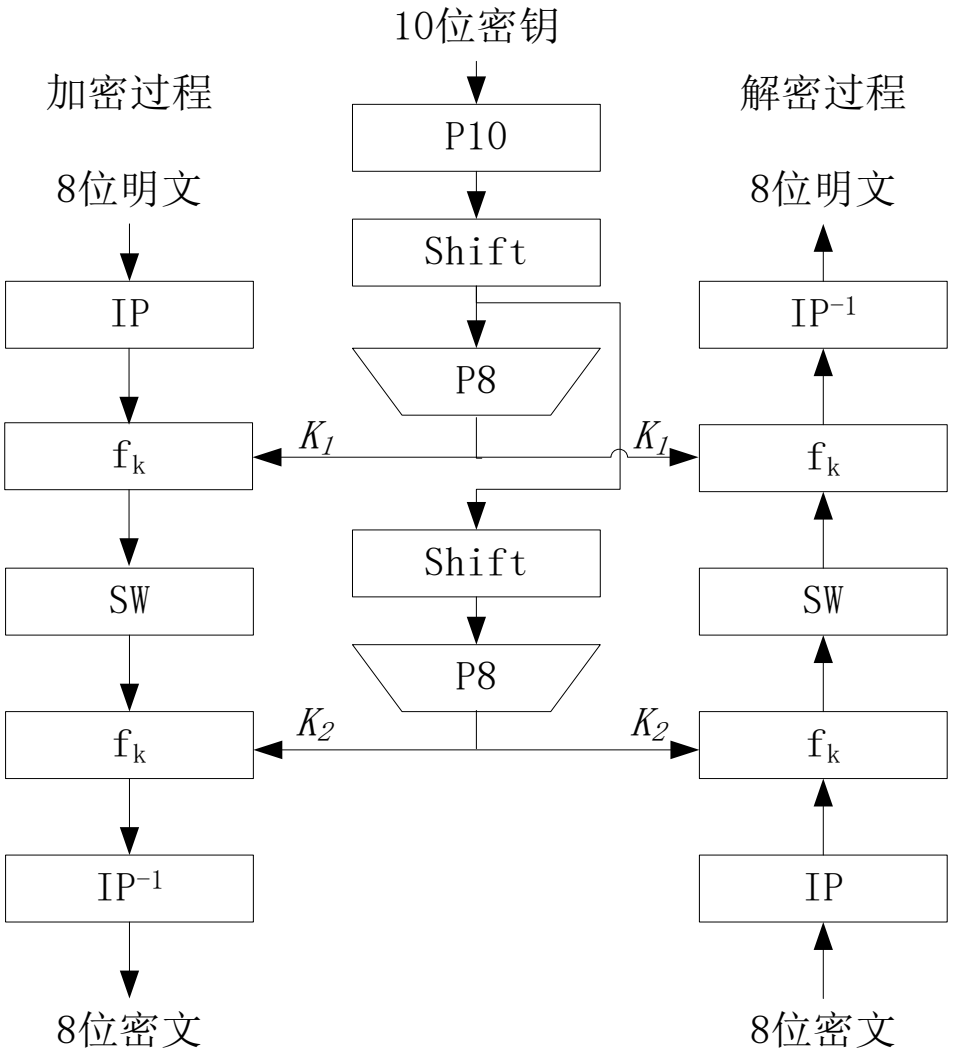
- 涉及的主要函数：
  - 与密钥变换有关的两个置换函数：P8、P10；与密钥变换有关的循环移位函数Shift。
  - 用于数据加密变换的4个基本函数：初始置换IP、复合函数f_k、转换函数SW、末尾置换IP^{-1}。

