应用密码学（第4讲）

------ 流密码与LFSR序列

密码学与编码理论培训班

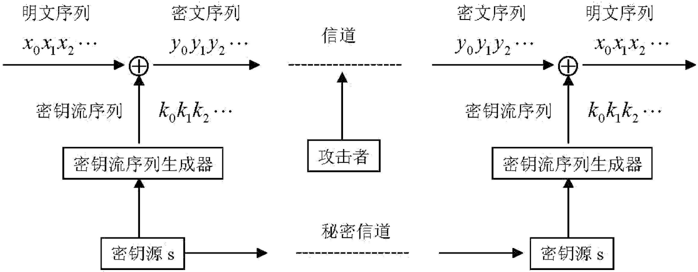
2018年8月1-10日@西宁

本节概要

从Vernam密码谈起...

$$
\begin{array}{rll}
 & 1001101001110100101001100101011001010100100010 & \textit{plaintext} \\
+ & 1011010100110101011010101011011101001011100101 & \textit{keystream} \\
\hline
 & 0010111101000001110011001110000100011111000111 & \textit{ciphertext}
\end{array}
$$

序列密码的加密模型


明文序列
$x_0x_1x_2\cdots$
密文序列
$y_0y_1y_2\cdots$
信道
密文序列
$y_0y_1y_2\cdots$
明文序列
$x_0x_1x_2\cdots$
密钥流序列
$k_0k_1k_2\cdots$
密钥流序列
$k_0k_1k_2\cdots$
密钥流序列生成器
攻击者
密钥流序列生成器
秘密信道
密钥源 s
密钥源 s

本节概要

定义 1

我们把有限域$\mathbf{F}_2$上的一个无限序列

$$\mathbf{a} = (a_0, a_1, a_2, a_3, \cdots), \quad a_i \in \mathbf{F}_2 \tag{1}$$

称为二元序列或简称序列。我们说一个二元序列$\mathbf{a}$是周期的，如果存在正整数l，使得

$$a_k = a_{k+l} \tag{2}$$

对一切非负整数k成立. 满足上述条件的最小的正整数l称为$\mathbf{a}$的周期，记作$p(\mathbf{a})$, 即$p(\mathbf{a}) = l$. 而把$(a_0, a_1, \cdots, a_{l-1})$叫做$\mathbf{a}$的一个周期。

定理 1

设$\mathbf{a} = (a_0, a_1, a_2, a_3, \cdots)$是周期为$p(\mathbf{a})$的二元序列。并设正整数$l$对任何的 非负整数$k$, $a_k = a_{k+l}$。则一定有$p(\mathbf{a})|l$。

定义 2 (游程)

设$\mathbf{a}$是$\mathbf{F}_2$上周期为l的二元序列。将$\mathbf{a}$的一个周期

$$(a_0, a_1, \cdots, a_{l-1})$$

依次排列在一个圆周上使a_{l-1}与a_0相接，我们把这个圆周上形如

$$0\underbrace{111\cdots111}_{\text{都是}1}0 \quad \text{或} \quad 1\underbrace{000\cdots000}_{\text{都是}0}1$$

的一连串两两相邻的项分别叫做$\mathbf{a}$的一个周期中的一个1游程和一个0游程，而 1游程中1的个数或0游程0的个数叫做游程的长度。

定义 3

设有$\mathbf{F}_2$上周期等于l的二元序列

$$\mathbf{a} = (a_0, a_1, a_2, \cdots).$$

a的自相关函数$c_{\mathbf{a}}(t)$是从非负整数 集合到整数集合的函数：

$$c_{\mathbf{a}}(t) = \sum_{i=0}^{l-1} \eta(a_i)\eta(a_{i+t}), t \geq 0,$$

其中$\eta(0) = 1, \eta(1) = -1$。

Golomb随机性公设

1. 在序列的一个周期中，当周期为偶数时，1的个数与0的个数相等；当周期为奇数时，1的 个数与0的个数差一。
2. 在序列的一个周期中，长为1的游程占总游程的$\frac{1}{2}$，长为2的游程占总游程的$\frac{1}{2^2}$，长为3的游程占总游程的$\frac{1}{2^3}$，$\cdots$. 在同样长度的游程中，1游程和0游程大致各占一半。
3. 自相关函数$c(t)$在$t=0$时最高，在$t\neq 0$迅速下降。

定义 4

设$\mathbf{a}=(a_0,a_1,a_2,\cdots)$是$\mathbf{F}_2$上一个周期等于$\nu$的二元序列。如果 对一切的$t\not\equiv 0 (\mathrm{mod}\,\nu)$，有

$$c_{\mathbf{a}}(t)=-1$$

我们就说$\mathbf{a}$是个伪随机序列。

定理 2

设$\mathbf{a}$是一个伪随机序列，那么$\mathbf{a}$的周期ν一定是奇数，而且在 $\mathbf{a}$的一个周期里，1出现的个数和0出现的个数相差1, 即1出现的个数 比0出现的个数多1或者少1。

证明： 设 $\mathbf{a}=(a_0,a_1,a_2,\cdots,a_{\nu-1},a_\nu,a_{\nu+1}\cdots)$, 根据伪随机序列的定义可知

$$c_{\mathbf{a}}(t)=\begin{cases}\nu, & \text{如果 } t=0;\\ -1, & \text{如果 } 0<t<\nu.\end{cases}$$

于是 $\sum_{t=0}^{\nu-1}c_{\mathbf{a}}(t)=1$. 另一方面，根据自相关函数的定义，

$$\begin{aligned}1=\sum_{t=0}^{\nu-1}c_{\mathbf{a}}(t) &= \sum_{t=0}^{\nu-1}\sum_{i=0}^{\nu-1}\eta(a_i)\eta(a_{i+t})\\ &= \sum_{i=0}^{\nu-1}\eta(a_i)\sum_{t=0}^{\nu-1}\eta(a_{i+t})\\ &= \left(\sum_{i=0}^{\nu-1}\eta(a_i)\right)^2\end{aligned}$$

因此 $\sum_{i=0}^{\nu-1}\eta(a_i)=1$ 或 -1, 所以$\eta(a_0),\eta(a_1),\cdots,\eta(a_{\nu-1})$中$-1$和1的个数相差1。于是，在 $\mathbf{a}$的一个周期中，1出现的个数和0出现的个数相差1, 因此ν一定为奇数。

定理 3

在$\mathbf{F}_2$上周期序列的一个周期中，0游程的个数一定等于1游程的个数。更进一步伪随机序列的周期 一定$\equiv 3 \pmod 4$，而在周期等于ν的伪随机序列的一个周期中， 0游程的个数和1游程的个数都等于$\frac{\nu+1}{4}$。

证明： 设$\mathbf{a}$是$\mathbf{F}_2$上的一个周期序列，周期等于ν。在$\mathbf{a}$的一个周期

$$(a_0, a_1, \cdots, a_{\nu-1})$$

中，形如01或10的两项分别叫做从0到1或从1到0的变化。显然，$\mathbf{a}$的一个 周期中从0到1变化的个数等于从1到0变化的个数，而且0游程的个数就等于从0 到1变化的个数，1游程的个数就等于从1到0变化的个数。因此，$\mathbf{a}$的一个周期 中，0游程的个数等于1游程的个数。
更进一步，设$\mathbf{a}$的一个周期中, 0游程的个数为m，那么1游程的个数, 从0到1变化的个数和从1到0变化的个数也都为m。于是

$$\begin{aligned} c_{\mathbf{a}}(1) &= \sum_{i=0}^{\nu-1} \eta(a_i)\eta(a_{i+1}) \\ &= m \cdot ((-1) \cdot 1) + m \cdot (1 \cdot (-1)) + (\nu - 2m) \\ &= \nu - 4m \end{aligned}$$

