序列密码中密钥流生成器的安全性研究与分析

孙菁，傅德胜
南京信息工程大学计算机与软件学院，南京 210044

摘要：序列密码设计的核心在于密钥流生成器的设计，本文分析了基于线性反馈移位寄存器的序列密码密钥流生成器的工作机制，在研究Geffe和收缩式非线性组合生成器模型的重构条件之基础上，给出了一般情况下非线性组合生成器整体的重构条件。本文的工作对构建安全序列密码体制有很好的指导意义。

关键词：序列密码，LFSR，组合生成器
中图分类号 TN918

Research and analysis on the security of key generator

SUN Jing, FU De-sheng
School of computer and software, Nanjing university of information science and technology, Nanjing 210044

Abstract: Design of key generator is the core process in stream cipher. This paper based on the research of stream cipher and the reconstruction condition of Geffe and shrinking generator, proposed the reconstruction condition for general nonlinear combining generator, which would give guiding significance to the design for the secure stream cipher system.

Key Words: stream cipher; LFSR; combining generator

0 引言

序列密码密钥流生成器有多种结构，多数是用一个或多个具有最大周期的线性反馈移位寄存器（Linear_feedback_shift, LFSR）作驱动器来产生一系列状态序列，使之周期最长、统计特性良好；然后这些状态序列经过一个非线性组合函数f后得到高线性复杂度的密钥序列$\{k_i\}_{i\geq 0}$。自从线性移位寄存器的有效综合算法提出后[1],，人们对单个LFSR的复杂度及相关攻击进行了大量研究，包括序列的周期，线性复杂度等代数特性和相关函数，0-1分布及游程分布等统计特性[2], [3], [4]，但是这些研究只反映了单个LFSR的安全状态，并没有涉及整个组合生成器的安全策略。本文主要研究密钥流生成器在多个LFSR下的整体安全状态，在研究Geffe和收缩钟控式非线性组合生成器模型的重构条件之基础上，给出了一般情况下非线性组合生成器整体的重构条件。

1 线性反馈移位寄存器

1.1 LFSR 的工作机制

线性反馈移位寄存器由 n 个存储器与一个线性反馈函数组成，移位寄存器每次向右移动一位，新的最左边的位根据反馈函数计算得到，移位寄存器输出的位是最低位[5]。反馈函数是寄存器中某些位的线性函数。每一存储器称为 LFSR 的一级（或抽头），这些级的内容构成该 LFSR 的状态，每一状态对应于 GF（2）上的一个 n 维向量，共有 2^n 种可能的状态。

理论上 n 级 LFSR 在重复之前最多能产生 2^n-1 位长的状态（除去全零状态），但只有具有一定抽头序列 LFSR 才能循环的通过所有 2^n-1 位长的状态，这种序列称为 m 序列。

为了使 LFSR 有最大周期，抽头必须符合一定规则或者 LFSR 的特征多项式必须是本原多项式，不同级数下本原多项式系数见下表 1。N 次本原多项式的个数 $\lambda(n)=\frac{\phi(2^n-1)}{n}$，其中 ϕ 为欧拉函数。已经证明，对于任意的正整数 n，至少存在一个 n 次本原多项式，所以对于任意的 n 级 LFSR，至少存在一种连接方式使其输出序列为 m 序列。m 序列有密码学很多优良特性，如平衡性、移位可加性，它是最常采用的扩谱码序列，也是攻击者们致力于分析的序列。

表 1 本原多项式系数表

n	本原多项式系数	n	本原多项式系数
20	1C6873	25	3FADF21
30	74B5F959	35	892ED AC73
40	15A02FC7587	45	27D99F781D27
50	73E876C9DCD97	55	BEC81 BD6CE836 B
60	1FB7 B391826 A7E39	65	21A762 CC0F6015D57
70	4E055031 A9E62660FF	75	824EC073F B02 B82 A49F
80	16428CF31D321 B4D B6 B1 B	85	3189E29C9F12E75 B82CDD5

1.2 LFSR 内部特性

定理 1：GF（p）上的 n 级 LFSR 产生的 m 序列 z^∞ 具有以下性质[5]：

（1） z^∞ 最小周期为 p^n-1，线性复杂度为 n。

（2） 若 LFSR 的初始状态为全零，则输出 z^∞ 为全零序列。

（3） 若 LFSR 的初始状态为全 1，则输出 z^∞ 与输出位的奇偶性相关。

（4） 对任意 $X \in GF(p), X \neq 0,$ 在 z^∞ 的一个周期内 X 出现 p^n 次，0 出现 p^n-1 次。

（5） 如果 $z_{j-1} \neq a, (z_j z_{j+1} ... z_{j+r-1}) = (aa...a), z_{j+r} \neq a,$ 则称 $z_j z_{j+1} ... z_{j+r-1}$ 是一个长为 r

