

# 广义互缩生成器

高军涛 董丽华 胡予濮

(西安电子科技大学计算机网络与信息安全教育部重点实验室 西安 710071)

**摘 要** 设计了一类称为广义互缩生成器的密钥流生成器. 研究表明该类密钥流生成器所产生的序列具有如下良好特性: (1) 大的周期; (2) 高的线性复杂度; (3) 生成的广义互缩序列族具有线性空间结构, 形成 Abel 群; (4) 广义互缩序列族内序列间互相关函数值可以由控制序列中 1 的数目来确定; (5) 在一定条件下, 序列的  $k$ -错线性复杂度显著增加. 另一方面对新序列进行的安全性分析结果表明, 与互缩序列相比, 由较少的密钥量可以获得更好的安全性.

**关键词** 流密码; 广义互缩序列;  $k$ -错线性复杂度; 互相关性; 密码分析

中图法分类号 TN918

## Generalized Shrinking Generator

GAO Jun-Tao DONG Li-Hua HU Yu-Pu

(Key Laboratory of Computer Networks & Information Security of Ministry of Education, Xidian University, Xi'an 710071)

**Abstract** This paper presents a new sequence generator called generalized shrinking generator. The new sequences have pseudorandom properties as follows: (1) Large period; (2) High linear complexity; (3) The sequence family composes a linear space and an Abel group; (4) In the sequence family, the correlation feature is determined by the number of 1 in the clock-control sequence; (5) The  $k$ -error linear complexity has a sharp increase if the clock-controlled sequence is chosen as generalized self-shrinking sequence. On the other hand, the authors give a security analysis for the new sequences. The result shows that the new sequences with fewer amounts of keys are more secure than the shrinking sequences with more keys.

**Keywords** stream cipher; generalized shrinking sequence;  $k$ -error linear complexity; correlation, cryptanalysis

## 1 引 言

伪随机序列在模拟仿真、软件测试、全球定位系统、CDMA 系统等方面有着广泛的应用. 找到一种好的序列生成机制, 能生成具有良好性质的序列是人们追求的目标. 在伪随机序列的发展过程中, 人们不断地引入新的数学工具, 如: 代数和谱分析理论等, 使得序列生成发展到了一个成熟的阶段, 相继有

了大量的序列生成器体制提出.

Coppersmith 等在文献[1]中提出了缩减生成器的概念. 这是一类仅由两个线性反馈移位寄存器(LFSR)组成的序列生成器. 这两个线性反馈移位寄存器分别记为 LFSR1 和 LFSR2, 其中 LFSR1 在 LFSR2 的控制下按照所定义的规则输出比特流. 其生成规则如下: 若 LFSR2 输出 1, 则输出 LFSR1 的对应比特; 否则, 放弃输出. 缩减序列的结构是非常简单的, 目前主要用于伪随机数的生成.

本文中,我们基于向量和不规则抽样概念提出一类新的缩减生成器,称之为广义缩减生成器,其结构图如图 1 所示.

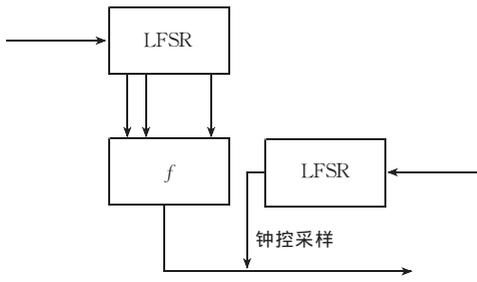


图 1 广义缩减生成器结构图

图 1 中函数  $f$  可以是线性函数或者非线性函数,在本文中我们选择  $f$  为线性函数. 线性函数易于分析,同时生成的序列经过采样以后可以得到较好的伪随机性质. 另一方面,利用不同的线性函数生成的序列组成的序列族具有良好的性质. 广义互缩序列生成器的定义如下.

**定义 1.** 设  $s = s_0 s_1 \dots$  是由 LFSR1 产生的二元  $n$  级  $m$  序列,  $t = t_0 t_1 \dots$  是由 LFSR2 产生的二元序列,  $G = (g_0, g_1, \dots, g_{n-1}) \in GF(2)^n$ , 序列  $v = v_0 v_1 \dots$ , 其中

$v_j = g_0 s_j + g_1 s_{j-1} + \dots + g_{n-1} s_{j-n+1}$ ,  $j = 0, 1, 2, \dots$  当  $t_j = 1$  时, 输出  $v_j$ , 否则不输出. 这样得到序列  $a = a_0, a_1, \dots$  称为  $GF(2)$  上的广义互缩序列. 称  $F(a, s, t) = \{a \mid G \in GF(2)^n\}$  为  $GF(2)$  上基于序列  $s$  和  $t$  的广义互缩序列族.

在所定义条件下,序列族内的序列之间有可控的相关性质(互相关函数值为钟控序列中元素 1 的个数,因此是可以根据实际情况选择的),构成 Abel 群;序列在单个符号替换下线性复杂度有明显的增加;在被控序列为广义自缩序列且控制序列为  $m$  序列的情况下,  $k$ -错线性复杂度有明显的增加. 这都是与互缩序列的主要不同之处. 同时生成的广义互缩序列和互缩序列一样具有大的周期和高的线性复杂度以及结构简单等特点;

本文第 2 节证明了广义互缩序列族具有线性空间、群结构性质,分析了族内序列的互相关性质;第 3 节讨论广义互缩序列的周期和线性复杂度;第 4 节论证了广义互缩序列在单符号替换以及  $k$  个符号替换下线性复杂度的稳定性;第 5 节是安全性讨论;最后是比较和总结.

## 2 广义互缩序列族的性质

设序列  $s$  的反馈多项式  $s(x)$  为  $n$  级本原多项

式,则序列  $s$  为  $m$  序列,  $s$  的项可以用迹函数表示为

$$s_i = \text{Tr}_n(\omega \alpha^i),$$

其中  $\omega \in GF(2^n)$ ,  $\alpha$  为  $s(x)$  在  $GF(2^n)$  上的本原根.