S-DES加密算法

- 解密与加密基本一致。

- 加密:
  - $IP^{-1}(f_{k2}(SW(f_{k1}(IP(明文)))))$

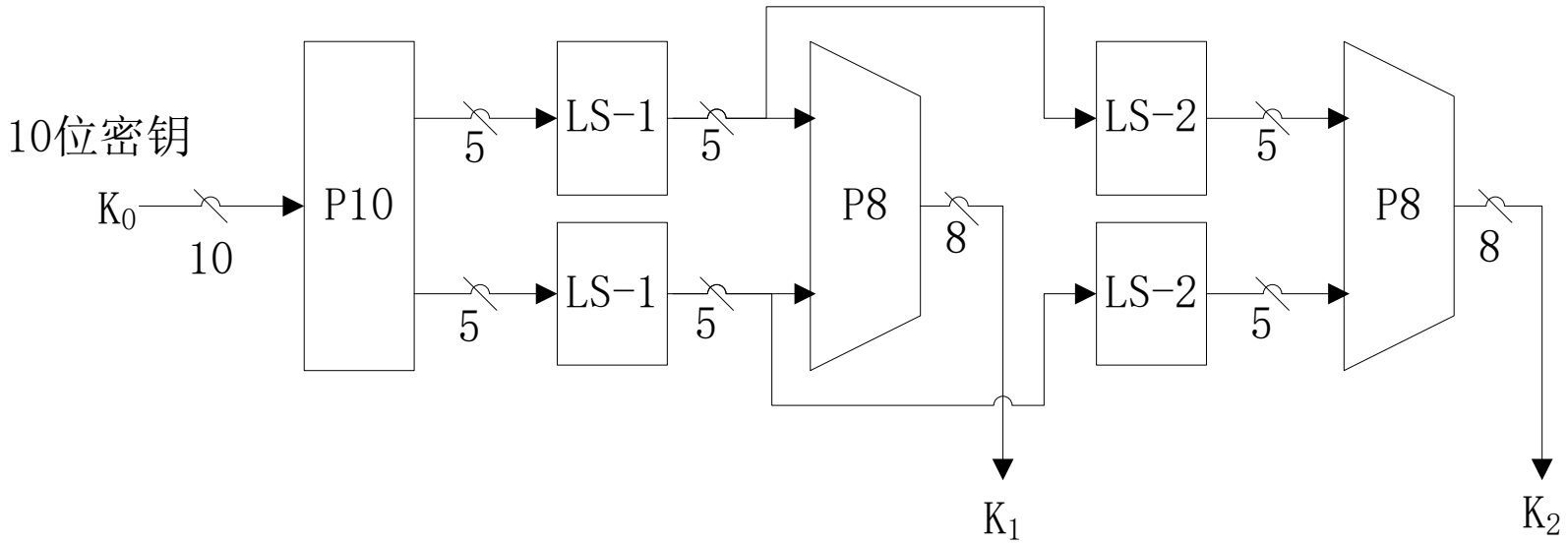
- 解密:
  - $IP^{-1}(f_{k1}(SW(f_{k2}(IP(密文)))))$



