21世纪高等学校计算机规划教材

现代密码学

Modern Cryptography

作　者：何大可 彭代渊 唐小虎 何明星 梅其祥

出版社：人民邮电出版社

现代密码学

Modern Cryptography

第2章 流密码

孙玉花

中国石油大学 理学院

Sunyuhua_1@163.com

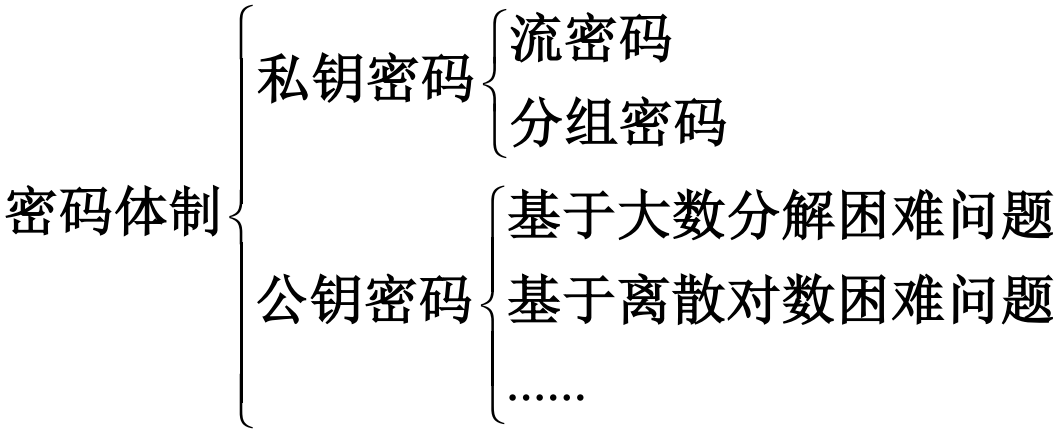
2019年9月

第2章 流密码

- 2.1 流密码一般模型
- 2.2 线性反馈移位寄存器序列
- 2.3 线性复杂度及B-M算法
- 2.4 非线性序列生成器
- 2.5 流密码算法

2.1 流密码一般模型

- **实用密码体制的分类**



2.1 流密码一般模型

- 流密码(stream cipher)(序列密码)体制模型

明文序列: $m= m_1\, m_2\, m_3\, \ldots;$

密钥序列: $z=z_1\, z_2\, z_3\, \ldots;$

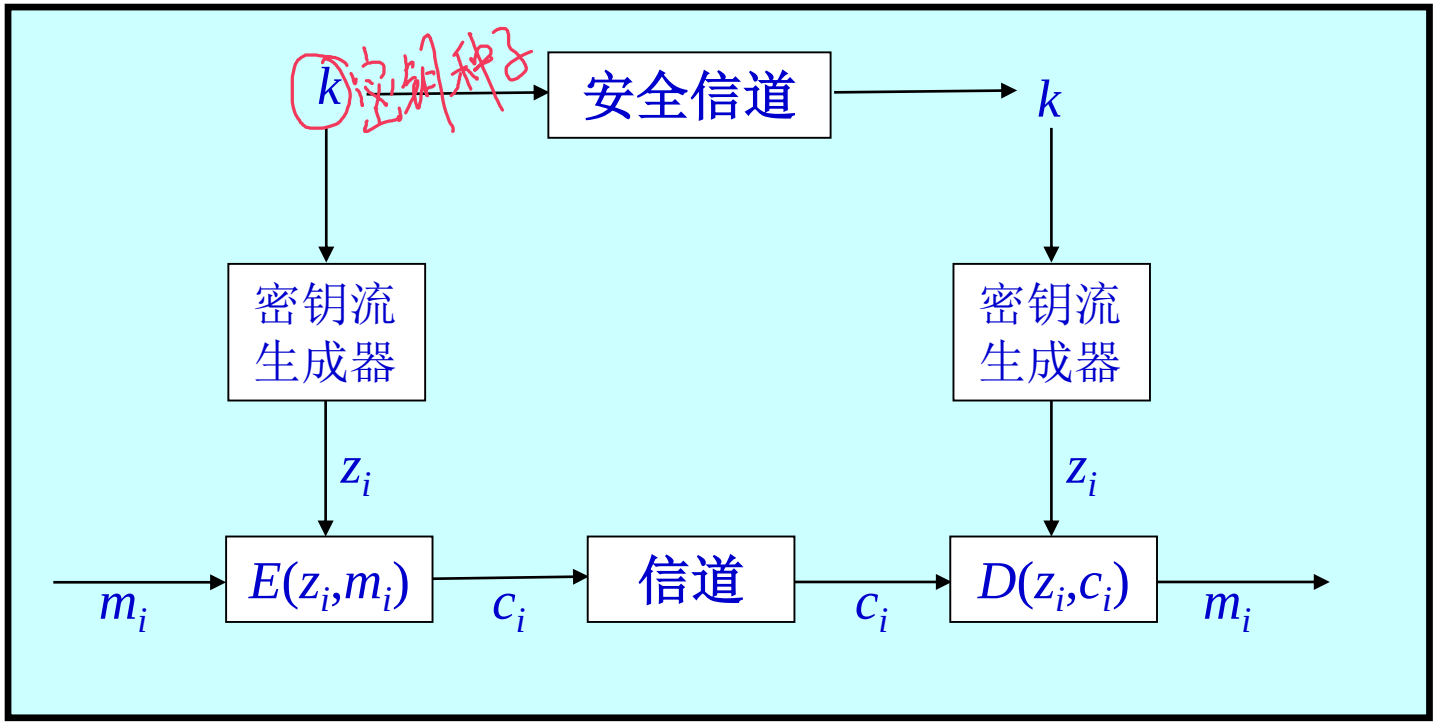
密文序列: $c= c_1 c_2\, c_3\, \ldots;$

加密变换: $c_i=E(z_i,m_i)\ \ (i=1,2,3,\ldots);$

解密变换: $m_i=D(z_i,\, c_i)\ \ (i=1,2,3,\ldots).$

2.1 流密码一般模型

- 流密码原理框图



2.1 流密码一般模型

- **流密码体制的安全性**

  当流钥序列是具有均匀分布的离散无记忆随机序列时, 在理论上是不可破译的.

- **实用的困难性**

  真正的具有均匀分布的随机序列是不可能重复产生的.

  密钥序列长(至少与明文序列一样长), 其管理(存储、分配)难.

- **设计流密码体制的关键问题**

  设计产生密钥序列的方法.

2.1 流密码一般模型

- 流密码的分类
  - 同步流密码(SSC: synchronous stream cipher)

    产生密钥序列的算法与明文、密文无关.

$$\sigma_i = F(\sigma_{i-1}, k),$$
$$z_i = f(\sigma_i, k),$$
$$c_i = E(z_i, m_i).$$

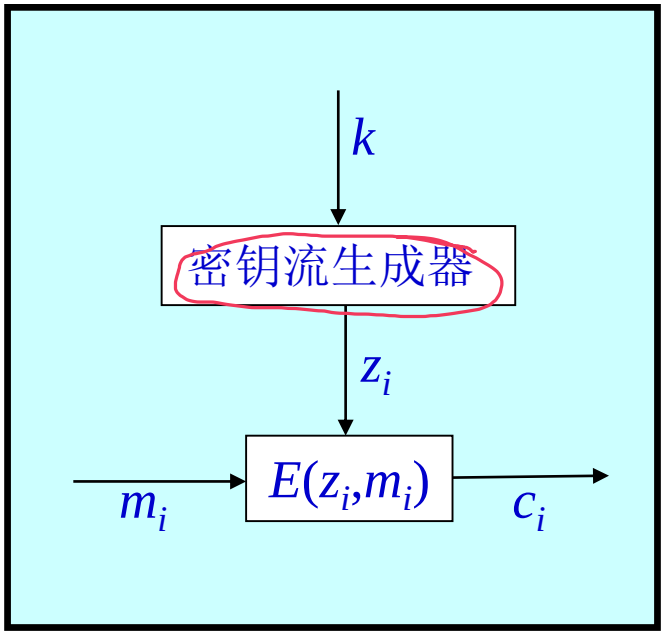
σ_i : 密钥流生成器的内部状态

σ_0 : 密钥流生成器的初始状态

k : 种子(初始)密钥

F : 状态转移函数

f : 密钥流生成函数



2.1 流密码一般模型

- 序列密码的分类
  - 同步流密码(SSC: synchronous stream cipher)
    - 只要通信双方的密钥序列产生器具有相同的“种子序列”和相同的“初始状态”,就能产生相同的密钥序列.
    - 通信双方必须保持精确同步,才能正确解密.
    - 容易检测插入、删除、重播等主动攻击.
    - 没有差错传播.