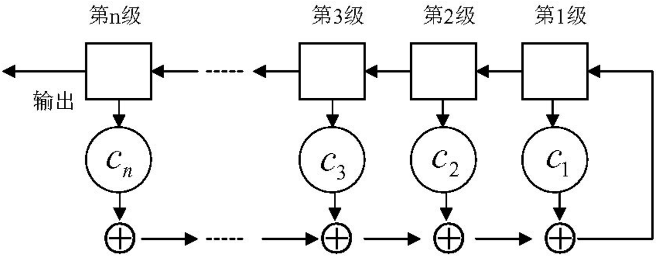
因为$\mathbf{a}$是伪随机序列，所以又有$c_{\mathbf{a}}(1)=-1$，因此

$$\nu-4m=-1$$

这样就证明了$\nu\equiv 3 \pmod 4$，而$m=\frac{\nu+1}{4}$。

本节概要

线性移位寄存器



一个n级线性移位寄存器有n个寄存器和一个反馈开关电路组成，从右到左分别称为 第1级寄存器、第2级寄存器，$\cdots$，第n级寄存器。每个寄存器的状态分别 用0和1表示，而0和1总可看成是有限域$\mathbf{F}_2$中的元素。

线性移位寄存器序列

当线性移位寄存器的n个初始值$a_0, a_1, \cdots, a_{n-1}$给定之后，不断地加移位脉冲，$n$级线性移位寄存器的 输出就会输出一序列

$$a_0, a_1, a_2, \cdots$$

满足线性递推关系(或**反馈逻辑**)

$$a_k = \sum_{i=1}^{n} c_i a_{k-i}, \ a_k + \sum_{i=1}^{n} c_i a_{k-i} = 0, \ k > n \tag{3}$$

这个序列就称为(n级)**线性移位寄存器序列**或线性递归序列。
n级线性移位寄存器序列中任何连续n个项都叫做该序列的一个状态，而形如：

$$(a_k, a_{k+1}, \cdots, a_{k+n-1}), k \geq 0 \tag{4}$$

的状态称为第k个状态，记为s_k. 状态s_0称为序列的初始状态。

特征多项式

从线性移位寄存器的反馈逻辑可构造的如下多项式

$$f(x) = x^n + \sum_{i=1}^{n} c_i x^{n-i} \text{ 和 } \tilde{f}(x) = 1 + \sum_{i=1}^{n} c_i x^i$$

分别称为n级线性移位寄存器的**特征多项式**和**联接多项式**，而把满足递推关系式3 的n级线性移位寄存器序列称作 由$f(x)$产生的（二元）n级线性移位寄存器序列，简称由$f(x)$产生的序列。我们通常 用符号$G(f)$表示由$f(x)$产生的所有序列的全体组成的集合。

定理 4

设$f(x) = x^n + \sum_{i=1}^{n} c_i x^{n-i}$，那么由$f(x)$产生的序列的总数是$2^n$，即$|G(f)| = 2^n$。更进一步，$G(f)$可以看成是二元域$\mathbf{F}_2$上的$n$维向量空间。

线性移位寄存器序列的周期 I

定理 5

非退化的n级线性移位寄存器序列一定是周期序列，而且它的周期$\leq 2^n - 1$.

证明：假设一个非退化的n级线性移位积存器的特征多项式为$f(x) = x^n + \sum_{i=1}^{n} c_i x^{n-i}$， $\mathbf{a} = (a_0, a_1, a_2, a_3, \cdots)$是由该线性移位寄存器生成的一个序列，那么$c_n \neq 0$且

$$a_{k+n} = \sum_{i=1}^{n} c_i a_{n+k-i}, \quad k \geq 0 \tag{5}$$

现在考察序列$\mathbf{a}$的状态$s_k (k \geq 0)$，它们都是一些$\mathbf{F}_2$上的n维行向量。注意到$\mathbf{F}_2$上不同的n维行向量最多只有2^n个，因此一定存在$0 \leq r < s \leq 2^n$，使得$s_r = s_s$. 由于序列的状态和线性递推关系完全决定了从该状态开始之后序列 的所有元素，所以对任给的$k \geq r$，

线性移位寄存器序列的周期 II

$s_{k+s-r}=s_k$. 令$l=s-r$, k_0是最小的非负整数使得对任给的$k\geq k_0$, $s_{k+l}=s_k$, 我们来证$k_0=0$.
假设$k_0\geq 1$, 则根据(5),

$$s_{k_0+n-1+l}=\sum_{i=1}^{n}c_is_{k_0+n-1+l-i}$$

但$c_n\neq 0$, 因此$c_n=1$, 所以

$$s_{k_0-1+l}=s_{k_0+n-1+l}-\sum_{i=1}^{n-1}c_is_{k_0+n-1+l-i}=s_{k_0+n-1}+\sum_{i=1}^{n-1}c_is_{k_0+n-1-i} \tag{6}$$

另一方面，根据(5), 我们可以直接得到

$$s_{k_0-1+n}=\sum_{i=1}^{n}c_is_{k_0-1+n-i}$$

线性移位寄存器序列的周期 III

再次注意到$c_n = 1$, 我们有

$$s_{k_0-1} = s_{k_0+n-1} - \sum_{i=1}^{n-1} c_i s_{k_0-1+n-i} \tag{7}$$

比较(6)式和(7)式知$s_{k_0-1} = s_{k_0-1+l}$. 这与k_0的选取相矛盾。 所以$k_0 = 0$，即对所有$k \geq 0$, $s_{k+l} = s_k$. 所以**a**是周期序列。 □

定义 5 (m-序列)

当n级线性移位寄存器序列的周期达到最大值$2^n - 1$时，就叫做**最长二元n级线性移位寄存器序列**，简称为**m序列**。

序列的生成多项式 I

定理 6

设$\mathbf{a}$是$\mathbf{F}_2$上的一个周期序列，$f(x), g(x)$是$\mathbf{F}_2$上的两个多项式。如果$\mathbf{a} \in G(f)$且$\mathbf{a} \in G(g)$，则$\mathbf{a} \in G(f \pm g)$。

证明： 设

$$\mathbf{a} = (a_0, a_1, a_2, \cdots)$$

$$f(x) = \sum_{i=0}^{n} c_i x^i$$

$$g(x) = \sum_{i=0}^{m} d_i x^i$$

序列的生成多项式 II

其中$c_0 \neq 0, d_0 \neq 0$. 那么

$$c_n a_k + c_{n-1} a_{k-1} + \cdots + c_0 a_{k-n} = 0, k \geq n$$

$$d_m a_k + d_{m-1} a_{k-1} + \cdots + d_0 a_{k-m} = 0, k \geq m$$

取$M = \max(m, n)$, 并令

$$c_{n+1} = c_{n+2} = \cdots = c_M = 0, \text{如果} M > n$$

$$d_{m+1} = d_{m+2} = \cdots = d_M = 0, \text{如果} M > m$$

于是有

$$(c_M \pm d_M) a_k + (c_{M-1} \pm d_{M-1}) a_{k-1} + \cdots + (c_M \pm d_M) a_{k-M} = 0, k \geq M$$

序列的生成多项式 III

由这个递推关系确定的多项式为

$$\sum_{i=0}^{M}(c_i \pm d_i)x^i = f(x) \pm g(x)$$

所以

$$\mathbf{a} \in G(f \pm g).$$

□

定理 7

设$\mathbf{a}$是$\mathbf{F}_2$上的一个周期序列，$f(x)$是$\mathbf{F}_2$上的一个多项式。如果$\mathbf{a} \in G(f)$，则对$\mathbf{F}_2$上的任何多项式$h(x)$，一定有$\mathbf{a} \in G(f \cdot h)$。