的 a 的游程。在序列z^{∞}的一个最小周期内：游程的总个数为 p^n-1；对于 1≤r≤n-2，每个$a \in GF(p)$，长为 r 的 a 的游程有（p-1）$^2p^{n-k-2}$个，长为 n-1 的 0 游程有 p-1 个；对于每个$a \neq 0$，长为 n-1 的 a 游程有 p-2 个，长为 n 的 a 游程有 1 个。

（6） 对于 1≤j≤p^n-1，令 x， y，…z 为抽头序列，b_{jx}， b_{jy}， b_{jz}为 j 轮抽头序列状态值，z_j输出 $b_{jx}\oplus b_{jy}\oplus \ldots \oplus b_{jz}$。

（7） 固定阶数和抽头的 LFSR 的输出序列中，输出序列的前一段是后一段的输入。

LFSR 安全分析

推论 1：对于单个阶长为 n 的 LFSR，假定能够获得输出序列中任意一段长度为 n 的连续串，根据 LFSR 的特性（6）、（7），就能够在不知道初始状态的情况下重构出该 LFSR 的所有输出。

如对于阶长为 4 的 LFSR，如图 2-1 所示，选取本原多项式 $p(x)=x^4+x+1$ 即选取抽头（4，1）使其成为一个 m 序列，周期为 2^4-1=15，设寄存器初始输入为 1010，经过往复循环运算，一个周期内它的输出序列是：101011001000111。

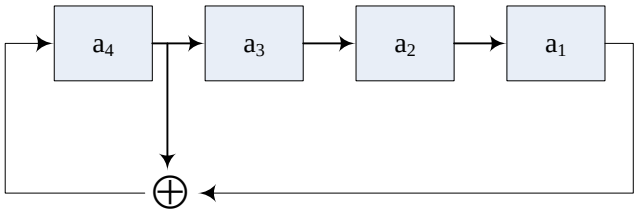

图 1 抽头为（4，1）的 4 级 LFSR

根据推论 1，将任意一段长度为 4 的连续输出序列放入阶长和本原多项式或抽头均与原始 LFSR 相同的线性反馈移位寄存器中进行运算，产生的输出序列与原序列相同。该 LSFR 的输出的表示为：$z_5= z_1\oplus z_4$，z_1 和 z_4分别表示当前轮 1 号、4 号寄存器中的值。由此可得，任何单个 LSFR 的输出都可由上述等式表示，即每一轮的输出都是当前轮抽头异或。

2 非线性组合生成器安全分析

真实设计的序列密码生成器往往采用多个 LFSR 级联相互反馈来获得更大周期序列和更好的伪随机性特征，加入了钟控或者复合控制器等等，分析起来十分困难。

2.1 Geffe 生成器

在 Geffe 发生器中，三个 LFSR 的长度分别为 n1，n2，n3。若 $LFSR_i$的输出序列为$\{s_k^i\}$，则输出序列$\{z_k\}$可以表示为$z_k = s_k^1 s_k^3 + s_k^2 \overline{s_k^3}$。如果将整个状态机看成一个特殊结构的 LFSR，令其为 LFSR4，根据输出模型，可以设定 LFSR1、 LFSR2、 LFSR3 的初始输入为

未知量，未知量的数量为 n1+n2+n3，那么只需获得长 n1+n2+n3 的输出序列就可以重构这三个 LFSR 的内部状态。

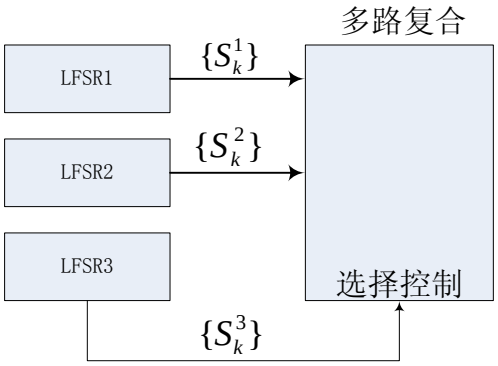

图 2 Geffe 发生器

2.2 收缩式反馈寄存器

收缩式反馈寄存器[6]采用不同类型的时钟控制。发生器使用两个 LFSR：LFSR1 和 LFSR2。它们分别按各自时钟运行，LFSR1 在时间 t-1 的输出为 1 时，LFSR2 在时间 t 输出为密钥流，否则舍去。

