

文章编号:1001-9081(2009)05-1334-02

[a, b]-缩减生成器

廖翠玲,余昭平

(信息工程大学 电子技术学院, 郑州 450004)

(230cuiling@163.com)

摘要:利用两个反馈移位寄存器(LFSR)构造了一类新型的缩减生成器——[a, b]-缩减生成器,证明了其输出序列的周期、线性复杂度、重量复杂度、k-错线性复杂度及其0、1个数。理论分析和局部随机性检验表明这类缩减生成器序列具有好的统计特性,适合流密码系统的使用。

关键词:线性反馈移位寄存器;周期;线性复杂度;局部随机性检验

中图分类号: TN918 **文献标志码:**A

[a, b]-shrinking generator

LIAO Cui-ling, YU Zhao-ping

(Institute of Electronic Technology, Information Engineering University, Zhengzhou Henan 450004, China)

Abstract: A new construction of a new generator, called the [a, b]-shrinking generator was investigated based on two Linear Feedback Shift Registers (LFSR). The period, linear complexity, weight complexity and the numbers of element of 1 and 0 of the output sequence of the [a, b]-shrinking generator were proved. Both the theoretic and the experimental results of local randomness tests show that the [a, b]-shrinking generator is suitable for stream cipher cryptosystems.

Key words: Linear Feedback Shift Register (LFSR); period; linear complexity; local randomness test

0 引言

线性移位寄存器因其结构简单、速度快等优点而成为构造密钥流生成器的最重要部件之一。目前比较成熟的流密码系统主要基于线性反馈移位寄存器(Linear Feedback Shift Register, LFSR)构造,我们把这种系统分为两大类:一是用LFSR和非线性布尔函数结合,如非线性组合流密码和前馈流密码;二是用一个LFSR去控制另一个LFSR,有钟控生成器和缩减型生成器。缩减生成器因为其结构简单而引人注目,目前还没有成功的攻击方法。国内外很多学者专家设计出很多缩减型生成器,包括互缩生成器和自缩生成器^[1],并在理论上证明了这些生成器的安全性和实际可用性。

文献[2]构造了[a, b]-自缩减生成器,从理论分析和局部随机性检验上表明其适合流密码系统的使用。文献[3]构造了广义互缩生成器的密钥流生成器,它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。本文综合这两种生成器的特点,利用两个反馈移位寄存器LFSR,构造了另一类新型缩减生成器——[a, b]-缩减生成器,通过分析得到[a, b]-缩减生成器的输出序列具有良好的性质:指教级周期,指教级线性复杂度和良好的统计特性。理论分析和局部随机性检验表明[a, b]-缩减生成器适合流密码系统的使用。

1 [a, b]-缩减生成器及其性质

1.1 [a, b]-缩减生成器的构造

[a, b]-缩减生成器的基本组成部分是两个二元的线性移位寄存器A_{LFSR}和B_{LFSR},设A_{LFSR}的级数为n_A,输出序列为{α_t}_{t≥0},B_{LFSR}的级数为n_B,输出序列为{β_t}_{t≥0}。设x_t=aα_t+b(α_t+1),定义累计函数G_A(t)= $\sum_{i<t} x_i$,G_A(0)=0,则[a,

b]-缩减生成器的输出序列按以下方式缩减产生:若α_t=1,则输出β_{G_A(t)},否则不输出。这样得到的序列为{Z(t)}_{t≥0}。等价地,定义:

$$G_s: \{0, 1, 2, \dots\} \rightarrow \{0, 1, 2, \dots\}$$

$$G_s(t-1) = G(t')$$

其中,α_{t'}=1且t为α₀,α₁,...,α_{t'}中1的个数,则[a, b]-缩减生成器的输出序列{Z(t)}_{t≥0}为:Z(t)=β_{G_s(t)}。