证明：

$$\mathbf{a} = (a_0, a_1, a_2, \cdots)$$

序列的生成多项式 IV

$$f(x) = \sum_{i=0}^{n} c_i x^{n-i}$$

$$h(x) = \sum_{i=0}^{m} c_i x^{m-i}$$

有定理假设知

$$c_0 a_k + c_1 a_{k-1} + \cdots + c_n a_{k-n} = 0, k \geq n$$

取

$$d_0 = c_0, d_1 = c_1, \cdots, d_n = c_n, d_{n+1} = 0$$

那么

$$d_0 a_k + d_1 a_{k-1} + \cdots + d_n a_{k-n} + d_{n+1} a_{k-(n+1)} = 0, k \geq n+1$$

序列的生成多项式 V

这个递推关系式所确定的多项式为

$$\sum_{i=0}^{n+1} d_i x^{n+1-i} = \sum_{i=0}^{n} c_i x^{n+1-i} = xf(x)$$

因此$\mathbf{a} \in G(xf(x))$。 从而利用数学归纳法可以证明对任何的i, $\mathbf{a} \in G(x^i f(x))$。所以根据定理6,

$$\mathbf{a} \in G(\sum_{i=0}^{m} c_i(x^{m-i} f(x))) = G(f(x)\sum_{i=0}^{m} c_i x^{m-i}) = G(f(x)h(x)).$$

□

序列的生成多项式 VI

定理 8

设$\mathbf{a}$是$\mathbf{F}_2$上的一个周期序列，那么存在着$\mathbf{F}_2$上的一个多项式$f(x)$ 具有性质：$\mathbf{a} \in G(f)$且$\mathbf{a} \in G(h(x))$当且仅当$f(x)|h(x)$。 更进一步，适合上述性质的多项式$f(x)$是唯一确定的，而且如果$\mathbf{a}$是非0周期序列， 那么$\deg(f(x)) \geq 1$。

证明： 令$S = \{t(x)|\mathbf{a} \in G(t(x))\}$，则$S$是一个非空集合，这是因为如果$\mathbf{a}$的周期 是$l$, $x^l + 1$一定在S中。由定理6和定理7知S是$\mathbf{F}_2[x]$ 中的一个理想，因此S是一主理想。设$f(x)$是S的一个生成元，则 $h(x) \in S \Leftrightarrow f(x)|h(x)$，因此$\mathbf{a} \in G(h(x)) \Leftrightarrow f(x)|h(x)$. 进一步，如果$\mathbf{a}$还是非零序列，则非零常数多项式不在$S$中，从而$\deg(f(x)) \geq 1$. 定理得证。 $\square$

序列的极小多项式 I

定义 6

设$\mathbf{a}$是$\mathbf{F}_2$上的一个周期序列，那么根据定理8, 存在$\mathbf{F}_2$ 上唯一的首一多项式$f(x)$使得$\mathbf{a} \in G(h(x))$当且仅当$f(x)|h(x)$。这个多项式$f(x)$称作 $\mathbf{a}$的极小多项式。

定理 9

任给$\mathbf{F}_2$上一个多项式$f(x)$, 则必有$\mathbf{F}_2$上的一个周期序列存在，它以$f(x)$为极小 多项式。

证明： 考察$G(f)$中由初始状态$(\underbrace{0,0,\cdots,0}_{\text{全为0}},1)$产生的序列$\mathbf{a}$。显然$\mathbf{a} \neq \mathbf{0}$。假设$\mathbf{a}$的极 小多项式是$h(x) \neq f(x)$, 则有$h(x)|f(x)$，所以$\deg h(x) < \deg f(x)$。这样$\mathbf{a}$将是$G(h)$中从0状态 得到的序列，因而是0序列。这与$\mathbf{a} \neq 0$矛盾，所以$\mathbf{a}$是以$f(x)$为极小多项式的序列。 □

序列的极小多项式 II

定理 10

设$f(x)$是$\mathbf{F}_2$上的一个次数大于1的多项式，那么以$f(x)$为极小多项式的线性移位寄存器序列的周期就等于$f(x)$的周期。

证明： 设$\mathbf{a}$是一以$f(x)$为极小多项式的 线性移位寄存器序列，周期为ν，那么$\mathbf{a} \in G(x^\nu - 1)$。 因$f(x)$是$\mathbf{a}$的极小多项式，所以$f(x)|(x^\nu - 1)$, $p(f)|\nu$。 另一方面，根据多项式周期的定义，我们有$f(x)|(x^{p(f)} - 1)$, 所以 $\mathbf{a} \in G(x^{p(f)} - 1)$, $\nu | p(f)$。因此$\nu = p(f)$。 □

推论 1

设$f(x)$是$\mathbf{F}_2$上的一个次数$n \geq 1$的多项式，那么对于$G(f)$中任一非零的n级线性移位寄存器序列$\mathbf{a}$，都有$p(\mathbf{a})|p(f)$。

证明： 设$\mathbf{a}$的极小多项式是$h(x)$，则$p(\mathbf{a}) = p(h)$且$h(x)|f(x)$。但$f(x)|(x^{p(f)} - 1)$，所以$h(x)|(x^{p(f)} - 1)$，于是$p(h)|p(f)$，所以$p(\mathbf{a})|p(f)$。 □

推论 2

设$f(x)$是$\mathbf{F}_2$上的不可约多项式，那么$G(f)$中任意非0线性 移位寄存器序列均以$f(x)$为极小多项式，而且他们的周期都等于$f(x)$的周期。

推论 3

设$f(x)$是$\mathbf{F}_2$上的n次多项式，则一个非0序列$\mathbf{a} \in G(f)$是 n级m-序列当且仅当$f(x)$是n次本原多项式。

证明：
充分性：设$f(x)$是本原多项式，则其周期为$2^n - 1$。根据推论2，$\mathbf{a}$的周期为$2^n - 1$，因而是m-序列。
必要性：如果$\mathbf{a}$是n级m-序列，则$\mathbf{a}$的周期为$2^n - 1$，所以$G(f)$中的 每一非0序列都以$\mathbf{a}$的某一状态为初始状态，因而其周期必为$2^n - 1$。 下面我们首先证明$f(x)$是不可约多项式。假设$f(x)$可约，$h(x)$是它的一个不可约因子， $\deg(h(x)) = k < n$，则$p(h) \le 2^k - 1 < 2^n - 1$，因此$G(h)$中的非0序列的周期为$p(h) < 2^n - 1$。 但由$h(x)|f(x)$知，$G(h) \subset G(f)$，即$G(h)$中的序列也在$G(f)$中，因此$G(h)$中的 非0序列的周期也应为$2^n - 1$。这与$p(h) < 2^n - 1$矛盾。所以$f(x)$是不可约的。从而 根据推论2知$f(x)$的 周期应为$2^n - 1$，所以是本原多项式。 □