S-DES的体制

S-DES的密钥产生

- **P10=(3,5,2,7,4,10,1,9,8,6)**
- **循环左移函数LS**
- **P8=(6,3,7,4,8,5,10,9)**



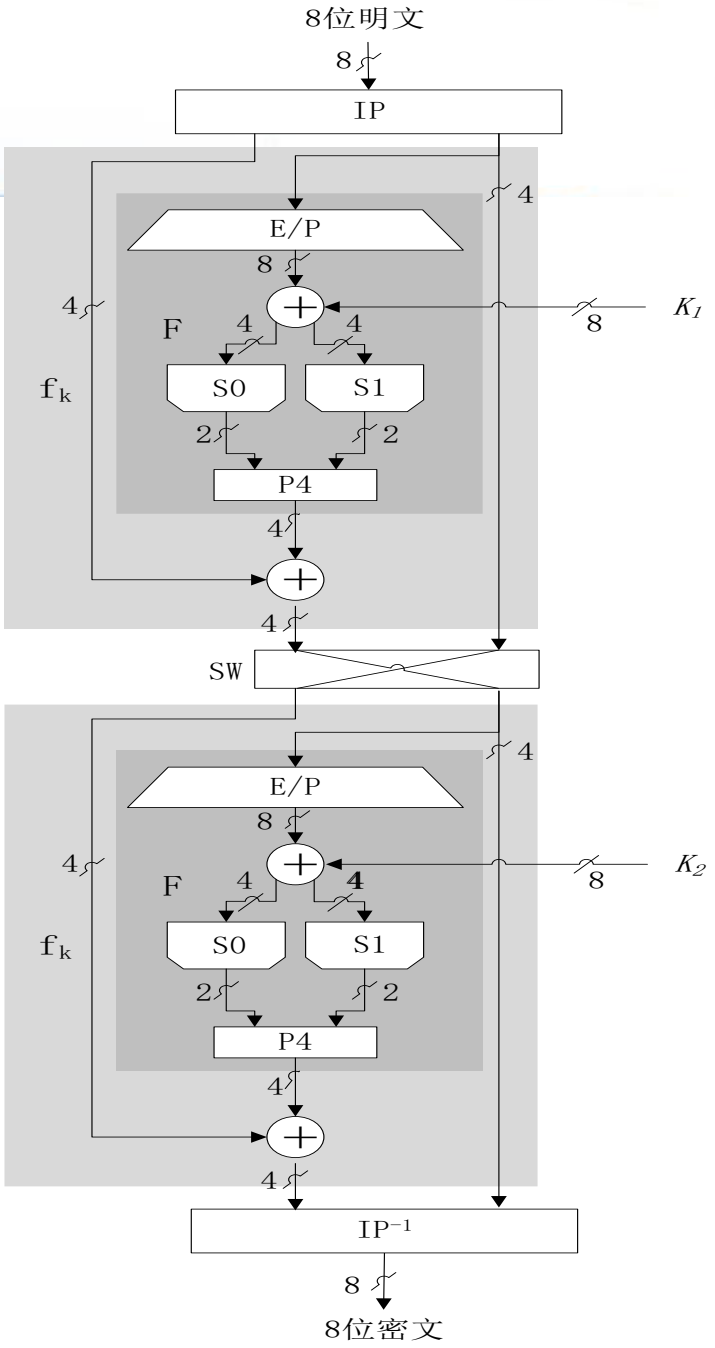
S-DES的密钥产生

S-DES的加密变换过程

- IP=(2,6,3,1,4,8,5,7)
- IP^{-1}=(4,1,3,5,7,2,8,6)

- E/P=（4,1,2,3,2,3,4,1）
- “⊕”：按位异或运算；
- S盒函数
  - S0和S1为两个盒子函数，将输入作为索引查表，得到相应的系数作为输出。