1.2 [a, b]-缩减生成器的性质

设A_{LFSR}的周期为P_A,一个周期内1的个数为N₁,B_{LFSR}的周期为P_B。由[a, b]-缩减生成器的构造方法可知,当A_{LFSR}输出P_AP_B个比特时,由于:

$$G_s(P_A P_B) = P_B G_A(P_A)$$

$$G_s(P_A P_B + 1) = P_B G_A(P_A) + x_0$$

$$G_s(P_A P_B + 2) = P_B G_A(P_A) + x_0 + x_1, \dots, G_A(t)$$

重复出现,因此[a, b]-缩减生成器的输出序列为周期序列。

当A_{LFSR}输出P_AP_B个比特时,[a, b]-缩减生成器输出P_BN₁个比特。因此[a, b]-缩减生成器输出序列的周期P_Z|P_BN₁。在一定条件下,[a, b]-缩减生成器输出序列的周期P_Z=P_BN₁。

定理1 设c=max{a, b},A_{LFSR}的级数为n_A<P_B/c,且(P_B, G_A(P_A))=1,则P_Z=P_BN₁。

证明 由上面的分析有P_Z|P_BN₁,只要再证明P_BN₁|P_Z即可。

由[a, b]-缩减生成器的构造有:Z(i+jN₁)=β_{G_s(i)+jG_A(P_A)},又由周期序列的定义知:Z(i+jN₁)=Z(i+jN₁+P_Z),即β_{G_s(i)+jG_A(P_A)}=β_{G_s(i+P_Z)+jG_A(P_A)},因此有P_B|G_s(i+P_Z)-G_s(i)成立。

设存在j_i使得G_s(i+P_Z)-G_s(i)=j_iP_B成立。当i=

收稿日期:2008-11-05;修回日期:2009-01-20。

作者简介:廖翠玲(1985-),女,福建泉州人,硕士研究生,主要研究方向:密码理论、控制序列分析; 余昭平(1962-),男,安徽宿松人,教授,主要研究方向:密码理论、信息安全。

$i+1$ 时, 存在 j_{i+1} 使得 $G_s(i+1+P_Z) - G_s(i+1) = j_{i+1}P_B$ 成立。因此: $G_s(i+1+P_Z) - G_s(i+P_Z) = G_s(i+1) - G_s(i) + (j_{i+1} - j_i)P_B$ 。

由于 n_A 级 A_{LFSR} 的输出序列不存在 n_A 长的 0 游程, 因此 $G_s(i+1) - G_s(i) \leq n_A c$, 从而 $j_{i+1} - j_i = 0$, 否则 $P_B \leq n_A c$, 这与已知条件矛盾。从而有 $G_s(i+1+P_Z) - G_s(i+P_Z) = G_s(i+1) - G_s(i)$ 成立, 即序列 $\{\alpha_t\}_{t \geq 0}$ 中以 α_i 为起点的序列和以 $\alpha_{(i+P_Z)}$ 为起点的序列相同。这说明 $P_A | (i+P_Z)' - i'$, 也就是 $\alpha_{i'}, \dots, \alpha_{(i+P_Z-1)'}$ 中元素为 1 的个数(有 P_Z 个)为 N_1 的倍数, 从而 $N_1 | P_Z$ 。

不妨设 $P_Z = \mu N_1$, 则对任意的 j , 有 $Z(0) = Z(jP_Z)$, 从而 $\beta_{G_s(0)} = \beta_{G_s(0)+j\mu G_A(P_A)}$ 。因此 $P_B | \mu G_A(P_A)$, 又 $(P_B, G_A(P_A)) = 1$, 于是 $P_B | \mu$, 从而 $P_B N_1 | \mu N_1 = P_Z$ 。因此, $P_Z = P_B N_1$ 。

定理2 设 A_{LFSR} 和 B_{LFSR} 为 m -序列, $c = \max\{a, b\}$, A_{LFSR} 的级数为 $n_A < P_B/c$ 且 $(P_B, G_A(P_A)) = 1$, 则 $[a, b]$ -缩减生成器的输出序列的线性复杂度 LC 满足: $P_B \cdot 2^{n_A-2} < LC \leq n_B \cdot 2^{n_A-1}$ 。