设  $H(G)$  表示向量  $G$  的汉明重量,即  $G = (g_0, g_1, \dots, g_{n-1})$  中分量  $g_{i_1}, g_{i_2}, \dots, g_{i_h}$  等于 1,其余的分量等于 0,其中  $0 \leq i_1 < \dots < i_h \leq n-1$ . 根据定义 1 和有限域的知识,可以得到

$$\begin{aligned} v_j &= \text{Tr}_n(\omega \alpha^{j-i_1}) + \text{Tr}_n(\omega \alpha^{j-i_2}) + \dots + \text{Tr}_n(\omega \alpha^{j-i_h}) \\ &= \text{Tr}_n(\omega \alpha^{j-i_h} (\alpha^{i_h-i_1} + \alpha^{i_h-i_2} + \dots + 1)) \\ &= \text{Tr}_n(\omega \alpha^{j+l}) \end{aligned} \quad (1)$$

其中  $l \in \{0, 1, \dots, P_s - 1\}$ ,  $P_s$  表示序列  $s$  的最小周期. 注意不同的向量  $G$  对应不同的值  $l$ ,这是由有限域中元素的向量表示唯一性决定的.

通过序列的迹表示来证明序列族的封闭性.

设序列  $t$  中第  $i$  个 1 的位置为  $k_i$ , 即  $t_{k_i} = 1$ , 则按照定义 1, 生成的序列  $a$  中第  $i$  项可以表示为

$$a_i = s_{k_i+l} = \text{Tr}_n(\omega \alpha^{k_i+l}) \quad (2)$$

所以序列  $a$  可以表示为

$$\begin{aligned} a &= (\text{Tr}_n(\omega \alpha^{k_0+l}), \text{Tr}_n(\omega \alpha^{k_1+l}), \\ &\quad \text{Tr}_n(\omega \alpha^{k_2+l}), \dots), \forall l_1, \\ &\quad l_2 \in \{0, 1, \dots, P_s - 1\}, \\ a^{(1)} &= (\text{Tr}_n(\omega \alpha^{k_0+l_1}), \text{Tr}_n(\omega \alpha^{k_1+l_1}), \\ &\quad \text{Tr}_n(\omega \alpha^{k_2+l_1}), \dots); \\ a^{(2)} &= (\text{Tr}_n(\omega \alpha^{k_0+l_2}), \text{Tr}_n(\omega \alpha^{k_1+l_2}), \\ &\quad \text{Tr}_n(\omega \alpha^{k_2+l_2}), \dots), \end{aligned}$$

则

$$\begin{aligned} a^{(1)} + a^{(2)} &= (\text{Tr}_n(\omega \alpha^{k_0} (\alpha^{l_1} + \alpha^{l_2})), \text{Tr}_n(\omega \alpha^{k_1} (\alpha^{l_1} + \alpha^{l_2})), \\ &\quad \text{Tr}_n(\omega \alpha^{k_2} (\alpha^{l_1} + \alpha^{l_2})), \dots) \end{aligned}$$

所以  $\exists l_3 \in \{0, 1, \dots, P_s - 1\}$ , 使得  $\alpha^{l_3} = \alpha^{l_1} + \alpha^{l_2}$ , 即

$$\begin{aligned} a^{(3)} &= a^{(1)} + a^{(2)} = (\text{Tr}_n(\omega \alpha^{k_0+l_3}), \text{Tr}_n(\omega \alpha^{k_1+l_3}), \\ &\quad \text{Tr}_n(\omega \alpha^{k_2+l_3}), \dots) \end{aligned}$$

所以  $\forall a^{(1)}, a^{(2)} \in F(a, s, t)$ ,  $\exists a^{(3)} \in F(a, s, t)$ , 使得  $a^{(3)} = a^{(1)} + a^{(2)}$ , 即序列族是封闭的.

**定理 1.** 若序列  $t$  构成的向量  $(t_0, t_1, \dots, t_{2^n-2})$  中含有至少  $2^{n-1} + 1$  个 1, 则广义互缩序列族  $F(a, s, t)$  构成一个  $GF(2)$  上的线性空间; 因此  $F(a, s, t)$  是一个 Abel 加群, 单位元为  $000\dots$ .

**证明.** 根据上述的推导过程可知  $F(a, s, t)$  满足封闭性质; 当  $G$  为零向量时得到单位元  $000\dots$ ; 每个元素的逆元是其本身. 所以  $F(a, s, t)$  构成一个 Abel 加群.

$\forall G_1, G_2 \in GF(2)^n$ ,  $a(G_1), a(G_2)$  分别是由  $G_1, G_2$  按照定义 1 方式生成的序列, 则显然有  $a(G_1) +$

$a(G_2) = a(G_1 + G_2)$ ; 下面证明  $G_1 \neq G_2 \Leftrightarrow a(G_1) \neq a(G_2)$ ,

$a(G_1) \neq a(G_2) \Rightarrow G_1 \neq G_2$  显然成立.

$G_1 \neq G_2 \Rightarrow a(G_1) \neq a(G_2)$ . 若上述关系不成立, 则存在非零向量  $G$ , 使得  $a(G) = 000\dots$ . 由于序列  $s$  是  $m$  序列且序列  $t$  构成的向量  $(t_0, t_1, \dots, t_{2^n-2})$  中含有至少  $2^{n-1} + 1$  个 1, 因此一定可以在序列  $s$  中找到  $n$  个线性无关的  $n$  维向量  $S^i$  使得这些向量对应的生成器的输出均为 0, 即  $S^i G^T = 0, i = 1, 2, \dots, n, G^T$  表示向量  $G$  的转置. 由该线性方程组可知, 向量  $G$  只有一个解, 即  $0$  向量. 所以  $G_1 \neq G_2 \Rightarrow a(G_1) \neq a(G_2)$ .

因此  $F(a, s, t)$  与  $GF(2)^n$  同构. 证毕.

### 3 广义互缩序列的周期和线性复杂度

#### 3.1 广义互缩序列的周期

由序列的迹函数表示可以看出证明广义互缩序列周期的方法和证明互缩序列周期的方法是类似的, 本文为了完整性, 给出证明过程.

**定理 2.** 设  $s = s_0 s_1 \dots$  是由 LFSR1 产生的二元  $m$  序列,  $t = t_0 t_1 \dots$  是由 LFSR2 产生的二元序列,  $a$  是按照定义 1 的生成方式产生的序列, 则

- (1) 当  $\gcd(P_s, P_t) = 1$  时,  $P_a = P_s | 1|_t$ ;
- (2) 当  $\gcd(P_s, P_t) = d > 1$  时,  $P_a = \frac{P_s}{d} | 1|_t$ ,

其中  $P_a, P_s, P_t$  分别表示序列  $a, s, t$  的最小周期,  $|1|_t$  表示序列  $t$  一个周期内 1 的个数.