序列的迹表达式 I

定理 11

设$\mathbf{a} = (a_0, a_1, a_2, \cdots)$是$\mathbf{F}_2$上的一个周期为$2^n - 1$的$m$-序列，它的极小多项式$f(x)$是 n次本原多项式。再设α是$f(x)$的任意一根，那么总有$\beta \in \mathbf{F}_{2^n}^*$使

$$a_k = Tr(\beta\alpha^k) = \sum_{j=0}^{n-1} (\beta\alpha^k)^{q^j}, k \geq 0.$$

反之，设$f(x)$是一n次本原多项式，α是它的任意一个根，那么对任意的$\beta \in \mathbf{F}_{2^n}^*$，

$$(Tr(\beta), Tr(\beta\alpha), Tr(\beta\alpha^2), \cdots,)$$

都是$G(f)$中的m-序列，而且这样就得到$G(f)$中全部非零序列。

序列的迹表达式 II

证明：先设$\mathbf{a}=(a_0,a_1,a_2,\cdots)$是$\mathbf{F}_2$上的一个周期为$2^n-1$的$m$-序列，它的极小多项式$f(x)$是 n次本原多项式, α是$f(x)$的任意一根。那么

$$\alpha,\alpha^q,\cdots,\alpha^{q^{n-1}}$$

就是$f(x)$的所有根，且都是$\mathbf{F}_{2^n}$的本原元. 写

$$f(x)=x^n+\sum_{i=1}^{n}c_ix^{n-i}.$$

则

$$f(\alpha)=\alpha^n+\sum_{i=1}^{n}c_i\alpha^{n-i}=0,$$

从而

$$\alpha^k+\sum_{i=1}^{n}c_i\alpha^{k+n-i}=0,k\geq n.$$

序列的迹表达式 III

这就是说，序列

$$(1, \alpha, \alpha^2, \alpha^3, \cdots)$$

满足线性递推关系式(3). 因此对任意$\beta \in \mathbf{F}_{2^n}$, 序列

$$(\beta, \beta\alpha, \beta\alpha^2, \beta\alpha^3, \cdots)$$

也满足线性递推关系式(3). 那么对任意$j, 0 \le j \le n-1$, 序列

$$(\beta^{q^j}, (\beta\alpha)^{q^j}, (\beta\alpha^2)^{q^j}, (\beta\alpha^3)^{q^j}, \cdots) \tag{8}$$

也满足线性递推关系式(3). 将上式对$j = 0, 1, \cdots, n-1$求和，所得序列

$$(Tr(\beta), Tr(\beta\alpha), Tr(\beta\alpha^2), \cdots,) \tag{9}$$

序列的迹表达式 IV

自然也满足线性递推关系式(3), 而且都是$\mathbf{F}_2$上的序列， 因此都属于$G(f)$. 下面 我们证明当β跑遍$\mathbf{F}_{2^n}$的所有元素时，(9)正好得到2^n个两两不同的 序列。设有$\beta_1, \beta_2 \in \mathbf{F}_{2^n}$使

$$Tr(\beta_1\alpha^k) = Tr(\beta_2\alpha^k),\ k \geq 0.$$

令$\beta_0 = \beta_1 - \beta_2$, 那么从上式可以推出

$$Tr(\beta_0\alpha^k) = 0,\ k \geq 0.$$

由于$1, \alpha, \alpha^2, \cdots, \alpha^{n-1}$正好构成$\mathbf{F}_{2^n}$在 $\mathbf{F}_2$上的一组基，所以由β_0定义的线性变换$L_{\beta_0} = 0$,所以有$\beta_0 = 0$, 从而$\beta_1 = \beta_2$. 这样我们就证明了，当β跑遍$\mathbf{F}_{2^n}$的所有元素时，(9)正好得到$G(f)$中全部2^n个序列。
当$\beta = 0$时， 我们得到零序列；当β跑遍$\mathbf{F}_{2^n}^*$时， 我们就得到$G(f)$中全部$2^n - 1$个非零序列。这实际上我们证明了定理的后半部。令$\mathbf{a} \in G(f)$, 故有$\beta \in \mathbf{F}_{2^n}^*$ 使(9)产生的序列就是$\mathbf{a}$. 定理的前半部分也得证。 □

左移变换 I

定义 7

设$\mathbf{a}=(a_0,a_1,a_2,\cdots)$是二元序列，定义作用在$\mathbf{a}$上的**左移变换**$L$如下：

$$L(\mathbf{a})=(a_1,a_2,\cdots)$$

并定义

$$L^0(\mathbf{a})=\mathbf{a}$$

$$L^t(\mathbf{a})=L(L^{t-1}(\mathbf{a}))$$

这里t是个正整数。

左移变换 II

定理 12

设有$\mathbf{F}_2$上首项系数为1的n次多项式:

$$f(x) = x^n + \sum_{i=1}^{n} c_i x^{n-i},$$

$\mathbf{a}$是$G(f)$中的非零序列。如果$\mathbf{a}$是m-序列，那么$\mathbf{a}$的左移都是$G(f)$中的m-序列，且

$$\mathbf{a}, L(\mathbf{a}), L^2(\mathbf{a}), \cdots, L^{2^n-2}(\mathbf{a})$$

正好是$G(f)$中的全部非零序列。更进一步，$\mathbf{a}$的任何连续2^n-1个状态恰是$\mathbf{F}_2$上 所有2^n-1个两两不同的非零n维向量的全体。

左移变换 III

推论 4

设有$\mathbf{F}_2$上首项系数为1的n次多项式:

$$f(x) = x^n + \sum_{i=1}^{n} c_i x^{n-i},$$

$\mathbf{a}$是$G(f)$中的非零序列。如果$\mathbf{a}$是m-序列，$t_1 > t_2 > 0$, 那么当$2^n - 1 \nmid t_1 - t_2$时，

$$L^{t_1}(\mathbf{a}) + L^{t_2}(\mathbf{a})$$

也是$G(f)$中的m-序列。

证明：根据m-序列的定义，当$2^n - 1 \nmid t_1 - t_2$时，$L^{t_1}(\mathbf{a})$和$L^{t_2}(\mathbf{a})$的 初始状态不等，因此$L^{t_1}(\mathbf{a}) + L^{t_2}(\mathbf{a}) \neq 0$. 但根据定理??, $G(f)$ 是$\mathbf{F}_2$上的向量空间，因此$L^{t_1}(\mathbf{a}) + L^{t_2}(\mathbf{a})$是$G(f)$中的非零序列。再 根据定理12知，$L^{t_1}(\mathbf{a}) + L^{t_2}(\mathbf{a})$也是$G(f)$中的$m$-序列。 □

移位相加特性 I

定理 13 (移位相加特性)

设$\mathbf{a}$是一周期为l的二元周期序列。如果对任意的$0 \le i, j \le l-1$, $L^i(\mathbf{a}) + L^j(\mathbf{a}) = 0$或$L^k(\mathbf{a})$(对某一$0 \le k \le l-1$), 那么, 存在一整数$n$, 使$l = 2^n - 1$, 而$\mathbf{a}$是周期为$2^n - 1$的$m$-序列。

证明：考察集合

$$V = \{L^i(\mathbf{a}|i = 0, 1, 2, \cdots, l-1\} \bigcup \{0\}$$

则$|V| = l + 1$. 定义

$$0 \cdot L^i(\mathbf{a}) = \mathbf{0}, 1 \cdot L^i(\mathbf{a}) = L^i(\mathbf{a}), i = 0, 1, 2, \cdots, l-1$$

由于$L^i(\mathbf{a}) + L^j(\mathbf{a}) = \mathbf{0}$或$L^k(\mathbf{a})$(对某$0 \le k \le l-1$), 因此可知$V$构成$\mathbf{F}_2$上的向量空间。令 $n = \dim_{\mathbf{F}_2}(V)$, 则$|V| = 2^n$, 所以$l = 2^n - 1$.