证明 由定理1知道, $P_Z = P_B N_1$ 。如果存在一个 d 次多项式 $p(x)$, 使得对任意的 $t \geq 0$, 有 $\sum_{i=1}^d p_i Z(i+t) = 0$, 则 $\{Z(t)\}_{t \geq 0}$ 的线性复杂度至多为 d 。设 $(Z_{k+t}^{N_1})$ 表示对 $\{Z(t)\}_{t \geq 0}$ 的 N_1 采样, 则它相当于对序列 $\{\beta_t\}_{t \geq 0}$ 的 $G_A(P_A)$ 采样, 由于 $(P_B, G_A(P_A)) = 1$, 因此 $(Z_{k+t}^{N_1})$ 也为 m -序列且线性复杂度为 n_B 。故存在 n_B 次多项式 $q(x) = q_0 + q_1 x + \dots + q_{n_B} x^{n_B}$, 使得对 $k = 0, 1, \dots, N_1 - 1$, 有: $\sum_{i=1}^d q_i Z(k + (i+t)N_1) = 0$, 因此序列 $Z_{tN_1}, Z_{tN_1+1}, \dots, Z_{N_1-1+(t+n_B)N_1}$ 满足递归关系式 $\sum_{i=0}^{nB} p_i Z(i+t) = 0$ 。这里: $P_j = \begin{cases} 0, & j \bmod N_1 \neq 0 \\ q_{j/N_1}, & j \bmod N_1 = 0 \end{cases}$, 故存在多项式 $p(x) = q(x)^{N_1}$, 满足 $p(Z(t)) = 0$ 。因此 $p(x)$ 的次数为 $n_B N_1$, $LC \leq n_B \cdot 2^{n_A-1}$ 。

设序列 $\{Z(t)\}_{t \geq 0}$ 的极小多项式为 $g(x)$ 。假设 $p(x) = q(x)^r, r \leq 2^{n_A-2}$, 则 $g(x) | [q(x)]^{2^{n_A-2}}$, 又 $q(x) | 1 - x^{P_B}$, 因此 $g(x) | [1 - x^{P_B}]^{2^{n_A-2}} = 1 - x^{P_B 2^{n_A-2}}$, 这与 $g(x)$ 的最小周期为 $P_B 2^{n_A-1}$ 矛盾。因此 $LC > P_B \cdot 2^{n_A-2}$ 。

综上, 有 $P_B \cdot 2^{n_A-2} < LC \leq n_B \cdot 2^{n_A-1}$ 。

线性复杂度具有不稳定性, 因此有必要研究密钥流序列的线性复杂度的稳定性^[4]问题, 下面分析 $[a, b]$ -缩减生成器的输出序列的重量复杂度和 k -错线性复杂度。

定理3 设 A_{LFSR} 和 B_{LFSR} 为 m -序列, $c = \max\{a, b\}$, A_{LFSR} 的级数为 $n_A < P_B/c$, 且 $(P_B, G_A(P_A)) = 1$, 则 1-重量复杂度 $WC_1(Z) \geq 2^{n_A-1}(2^{n_B} - n_B - 1)$ 。

证明 设 $f(x)$ 为 $[a, b]$ -缩减生成器的输出序列 $\{Z(t)\}$ 的极小多项式, $Z^{p_Z}(x) = Z(0) + Z(1)x + \dots + Z(p_Z-1)x^{(p_Z-1)}$, 则 $f(x) = \frac{1 - x^{p_Z}}{\gcd(Z^{p_Z}(x), 1 - x^{p_Z})}$ 。