**证明.** 为证明方便, 本文认为序列是首尾相连的, 同时假设  $LC(t) \leq P_s$ , 其中  $LC(t)$  表示序列  $t$  的线性复杂度. 按照定义 1 中序列的生成原理, 当  $\gcd(P_s, P_t) = 1$  时, 所生成序列  $a$  在  $P_s | 1|_t$  个元素之后重复, 所以  $P_a | P_s | 1|_t$ .

把  $t$  中第  $i$  个 1 的位置记为  $k_i$ , 即  $a_i = s_{k_i+l}, i = 0, 1, \dots$ . 在此种标记下序列  $a$  向前移动  $|1|_t$  个位置就使得序列  $s$  向前移动  $P_t$  个元素. 即对任意的  $i$  和  $j$  都有  $a_{i+j|1|_t} = s_{k_i+l+jP_t}$ , 同样对于所有的  $i$  和  $j$  都有  $a_{i+j|1|_t} = a_{i+P_a+j|1|_t}$ , 所以  $s_{k_i+l+jP_t} = s_{k_i+P_a+l+jP_t}, i = 0, 1, \dots$ .

根据序列的性质可知:  $P_s | (k_{i+P_a} - k_i)$ , 即  $P_s | ((k_{i+P_a} + l) - (k_i + l))$ . 这等价于  $\forall i, \exists j_i$  使得  $k_{i+P_a} + l = k_i + l + j_i P_s$ ; 对于  $i+1$  该式也成立, 即:  $k_{i+1+P_a} + l = k_{i+1} + l + j_{i+1} P_s$ , 所以有  $(k_{i+1+P_a} + l) - (k_{i+P_a} + l) = (k_{i+1} + l) - (k_i + l) + (j_{i+1} - j_i) P_s$ . 注意到:  $(k_{i+1+P_a} + l) - (k_{i+P_a} + l)$  与  $(k_{i+1} + l) - (k_i +$

$l)$  都是序列  $t$  中相邻的两个 1 之间 0 的个数. 如果  $j_{i+1} - j_i \neq 0$ , 则可知在序列  $t$  中有至少连续  $P_s$  个元素中不含有元素 1, 这与假设  $LC(t) \leq P_s$  矛盾. 所以, 我们有  $j_{i+1} - j_i = 0$ , 因此  $(k_{i+1+P_a} + l) - (k_{i+P_a} + l) = (k_{i+1} + l) - (k_i + l), i = 0, 1, \dots$ . 该式说明序列  $t$  从  $t_{k_i+l}$  开始的子序列与从  $t_{k_{i+P_a}+l}$  开始的子序列是一致的. 也就说明  $P_t | ((k_{i+P_a} + l) - (k_i + l))$ , 即从  $t_{k_i+l}$  开始到  $t_{k_{i+P_a}+l}$  结束的序列的长度是  $P_t$  的倍数. 因而这一段元素 1 的个数也是  $|1|_t$  的倍数. 而这一段元素 1 的个数正好是  $P_a$ , 所以  $P_a = r | 1|_t$ .

又  $\forall j, s_{k_0+l} = a_0 = a_{jP_a} = a_{jr|1|_t} = s_{k_0+l+jrP_t}$ , 所以  $P_s | rP_t$ .

由于  $\gcd(P_s, P_t) = 1$  时,  $P_s | r$ ; 结合  $P_a = r | 1|_t$ , 得到  $P_s | 1|_t$  整除  $P_a$ ; 结合  $P_a | P_s | 1|_t$ , 可知当  $\gcd(P_s, P_t) = 1$  时,  $P_a = P_s | 1|_t$ .

同理当  $\gcd(P_s, P_t) = d > 1$  时, 所生成序列  $a$  在  $\frac{P_s}{d} | 1|_t$  个元素之后重复, 即  $P_a | \frac{P_s}{d} | 1|_t$ . 由上述的证明过程我们知道, 证明命题“ $P_a = r | 1|_t$ ”的过程与  $\gcd(P_s, P_t)$  的值是没有关系的. 因此, 当  $\gcd(P_s, P_t) = d$  时, 也可以得到  $P_a = r | 1|_t$ .

又  $\forall j, s_{k_0+l} = a_0 = a_{jP_a} = a_{jr|1|_t} = s_{k_0+l+jrP_t}$ , 故  $P_s | rP_t$  所以当  $\gcd(P_s, P_t) = d > 1$  时,  $\frac{P_s}{d} | r$ ; 结合  $P_a = r | 1|_t$ , 得到  $\frac{P_s}{d} | 1|_t | P_a$ , 同时  $P_a | \frac{P_s}{d} | 1|_t$ , 因此  $P_a = \frac{P_s}{d} | 1|_t$ . 证毕.

由证明过程可以知道:  $\forall l \in \{0, 1, \dots, P_s - 1\}$ , 由序列  $s$  和  $t$  产生的序列  $a$  的周期都满足上述结论. 这里只需要  $s$  为  $m$  序列,  $t$  是以 LFSR 为基础产生的序列且序列  $t$  中构成的向量  $(t_0, t_1, \dots, t_{2^n-2})$  中含有至少  $2^{n-1} + 1$  个 1, 就会生成  $F(a, s, t)$  中  $2^n$  个序列, 这些序列构成线性空间和 Abel 加群. 下面考虑族内序列的互相关性质:

当  $\gcd(P_s, P_t) = 1$  时, 考虑  $s$  和  $t$  的元素对序列  $(s_{j+l}, t_j), j = 0, 1, 2, \dots$ , 周期为  $P_s \cdot P_t$ , 再考虑子序列  $(s_{k_i+l}, t_{k_i}), i = 0, 1, 2, \dots, k_i$  表示序列  $t$  中第  $i$  个 1 的位置. 这个子序列的周期是  $P_s \cdot |1|_t$ , 与  $a$  的最小周期  $P_a$  相同. 对于每一个  $i$  有  $(a_i, a_{i+|1|_t}, \dots, a_{i+P_s|1|_t}) = (s_{k_i+l}, s_{k_i+l+P_t}, \dots, s_{k_i+l+P_sP_t})$ , 因此  $a$  的一个周期元素正是  $s$  的  $|1|_t$  个周期元素的重排. 所以当  $s$  是  $m$  序列时,  $a$  中 1 出现  $2^{LC(s)-1} | 1|_t$  次, 0 出现  $(2^{LC(s)-1} - 1) | 1|_t$  次.