移位相加特性 II

再考察V中的向量$L^0(\mathbf{a}), L^1(\mathbf{a}), \cdots, L^n(\mathbf{a})$. 由于$V$是$n$维向量空间，因此$L^0(\mathbf{a}), L^1(\mathbf{a}), \cdots, L^n(\mathbf{a})$是线性相关的，从而一定存在不全为零的数$c_0, c_1, \cdots, c_n \in F_2$, 使得

$$c_0 L^n(\mathbf{a}) + c_1 L^{n-1}(\mathbf{a}) + \cdots + c_n L^0(\mathbf{a}) = \mathbf{0}.$$

查看上述向量等式中的每个分量，我们可得到

$$c_0 a_{n+k} + c_1 a_{n+k-1} + \cdots + c_n a_k = 0, k = 0, 1, 2, \cdots$$

所以$\mathbf{a}$是个n级线性移位寄存器序列（注意，$c_0 \neq 0$, 因为小于n级的线性移位寄存器序列周期不能达到$2^n - 1$）。所以，$\mathbf{a}$是m-序列。 □

定理 14

设$\mathbf{a}$是$\mathbf{F}_2$上周期为$2^n - 1$的m-序列，那么1在$\mathbf{a}$的一个周期中恰出现2^{n-1}次，而0在$\mathbf{a}$的一个周期中恰出现$2^{n-1} - 1$次。

m-序列的相关特性 I

定理 15

设$\mathbf{a}$是$\mathbf{F}_2$上周期为2^n-1的m-序列, 那么

$$c_{\mathbf{a}}(t)=\begin{cases} 2^n-1, & \text{如果 } t\equiv 0 \bmod (2^n-1), \\ -1, & \text{如果 } t\not\equiv 0 \bmod (2^n-1). \end{cases}$$

证明： 首先我们有

$$c_{\mathbf{a}}(t)=\sum_{i=0}^{2^n-2}\eta(a_i)\eta(a_{i+t})=\sum_{i=0}^{2^n-2}\eta(a_i+a_{i+t}).$$

且$(a_0+a_t, a_1+a_{1+t}, a_2+a_{2+t},\cdots)=L^0(\mathbf{a})+L^t(\mathbf{a})$。

m-序列的相关特性 II

若$t \not\equiv 0 \bmod (2^n-1)$，则由根据推论4知，$L^0(\mathbf{a})+L^t(\mathbf{a})$也 是周期为$2^n-1$的$m$-序列．所以根据定理14，在$L^0(\mathbf{a})+L^t(\mathbf{a})$的一个周期中，1的个数应比0的个数多1，因此

$$c_{\mathbf{a}}(t)=\sum_{i=0}^{2^n-2}\eta(a_i+a_{i+t})=-1$$

若$t \equiv 0 \bmod (2^n-1)$，则$L^0(\mathbf{a})+L^t(\mathbf{a})=\mathbf{0}$，所以有

$$c_{\mathbf{a}}(t)=\sum_{i=0}^{2^n-2}\eta(a_i+a_{i+t})=2^n-1$$

定理证毕。 □

m-序列的伪随机特性 I

定理 16

设$\mathbf{a}$是一个周期等于2^n-1的二元m-序列。那么在$\mathbf{a}$的一个周期中， 游程的总数等于2^{n-1}，其中0游程的个数等于1游程的个数都等于2^{n-2}，长度为$r(1\leq r<n-1)$的0游程的个数和1游程的个数都等于2^{n-2-r}，长为$n-1$的0游程的个数等于1，长为$n-1$的1游程的个数等于0，长为n的0游程的个数等于0，长为n的1游程的个数等于1， 长度大于 n的游程的个数均为0。

证明： 设$1\leq r<n-1$。那么在$\mathbf{a}$的一个周期中长为r的0游程的个数就等于$\mathbf{a}$中 形如

$$1\quad \underbrace{00\cdots0}_{\text{r个0}}\quad 1\quad b_{r+3}b_{r+4}\cdots b_n, b_i\in\mathbf{F}_2$$

的状态的个数，而后者显然等于2^{n-2-r}。同理，$\mathbf{a}$中长为r的1游程的个数 也等于2^{n-2-r}。

m-序列的伪随机特性 II

根据定理??，**a**的一个周期中，游程的总数为

$$2 \cdot \frac{2^n - 1 + 1}{4} = 2^{n-1}$$

所以，**a**的一个周期中长$\geq n-1$的游程只有2个。因

$$\underbrace{1 \cdots 1}_{\text{n个1}}$$

是**a**中的一个状态，而**a**的每个状态在**a**的一个周期中只出现一次，所以这个状态 之前与之后的符号都是0，因此**a**恰有一个长为n的1游程。又因

$$\underbrace{0 \cdots 0}_{\text{n个0}}$$

m-序列的伪随机特性 III

不是**a**的状态，所以**a**不能有长为n的0游程。又

$$1\underbrace{0\cdots0}_{\text{n-1个0}}$$

是**a**的一个状态。显然，这个状态之后的符号必为1。这样

$$1\underbrace{0\cdots0}_{\text{n-1个0}}1$$

就是**a**的一个长为$n-1$的0游程。由此推出**a**没有长为$n-1$的1游程和 长度$\geq n$的游程。 □

本节概要

线性综合解 I

设N是一个正整数，$a^{(N)} = a_0a_1\cdots a_{N-1}$是有限域$F_q$上的一个长度为$N$的序列，$f_N(x)$是一个能产生 $a^{(N)}$且阶数最小的线性移位寄存器的联接多项式，l_N是该线性移位寄存器的阶数。我们称二元组$(f_N(x), l_N)$为$a^{(N)}$的线性综合解。应当指出，$\deg(f_N(x)) \leq l_N$, 这是因为产生$a^{(N)}$且阶数最小的线性移位寄存器 可能是退化的，这时我们有$\deg(f_N(x)) < l_N$. 另外我们约定0阶线性移位寄存器的联接多项式为$f(x) = 1$，且长度为$n(\leq N)$的零序列$\underbrace{000\cdots0}_{n}$由0阶的线性移位寄存器产生。

Berlekamp-Massey算法 I

(1) 设n_0是一个非负整数，满足

$$a_0 = a_1 = \cdots = a_{n_0-1} = 0, a_{n_0} \neq 0$$

我们则取

$$d_0 = d_1 = \cdots = d_{n_0-1} = 0, d_{n_0} = a_{n_0},$$

并令

$$f_1(x) = f_2(x) = \cdots = f_{n_0}(x) = 1$$

$$l_1 = l_2 = \cdots = l_{n_0} = 0$$

同时可以取任意一个n_0+1级的线性移位寄存器作为$(f_{n_0+1}(x), l_{n_0+1})$，但为了确定起见，我们令

$$f_{n_0+1}(x) = 1 - d_{n_0}x^{n_0+1},\ l_{n_0+1} = n_0 + 1.$$

Berlekamp-Massey算法 II

(2) 假设$(f_i(x), l_i)$, $1 \le i \le n \le N$已经求得。而

$$l_1 = l_2 = \cdots = l_{n_0} < l_{n_0+1} \le l_{n_0+1} \le \cdots \le l_n$$

令 $f_n(x) = 1 + c_{n,1}x + \cdots + c_{n,l_n}x^{l_n}$, 并计算

$$d_n = a_n + c_{n,1}a_{n-1} + \cdots + c_{n,l_n}a_{n-l_n}.$$

如果$d_n = 0$, 则取

$$f_{n+1}(x) = f_n(x), \; l_{n+1} = l_n$$

如果$d_n \ne 0$, 这时一定存在$1 \le m < n$, 使

$$l_m < l_{m+1} = l_{m+2} = \cdots = l_n,$$

取

$$f_{n+1}(x) = f_n(x) - d_n d_m^{-1} x^{n-m} f_m(x),$$
$$l_{n+1} = \max\{l_n, n+1-l_n\}.$$