再设 $g(x) = \gcd(Z^{p_Z}(x), 1 - x^{p_Z})$, $g_1(x) = \gcd(Z^{p_Z}(x) + x^i, 1 - x^{p_Z})$, 则:

$$g_1(x) = \gcd(Z^{p_Z}(x) + x^i, 1 - x^{p_Z}) =$$

$$\gcd(Z^{p_Z}(x) + x^i, f(x)g(x)) =$$

$$\gcd(Z^{p_Z}(x) + x^i, f(x))$$

由 1-重量线性复杂度的定义有:

$$WC_1(Z) = \min_{0 \leq i < p_Z} \left\{ \deg \left(\frac{1 - x^{p_Z}}{g_1(x)} \right) \right\} \geq \deg \left(\frac{1 - x^{p_Z}}{f(x)} \right) = 2^{n_A-1} \cdot (2^{n_B} - 1) - n_B \cdot 2^{n_A-1} = 2^{n_A-1}(2^{n_B} - n_B - 1)$$

当 $n \geq 3$ 时, 有: $WC_1(Z) \geq n_B \cdot 2^{n_A-1} \geq LC(Z)$ 。

定理4 设 A_{LFSR} 和 B_{LFSR} 为 m -序列, $c = \max\{a, b\}$, A_{LFSR} 的级数为 $n_A < P_B/c$ 且 $(P_B, G_A(P_A)) = 1$, 则当 k 为奇数时, $LC_k(Z) = P_Z$, 当 k 为偶数且 $\gcd(x^{i_1} + x^{i_2} + \dots + x^{i_k}, 1 - x^{p_Z}) \neq (1 - x)^{P_Z - LC(Z)}$ 时, $LC_k(Z) \geq LC(Z)$ 。

证明 由定理1知序列 $P_Z = P_B N_1$, 根据有限域的知识有: $\sum_{i \geq 0} Z(i)x^i = Z^{p_Z}(x)/(1 - x)^{p_Z} = u(x)/f_Z(x)$, 若 $(u(x), f_Z(x)) = 1$, $f_Z(x)$ 就是序列 $\{Z(t)\}$ 的极小多项式, $f_Z(x)$ 是 $(1 - x)$ 的幂次形式。

1) 当序列 $\{Z(t)\}$ 改变一个符号时, $\sum_{i \geq 0} \tilde{Z}(i)x^i = Z^{p_Z}(x) + x^i / (1 - x)^{p_Z}$, 同时 $(1 - x) | Z^{p_Z}(x)$, 否则序列 $\{Z(t)\}$ 的线性复杂度达到最大。 $\gcd(Z^{p_Z}(x) + x^i, (1 - x)^{p_Z}) = 1$, 因此 $LC_1(Z) = P_Z$ 。

2) 当序列 $\{Z(t)\}$ 改变两个比特时, $\sum_{i \geq 0} \tilde{Z}(i)x^i = Z^{p_Z}(x) + x^{i_1} + x^{i_2} / (1 - x)^{p_Z}$, 若: $x^{i_1} + x^{i_2} = (1 - x)^m h(x)$, 则:
 $\gcd(Z^{p_Z}(x) + x^{i_1} + x^{i_2}, (1 - x)^{p_Z}) = \begin{cases} (1 - x)^{P_Z - LC(Z)}, & P_Z - LC(Z) < m \\ (1 - x)^m, & P_Z - LC(Z) \geq m \end{cases}$

从而当: $\gcd(Z^{p_Z}(x) + x^{i_1} + x^{i_2}, (1 - x)^{p_Z}) \neq (1 - x)^{P_Z - LC(Z)}$ 时, $LC_2(Z) \geq LC(Z)$ 。

根据上面的推导可以得到: 当 k 为奇数时, $LC_k(Z) = P_Z$; 当 k 为偶数且 $\gcd(x^{i_1} + x^{i_2} + \dots + x^{i_k}, 1 - x^{p_Z}) \neq (1 - x)^{P_Z - LC(Z)}$ 时, $LC_k(Z) \geq LC(Z)$ 。