因此可以得到下面关于互相关性的推论.

推论 1.  $\forall a^{(1)}, a^{(2)} \in F(a, s, t)$ , 定义  $C(a^{(1)},$

$$a^{(2)}) = \sum_{i=0}^{P_a-1} (-1)^{a_i^{(1)}+a_i^{(2)}}, \text{ 若 } a^{(1)} = a^{(2)}, \text{ 则 } C(a^{(1)}, a^{(2)}) = P_a; \text{ 若 } a^{(1)} + a^{(2)} = 111\cdots, \text{ 则 } C(a^{(1)}, a^{(2)}) = -P_a; \text{ 其它情况下, } C(a^{(1)}, a^{(2)}) = |1|_t.$$

注:  $|1|_t$  指的是序列  $t$  一个周期内 1 的个数, 定理 2 的假设: “序列  $t$  构成的向量  $(t_0, t_1, \dots, t_{2^n-2})$  中含有至少  $2^{n-1} + 1$  个 1” 并不矛盾. 另外, 在序列族  $F(a, s, t)$  中可能含有  $111\cdots$ , 这并不是确定的.

### 3.2 广义互缩序列的线性复杂度

定理 3. 设  $s = s_0 s_1 \cdots$  是由 LFSR1 产生的二元  $m$  序列,  $t = t_0 t_1 \cdots$  是由 LFSR2 产生的二元序列, 序列  $a$  是按照定义 1 的生成方式所产生的序列, 如果  $|1|_t = 2^r$ , 且  $\gcd(P_s, P_t) = 1$ , 则  $LC(s) 2^{r-1} \leq LC(a) \leq LC(s) 2^r$ , 其中  $r \in Z^+$ ,  $LC(s), LC(a)$  分别代表序列  $s$  和  $a$  的线性复杂度;  $P_s, P_t$  分别表示序列  $s, t$  的周期,  $|1|_t$  表示序列  $t$  一个周期内 1 的个数.

证明. 只需要找到序列  $a$  的极小多项式并估计其次数就可以了.

把序列  $a_{j|1|_t}$  记作  $a^{||1|}$ ,  $j = 0, 1, \dots$ , 则可用  $s$  的项表示为  $s_{k_0+t+jP_t}$ , 因为  $\gcd(P_s, P_t) = 1$ , 所以这也是一个最大长度序列, 设它的极小多项式为  $f(x) = c_0 + c_1 x + \dots + c_{n-1} x^{n-1} + x^n, n = LC(s)$ .

因而有  $c_0 a_{j|1|_t} + c_1 a_{(j-1)|1|_t} + \dots + c_{n-1} a_{(j-n+1)|1|_t} + a_{(j-n)|1|_t} = 0, j \geq n$ . 由此可知  $a$  的一个特征多项式为  $f(x^{||1|})$ , 次数为  $n|1|_t = LS(s) 2^r$ . 因此  $a$  的最大值为  $LS(s) 2^r$ .

设  $f_a(x)$  是  $a$  的极小多项式, 则  $f_a(x) | f(x^{||1|}) = f(x)^{2^r}$ . 由于  $f(x)$  是不可约多项式, 所以可以假设  $f_a(x) = (f(x))^k$ . 假设  $k \leq 2^{r-1}$ , 则  $f_a(x) | f(x)^{2^{r-1}}$ , 由于序列  $s_{k_0+t+jP_t} (j = 0, 1, \dots)$  的周期为  $P_s$ , 所以  $f(x) | (1-x^{P_s}), f_a(x) | (1-x^{P_s})^{2^{r-1}}$ , 这表明  $a$  的周期最大为  $P_s 2^{r-1}$ . 但是在本定理的假设条件下, 根据定理 1 得到的序列  $a$  的周期为  $P_s 2^r$ . 矛盾. 证毕.

下面讨论  $\gcd(P_s, P_t) = d > 1$  时序列线性复杂度的界, 下述命题成立.

命题 1. 对于  $GF(q)$  上周期  $P(v) = kq^r$  的序列  $v$ , 其线性复杂度  $LC(v) > q^{r-1}$ , 其中  $q = p^m, p$  为素数,  $k \in Z^+$ .

证明. 设  $f(x)$  是序列  $v$  的最小多项式, 则  $f(x) | (1-x^{kq^r})$ , 即  $f(x) | (1-x^k)^{q^r}$ ;

而  $1-x^k = (1-x)(1+x+\dots+x^k)$ , 所以有:

$$f(x) | (1-x)^{q^r} (1+x+\dots+x^{k-1})^{q^r};$$

因而  $f(x) = (1-x)^{\omega_1} (1+x+\dots+x^{k-1})^{\omega_2}$ , 其中  $\omega_1, \omega_2 \leq q^r$ ,

若  $\omega_1 \leq q^r$  且  $\omega_2 \leq q^{r-1}$ , 则  $f(x) | (1-x)^{q^{r-1}} (1+x+\dots+x^{k-1})^{q^{r-1}}$ ; 说明序列  $v$  的周期可以为  $kq^{r-1}$ , 这与假设矛盾, 所以  $\omega_1, \omega_2$  不能同时小于等于  $q^{r-1}$ . 所以  $LC(v) > q^{r-1}$ . 证毕.

根据命题并结合定理 1 的结论可得到下面定理.

定理 4. 设  $s = s_0 s_1 \cdots$  是由 LFSR1 产生的二元  $m$  序列,  $t = t_0 t_1 t_2 \cdots$  是由 LFSR2 产生的二元序列, 序列  $a$  是按照定义 1 的生成方式所产生的序列, 如果  $|1|_t = 2^r$ , 且  $\gcd(P_s, P_t) = d > 1$ , 则  $LC(s) 2^r > LC(a) > 2^{r-1}$ . 其中  $r \in Z^+$ ;  $LC(s), LC(a)$  分别代表序列  $s$  和  $a$  的线性复杂度.  $P_s, P_t$  分别表示序列  $s, t$  的周期,  $|1|_t$  表示序列  $t$  一个周期内 1 的个数.