BM算法举例 I

求周期为8的序列$a^{(8)} = 00101101$的线性综合解。在这儿$a_0 = 0, a_1 = 0, a_2 = 1, a_3 = 0, a_4 = 1, a_5 = 1, a_6 = 0, a_7 = 1$.
首先$n_0 = 2$，因此

$$d_0 = d_1 = 0, d_2 = 1$$

$$f_1(x) = f_2(x) = 1, f_3(x) = 1 - x^3$$

$$l_1 = l_2 = 1, l_3 = 3$$

计算$d_3 = a_3 - a_0 = 0 + 0 = 0$，因此取

$$f_4(x) = f_3(x) = 1 - x^3, l_4 = l_3 = 3$$

计算$d_4 = a_4 - a_1 = 1 - 0 = 1 \neq 0$，这时$l_2 < l_3 = l_4$，因此$m = 2$,

$$f_5(x) = f_4(x) - d_4 d_2^{-1} x^{4-2} f_2(x) = 1 - x^3 - x^2 = 1 + x^2 + x^3$$

$$l_5 = \max\{l_4, 4 + 1 - l_4\} = 3$$

BM算法举例 II

计算$d_5 = a_5 + a_3 + a_2 = 1 + 0 + 1 = 0$, 因此

$$f_6(x) = f_4(x) = 1 + x^2 + x^3, l_6 = l_5 = 3$$

计算$d_6 = a_6 + a_4 + a_3 = 0 + 1 + 0 = 1 \neq 0$, 这时$l_2 < l_3 = l_4 = l_5$, 因此$m = 2$,

$$f_7(x) = f_6(x) - d_6 d_2^{-1} x^{6-2} f_2(x) = 1 + x^2 + x^3 - x^4 = 1 + x^2 + x^3 + x^4$$

$$l_7 = \max\{l_6, 7 - l_6\} = 4$$

计算$d_7 = a_7 + a_5 + a_4 + a_3 = 1 + 1 + 1 + 0 = 1 \neq 0$, 这时$l_6 < l_7$, 因此$m = 6$,

$$f_8(x) = f_7(x) - d_7 d_6^{-1} x^{7-6} f_6(x) = 1 + x^2 + x^3 + x^4 - x(1 + x^2 + x^3) = 1 + x + x^2$$

$$l_8 = \max\{l_7, 8 - l_7\} = 4$$

所以$a^{(8)} = 00101101$的线性综合解为$(1 + x + x^2, 4)$.

引理 1

设有一个长为$n+1$的二元序列

$$a_0a_1\cdots a_{n-1}a_n \tag{10}$$

并假定线性移位寄存器$(f(x), l)$能产生上述序列的前n项，但不能产生整个序列。那么任何一个能产生 整个序列的线性移位寄存器的级数l'一定满足

$$l' \geq n+1-l$$

证明：我们说$l \leq n$，因为如果$l > n$时，则任一l级线性移位寄存器都能产生序列(10), 这与定理的 条件不符。
当$l = n$时，$n+1-l = 1$。因此我们只要证$l' \neq 0$即可。如若$l' = 0$, 则序列(10)必须是零序列，因此$(f(x), l)$ 从全零状态出发也能产生(10)，与定理的假设矛盾。因此$l' \neq 0$, 从而$l' \geq 1$.

现在假设$l < n$，并假定$(f_1(x), l')$能产生(10)，设

$$f(x) = 1 + c_1 x + \cdots + c_l x^l,$$

$$f_1(x) = 1 + c'_1 x + \cdots + c'_{l'} x^{l'}$$

那么

$$-\sum_{i=1}^{l} c_i a_{k-i} = \begin{cases} = a_k, & \text{如果} k = l, l+1, \cdots, n-1; \\ \neq a_n, & k = n. \end{cases} \tag{11}$$

$$-\sum_{i=1}^{l'} c'_i a_{k-i} = a_k, \quad \text{如果} k = l', l'+1, \cdots, n-1, n. \tag{12}$$

如果本引理不成立，那么$l' \leq n - l$，这时$\{a_{n-l}, a_{n-l+1}, \cdots, a_{n-1}\}$就是$\{a_{l'}, a_{l'+1}, \cdots, a_{n-1}, a_n\}$的子集。于是利用(12)式有

$$-\sum_{i=1}^{l} c_i a_{n-i} = \sum_{i=1}^{l} c_i \sum_{j=1}^{l'} c'_j a_{n-i-j} = \sum_{j=1}^{l'} c'_j \sum_{i=1}^{l} c_i a_{n-i-j}$$

从$l' \leq n-l$我们又推出$l \leq n-l'$，因此$\{a_{n-l'}, a_{n-l'+1}, \cdots, a_{n-1}\}$就是$\{a_l, a_{l+1}, \cdots, a_{n-1}, a_n\}$的子集。于是利用(11)式我们又可将上式右侧写成

$$\sum_{j=1}^{l'} c_j' \sum_{i=1}^{l} c_i a_{n-i-j} = -\sum_{j=1}^{l'} c_i' a_{n-j}$$

再根据(12)式，

$$-\sum_{j=1}^{l'} c_i' a_{n-j} = a_n$$

因此

$$-\sum_{i=1}^{l} c_i a_{n-i} = a_n$$

这就是说，$(f(x), l)$可以产生序列(10)，这与引理的假设相矛盾。因此$l' \leq n-l$这一假设不成立。所以一定有 $l' \geq n+1-l$. 定理证毕。 □

定理 17

设$a^{(N)} = (a_0a_1\cdots a_{N-1})$是一个长为$N$的二元序列，那么按*Berlekamp-Massey*算法求出的每一个$(f_n(x), l_n)(n = 1, 2, \cdots, N)$确实是产生$a^{(N)}$前$n$项的最短的线性移位寄存器。

证明：我们对n用数学归纳法来证明定理。当$n \le n_0$时，定理显然成立。现在我们考察$n = n_0 + 1$情况。根据 算法，

$$f_{n_0+1}(x) = 1 - d_{n_0}x^{n_0+1}, l_{n_0+1} = n_0 + 1$$

显然，$n_0 + 1$级的线性移位寄存器$(f_{n_0+1}(x), l_{n_0+1})$能产生$a^{(N)}$的前$n_0 + 1$项，而且级数$< n_0 + 1$的 任一线性移位寄存器都不能产生$a^{(N)}$的前$n_0 + 1$项，这是因为当一个序列的前$n_0$全为零时，级数不大于$n_0$的 线性移位寄存器只能产生0序列。所以当$n = n_0 + 1$时，定理成立。

现在假设$n_0 \le n < N$, 而且定理对于$1, 2, \cdots, n$都成立，我们要证定理对于$n+1$也成立。根据算法，我们知道

$$l_1 \le l_2 \le \cdots \le l_n$$

而且

$$d_n = a_n + c_{n,1}a_{n-1} + \cdots + c_{n,l_n}a_{n-l_n}$$

如果$d_n = 0$, 那么根据算法，

$$f_{n+1}(x) = f_n(x), \; l_{n+1} = l_n$$

根据归纳假设，$(f_n(x), l_n)$是产生$a^{(N)}$的前n项的一个最短的线性移位寄存器。又因为$d_n = 0$, 所以$(f_n(x), l_n)$ 也产生$a^{(N)}$的前$n+1$项，因此$(f_{n+1}(x), l_{n+1}) = (f_n(x), l_n)$也是产生$a^{(N)}$的前$n+1$项的一个最短的线性 移位寄存器。

如果$d_n \ne 0$, 那么一定存在$1 \le m < n$, 使

$$l_m < l_{m+1} = l_{m+2} = \cdots = l_n,$$

这时我们一定

有$d_m \neq 0$且$l_n = l_{m+1} = \max\{l_m, m+1-l_m\} = m+1-l_m$，从

而$l_m = m+1-l_n$. 根据算法

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

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

所以

$$\begin{aligned}\deg(f_{n+1}(x)) &\leq \max\{\deg(f_n(x)), n-m+\deg(f_m(x))\} \\ &\leq \max\{l_n, n-m+l_m\} \\ &\leq \max\{l_n, n+1-l_n)\} \\ &= l_{n+1}\end{aligned}$$