- P4=（2,4,3,1）
- SW：将左4位和右4位交换。



S-DES的加密过程

S盒函数

S0=

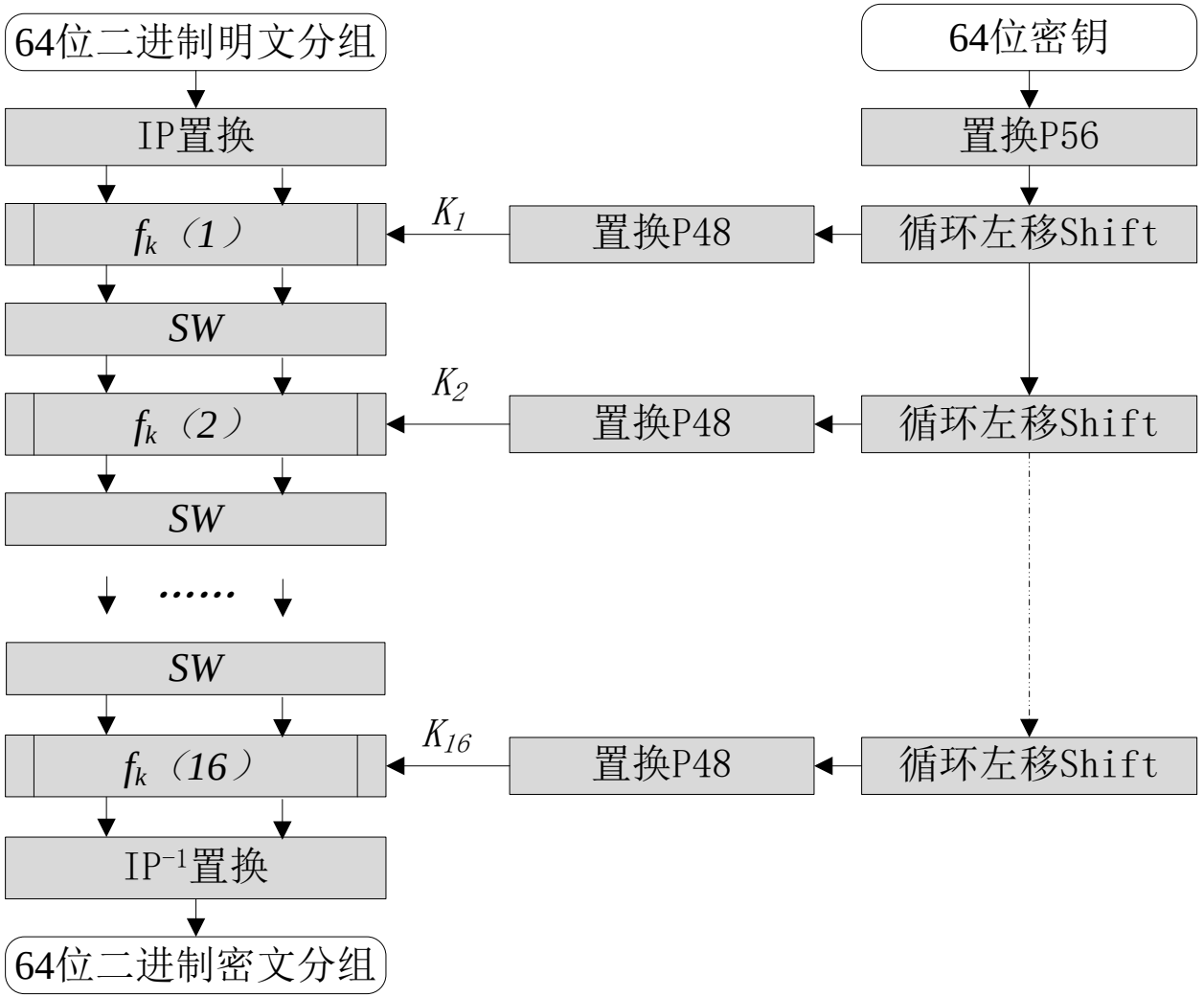
	0	1	2	3
0	1	0	3	2
1	3	2	1	0
2	0	2	1	3
3	3	1	3	2

S1=

	0	1	2	3
0	0	1	2	3
1	2	0	1	3
2	3	0	1	0
3	2	1	0	3

- S盒函数按下述规则运算：
  - 输入的第1位和第4位二进制数合并为一个两位二进制数，作为S盒的行号索引*i*；
  - 将第2位和第3位同样合并为一个两位二进制数，作为S盒的列号索引*j*；
  - 确定S盒矩阵中的一个系数（*i*，*j*）。
  - 此系数以两位二进制数形式作为S盒的输出。
  - 例如：
    - $L'=(l_0, l_1, l_2, l_3)=(0, 1, 0, 0)$，　$(i, j)=(0, 2)$
    - 在S0中确定系数3，则S0的输出为11B。