上述这些结果表明:  $d$  越小, 广义互缩序列周期越大; 在定理所假设条件下, 序列具有指数级的线性复杂度.

选择  $s$  为  $m$  序列是有原因的, 一个是因为其生成方式简单, 另一个是因为这样得到的序列族以及序列之间具有良好的伪随机性质.

## 4 广义互缩序列线性复杂度的稳定性

### 4.1 单符号替换下的线性复杂度的稳定性

设  $GF(2)$  上的周期序列  $b(j) = b_0(j), b_1(j) \cdots$ , 满足  $b_i(j) = \begin{cases} 1, & i \equiv j \pmod{P_a} \\ 0, & \text{其它} \end{cases}$  其生成函数为

$S^{P_a}(x) = x^j, j \in \{0, 1, \dots, P_a - 1\}$ . 同时定义序列  $a$  的极小多项式为  $f_a(x)$ , 其生成函数可以记为  $S^{P_a}(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{P_a-1} x^{P_a-1}$ . 其中  $a_0, a_1, \dots, a_{P_a-1} \in GF(2)$ . 利用文献[2, 3]的证明思想, 得到下面的定理.

定理 5. 设  $s = s_0 s_1 \cdots$  是由 LFSR1 产生的二元  $m$  序列,  $t = t_0 t_1 \cdots$  是由 LFSR2 产生的二元序列,  $a = a_0 a_1 \cdots$  是由 LFSR1 和 LFSR2 按照定义 1 所述方式生成的序列. 如果  $\gcd(P_s, P_t) = d$ , 则:  $|1|_t = 2^r$  时,  $LC_1(a) \geq \left( \frac{2^{LC(s)} - 1}{d} - LC(s) \right) 2^r$ ; 其中  $r \in Z^+$ ,  $|1|_t$  表示序列  $t$  中 1 的个数,  $LC_1(a)$  表示序列  $a$  在单符号替换下的线性复杂度.

证明. 根据前面假设, 单符号替换序列  $a$  的第  $j$  项所得序列  $e$  的生成函数可以记为

$$S^{P_e}(x) = S^{P_a}(x) + S^{P_b}(x) = a_0 + a_1 x + \dots +$$

$$(a_j + 1)x^j + \cdots + a_{p_a-1}x^{p_a-1}.$$

设  $f_a(x) = \frac{1-x^{p_a}}{\gcd(1-x^{p_a}, S^{p_a}(x))}$ , 记  $g(x) = \gcd(1-x^{p_a}, S^{p_a}(x))$ , 显然有  $1-x^{p_a} = f_a(x)g(x)$ .

设  $f_e(x) = \frac{1-x^{p_a}}{\gcd(1-x^{p_a}, S^{p_e}(x))}$ , 记  $e(x) = \gcd(1-x^{p_a}, S^{p_e}(x))$ , 则

$$\begin{aligned} e(x) &= \gcd(f_a(x)g(x), S^{p_e}(x)) \\ &= \gcd(f_a(x)g(x), S^{p_a}(x) + x^j). \end{aligned}$$

下面用反证法证明  $\deg(e(x)) \leq \deg(f_a(x))$ :

若  $\deg(e(x)) > \deg(f_a(x))$ , 则必有  $\deg(\gcd(g(x), S^{p_a}(x) + x^j)) > 1$ , 即  $g(x) | S^{p_a}(x)$  且  $g(x) | x^j$ ,  $j \in \{0, 1, \dots, p_a - 1\}$ . 同时注意到  $g(x) | 1 - x^{p_a}$ , 即  $g(0) \neq 0$ . 矛盾. 所以有  $\deg(e(x)) \leq \deg(f_a(x))$ .

所以有  $LC_1(a) = LC(e) = \deg[(1-x^{p_a})/e(x)] \geq \deg[(1-x^{p_a})/f_a(x)]$ , 根据定理 2 我们知道

$$\begin{aligned} LC_1(a) &\geq P_a - LC(s)2^r = \frac{(2^{LC(s)} - 1)}{d} 2^r - LC(s)2^r \\ &= \left( \frac{(2^{LC(s)} - 1)}{d} - LC(s) \right) 2^r. \quad \text{证毕.} \end{aligned}$$

注: 若  $s$  和  $t$  都是  $m$  序列且  $\gcd(P_s, P_t) = 1$  时, 那么我们可以得到比较具体的结果:

$$LC_1(a) \geq (2^{LC(s)} - 1 - LC(s))2^{LC(t)-1}.$$

#### 4.2 $k$ -错线性复杂度

根据定义 1 可知对于驱动序列  $s$  和  $t$  的基本要求是: (i) 生成方式简单; (ii) 以 LFSR 为基础; (iii) 所生成的广义互缩序列具有良好的伪随机性质. 因此, 在实际应用中可以选择满足上述 3 条件的任何序列作为驱动序列.

在本小节中,  $s$  选定为广义自缩序列生成器,  $t$  为  $m$  序列生成器. 这样选择的原因是  $s$  和  $t$  两个序列的周期是互素的, 且满足上面的 3 条. 关于广义自缩序列性质请参看文献[4, 5]. 下面给出广义自缩序列的定义.

定义 2. 设  $s = s_0 s_1 \cdots$  是  $GF(2)$  上的  $n$  级  $m$ -序列, 同时设  $G = (g_0, g_1, \dots, g_{n-1}) \in GF(2)^n$ , 序列  $v = v_0 v_1 \cdots$ , 其中

$$v_j = g_0 s_j + g_1 s_{j-1} + \cdots + g_{n-1} s_{j-n+1}, \quad j = 0, 1, 2, \dots.$$

当  $s_j = 1$  时, 输出  $v_j$ , 否则不输出. 这样得到序列  $b = b_0, b_1, \dots$ , 称为  $GF(2)$  上的广义自缩序列.