这就证明了(f_{n+1}, l_{n+1})是个l_{n+1}级线性移位寄存器。又因为$(f_n(x), l_n)$产生$a^{(N)}$的前n项，而$(f_m(x), l_m)$ 产生$a^{(N)}$的前m项，因此

$$a_k + \sum_{i=1}^{l_{n+1}} c_{n+1,i} a_{k-i} = a_k + \sum_{i=1}^{l_n} c_{ni} a_{k-i} - d_n d_m^{-1}(a_{k-n+m} + \sum_{i=1}^{l_m} c_{mi} a_{k-n+m-i})$$

$$= \begin{cases} 0, & \text{如} k = l_{n+1}, l_{n+1}+1, \cdots, n-1; \\ d_n - d_n d_m^{-1} d_m = 0, & \text{如果} k = n. \end{cases}$$

这就证明了$(f_{n+1}(x), l_{n+1})$产生$a^{(N)}$的前$n+1$项. 但因$d_n \neq 0$, $(f_n(x), l_n)$不能产生$a^{(N)}$的前$n+1$项, 所以根据定理1, 产生 $a^{(N)}$的前$n+1$项的最短线性移位寄存器的级数一定$\geq \max\{l_n, n+1-l_n\}$. 所以 $(f_{n+1}(x), l_{n+1})$是产生$a^{(N)}$的前$n+1$项的最短线性移位寄存器。
这就证明了$n+1$定理也成立。因此定理成立。 □

定理 18

设$a^{(N)} = (a_0 a_1 \cdots a_{N-1})$是一个长为$N$的二元序列，$(f_N(x), l_N)$是它的一个线性综合解。那么这个解是唯一线性综合解的充要条件是$2l_N \le N$.

证明：先假定$2l_N > N$, $(f_N(x), l_N)$是用Berlekamp-Massy算法求出的产生$a^{(N)}$的最短的线性移位寄存器。写

$$f_N(x) = 1 + c_1 + \cdots + c_{l_N} x^{l_N}$$

那么

$$a_k = -\sum_{i=1}^{l_N} c_i a_{k-i}, k = l_N, l_{N+1}, \cdots, N-1$$

任选

$$a_N \neq -\sum_{i=1}^{l_N} c_i a_{N-i}$$

那么$(f_N(x), l_N)$不能产生序列长为$N+1$的序列

$$(a_0 a_1 \cdots a_{N-1} a_N)$$

因此用Berlekamp-Massy算法求出的产生上述序列的最短线性移位寄存器$(f_{N+1}(x), l_{N+1}) \neq (f_N(x), l_N)$，而且 $l_{N+1} = \max\{l_N, N+1-l_N\}$. 我们说$l_{N+1} = l_N$，因为否则的话，$\max\{l_N, N+1-l_N\} = l_{N+1} > l_N$，所以 $l_N < N+1-l_N$，因此$2l_N < N+1$，即$2l_N \leq N$. 这与假定相矛盾。所以$(f_{N+1}(x), l_{N+1})$也是一个 $l_{N+1} = l_N$级的能够产生$a^{(N)}$的线性移位寄存器。因此产生$a^{(N)}$的最短线性移位寄存器不唯一。
我们再假定$2l_N \leq N$. 我们要证明这时产生$a^{(N)}$的最短线性移位寄存器是唯一的。设$(f(x), l_N)$和 $(f_1(x), l_N)$都是产生$a^{(N)}$的最短的线性移位寄存器，我们要证明$f(x) = f_1(x)$. 写

$$f(x) = 1 + c_1 x + \cdots + c_{l_N} x^{l_N},$$

$$f_1(x) = 1 + c_1' x + \cdots + c_{l_N}' x^{l_N}$$

那么

$$a_k = -\sum_{i=1}^{l_N} c_i a_{k-i} = \sum_{i=1}^{l_N} c_i' a_{k-i}, k = l_N, l_{N+1}, \cdots$$

归纳地定义

$$a_k = -\sum_{i=1}^{l_N} c_i a_{k-i}, k = N, N+1, \cdots ,$$

那么$(f(x), l_N)$也产生无限序列

$$a_0, a_1, a_2, \cdots , a_{N-1}, a_N, a_{N+1}, \cdots \tag{13}$$

而且对任给的$n \geq N$, $(f(x), l_N)$也是产生(13)的前n项的一个最短线性移位积存器。因此如果用 l_n表示产生(13)的前n项的最短线性移位寄存器的级数，那么

$$l_n = l_{n+1} = l_{n+2} = \cdots \tag{14}$$

下面我们用反证法证明$(f_1(x), l_N)$也产生(13). 假设不然的话，我们用N'表示最小的正整数使得 $(f_1(x), l_N)$能产生(13)的前N'项，但不能产

生(13)的前$N'+1$项. 那么显然有$N' \geq N$, 且根据引理1, $l_{N'+1} \geq N'+1-l'_N$. 于是从$2l_N \leq N$推出$2l_{N'} = 2l_N \leq N \leq N'$. 所以$l_{N'} \leq N'-l_{N'} < N'+1-l_{N'} \leq l_{N'+1} = l_N$. 这与 (14)相矛盾。因此$(f_1(x), l_N)$一定也产生(13).

下面我们再证明$f(x) = f_1(x)$. 仍然用反证法。假定$f(x) \neq f_1(x)$. 那么可设m是最小正整数使$c_m \neq c'_m$, 即

$$c_1 = c'_1, c_2 = c'_2, \cdots, c_{m-1} = c'_{m-1}, c_m \neq c'_m \tag{15}$$

自然有$1 \leq m \leq l_N$. 因为$(f(x), l_N)$和$(f_1(x), l_N)$都产生(13), 所以

$$\begin{array}{ll} a_k = -\sum_{i=1}^{l_N} c_i a_{k-i} & \\ a_k = -\sum_{i=1}^{l_N} c'_i a_{k-i} & \end{array} \quad k = l_N, l_N+1, \cdots$$

将以上两式相减, 并注意到(15), 得

$$\sum_{i=m}^{l_N} (c_i - c'_i) a_{k-i} = 0, \ k = l_N, l_N+1, \cdots \tag{16}$$

因此

$$a_{k-m} = -\sum_{i=m+1}^{l_N} (c_m - c'_m)^{-1}(c_i - c'_i)a_{k-i}, \; k = l_N, l_N+1, \cdots$$

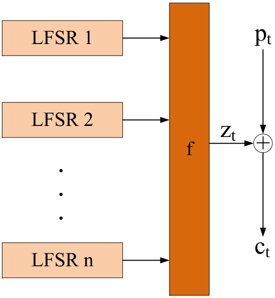
所以

$$a_k = -\sum_{i=1}^{l_N-m} (c_m - c'_m)^{-1}(c_{m+i} - c'_{m+i})a_{k-i}, \; k = l_N - m, l_N - m + 1, \cdots$$

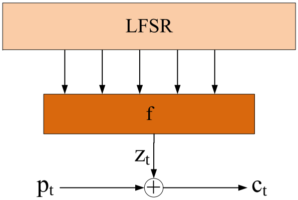
令$g(x) = 1 + \sum_{i=1}^{l_N-m}(c_m - c'_m)^{-1}(c_{m+i} - c'_{m+i})x^i$, 那么 $(g(x), l_N - m)$也产生(13). 特别, $(g(x), l_N)$也产生(13)的前N项，即产生序列$a^{(N)}$. 但$(f(x), l_N)$是产生$a^{(N)}$的一个最短线性移位寄存器，而$l_N - m < l_N$. 这是一个矛盾。所以$f(x) \neq f_1(x)$ 的假设不成立，因此$f(x) = f_1(x)$. 这就证明了产生$a^{(N)}$的最短线性移位寄存器是唯一的。 定理证毕。 □