DES算法



DES算法框图

DES的安全问题

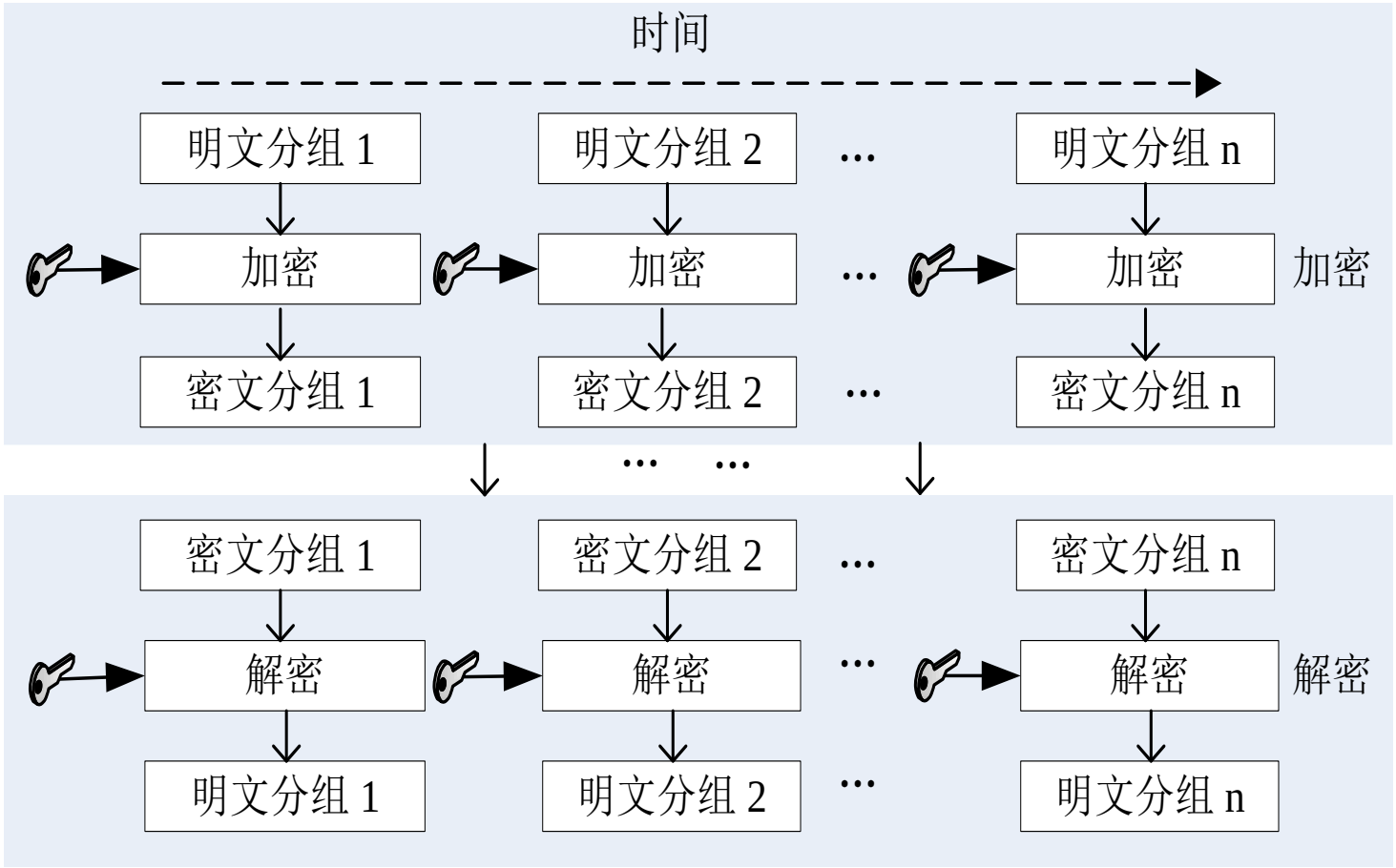
- 1977年，人们估计要耗资两千万美元建成一个专门计算机用于DES的破译，而且需要12个小时的破解才能得到结果。
- 1994年世界密码大会，M. Matsui提出线性分析方法，利用243个已知明文，成功破译DES。
- 1997年，首届“向DES挑战”的竞技赛。罗克·维瑟用了96天时间破解了用DES加密的一段信息。
- 2000年1月19日，电子边疆基金会组织25万美元的DES解密机，以22.5小时成功破解DES加密算法。