定理 6. 设  $s = s_0 s_1 \cdots$  是由 LFSR1 生成的  $GF(2)$  上的广义自缩序列, 周期为  $2^{n-1}$ ;  $t = t_0 t_1 \cdots$  是由 LFSR2 产生的  $m$  序列, 且  $|t| = 2^r$ ,  $a = a_0 a_1 \cdots$  是由 LFSR1 和 LFSR2 按照定义 1 所述方式生成的序

列, 其中  $H(G) = 1$ , 那么, 若  $k$  为奇数,  $LC_k(a) = P_a = 2^{n-1} \cdot 2^r$ ; 若  $k$  为偶数, 且  $\deg(\gcd(1+x^{p_a}, x^{i_1} + x^{i_2} + \cdots + x^{i_k})) \neq P_a - LC(a)$ , 则  $LC_k(a) \geq LC(a)$ . 其中  $LC_k(a)$  表示序列  $a$  的  $k$ -错线性复杂度;  $H(G)$  表示定义 1 中向量  $G$  的 Hamming 重量.

证明. 由定理 1 知: 序列  $a$  的周期为  $P_a = 2^{n-1} \cdot 2^r$ , 根据有限域的知识,  $\sum_{i \geq 0} a_i x^{-(i+1)} = \frac{a^{P_a}(x)}{x^{P_a} - 1} = \frac{a^{P_a}(x)}{(x-1)^{P_a}} = \frac{u(x)}{f_a(x)}$ , 若  $\gcd(u(x), f_a(x)) = 1$ ,  $f_a(x)$  就为序列  $a$  的极小多项式, 显然有  $(x-1) \nmid u(x)$ ; 同时有:  $(x-1)^{P_a} = f_a(x)g(x)$ .  $f_a(x) = (x-1)^{LC(a)}$ , 因此  $g(x)$  是  $(x-1)$  的幂次形式.

(1) 当序列  $a$  改变一个比特时,  $\sum_{i \geq 0} \hat{a}_i x^{-(i+1)} = \frac{\hat{a}^{P_a}(x)}{x^{P_a} - 1} = \frac{a^{P_a}(x) + x^i}{(x-1)^{2^{n-1}+r}}$ ; 同时  $(x-1) \mid a^{P_a}(x)$ , 否则序列  $a$  的线性复杂度达到最大.  $u_1(x) = \gcd((x-1)^{P_a}, a^{P_a}(x) + x^i) = 1$ ,  $LC_1(a) = P_a$ .

(2) 当序列  $a$  改变两个比特时,  $\sum_{i \geq 0} \hat{a}_i x^{-(i+1)} = \frac{\hat{a}(x)}{x^{P_a} - 1} = \frac{a^{P_a}(x) + x^{i_1} + x^{i_2}}{(x-1)^{2^{n-1}+r}}$ , 其中  $0 \leq i_1 < i_2 \leq P_a - 1$ .

若  $x^{i_1} + x^{i_2} = h(x)(x-1)^m$ , 则有

$$\begin{aligned} u_2(x) &= \gcd((x-1)^{P_a}, a^{P_a}(x) + x^{i_1} + x^{i_2}) = \\ &= \begin{cases} u(x), & P_a - LC(a) < m; \\ (x-1)^m, & P_a - LC(a) > m. \end{cases} \end{aligned}$$

所以  $LC_2(a) \geq LC(a)$ .

根据上面的推导可以得到: 当序列  $a$  改变奇数个比特时,  $LC_k(a) = P_a$ , 达到最大. 而当  $a$  改变偶数个比特时, 且  $\deg(\gcd(1+x^{p_a}, x^{i_1} + x^{i_2} + \cdots + x^{i_k})) \neq P_a - LC(a)$  时,  $LC_k(a) \geq LC(a)$ . 证毕.

注: 此时, 生成的序列的周期  $P_a = 2^{n-1} \cdot 2^r$ , 因此线性复杂度  $LC(a) > 2^{n-1} \cdot 2^{r-1}$ , 但生成的序列族不具有群结构;  $H(G) > 1$  时, 周期的情况较复杂, 需做进一步探讨.

## 5 安全性讨论

在互缩序列的设计中, 作者建议用 LFSR1 和 LFSR2 的初始状态以及线性反馈移位寄存器的反馈多项式作为密钥. 而广义互缩序列可以选取 LFSR1 和 LFSR2 的初始状态和向量  $G$  作为密钥, 密钥量少. 本部分将表明新型的序列可以用较少的密钥量实现更高的安全性来抵抗已知的攻击.

已知的对互缩序列的攻击主要有: 分别征服攻

击、概率相关攻击、低复杂度相关攻击、可区分攻击等. 本部分将对所有的攻击进行分析.

### 5.1 分别征服攻击<sup>[1]</sup>

Coppersmith 等在提出互缩序列的同时, 提出了两种针对互缩序列的分别征服攻击. 两种攻击都属于已知明文攻击.

第一种攻击是假定已知 LFSR1 和 LFSR2 的反馈多项式和度数. 攻击者需要猜测 LFSR2 的初始状态, 在选定一个 LFSR2 的初始状态以后, 则 LFSR2 生成的序列  $t$  被恢复出来. 利用序列  $t$  和已知的一些密钥流序列  $a$  的信息, 能得到序列  $s$  的非连续的比特片段. 利用得到的这些比特信息, 我们能构造关于 LFSR1 初始状态的线性方程, 因此能够在多项式时间内求解. 这种攻击需要穷举搜索 LFSR2 的初始状态, 其计算复杂度是 LFSR2 的初始状态的指数级. 因此当 LFSR2 的初始状态比较大的时候, 这种攻击方法是不实用的.

第二种攻击假定反馈多项式未知, 攻击者需要猜测反馈多项式和 LFSR2 的初始状态. 假设已知密钥流序列  $a$  的  $N$  个比特, 攻击者能够获得乘积序列  $p_n = t_n s_n$ , 该序列的线性复杂度至多为  $l_1 l_2$ , 这里  $l_1, l_2$  分别表示 LFSR1 和 LFSR2 的级数. 由  $N = 2l_1 l_2$  个比特, 用 BM 算法, 全部乘积序列可以被计算出来. 利用得到的序列流可以重构出序列  $s$ , 因此也能得到 LFSR1 的初始状态和反馈多项式. 这种攻击的复杂度是  $O(l_1 2^{l_2})$ . 因此当 LFSR2 的初始状态比较大的时候, 这种攻击方法也是不实用的.

广义互缩序列作为密钥流时密钥种子是 LFSR1 和 LFSR2 的初始状态和向量  $G$ , 因此第二种分别征服攻击中, 就不需要来穷举反馈多项式, 但仍然需要来穷搜索 LFSR2 的初始状态. 从上面的分析来看, 如果选用分别征服攻击来对广义互缩序列实施攻击, 则都需要穷举搜索 LFSR2 的初始状态, 显然这是不实用的.