2.1 流密码一般模型

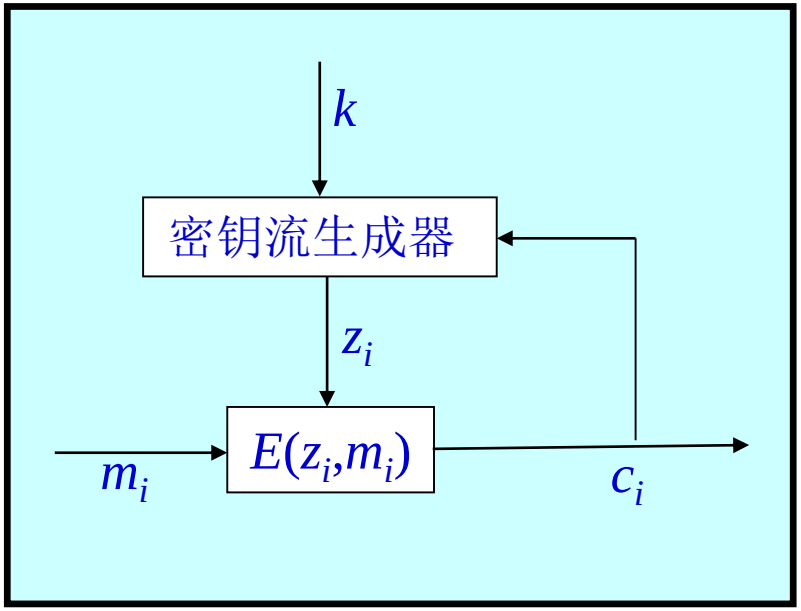
- 流密码的分类
  - 自同步流密码(SSSC: self-synchronous stream cipher)

    产生密钥序列的算法与以前的秘文有关.

$$\sigma_i = F(\sigma_{i-1}, c_{i-1}, \ldots, c_{i-l}, k)$$
$$= F(\sigma_{i-1}, m_{i-1}, \ldots, m_{i-l}, k)$$
$$z_i = f(\sigma_i, k),$$
$$c_i = E(z_i, m_i).$$

F：状态转移函数

f：密钥流生成函数



2.1 流密码一般模型

- 自同步流密码(SSSC)
  - 密钥流生成器是一种有记忆变换器
  - 密钥流与明文符号有关：

    i 时刻的密文不仅取决于i 时刻的明文，而且与i 时刻之前的l个明文符号有关
  - 具有有限的差错传播
  - 具有自同步能力
  - 把明文每个字符扩散在密文多个字符中,强化了抗统计分析的能力

2.1 流密码一般模型

- 二元加法序列密码

明文序列: $m = m_1\, m_2\, m_3 \ldots;$

密钥序列: $z = z_1\, z_2\, z_3 \ldots;$

密文序列: $c = c_1\, c_2\, c_3 \ldots;$

加密变换: $c_i = z_i \oplus m_i \quad (i = 1,2,3,\ldots);$

解密变换: $m_i = z_i \oplus c_i \quad (i = 1,2,3,\ldots).$

第2章 流密码

- 2.1 流密码一般模型
- 2.2 线性反馈移位寄存器序列
- 2.3 线性复杂度及B-M算法
- 2.4 非线性序列生成器
- 2.5 流密码算法

2.2 线性反馈移位寄存器序列

- 伪随机序列

  考虑二元序列: $a=\{a_i\}=a_0a_1a_2a_3\ldots.$

  - 周期序列

定义2.1 设$a=(a_0, a_1,\ldots,a_i,\ldots)$是一个二元序列，若存在正整数$N$和非负整数$m$，使得$a_{i+N}=a_i$对于任意$i\geq m$成立，则称二元序列$a$是终归周期序列。如果$m=0$，则称序列$a$是严格周期序列，简称周期序列。而满足$a_{i+N}=a_i\ (i\geq m)$的最小正整数$N$被称为序列$a$的周期。

2.2 线性反馈移位寄存器序列

- 伪随机序列
  - 序列的游程

定义2.2 设$a=(a_0, a_1,\ldots,a_i,\ldots)$是一个周期为$N$的二元序列，在一个周期内连续出现的最多的符号“0”（或1）的串，称为0（或1）的一个游程。在一个游程中，0（或1）的个数称为该游程的长度。

例：在序列 $k=\{k_i\}$=00111010000111100中，有

长为1的0游程一个；长为4的0游程一个；长为5的0游程一个；长为1的1游程一个；长为3的1游程一个；长为4的1游程一个.

2.2 线性反馈移位寄存器序列

- 伪随机序列
  - 序列的相关函数

定义2.3 设$a=(a_0,a_1,\ldots,a_{N-1})$和$b=(b_0,b_1,\ldots,b_{N-1})$是两个周期为$N$的二元周期序列，其相关函数定义为

$$R_{a,b}(\tau)=\sum_{i=0}^{N-1}(-1)^{a_i+b_{i+\tau}},\quad -N<\tau<N$$

特别地，如果$a=b$，则$R_{a,a}(\tau)$被称为自相关函数，其中当$\tau=0$，$R_{a,a}(0)$被称为同相自相关函数，而当$\tau\neq 0$，$R_{a,a}(\tau)$被称为异相自相关函数。

2.2 线性反馈移位寄存器序列

◆例 2.1 已知序列a=0101110 0101110…，则

a是周期为7的周期序列;

a一共有4个游程：00,1,0,111，长度分别为2,1,1,3;

求a的自相关函数:

a=0101110, a =1011100…,

$R_{a,a}(0)=(-1)^{0+0}+(-1)^{1+1}+(-1)^{0+0}+(-1)^{1+1}+(-1)^{1+1}+(-1)^{1+1}+(-1)^{0+0}=7$,

a=0101110, Ta =1011100…,

$R_{a,a}(1)=(-1)^{0+1}+(-1)^{1+0}+(-1)^{0+1}+(-1)^{1+1}+(-1)^{1+1}+(-1)^{1+0}+(-1)^{0+0}=-1$,

a=0101110, T^2a =0111001…,

$R_{a,a}(2)=(-1)^{0+0}+(-1)^{1+1}+(-1)^{0+1}+(-1)^{1+1}+(-1)^{1+0}+(-1)^{1+0}+(-1)^{0+1}=-1$,

$R_{a,a}(3)= R_{a,a}(4)= R_{a,a}(5)= R_{a,a}(6)=-1$.

伪随机序列

- 哥伦布(Golomb, 1955)随机性假设

(G1)：在一个周期内，0与1出现的个数至多相差1。也即，如果N为偶数，则在一个周期内0与1的数目各占$N/2$；如果N为奇数，则在一个周期内0的数目为$(N+1)/2$或者$(N-1)/2$，相应地1的数目为$(N-1)/2$或者$(N+1)/2$。

(G2)：在一个周期内，长度为i 的游程个数占游程总数的$1/2^i$，$i=1,2,\ldots$。且在长度为i的游程中，0的游程与1的游程数目相等或至多相差一个。

(G3)：序列的异相自相关函数是一个常数。

长度为i的游程数目分布

- 满足上述三个条件的序列称为拟噪声序列, 或伪噪声序列(pseudo noise sequence), 简记为: PN序列.
- PN序列在CDMA,通信同步, 导航, 雷达测距等领域有重要应用.

Solomon W. Golomb: Shimonoseki, Japan, October 10-14, 200

伪随机序列

- 密钥序列$k=\{k_i\}=k_0k_1k_2k_3\ldots$.满足的条件G1,G2,G3和以下三个条件:

(C1) 周期p要长. 如$p>10^{50}$.

(C2) 生成容易.

(C3)具有不可预测性(unpredictability):

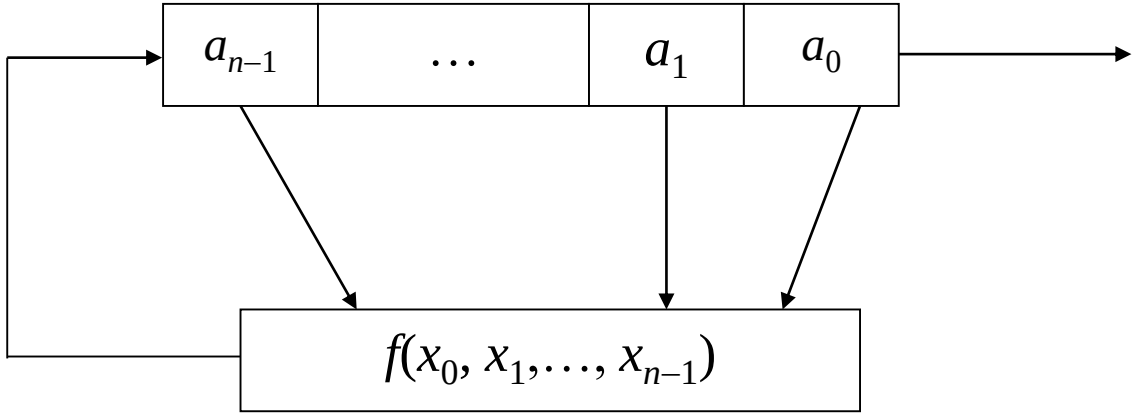
当密钥序列k的 任何部分泄露时,要分析整个密钥序列, 在计算上是不可行的. 线性复杂度高

- C3决定了密码的强度, 是序列密码理论的核心.
- 主要研究问题:

  线性复杂度, 相关免疫性,不可预测性等.