- DES的最近一次评估是在1994年，同时决定1998年12月以后，DES将不再作为联邦加密标准。

- DES的密钥长度仅有56bit（64bit中有8位用于奇偶校验）。

分组密码的工作模式

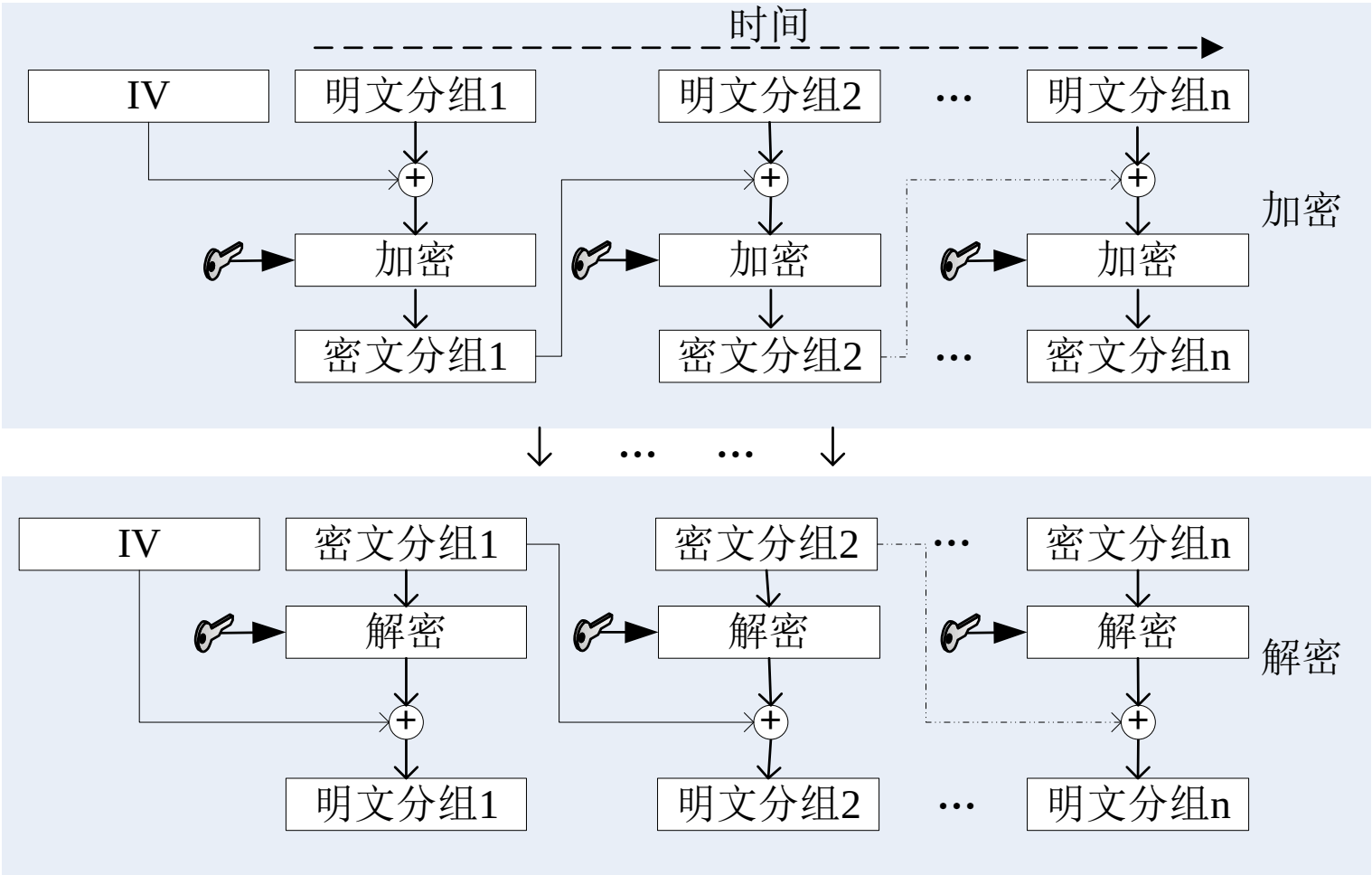
- 电子编码本模式（Electronic Code Book，ECB）
  - 明文分组，分块加密；不是m位的整数倍则填充规定字符。