本节概要

LFSR型流密码结构



组合生成器



滤波生成器

LFSR型流密码的代数攻击

- 建立密钥与密钥流之间的代数方程组

$$\begin{cases} z_0 = f(K) \\ z_1 = f(L(K)) \\ z_2 = f(L^2(K)) \\ \cdots\cdots \\ z_t = f(L^t(K)) \\ \cdots\cdots \end{cases} \tag{17}$$

- 求解方程组获得密钥 K

方程的特点：

- 一般情况下，这是一个超定的方程；
- 每个方程的次数应很高，但又都是一样的；

Toyocrypt流密码的代数攻击

Toyocrypt是2000年由东洋通信机公司设计的一个流密码算法，曾提交给日本密码研究与评估委员会的CRYPTREC项目 参与日本电子政府推荐算法的遴选工作并进入到第二轮. 它的滤波函数为:

$$f(x_1, x_2, \cdots, x_{128}) = x_{128} + \sum_{i=1}^{63} x_i x_{\pi(i+63)} + x_{11}x_{24}x_{33}x_{43} + x_2x_3x_{10}x_{13}x_{19}x_{21}x_{24}x_{26}x_{27}x_{29}x_{34}x_{39}x_{42}x_{43}x_{52}x_{54}x_{59} + \prod_{i=1}^{63} x_i$$

Toyocrypt流密码的代数攻击

Toyocrypt是2000年由东洋通信机公司设计的一个流密码算法，曾提交给日本密码研究与评估委员会的CRYPTREC项目 参与日本电子政府推荐算法的遴选工作并进入到第二轮. 它的滤波函数为:

$$f(x_1,x_2,\cdots,x_{128}) = x_{128} + \sum_{i=1}^{63} x_i x_{\pi(i+63)} + x_{11}x_{24}x_{33}x_{43}+$$

$$x_2x_3x_{10}x_{13}x_{19}x_{21}x_{24}x_{26}x_{27}x_{29}x_{34}x_{39}x_{42}x_{43}x_{52}x_{54}x_{59} + \prod_{i=1}^{63} x_i$$

在提出初期，被认为是一种实际上不可能破译的密码. 但使用代数攻击，2005年被日本情报处理推进机构（IPA）利用其 专用的网格计算系统（128个2GHz的64位CPU）成功破解，用时仅27分钟.

Toyocrypt 的攻击复杂度

令人惊奇的是，Toyocrypt所给出的方程并不像看起来那么困难...

注意到f 的4次项、17次项、63次项都含有公因子 $x_{24}x_{43}$. 因此若令$h(x) = f(x)(x_{24}+1)$, 我们 则有$\deg(h) = 3$, 从而

$$\boldsymbol{f(L^t(K)) = z_t \Rightarrow h(L^t(K)) = z_t(L^t(K)_{24} + 1)}$$

攻击复杂度

LFSR长度=**128**, 方程次数=**3**

- 已知密文：$\frac{1}{2}\binom{128}{3} \approx 2^{18}$
- 空间复杂度：$\binom{128}{3}^2 \approx 2^{37}$
- 时间复杂度：$\binom{128}{3}^\omega \approx 2^{49}$, $\omega = log_2 7 \approx 2.807$

上述攻击的关键...

是否存在 $g(x)$ 使得方程

$$h(x) = f(x)g(x) = z_t g(x)$$

次数较低(即$g(x)$和$h(x)$的次数都很低)?

定理 19

设 f 为 n 元布尔函数, 则存在非零布尔函数 g 使得

- $\deg(g) \leq \frac{n}{2}$
- $\deg(fg) \leq \frac{n+1}{2}$

返回

$$A = \{f(x)x^c : \deg(x^c) \leq \frac{n}{2}\}$$

$$B = \{x^c : \deg(x^c) \leq \frac{n+1}{2}\}$$

$$\left.\begin{array}{l} |A| = \displaystyle\sum_{0 \leq i \leq \frac{n}{2}} \binom{n}{i} \geq 2^{n-1} \\ \\ |B| = \displaystyle\sum_{0 \leq i \leq \frac{n+1}{2}} \binom{n}{i} > 2^{n-1} \end{array}\right\} \Rightarrow |A| + |B| > 2^n$$

$$\Rightarrow A \cup B \text{ 线性相关} \Rightarrow \boldsymbol{f(x)g(x) = h(x)}$$

LFSR型流密码的攻击复杂度

攻击步骤：

1. 寻找$g(x)$使得$g(x)$和$f(x)g(x)$的次数都不超过$\binom{L}{\lceil n/2 \rceil}$.
2. 带入足够多的密钥流比特得到方程：

$$h(L^t(K)) = z_t g(L^t(K)),\ t = 1, 2, 3, \cdots$$

3. 用线性化方法求解得到的方程组.

设LFSR的长度=L, 布尔函数变元的个数=n, 则按上述方法我们可以得到次数至多为$\lceil \frac{n}{2} \rceil$的方程, 因此单项式的个数最多为$\binom{L}{\lceil n/2 \rceil}$

- 已知密文：$\binom{L}{n/2}$
- 空间复杂度：$\binom{L}{n/2}^2$
- 时间复杂度：$\binom{L}{n/2}^\omega$

代数攻击的有效情形

S1 布尔函数f是一个低次的布尔函数;

S3 存在非零布尔函数g, 使得fg是一个低次的布尔函数(标准代数攻击)

代数攻击的有效情形

S1 布尔函数f是一个低次的布尔函数；

S3 存在非零布尔函数g, 使得fg是一个低次的布尔函数(标准代数攻击)

情形	$\deg(f)$	$\deg(g)$	**$\deg(fg)$**	方程	z_t	方程数
S1	低	$g=1$	低	$f(s)=z_t$	0/1	m
S3a	高	低	低	$f(s)g(s)=z_tg(s)$	0/1	m
S3b	高	低	$\boldsymbol{fg=0}$	$g(s)=0$	1	$m/2$
S3c	高	高	低	$f(s)g(s)=0$	0	$m/2$

代数攻击的有效情形

S1 布尔函数f是一个低次的布尔函数;

S3 存在非零布尔函数g, 使得fg是一个低次的布尔函数(标准代数攻击)

情形	$\deg(f)$	$\deg(g)$	**$\deg(fg)$**	方程	z_t	方程数
S1	低	$g=1$	低	$f(s)=z_t$	0/1	m
S3a	高	低	低	$f(s)g(s)=z_tg(s)$	0/1	m
S3b	高	低	$\boldsymbol{fg=0}$	$g(s)=0$	1	$m/2$
S3c	高	高	低	$f(s)g(s)=0$	0	$m/2$

- S3a $fg=h \Rightarrow fh=f\cdot fg=fg=h \Rightarrow (f+1)h=0$
- S3b $fg=0$
- S3c $fg=h \Rightarrow fh=h \rightarrow$ S3a

零化子与代数免疫度

代数免疫度:

$$\mathcal{AI}(f) = \min\{\deg(g) : fg = 0 \text{ 或 } (f+1)g = 0, g \neq 0\}$$

推论 5 (▸ 回顾定理)

$$\mathcal{AI}(f) \leq \lfloor \frac{n+1}{2} \rfloor$$

$$\mathcal{AI}(f) \leq \deg(f)$$

研究热点

1. 一些实用或特殊类型布尔函数的代数免疫度研究;
2. 代数免疫度与其他密码学性质之间的关系;
3. 具有最优代数免疫度布尔函数的性质与构造

谢谢大家！