伪随机序列

- 反馈移位寄存器(FSR: Feedback Shift Register)
  - n个寄存器: 从右至左依次称为第1,2,…,n 级
  - 反馈函数 $f(x_0, x_1,\ldots,x_{n-1})$: $GF(2)^n \rightarrow GF(2)$.
  - 工作原理: 当一个时钟脉冲来到时, 第i 级寄存器的内容传送给第i-1级寄存器(i=2,3,…,n),第1 级寄存器的内容为反馈移位寄存器的输出. 反馈函数 $f(x_0, x_1,\ldots,x_{n-1})$的值传送给第$n$级寄存器.
  - FSR的输出序列: $a_0, a_1,a_2,\ldots,a_n,\ldots$ 称为反馈移位寄存器序列(FSR序列).



反馈移位寄存器(FSR)

- 在任意时刻t, 第1至n级寄存器的内容

$$s_t=(a_t, a_{t+1},\ldots,a_{t+n-1})\in \mathrm{GF}(2)^n$$

称为FSR在时刻t 的状态(state).

$s_0=(a_0, a_1,\ldots,a_{n-1})$ 称为FSR的初始状态.决定寄存器级数

在时刻$t+1$ 的状态为:

$$s_{t+1}=(a_{t+1}, a_{t+2},\ldots,a_{t+n}),$$

$$a_{t+n}=f(a_t, a_{t+1},\ldots,a_{t+n-1}).$$

- 共有2^n个状态.
- 反馈函数 $f(x_1, x_2,\ldots,x_n)$是n个变量的布尔函数(Boolean function).

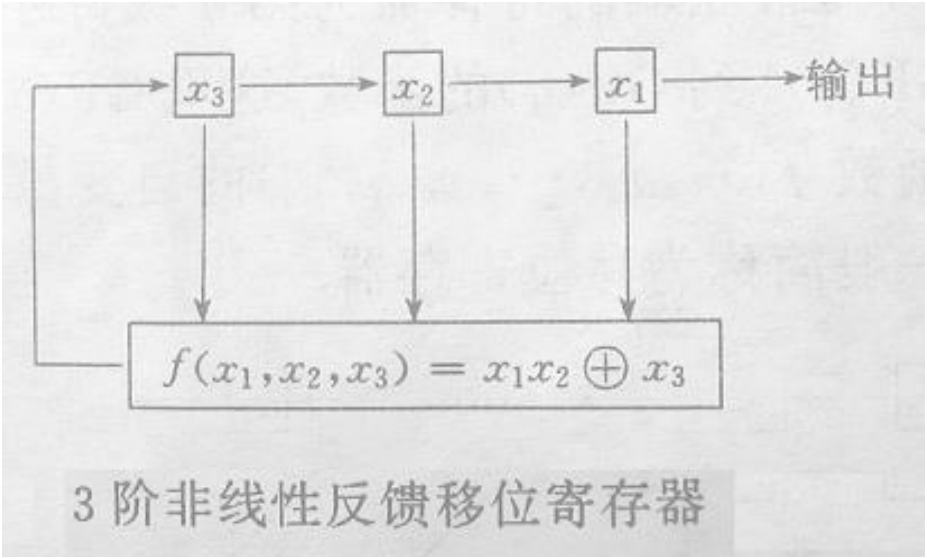
反馈移位寄存器(FSR)

- 例2.2 设有限域GF(2)上的3级FSR的反馈函数为:

$$f(x_1, x_2, x_3)=x_1 \otimes x_2 \oplus x_3$$

初始状态为$s_0=(1,0,1)$. 求FSR序列.

解: 反馈移位寄存器序列: a =1011…; 周期q=4.



3 阶非线性反馈移位寄存器

时刻	状态			输出
	3级	2级	1级	
0	1	0	1	1
1	1	1	0	0
2	1	1	1	1
3	0	1	1	1
4	1	0	1	1
5	1	1	0	0

线性反馈移位寄存器(LFSR)

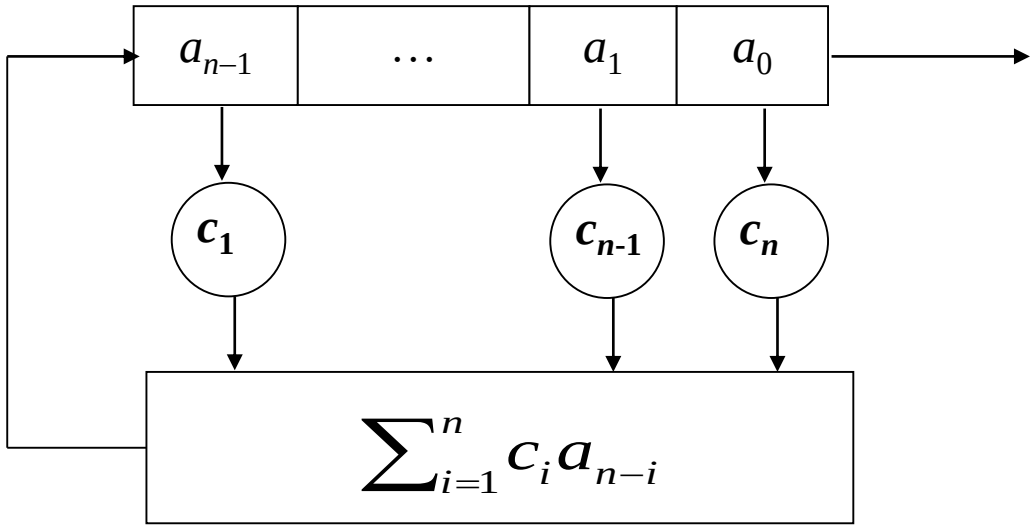
- 如果反馈函数 $f(x_1, x_2,\ldots,x_n)$是n个变量的线性函数:

$$f(x_1, x_2,\ldots,x_n) = c_1x_n + c_2x_{n-1} + \ldots + c_nx_1 \ (c_i \in \{0,1\})$$

则称为线性反馈移位寄存器(LFSR: linear feedback shift register). 输出的序列称为线性反馈移位寄存器序列, 记为LFSR序列。

- LFSR序列$a = (a_0, a_1,\ldots,a_{n-1},\ldots)$ 满足递推关系式:

$$a_{n+k} = c_1a_{n+k-1} + c_2a_{n+k-2} + \cdots + c_na_k \ (k \geq n)$$



线性反馈移位寄存器(LFSR)

- 反馈函数:

$$f(x_1, x_2,\ldots,x_n) = c_1x_n + c_2x_{n-1} + \ldots + c_nx_1 \quad (c_i \in \{0,1\})$$

- 如果$c_n=0$，则称LFSR是退化的；否则称LFSR是非退化的。
- 把多项式:

$$f(x)=1+c_1x+c_2x^2+\ldots+c_nx^n$$

称为LFSR的特征多项式(characteristic polynomial)，或级连多项式、或生成多项式。

线性反馈移位寄存器(LFSR)

◆例2.3　已知如图所示的3级LFSR.

特征多项式为：　$f(x)=1+x^2+x^3$

LFSR序列$a=(a_0, a_1,\ldots,a_{n-1},\ldots)$ 满足递推关系式:

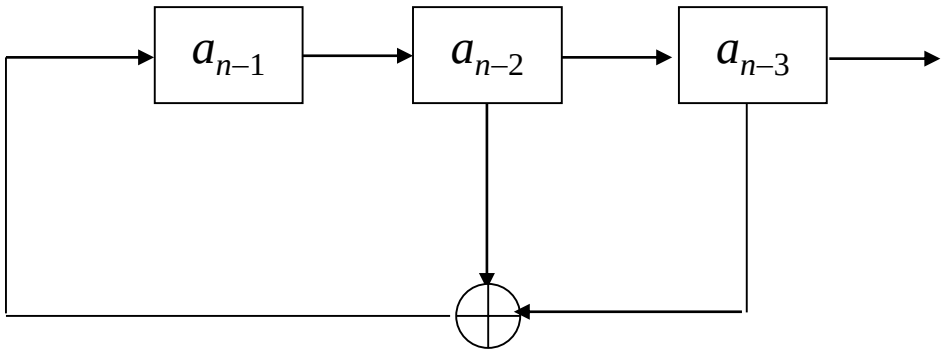
$$a_n=a_{n-2}+a_{n-3}.$$

如果设初始状态为: (0,0,1)

即$a_0=0,a_1=0,a_2=1$,

输出序列为: 0010111

周期为7.

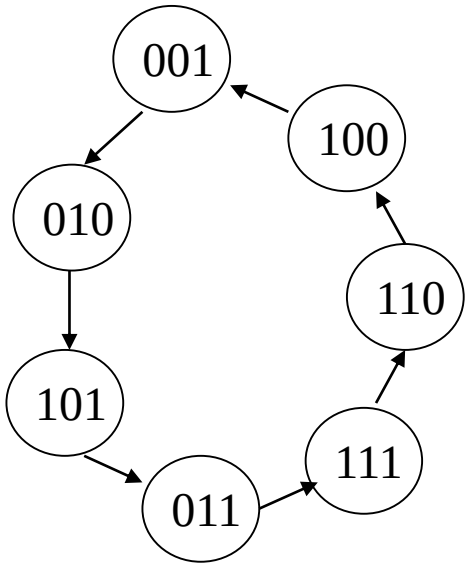


时刻	状态 3级	2级	1级	输出
0	1	0	0	0
1	0	1	0	0
2	1	0	1	1
3	1	1	0	0
4	1	1	1	1
5	0	1	1	1
6	0	0	1	1

线性反馈移位寄存器(LFSR)

◆例2.3　已知如图所示的3级LFSR.