ECB工作模式

密码分组链接模式（CBC）

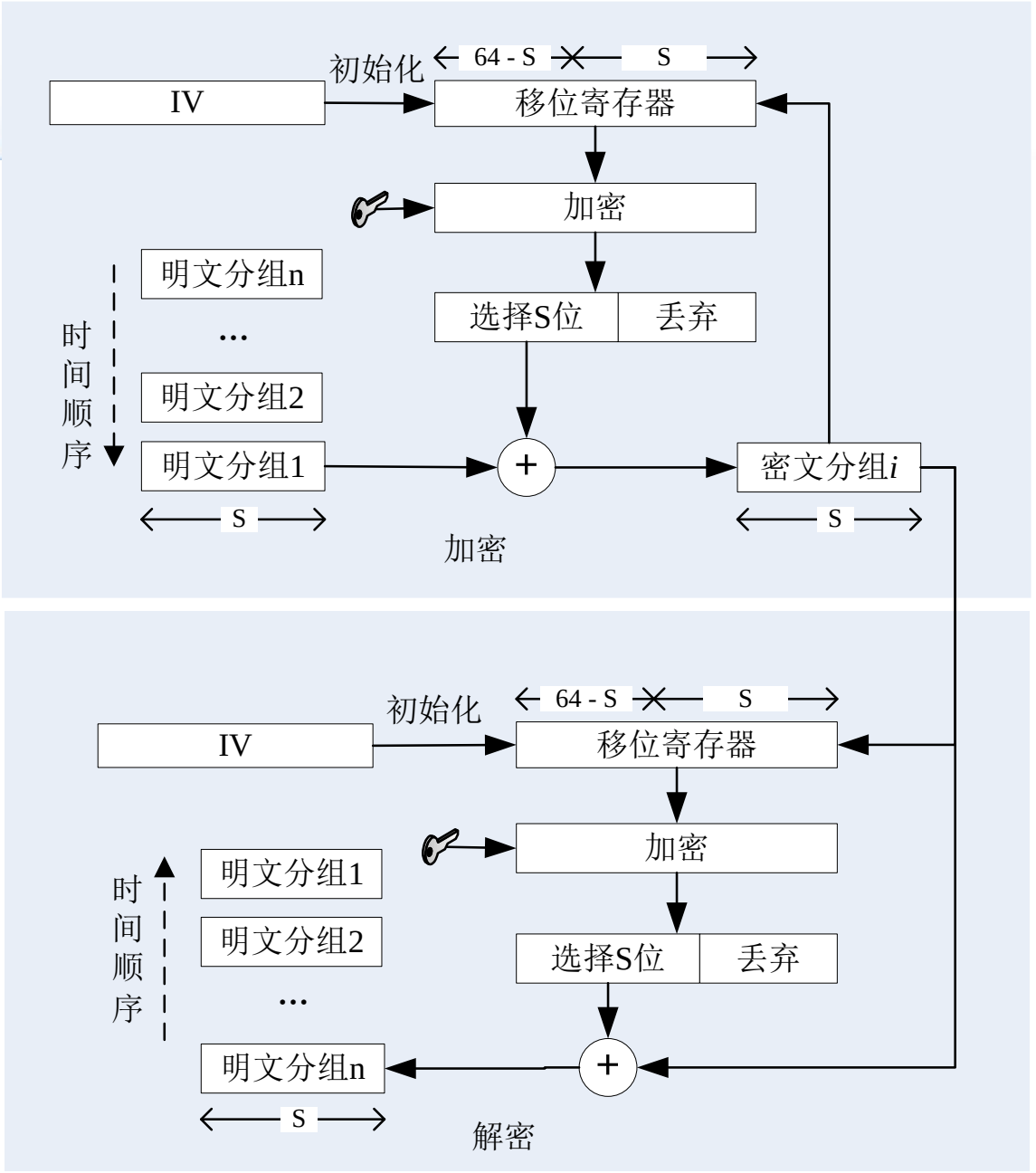
- **CBC：Cipher Block Chaining；IV：Initial Vector。**