定理5 设 $c = \max\{a, b\}$, A_{LFSR} 的级数为 $n_A < P_B/c$ 且 $(P_B, G_A(P_A)) = 1$, 记 B_{LFSR} 一个周期内元素 1 的个数为 N'_1 , 元素 0 的个数为 N'_0 , 则在 $[a, b]$ -缩减生成器的输出序列的一个周期内有 $N'_1 N_1$ 个 1, $N'_0 N_1$ 个 0。

证明 $[a, b]$ -缩减生成器的输出序列的周期 $P_Z = P_B N_1$, 而 B_{LFSR} 一个周期内元素 1 的个数为 N'_1 , 元素 0 的个数为 N'_0 , 当 $\alpha_t = 1$ 才有输出, 因此 $[a, b]$ -缩减生成器的输出序列的一个周期内有 $N'_1 N_1$ 个 1, $N'_0 N_1$ 个 0。

2 局部随机性检测

$[a, b]$ -缩减生成器的输出序列具有良好的整体随机性质, 但是仅有整体随机性还是不够的, 必须同时考虑局部随机性质。当序列的两个方面的随机性较好时, 才能以较大的概率断定伪随机序列“接近”真随机序列, 适合流密码的使用。文献[5]介绍了局部随机性检测的各项指标, 包括频数检测、序偶检测、扑克检测、游程检测、自相关检测。

随机选取 50 条序列进行检验。这里以 m -序列为为例。

(下转第 1338 页)

小值。

以随机方式抽取三组实验数据,每组数据包括 10 000 个正常连接数据和 1 000 个入侵数据,攻击记录所占比例为 10%,为检验未知攻击能力,三组数据中均有不同的攻击类型,三组数据如表 1 所示。

表 1 实验数据组成

数据集	实例数	正常实例数	入侵实例数	攻击类型数	分类攻击数
1	11 000	10 000	1 000	21	4
2	11 000	10 000	1 000	15	4
3	11 000	10 000	1 000	9	3

表 2 是改进后的 FCM 算法和原始 FCM 算法对上述数据集的仿真实验对比结果,在此用以下两个参数来描述系统性能:

检测率(DR) = 已检测出来的异常连接数目/异常连接总数目

误检率(FPR) = 正常连接误报为异常连接的数目/正常连接总数目

表 2 攻击记录占 10% 时两种算法的比较^[1] %

数据集	改进的 FCM 算法		原始 FCM 算法	
	检测率	误检率	检测率	误检率
1	92.75	1.06	87.65	1.68
2	91.31	1.11	85.39	1.35
3	96.56	0.95	90.23	1.06

为进一步检验改进算法的效果,同样采用表 1 所示的三组实验数据,正常实例数取 10 000、入侵实例数取 100,攻击类型数和分类攻击数不变,攻击记录所占比例为 1%,实验结果如表 3 所示。

从表 2、3 可以看出,改进后的 FCM 算法相对于原始 FCM 算法,有着较高的检测率和较低的误检率。在实验中,改进算法检测速度提升不是很明显,大约提高 1.56% ~ 2.05%。实

(上接第 1335 页)

C_{LFSR} 为 12 级 m - 序列,反馈多项式为 $f(x) = x^{12} + x^7 + x^4 + x^3 + 1$, D_{LFSR} 为 11 级 m - 序列,反馈多项式为 $g(x) = x^{11} + x^2 + 1$ 。

C_{LFSR} 和 D_{LFSR} 的初态分别为 [0 1 1 0 1 1 0 0 0 1 0 1] 和 [1 1 0 1 1 1 0 1 0 0 1]。取 $a = 3, b = 5$,序列长度 $N = 5000$,决策门限值 $\alpha = 5\%$ 。局部随机性检测结果如表 1 所示。