LFSR序列a= 0010111的状态转移图



时刻	状态			输出
	3级	2级	1级	
0	1	0	0	0
1	0	1	0	0
2	1	0	1	1
3	1	1	0	0
4	1	1	1	1
5	0	1	1	1
6	0	0	1	1

线性反馈移位寄存器(LFSR)

◆LFSR序列的性质

定理2.1 任何n级LFSR序列都是终归周期序列，且其周期至多是2^n-1。

◆m-序列

定义2.4 周期为2^n-1的n级线性LFSR序列称为最大长度(Maximal length)序列，简称为m-序列。

定理2.2 a是周期为2^n-1的m-序列的充分必要条件是其特征多项式$f(x)$为n阶本原多项式。

m-序列

- m-序列的个数

定理2.3 设$f(x)$是GF(2)上的n次本原多项式,则对任意非0的初始状态,由$f(x)$生成的m-序列是循环等价(cyclically equivalent)的.

即: 一个n次本原多项式只能生成一个m-序列.

定理2.4 二元域GF(2)上的n级m序列共有$\varphi(2^n-1)/n$个.

$Z_p^* = \{1, 2, \cdots, p-1\}$ 循环 群、交换群 "寻求" + "与".

$x^2+1 = (x+1)^2$ $(x+1)^p = x^p+1$

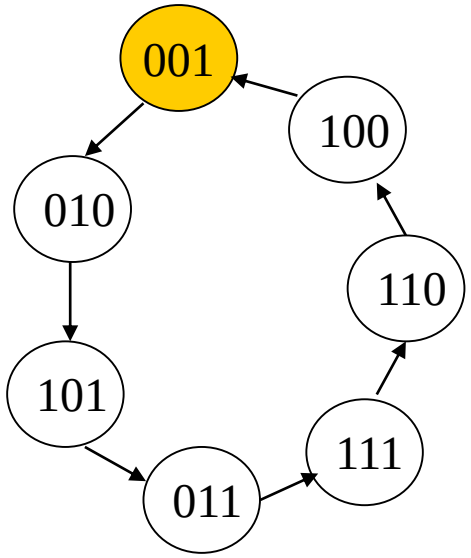
x^2+x+1 x^2+x+1 不可约多项式

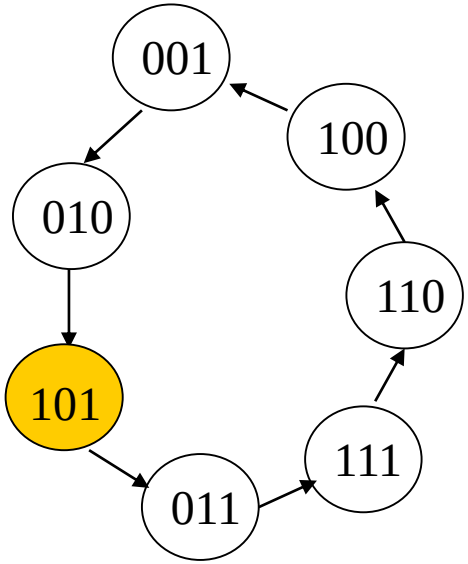
x^2+1 可约多项式

m-序列

◆例2.3　3级LFSR的特征多项式为:　$f(x)=1+x^2+x^3$

初始状态为: (0,0,1),
输出序列为: a=0010111





初始状态为: (1,0,1),
输出序列为: a=1011100.

m-序列

m-序列的伪随机性

定理2.5 设a是一个n级m序列，周期为2^n-1，则

(1) 在一个周期内，0、1出现的次数分别为$2^{n-1}-1$和2^{n-1}。

(2) 在一个周期内，总游程数为2^{n-1}。其中，对$1\le i\le n-2$，长为i的0游程、1游程各有2^{n-i-2}个；长为$n-1$的0游程1个，长为n的1游程1个。

(3) a的自相关函数为:

$$R_{a,a}(\tau)=\begin{cases}2^n-1, & \tau=0\\ -1, & \tau\neq 0\end{cases}$$

m-序列

- m-序列的伪随机性
  - **例2.4** 已知4级*m*序列*a*=100010011010111, 有
    - *n*=4
    - 7个0, 8个1
    - 游程总数为8
    - 长为1的0游程2个, 长为1的1游程2个
    - 长为2的0游程1个, 长为2的1游程1个
    - 长为3的0游程1个
    - 长为4的1游程1个.

m-序列

- m-序列的密码学性质
  - (C1) 周期长: $p=2^n-1$. 如$n=166$时,

    $p=10^{50}$ $(9.353610465\times10^{49})$.
  - (C2) 生成容易: 只要已知n次本原多项式,容易生成m序列.
  - (C3) m序列极不安全: 只要泄露$2n$位连续数字,就可以完全确定反馈多项式的系数, 从而得到m序列.

    已知$a_i, a_{i+1},\ldots, a_{i+2n-1}$, 由以下方程组可唯一解得

    $c_0, c_1,\ldots, c_{n-1}$.

$$\begin{bmatrix} a_i & a_{i+1} & \cdots & a_{i+n-1} \\ a_{i+1} & a_{i+2} & \cdots & a_{i+n} \\ \vdots & \vdots & & \vdots \\ a_{i+n-1} & a_{i+n} & \cdots & a_{i+2n-2} \end{bmatrix} \begin{bmatrix} c_{n-1} \\ c_{n-2} \\ \vdots \\ c_0 \end{bmatrix} = \begin{bmatrix} a_{i+n} \\ a_{i+n+1} \\ \vdots \\ a_{i+2n-1} \end{bmatrix}.$$

第2章 流密码

- 2.1 流密码一般模型
- 2.2 线性反馈移位寄存器序列
- 2.3 线性复杂度及B-M算法
- 2.4 非线性序列生成器
- 2.5 流密码算法

2.3 序列的线性复杂度

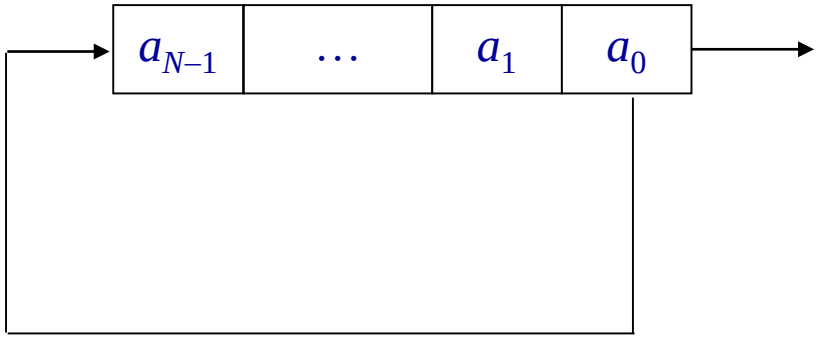
- 序列的密码分析问题

给定一个长度为N的二元密钥流序列a,假定捕获了a的一个长度为m的部分,不失一般性设为$(a_0,a_1,\ldots,a_{m-1})$,能否找到一个级数最短的LFSR,生成整个密钥流序列a？

- 问题一: 是否存在LFSR生成整个序列a?
- 问题二: 捕获序列a多长的部分，才能找到LFSR生成整个序列a？怎样确保得到的LFSR最短?

2.3 序列的线性复杂度

- 序列的LFSR
  - 设$a=(a_0,a_1,\ldots,a_{N-1})$是一个长度为$N$的序列，那么存在$N$级LFSR，生成整个序列$a$。



  - 当a是LFSR序列时，存在着更短的LFSR生成整个序列
  - 当a是非LFSR序列时，也可能存在着更短的LFSR生成整个序列a。

2.3 序列的线性复杂度

- B-M算法
  - 设$a^{(N)}=(a_0,a_1,\ldots,a_{N-1})$是一个长度为$N$的序列,$f_N(x)$是一个能生成$a^{(N)}$且级数最小的LFSR的特征多项式, l_N是LFSR的级数, 则把 $< f_N(x), l_N >$称为$a^{(N)}$的线性综合解.
  - BerleKamp-Massey (1969)提出了求解$< f_N(x), l_N >$的迭代算法.

B-M算法

(1) 取初始值：$f_0(x)=1,\ l_0=0$

(2) 假设$<f_0(x),l_0>,<f_1(x),l_1>,\ldots,<f_n(x),l_n>$已经求出. 设

$$f_n(x)=1+c_1x+c_2x^2+\ldots+c_{l_n}x^{l_n},$$

计算：$d_n=a_n+c_1a_{n-1}+c_2a_{n-2}+\ldots+c_{l_n}a_{n-l_n}.$

如果$d_n=0$, 则取：$f_{n+1}(x)=f_n(x),\ l_{n+1}=l_n.$

如果$d_n\neq 0$, 则区分以下两种情况：

(i) 若$l_0=l_1=\ldots=l_n=0$时，

取$f_{n+1}(x)=1-d_nx^{n+1},\ l_{n+1}=n+1.$

(ii) 当存在m满足：$l_m<l_{m+1}=l_{m+2}=\ldots=l_n.$取

$$f_{n+1}(x)=f_n(x)+d_nd_m^{-1}x^{n-m}f_m(x),$$

$$l_{n+1}=\max\{l_n,n+1-l_n\}.$$

(3) *If* $n<N-1, n=n+1,$ *goto* (2);

else output $<f_n(x),l_n>$, *end*.