### 5.2 相关攻击<sup>[6~9]</sup>

1994 年 Golić 和 O'Connor 在文献[6]中对一般钟控序列提出了概率相关攻击. 其基本思想就是考虑序列流  $s$  和  $a$  的联合概率. 作者需要考虑密钥流  $a$  和每一个可能产生密钥流  $a$  的 LFSR1 的初始状态之间的联合概率. 作者指出观测到的密钥流比特与删除信道的信道容量是相关的, 如果攻击者选取的初始状态等于原来的 LFSR1 的初始状态, 那么就会得到最大的联合概率, 由此, 其初始状态得以恢复. 1998 年, Simpson, Golić 和 Dawson 在文献[7]

中对于互缩序列实施了概率相关攻击. 作者假定 LFSR1 的反馈多项式是已知的, 在这种情况下, 概率相关攻击需要知道大概  $20l_1$  个密钥流比特, 计算复杂度为  $O(2^{l_1} l_1^2)$ .

对于广义互缩序列来说, 在其向量  $G$  未知情况下, 这种攻击是失效的. 因为向量  $G$  决定了生成器互缩的规则. 即使利用上述的方法得到 LFSR1 的初始状态, 由于向量  $G$  未知, 即互缩规则未知, 攻击者无法恢复出 LFSR2 的初始状态. 而攻击者想要获得向量  $G$  就必须进行穷举搜索. 因此概率相关攻击对于广义互缩序列来说也是不实用的.

1998 年, Johansson 提出了一种低复杂度的相关攻击算法. 该算法是基于删除信道的一种译码方法. 攻击者通过搜索输出流  $s$  中特殊的比特串以求得对于序列  $s$  中比特的后验概率(相对是较高的), 并利用译码算法来求得 LFSR1 的初始状态. 该类攻击的复杂度仍然是 LFSR1 长度的指数级. 具体内容参见文献[8].

2001 年 Golić 在文献[9]中又提出了利用后验概率作为工具实施的相关攻击. 相比于上述的方法, 其优势在于在某些假设下, 不需要穷举搜索序列的初始状态, 并且即使 LFSR 很长, 复杂度也不会很大.

对于广义互缩序列来说, 与互缩序列有部分密钥是相同的. 而且向量  $G$  的应用提高了序列抵抗攻击的能力, 下面分析  $G$  在抵抗后验概率相关攻击中的作用.

在攻击过程中, 关键的步骤是计算 LFSR1 的比特块的后验概率以及 LFSR1 和 LFSR2 单独比特的后验概率, 得到这些后验概率以后, 利用迭代译码算法来攻击互缩序列, 以得到 LFSR1 的初始状态. 因此计算这些后验概率可以大大提高攻击的效率.

下面的论述表明, 对于广义互缩序列, 求 LFSR1 的后验概率是困难的. 攻击的第一步是求概率  $Pr\{s^n | a^n\}$ , 实际上求的是  $Pr\{s^n | a^n\}$ , 这里  $s^n = s_1 s_2 \cdots s_n$ ,  $a^n = a_1 a_2 \cdots a_n$ .

求  $Pr\{s^n | a^n\}$  的递归公式:  $Pr\{s^n | a^n\} = Pr\{a^n | s^n\} = \sum_{l=0}^n 2^{-l} Q(l, n-l)$ , 这里

$$Q(l, n-l) = Pr\{a^{n-l}, |0|_{l^n} = l | s^n\}$$

且  $|0|_{l^n}$  表示序列  $t^n = t_0 t_1 t_2 \cdots t_n$  中 0 的个数.

显然, 求广义互缩序列的  $Q(l, n-l)$  并不是容易的. 因为在广义互缩序列当中  $Q(l, n-l) = Pr\{a^{n-l+j}, |0|_{l^{n+j}} = l | s^n\}$ ,  $-P_a \leq j \leq P_a$ ,  $j$  是由向量  $G$  来确定的, 这和上面讨论的概率相关攻击时向

量  $G$  未知是一致的. 攻击者在采用这种攻击的时候必须穷举搜索向量  $G$ , 其复杂度是  $O(2^n)$ . 因此该类攻击对于新序列来说是无效的.

### 5.3 可区分攻击<sup>[10~12]</sup>

可区分攻击的目的不是恢复序列生成器的密钥, 而是将得到的密钥流序列与一般的随机序列区分开来, 目的是知道这种密钥流序列是由哪一种生成器产生的. 然后再结合针对该类生成器的攻击方法对生成器实施攻击.

1995 年, Golić 针对一般的不规则钟控序列生成器提出了一种可区分攻击. 该类攻击的思想在文献<sup>[10, 11]</sup>中得到了应用. 研究表明这类攻击需要生成器中的 LFSR 具有特殊反馈多项式, 即反馈多项式  $f(x)$  的重量要足够小或者存在一个多项式  $g(x)$  使得  $f(x)g(x)$  有足够小的重量. 文中给出了两个例子: 第一个是互缩生成器的反馈多项式  $f(x)$  可以被化成重量为 4 的多项式(度数为 10000), 攻击者需要大概  $2^{48}$  个密钥流比特来进行区分攻击; 第二个例子是互缩生成器的反馈多项式  $f(x)$  可以被化成重量为 3 的多项式(度数为 40000), 攻击者需要大概  $2^{39}$  个密钥流比特来进行区分攻击.

2003 年, Ekdahl 在其博士论文<sup>[12]</sup>中针对互缩生成器提出了一种新的区分攻击, 并与上述的方法进行了比较. 结果表明对于上述两种情况分别需要  $2^{32}$  个和  $2^{23}$  个密钥流比特来进行区分攻击. 因此文献<sup>[12]</sup>中的工作要比文献<sup>[10, 11]</sup>中的工作更加优秀.

可以看出, 区分攻击对于互缩生成器有较大的局限性. 因为这种攻击仅是对于具有特定的反馈多项式的互缩生成器有效. 目前, 对“任意一个多项式  $f(x)$ , 找到一个多项式  $g(x)$ , 使得  $f(x)g(x)$  有足够小的重量”这个问题还没有有效的算法来实现. 因此, 对于给定的广义互缩生成器, 如果其反馈多项式  $f(x)$  的重量足够大, 攻击者找到  $g(x)$ , 使得  $f(x)g(x)$  有足够小的重量并不是容易的. 因此广义互缩序列可以抵抗这种可区分攻击.