表 1 局部随机性检测结果

检测方法	结果	检测方法	结果
指频率检测	0.7688	$d = 1$	- 1.1640
序偶检测	5.4522	$d = 2$	1.9520
$m = 5$	23.1000	$d = 3$	- 0.8063
扑克检测		$d = 4$	- 0.7074
$m = 8$	267.5184	$d = 5$	1.0612
游程检测	11.2351	$d = 6$	0.2547
		$d = 7$	0.3255
		$d = 8$	0.2831
		$d = 9$	- 1.7694
		$d = 10$	0.6512

改变参数 a, b, n_1, n_2 , 对产生的序列进行局部检验。结果发现, $[a, b]$ - 缩减生成器能通过局部随机性检测。

验结果表明,改进后的 FCM 算法对于入侵检测具有良好的可靠性和可行性。

表 3 攻击记录占 1% 时两种算法的比较^[2] %

数据集	改进的 FCM 算法		原始 FCM 算法	
	检测率	误检率	检测率	误检率
1	98.71	0.042	95.28	0.232
2	97.65	0.073	92.63	0.137
3	98.02	0.034	91.57	0.216

4 结语

本文提出了一种具有两阶段的模糊 FCM 聚类改进算法,通过对引入点密度函数加权系数,解决了样本点周围分布不均匀的问题,同时又借助加入样本特征矢量权重向量 W ,优化了样本特征矢量对分类贡献不均衡的情况,算法通过两个阶段优化了数据聚类结果,可应用于解决大数据集、情况复杂和实时性要求高的场合。最后结合网络入侵,利用 KDD CUP 99 数据集对该算法进行了仿真实验,结果表明该算法有效地提高了检测效率,算法具有良好的可靠性和可行性。

参考文献:

- [1] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms[M]. Norwell, MA, USA: Kluwer Academic Publishers, 1981.
- [2] 付辉. 模糊 C-均值(FCM)聚类算法的改进[J]. 科学技术与工程, 2007, 7(13): 3121~3123.
- [3] 高新波. 模糊聚类分析及其应用[M]. 西安: 西安电子科技大学出版社, 2004.
- [4] 朱卫夫, 王卫平, 梁亮. 基于模糊聚类分析的入侵检测方法[J]. 系统工程与电子技术, 2006, 28(3): 474~477.
- [5] PAL N R, BEZDEK J C. On clustering for the fuzzy c-means model [J]. IEEE Transactions on Fuzzy Systems, 1995, 3(3): 370~379.
- [6] KDD CUP 99 [DB/OL]. [2008-08-12]. <http://kdd.ics.uci.edu/databases/kddcup99/kdd-cup99.html>.

3 结语

本文利用两个反馈移位寄存器 LFSR 构造了一类新型的缩减生成器—— $[a, b]$ - 缩减生成器,通过分析得到 $[a, b]$ - 缩减生成器的输出序列具有良好的密码学性质:指教级周期,指教级线性复杂度和良好的统计特性。理论分析和局部随机性检验表明 $[a, b]$ - 缩减生成器适合流密码系统的使用。对 $[a, b]$ - 缩减生成器是否能抵抗现有的流密码攻击方法有待进一步的研究。

参考文献:

- [1] MEIER W, STAFFELBACH O. The self-shrinking generator [C]// Advances in Cryptology-Eurocrypt' 94, LNCS 950. Berlin: Springer-Verlag, 1995: 205~214.
- [2] 白恩健, 王静, 肖国镇. $[a, b]$ -自缩减生成器[J]. 计算机科学, 2004, 31(5): 107~110.
- [3] 高军涛, 董丽华, 胡予濮. 广义互缩减生成器[J]. 计算机学报, 2006, 29(6): 936~944.
- [4] 白恩健, 张斌, 肖国镇. 二元周期序列的线性复杂度和 k-错复杂度的关系[J]. 电子与信息学报, 2002, 24(12): 1821~1825.
- [5] RUKHIN A, SOTO J, NECHVATAL J, et al. A statistical test suite for random and pseudorandom number generators for cryptographic applications[M]. [S. l.]: NIST Special Publication, 2001: 13~98.