B-M算法

- 例2.5 设$a^{(10)}$=(0001101111), N=10, 求其线性综合解.

解:

取初始值: $f_0(x)=1,\ l_0=0,$

$\Rightarrow d_0=a_0=0,$ 故 $f_1(x)=1,\ l_1=0;$

$\Rightarrow d_1=a_1=0,$ 故 $f_2(x)=1,\ l_2=0;$

$\Rightarrow d_2=a_2=0,$ 故 $f_3(x)=1,\ l_3=0;$

$\Rightarrow d_3=a_3=1\neq 0,$

因为$l_0=l_1=l_2=l_3=0,$ 取 $f_4(x)=1-x^4,\ l_4=4.$

$\Rightarrow d_4=1\neq 0,\ m=3,\ n=4,$

已有 $<1,0>,<1,0>,<1,0>,<1,0>,<f_4(x)=1-x^4,4>.$

$\Rightarrow f_5(x)=f_4(x)+x=1+x-x^4,\ l_5=\max\{l_4,4+1-l_4\}=4.$

B-M算法

- 例2.5　设$a^{(10)}$=(0001101111), N=10, 求其线性综合解.

解:

$\Rightarrow d_5 = 1 \neq 0,\ n = 5, m = 3$

已有 $<1,0>, <1,0>, <1,0>, <1,0>, <f_4(x), 4>, <f_5(x), 4>,$

$f_5(x) = f_4(x) + x = 1 + x - x^4.$

$\Rightarrow d_6 = a_6 + a_5 + a_4 - a_2 = 0,$

$f_6(x) = f_5(x) + x^2 = 1 + x + x^2 - x^4,\ l_6 = \max\{l_5, 5 + 1 - l_5\} = 4.$

$\Rightarrow d_7 = a_7 + a_6 + a_5 - a_3 = 1 \neq 0,\ n = 7, m = 3,$

$f_7(x) = f_6(x) = 1 + x + x^2 - x^4,\ l_7 = l_6 = 4.$

$\Rightarrow d_8 = a_8 + a_7 + a_6 = 1 \neq 0,\ n = 8, m = 3,$

$f_8(x) = f_7(x) + x^4 = 1 + x + x^2,\ l_8 = \max\{l_7, 7 + 1 - l_7\} = 4.$

$\Rightarrow d_9 = a_9 + a_8 + a_7 + a_4 = 0,$

$f_9(x) = f_8(x) + x^5 = 1 + x + x^2 + x^5,\ l_9 = \max\{l_8, 8 + 1 - l_8\} = 5.$

B-M算法

- 例2.5　设$a^{(10)}$=(0001101111), N=10, 求其线性综合解.

解:

- $a^{(10)}$的线性综合解为:　$f_{10}(x)=1+x+x^2+x^5$,　$l_{10}=5$.
  若取初值:　$a^{(0)}$=00011, 则$f_{10}(x)$的LFSR序列
  a=0001101111 0011101000…,
  周期为:　$2^5-1=31$.

B-M算法

- B-M算法的性质
  - B-M算法的时间复杂度为$O(N^2)$.

定理2.6 给定长为N的序列$a^{(N)}=(a_0,a_1,\ldots,a_{N-1})$，如果用B-M算法得到的线性综合解为$(f_N(x), l_N)$，则以$f_N(x)$为生成多项式，产生的长为$l_N$的LFSR就是生成序列$a^{(N)}$的最短LFSR。

定理2.7 给定长为N的序列$a^{(N)}=(a_0,a_1,\ldots,a_{N-1})$，用B-M算法得到的线性综合解为$(f_N(x), l_N)$是唯一解的充要条件是$2l_N \le N$。

B-M算法

- B-M算法的性质

定理2.8 设$a^{(N)}=\{a_0, a_1,..,a_{N-1}\}$是一个长为$N$的序列, l_N是能产生$a^{(N)}$并且阶数最小的LFSR的阶数. 则当$2l_N > N$时, $a^{(N)}$ 线性综合解的个数为:

$$q^{2l_N - N}$$

2.3 序列的线性复杂度

- 序列的线性复杂度

给定序列a，生成它的最短LFSR的长度l_N就确定了。如果$2l_N \le N$，只需要捕获序列a连续的$2l_N$个比特，就能得到它的唯一解$(f_N(x), l_N)$，以$f_N(x)$为特征多项式的l_N级LFSR就可以生成整个序列a。

特别地，对于周期N很大，但l_N很小的序列a，比如周期为2^n-1的n级m-序列，利用B-M算法，只要捕获序列a连续的$2l_N$个比特，即序列很小一部分，就可以重构整个序列。因此，l_N实际上度量了序列a的线性的不可预测性。

2.3 序列的线性复杂度

- 序列的线性复杂度

定义2.5 生成长为N的序列$a=(a_0,a_1,\ldots,a_{N-1})$的LFSR的最短长度$l_N$被称为序列$a$的线性复杂度。

- n阶m序列的线性复杂度=n.

第2章 流密码

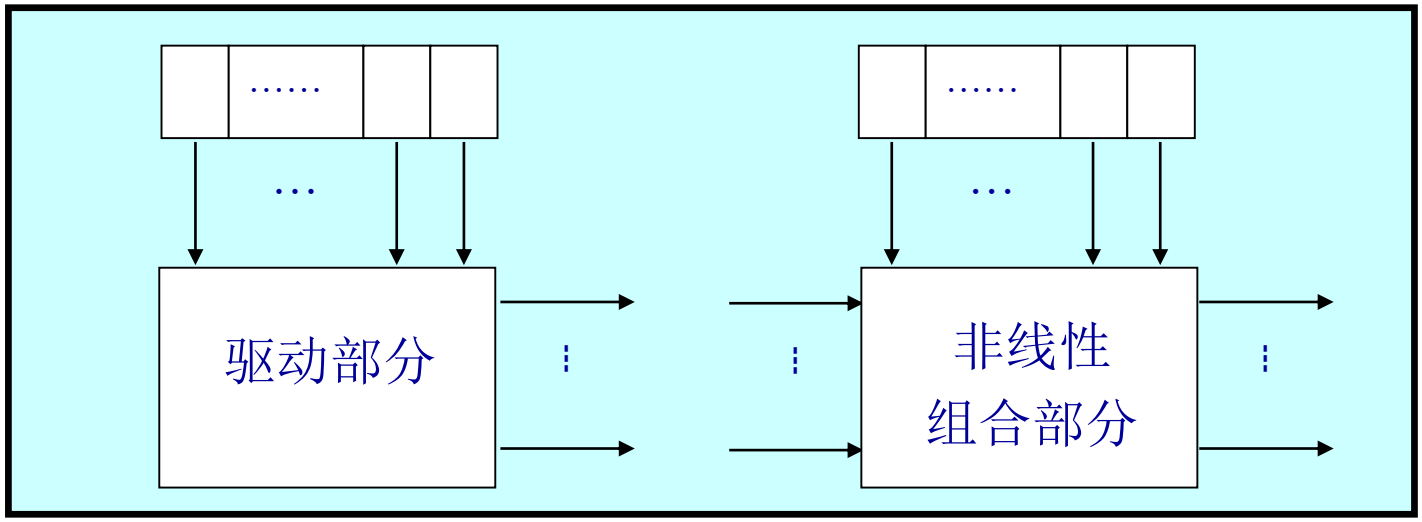
- 2.1 流密码一般模型
- 2.2 线性反馈移位寄存器序列
- 2.3 线性复杂度及B-M算法
- 2.4 非线性序列生成器
- 2.5 流密码算法

2.4 非线性序列生成器

- 密钥流生成器的分解

  Ruppe将密钥流生成器分成两部分：驱动部分和非线性组合部分

  - 驱动部分：可由m-序列或其它长周期的LFSR序列组成，用于控制密钥流生成器的状态序列，并为非线性组合部分提供伪随机性质良好的序列
  - 非线性组合部分：利用驱动部分生成的状态序列生成满足要求的密码特性好的密钥流序列



2.4 非线性序列生成器

- 非线性准则

  非线性组合部分可由布尔函数表示

  - n元布尔函数$f(x)$的代数正规型：

$$f(x):\mathrm{Z}_2^{\mathrm{n}}\to\mathrm{Z}_2,\ x=(x_1,x_2,\cdots,x_n),$$

$$\begin{aligned}f(x)&=f(x_1,x_2,\cdots,x_n)\\&=a_0+a_1x_1+a_2x_2+\cdots a_nx_n+\\&\quad a_{1,2}x_1x_2+\cdots+a_{n-1,n}x_{n-1}x_n+\\&\quad\cdots+\\&\quad a_{1,2,\cdots,n}x_1x_2\cdots x_n.\end{aligned}$$

2.4.1 非线性准则

◆代数次数

定义2.6 设$f(x)$是一个n元布尔函数，在$f(x)$的代数正规型中，一个乘积项中变量的个数称为该乘积项的次数。 $f(x)$的代数正规型中，全体非零系数乘积项次数的最大值称为$f(x)$的代数次数 。

- 当$f(x)$的代数次数为1时，$f(x)$称为线性布尔函数
- 当$f(x)$的代数次数大于1时，$f(x)$称为非线性布尔函数
- 对于非线性组合部分的布尔函数，应该具有尽可能大的代数次数

2.4.1 非线性准则

◆非线性度

定义2.7 设L是Z_2上所有线性函数的集合，即$L=\{u\cdot x+v\}|\ u\in Z_2^n, v\in Z_2\}$。则布尔函数$f(x)$的非线性度定义为

$$N_f = \min_{l(x)\in L} d_H(f(x), l(x))$$

其中$d_H(\cdot)$是汉明距离.