最近, 出现了针对 Bit-search 生成器的攻击<sup>①</sup>, 而 Bit-search 生成器可以看作是互缩生成器和自缩生成器的变种. 相比于以前的分析方法, 文中提出的方法有了较大的进步, 但攻击复杂度仍然是 LFSR 长度的指数级. 该方法是否对于互缩生成器有效, 仍然值得进一步研究.

总之, 从以上的分析可以看出, 现有的针对互缩生成器的攻击方法对于广义自缩生成器来说效率都是很低的, 或者等价于穷举搜索. 相比于互缩序列来

说, 广义互缩生成器用了较少的密钥量实现了更高的安全性.

## 6 比较和结论

最近胡予濮等提出了广义自缩序列, 设计思想和广义互缩生成器类似, 相比于广义自缩序列, 广义互缩序列具有以下优点:

(1) 最小周期是确定的. 广义自缩序列的最小周期是由向量  $G$  来确定的, 并不全是达到最大, 而本文提出的序列除由  $G=(00\cdots 0)$  生成的序列以外, 所有的序列最小周期都达到最大  $P_s | 1 |_t$ .

(2) 序列的多样性. 广义互缩序列的抽样序列  $t$  是可以任意选取的, 只要满足  $LC(t) \leq P_s$ . 因此序列族中序列之间的互相关函数值可以根据实际需要选取适当的  $t$  来获得.

相比于互缩序列来说, 该类序列用较少的密钥量实现了更高的安全性.

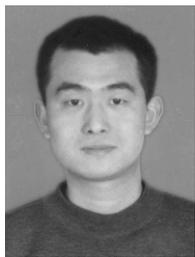
本文基于互缩生成器提出了广义互缩生成器. 研究表明: 在被控序列  $s$  为  $m$  序列的条件下, 该类生成器所产生的序列具有大的周期以及高的线性复杂度, 广义互缩序列族构成了线性空间以及 Abel 群, 序列之间的互相关性质是可控的, 所生成的序列在单符号替换下, 其线性复杂度有明显的增加. 在  $s$  为广义自缩序列且  $t$  为  $m$  序列的条件下, 序列同样具有周期大、线性复杂度高且生成简单的特点, 同时在  $k$  个符号替换下, 若  $k$  为奇数, 则  $k$  错线性复杂度达到最大; 若  $k$  为偶数, 则  $k$  错线性复杂度达到最大或者大于等于原来序列的线性复杂度. 这些性质都是与互缩序列的主要不同之处. 另外, 我们分析了序列在各种攻击下的安全性, 结果表明广义互缩序列的安全性质要优于互缩序列.

## 参 考 文 献

- 1 Coppersmith D., Krawczyk H., Mansour Y.. The shrinking generator. In: Proceedings of the Cryptology-CRYPT' 93, Santa Barbara, USA, 1994, 22~39
- 2 Blackburn S. R.. The linear complexity of the self-shrinking generator. IEEE Transactions on Information Theory, 1999, 45(6): 2073~2077
- 3 Dai Zon-Guo, Imamura K.. Linear complexity for one-symbol substitution of a periodic sequence over  $GF(q)$ . IEEE Transac-

① Hell M., Johansson T.. Some attack on the Bit-search generator, www.it.lth.se/martin/fse2005.pdf.

- tions on Information Theory, 1998, 44(3): 1328~1331
- 4 Hu Yu-Pu, Zhang Yu-Qing, Xiao Guo-Zhen. Symmetric Key Cryptography. Beijing: China Machine Press, 2001 (in Chinese)  
(胡予濮, 张玉清, 肖国镇. 对称密码学. 北京: 机械工业出版社, 2001)
  - 5 Hu Yu-Pu, Xiao Guo-Zhen. The generalized self-shrinking generator. IEEE Transactions on Information Theory, 2004, 50(4): 714~719
  - 6 Golić J. D., O'Connor L.. Embedding and probabilistic correlation attacks on clock-controlled shift registers. In: Proceedings of the Cryptology-Eurocrypt'94, Perugia, Italy, 1995, 230~243
  - 7 Simpson L., Golić J. D., Dawson E.. A probabilistic correlation attacks on the shrinking generator. In: Proceedings of the Information Security and Privacy'98, Brisbane, Australia, 1998, 147~158
  - 8 Johansson T.. Reduced complexity correlation attacks on two clock-controlled generators. In: Proceedings of the Cryptology-Asiacrypt'98, Beijing, 1998, 342~357
  - 9 Golić J. D.. Correlation analysis of the shrinking generator. In: Proceedings of the Cryptology-CRYPTO 2001, Santa Barbara, USA, 2001, 440~457
  - 10 Golić J. D.. Linear models for keystream generators. IEEE Transactions on Computers, 1996, 45(1): 41~49
  - 11 Golić J. D., Menicocci R.. Edit probability correlation attacks on stop/go clocked keystream generators. Journal of Cryptology, 2003, 16(1): 41~68
  - 12 Ekdahl P.. On LFSR based stream ciphers-analysis and design [Ph. D. dissertation]. Lund University, Sweden, 2003



**GAO Jun-Tao**, born in 1979, Ph. D. candidate. His major research interests focus on stream cipher and pseudorandom sequences.

**DONG Li-Hua**, born in 1977, Ph. D. candidate. Her major research interest is computer cryptography.

**HU Yu-Pu**, born in 1955, Ph. D., professor, Ph. D. supervisor. His major research interests focus on cryptography, networks and information security.

## Background

The major research interests of this group are computer cryptography, especially the stream cipher and pseudorandom sequences. This work is supported by the National Natural Science Foundation of China (60273084) and the Doctorial Foundation (20020701013).

Shrinking generator was presented by Coppersmith *et al.* in 1993, which is usually used for pseudorandom num-

ber generator. This paper proposes generalized shrinking generator based on the shrinking generator. The new generator can produce pseudorandom sequences with good pseudorandom properties such as large period, large linear complexity, and good stability of linear complexity. All of the properties guarantee that the generalized shrinking sequences can be used in stream ciphers.