CBC工作模式

密码反馈模式（CFB）

Cipher Feedback



CFB 工作模式

输出反馈模式（OFB）

Output Feedback

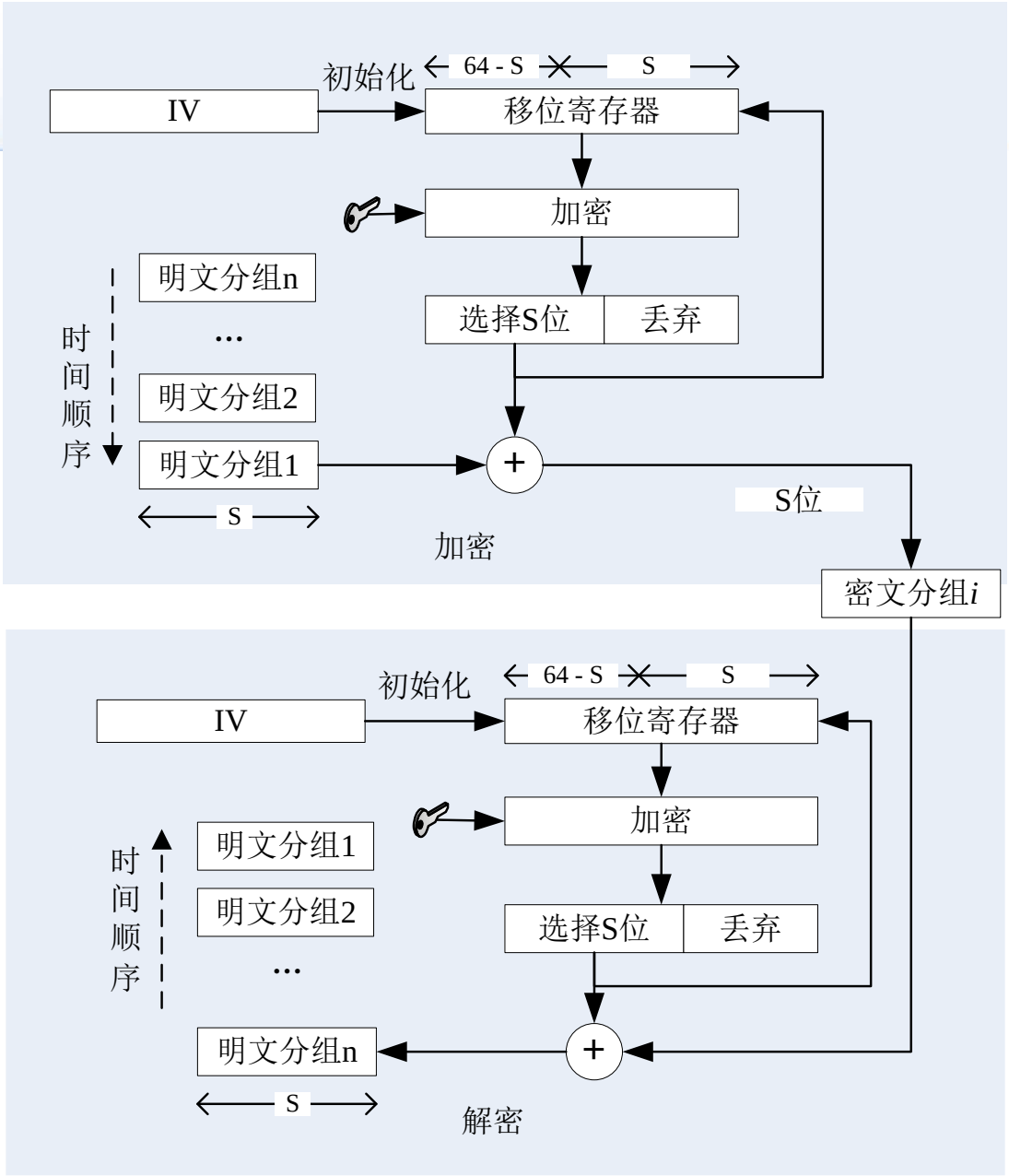


图 2.14 OFB 工作模式

其他对称密码简介

- 为提高安全性，主要有两种研究思路。
  - 对DES进行复合变换，强化它的抗攻击能力；
  - 开辟新的算法。

- 三重DES
- RC5
- IDEA
- AES算法

作业

- 阅读教材2.3.4。

1. 习题2（1）：密码体制五要素是什么？
2. 习题3（1）：俄语共有32个字母，设计一个乘数密码来加密俄语信息，并计算一下潜在的加密密钥有多少个，并列举。
3. 乘数密码中，当gcd(k, q)=1时，加密变换才是一一映射的。试证明之。
4. 乘数密码中，如何计算k^{-1}？此处k^{-1}为k在模q下的乘法逆元？请给出算法伪代码。
5. 给定密钥“11 1111 1111”，明文“00000000”，计算S-DES的密文。请按给出主要计算过程。