- 是密码系统为抵抗线性攻击而提出的指标
- 对于非线性组合部分的布尔函数，应该具有尽可能大的非线性度

2.4.1 非线性准则

◆退化布尔函数

定义2.8 设$f(x)$是一个n元布尔函数，如果存在Z_2上一个$k\times n$（$k<n$）的矩阵D，使得

$$f(x)= g(Dx)= g(y),$$

则称$f(x)$是退化的。

- 自变量经过线性变换后，n元布尔函数$f(x)$就简化为k元布尔函数$g(x)$
- 作为非线性组合部分的布尔函数应该避免退化性

2.4.1 非线性准则

◆布尔函数的相关免疫性

定义2.9 设$f(x_1,x_2,\ldots,x_n)$是n个彼此独立、对称的二元随机变量的布尔函数，称$f(x)$是m阶相关免疫的当且仅当$f=f(x_1,x_2,\ldots,x_n)$与$x_1,x_2,\ldots,x_n$中的任意m个随机变量

$(x_{i_1}, x_{i_2}, ..., x_{i_m})$统计无关。

即对于任意的$(a_1, a_2, ..., a_m) \in Z_2^m$和$a \in Z_2$，$f(x_1, x_2, ..., x_m)$满足

$$p(f = a, x_{i_1} = a_1, x_{i_2} = a_2, ..., x_{i_m} = a_m) = 2^{-m} p(f = a).$$

- 相关免疫性是为防止攻击者对密码系统进行相关攻击而提出的指标
- 希望作为非线性组合部分的布尔函数应该具有m阶相关免疫性，m尽可能地大

2.4.1 非线性准则

◆布尔函数的相关免疫性

定义2.10 设 $f(x_1,x_2,\ldots,x_n)$是一个n元布尔函数，$f(x)$的Walsh变换定义为

$$S_f(\omega)=2^{-n}\sum_{x\in Z_2^n} f(x)(-1)^{\omega\cdot x},$$

其逆变换为：$f(x)=\sum_{\omega\in Z_2^n} S_f(\omega)(-1)^{\omega\cdot x}.$

其中$x=(x_1,x_2,\ldots,x_n)\in Z_2^n,\ \omega=(\omega_1,\omega_2,\ldots,\omega_n)\in Z_2^n,$

$\omega\bullet x=\omega_1x_1+\omega_2x_2+\ldots+\omega_nx_n.$

*Walsh*变换：$f(x)\to S_f(\omega),$

*Walsh*逆变换：$S_f(\omega)\to f(x).$

$S_f(\omega)$也称为$f(x)$的*Walsh*谱.

2.4.1 非线性准则

◆布尔函数的相关免疫性

定理2.9 设$1\le m\le n$，n元布尔函数$f=f(x_1,x_2,\ldots,x_n)$是m阶相关免疫的当且仅当对于任意满足$1\le w_H(\omega)\le m$的$\omega=(\omega_1,\omega_2,\ldots,\omega_n)\in Z_2^n$，$f(x)$的Walsh谱都为0，即

$$S_f(\omega)=0.$$

这里$w_H(\cdot)$是汉明重量.

◆雪崩准则

定义2.11 设$f(x)$是一个n元布尔函数，如果对于任意满足: $w_H(e)=1$的$e=(e_1,e_2,\ldots,e_n)\in Z_2^n$，$f(x)+f(x+e)$是平衡函数，则称$f(x)$为满足严格雪崩准则.

2.4.1 非线性准则

◆扩散准则

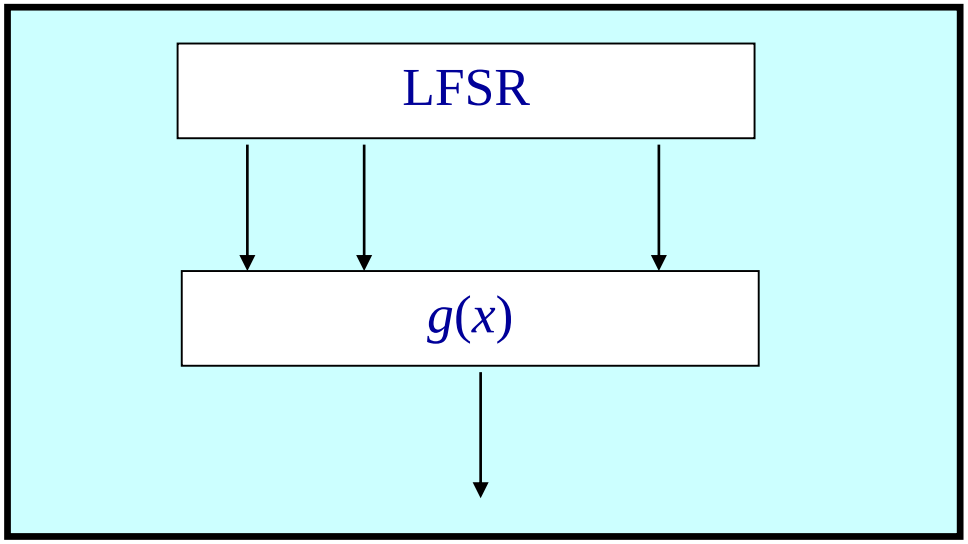
定义2.12 设$f(x)$是一个n元布尔函数，$1 \le m \le n-2$，如果对于任意满足: $1 \le w_H(e) \le m$ 的$e=(e_1, e_2, \ldots, e_n) \in Z_2^n$，$f(x)+f(x+e)$是平衡函数，则称 $f(x)$为满足m次扩散准则.

◆非线性序列设计准则

- 代数次数
- 非线性度
- 退化性
- 相关免疫性
- 雪崩准则
- 扩散准则

2.4.2 非线性序列生成器

- 滤波生成器（Filter geneator）
  - 滤波生成器又叫前馈生成器，由几个LFSR和滤波（前馈）函数$g(x)$两部分组成
  - 滤波函数要求具有很好的非线性性质，以增强生成器的抗攻击能力



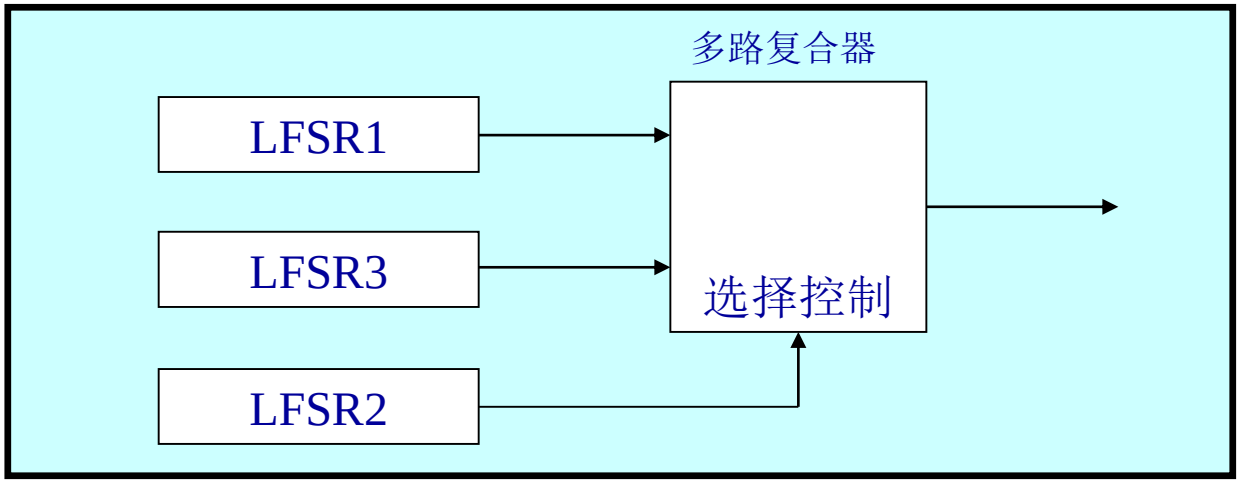
滤波生成器

- Geffe生成器
  - 使用三个级数两两互素的LFSR，其中LFSR1和LFSR3作为多路复合器的输入，LFSR2控制多路复合器的输出
  - 滤波函数

$$g(x_1, x_2, x_3) = x_1 x_2 + \overline{x_2} x_3$$

设a_1、a_2和a_3分别是LFSR、LFSR2和L FSR3的输出，则Geffe生成器的输出b为:

$$b = a_1 a_2 + \overline{a_2} a_3$$



滤波生成器

- Geffe生成器
  - 大的周期和线性复杂度

    设LFSR1、LFSR2和L FSR3周期分别为T_1,T_2和T_3，级数分别为n_1, n_2和n_3，则Geffe生成器的周期为$T_1T_2T_3$，线性复杂度为 $(n_1 + 1)n_2 + n_1n_3$.
  - 不安全

    由于生成器的输出与LFSR2的输出有75%是相同的，通过观察输出序列可以获得LFSR的初态和输出序列，即所谓的相关攻击破译Geffe生成器。因此Geffe生成器是不安全的.

滤波生成器

- J-K触发器
  - LFSR$_1$输出序列$\{a_k\}$, 周期为2^m-1.
  - LFSR$_2$输出序列$\{b_k\}$, 周期为2^n-1.
  - J-K触发器输出序列$\{c_k\}$

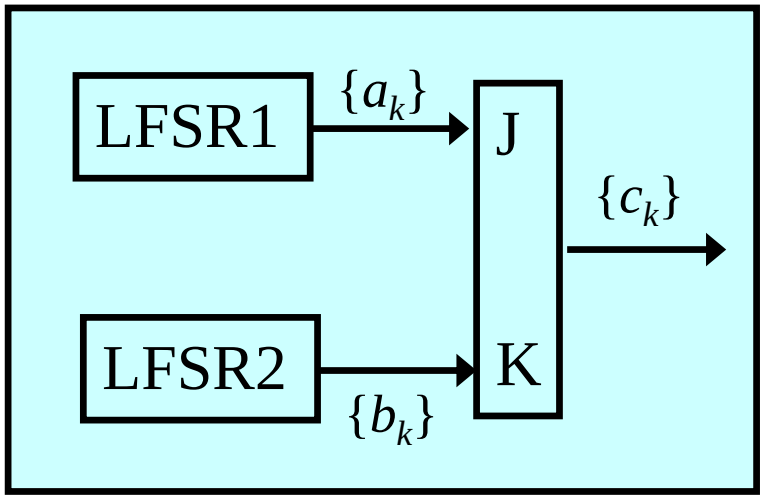
$$c_k = (\overline{a_k + b_k})c_{k-1} + a_k.$$

令c_{-1} =0,有

c_0=a_0,

c_1=(a_1+b_1+1)a_0+a_1,

c_2=(a_2+b_2+1) c_1+a_2 ,…



J	K	c_k
0	0	c_{k-1}
0	1	0
1	0	1
1	1	$\overline{c_{k-1}}$

J-K触发器

- 如果gcd(m, n)=1，且$a_0 + b_0 = 1$，则输出序列$\{c_k\}$的周期为:

  $(2^m-1)(2^n-1)$.

- J-K触发器输出序列$\{c_k\}$随机性好
- 不安全

  已知c_n与c_{n+1}, 便能对a_{n+1}与b_{n+1}的一个作出判断.

  $c_n=c_{n+1}=0 \Rightarrow a_{n+1}=0$;

  $c_n=0, c_{n+1}=1 \Rightarrow a_{n+1}=1$;

  $c_n=1, c_{n+1}=0 \Rightarrow b_{n+1}=1$;

  $c_n=c_{n+1}=1 \Rightarrow b_{n+1}=0$.

J-K触发器

- 例2.4.1 令m=2, n=3, 且$a_0 + b_0 = 1$,
  LFSR$_1$输出序列

  $\{a_t\}$=011…,

  LFSR$_2$输出序列

  $\{b_t\}$=1001011….

  有

  $c_0=a_0=0$,

  $c_1=(a_1+b_1+1)a_0+a_1=(1+0+1)0+1=1$,

  $c_2=(a_2+b_2+1)\,c_1+a_2=(1+0+1)1+1=1$,…

  $\{b_k\}$=0110 1001 1101 0100 1001 0….

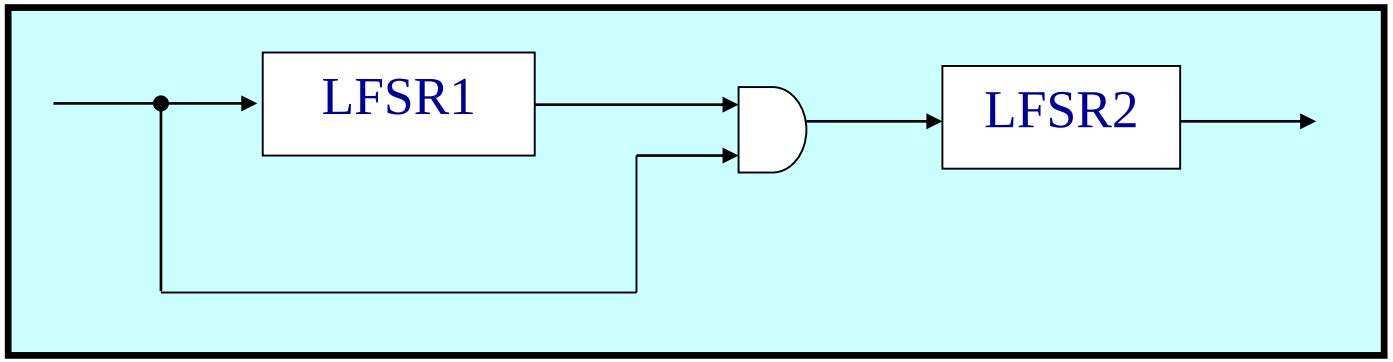
  周期为:

  $L=(2^m-1)\,(2^n-1)=(2^2-1)\,(2^3-1)=21$.

J	K	c_k
0	0	c_{k-1}
0	1	0
1	0	1
1	1	$\overline{c_{k-1}}$

2.4.2 非线性序列生成器

- 钟控序列生成器
  - 钟控生成器（Clock controlled generator）是由一个或几个FSR输出序列，控制另一个FSR的时钟。
  - 走停生成器（Stop-and-Go generator）
    - 当LFSR1输出1时，时钟脉冲通过与门使LFSR2进行一次移位，从而生成下一位；
    - 当LFSR1输出0，时钟脉冲无法通过与门使LFSR2移位（走），从而LFSR2重复输出前一位（停）



钟控序列生成器

- 钟控序列的周期

  设LFSR_1输出序列$\{a_k\}$, 周期为2^m-1，LFSR_2输出序列$\{b_k\}$, 周期为2^n-1，则钟控序列$\{c_k\}$的周期为：

  $$(2^m\text{-}1)(2^n\text{-}1).$$

- 钟控序列 $\{c_k\}$的线性复杂度为：$n(2^m\text{-}1)$.

钟控序列生成器

- **例2.6** 设$LFSR_1$为一个3级m-序列，其特征多项式为：$f_1(x)=1+x+x^3$，取初始值为$a_0=a_1=a_2=1$, 则输出序列$\{a_k\}=1110100$, 周期为$2^3-1=7$.

  设$LFSR_2$为一个3级m-序列，其特征多项式为：$f_1(x)=1+x^2+x^3$，取初始值为$b_0=b_1=b_2=1$, 则输出序列$\{b_k\}=1110010$, 周期为$2^3-1=7$.

  钟控序列:

  $\{c_k\}$=1110 0000 1011 1111 0001 1101 1111
  1001 1000 1111 0000 1001 1…

  周期为：$(2^m-1)(2^n-1)=(2^3-1)(2^3-1)=49$.

  线性复杂度为：$n(2^m-1)=3(2^3-1)=21$.

第2章 流密码

- 2.1 流密码一般模型
- 2.2 线性反馈移位寄存器序列
- 2.3 线性复杂度及B-M算法
- 2.4 非线性序列生成器
- 2.5 流密码算法

2.5 流密码算法

- RC4算法
  - RC4是由Rivest于1987年开发的一种序列密码,它已被广泛应用于Windows, Lotus Notes和其它软件,还被用于安全套接字(SSL)和无线通信系统等.
  - RC4优点是算法简单、高效，特别适于软件实现，加密速度比DES大约快10倍。
  - RC4可以支持不同密钥长度，美国政府特别限定，用于出口的RC4的密钥长度不得超过40位。

2.5 流密码算法

- RC4使用了一个2^8字节大小的非线性数据表(简称S表), S表的值$S_0,S_1,\ldots,S_{255}$是数字0到255的一个排列。对S表进行非线性变换,得到密钥流。
- RC4对S表的初始化算法(两个计数器I和J, $I=0$, $J=0$)

1. 对S表进行线性填充：$S_I=I$, $0\le I<255$;
2. 用密钥填充另一个256字节的数组K，如果密钥长 度小于256字节，则依次重复填充，直至填满这个数组中：$K_0, K_1,\ldots, K_{255}$;
3. 对于$I=0$到255重复以下步骤:
   (1) $J=J+S_I+K_I$ （mod 256）；
   (2) 交换S_I和S_J。

RC4算法

- RC4输出密钥流字节z的算法

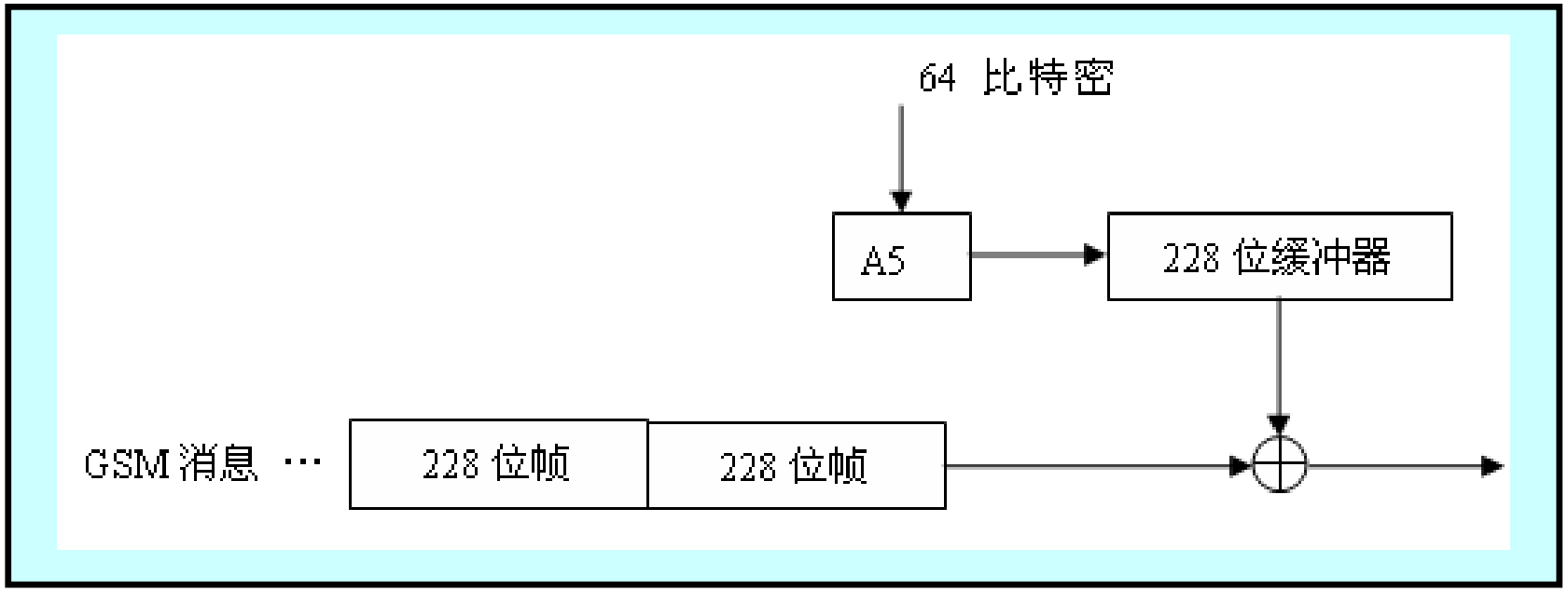
1. $I=0, J=0$
2. $I=I+1 \pmod{256}$;
3. $J=J+S_I \pmod{256}$;
4. 交换S_I和S_J;
5. $t=S_I+S_J \pmod{256}$;
6. $z=S_t$.

2.5 流密码算法

- A5算法
  - A5有两个版本：A5/1和A5/2，前者有更高的安全性，根据相关法规限制被仅用于欧洲范围，而后者用于其它地区。A5算法从未公布于众，但因为一些疏漏，该算法被Bradford大学研究人员泄密，我国学者徐胜波、何大可和王新梅也由此于1994年率先实现A5算法。
  - A5是欧洲数字蜂窝移动电话系统（GSM）采用的流密码算法，用于加密从移动台到基站的连接。

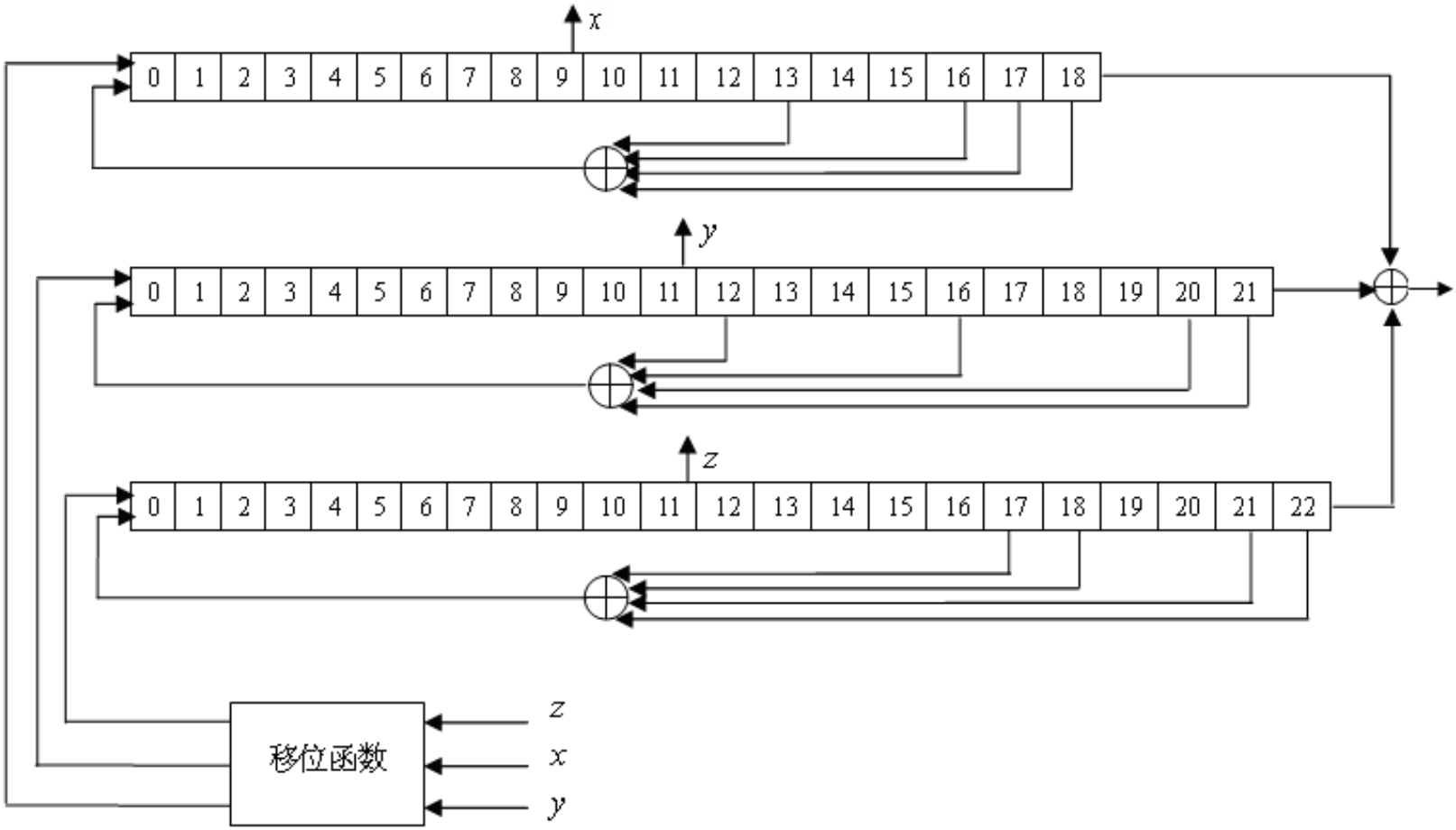
2.5 流密码算法

- A5算法
  - GSM 会话每帧有228bit
  - A5算法的密钥长64bit
  - 有一个22bit表征会话帧数
  - 每次产生228bit会话密钥。



A5算法

- A5算法


x
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
y
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
z
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
移位函数
z
x
y

A5算法

- LFRS1: $g_1(x)=x^{19}+x^{18}+x^{17}+x^{14}+1$;

  LFRS2: $g_2(x)=x^{22}+x^{21}+x^{17}+x^{13}+1$;

  LFRS3: $g_3(x)=x^{23}+x^{22}+x^{19}+x^{18}+1$.
- 时钟控制系统
  - 输入: LFRS1(10)= x, LFRS2(12)= y, LFRS3(12)= z
  - 控制逻辑: 如果LFRS1(10)=LFRS2(12)=LFRS3(12), 则3个LFRS都移1位; 否则相等的2个LFRS移1位, 另1个LFRS不移位.

    钟控函数: $g(x, y, z)= xy + xz + yz$
- 输出序列: LFRS1+LFRS2+LFRS3

2.5 序列密码实例

- A5算法工作过程
  - （1）将64比特密钥输入LFSR；
  - （2）将22比特帧数与LFSR反馈值模2加，再输入LFSR；
  - （3）LFSR开始停走钟控；
  - （4）舍去产生的100比特输出；
  - （5）产生114比特作为密钥流；
  - （6）舍去产生的100比特输出；
  - （7）产生114比特作为密钥流。

第2章　　习　题

- **P50: 习题1-7.**