这种钟控序列具有许多良好的性质，但它局部性质不理想，可测度高[7]。整个状态机可以看作一个特殊设计的 LFSR，令其为 LFSR3，可知 LFSR3 的输出与 LFSR1 和 LFSR2 相关。若 LFSR2 的输出属于集合 N，LFSR3 的输出属于集合 S，则 S∈N。相同的 LFSR 不同的初始条件输出的两个序列，其中一个必是另一个的位移。位置值等于初始条件在序列 1 和序列 2 中出现的位置差，正数则右移，负数左移。

例如，一个 n=5，抽头序列为(5, 2)的 LFSR，对于两种不同的初始输入，其中第二组是第一组左移 1 位，所得到的输出序列见表 2-2。另第一组序列为参考序列，则第二组只需左移 1 位即可和第一组序列重叠。

表 2　n=5，抽头序列为(5, 2)的 LFSR 输入输出

输入	输出
10101	10101110110001111100110100100000
01010	01010111011000111110011010010000

根据这一分析，可以任意取 LFSR1 和 LFSR2 的初始状态，获得三组输出，令 LSFR1 的输出为 M1，LSFR2 的输出为 N1，LSFR1 和 LSFR2 收缩反馈后最终输出为 S1。检查 S1 是否属于 S，如果不属于，让 LFSR1 时钟延迟一位再计算，得到的输出元素属于集合 S2，检查集合 S2 是否属于集合 S，如此循环最多需要 x（$x=2^n-1$，n 为 LFSR 的阶数）次就可以重构收缩式反馈发生器。

2.3 一般情况下非线性组合生成器攻击分析

设$LFSR_i$长为l_i，初始状态为$K_i=(a_1^i,a_2^i,...,a_{l_i}^i)$，输出序列为$(s_1^i,s_2^i,...,s_{l_i}^i)$，非线性组

合生成器f（x）的输出即密钥流为=$(z_1^i, z_2^i, ..., z_{l_i}^i)$，而且P($s_k^i \neq z_k^i$)= P_i=0.5-ε $_i$ (i =1，2，…m)。一般情况下，设0.5≥ε $_i$≥0，称ε $_i$为LFSR 的相关系数.

为了方便研究，假设LFSR的特征多项式是已知或者公开的，那么密码系统的密钥就是各个LFSR的初始状态。对LFSR的特征多项式是未知的情形，搜索所有可能的特征多项式。假定能获得足够长的明密文比特流，对明密文比特流的对应位置比特相异或，就得到了最终用来分析LFSR的初始状态的部分密钥流。攻击条件如下：

（1）已知N长密钥流序列Z=($z_1, z_2, ..., z_N$)，设LFSR$_i$ 对应的输出序列是$(s_1^i, s_2^i, ..., s_{n_i}^i)$。当N接近惟一解距离l / (1-H($p_i$))（其中H(・)是熵函数）时，攻击的计算复杂度趋于一般的相关攻击的计算复杂度，故假设N≥l / (1-H(p_i))。

（2）已知非线性组合函数f（x）的输出与LFSR $_i$ (i=l，2，…，p)输出的相关性：P($s_k^i \neq z_k^i$)= P_i=0.5-ε $_i$。

（3）已知LFSR$_i$的特征多项式g_i (x)及其长度l_i。攻击目的是从密钥流截断Z=($z_1, z_2, ..., z_N$) 恢复出对应的LFSR$_i$的输出，再由LFSR$_i$的递推式得到初始状态$K_i = (a_1^i, a_2^i, ..., a_{l_i}^i)$。其实仅需要得到LFSR$_i$输出的$l_i$个位置上的正确值就可以确定其初始状态，一般是直接恢复对应的LFSR$_i$的输出 $(s_1^i, s_2^i, ..., s_{l_i}^i)$[8]。

2.1节中已谈到，由LFSR$_i$的线性递推式，容易得到$s_j^i (j > l_i)$ 可以由$(s_1^i, s_2^i, ..., s_{l_i}^i)$ 线性表示，则对于$(s_1^i, s_2^i, ..., s_N^i)$ = $(s_1^i, s_2^i, ..., s_{l_i}^i) G^i$

其中
$$G^i = \begin{bmatrix} w_{11}^i & w_{21}^i & \cdots & w_{N1}^i \\ w_{12}^i & w_{22}^i & \cdots & w_{N2}^i \\ \cdots & \cdots & \cdots & \cdots \\ w_{1l}^i & w_{2l}^i & \cdots & w_{Nl}^i \end{bmatrix}$$

那么 $(s_1^i, s_2^i, ..., s_N^i)$ 就可以看成一个[N, li]线性分组码，信息组是$(s_1^i, s_2^i, ..., s_N^i)$，

Z=($z_1, z_2, ..., z_N$)可以看成 $(s_1^i, s_2^i, ..., s_N^i)$ 通过一个相关概率为P_i =0.5-ε $_i$的二进制对称信道(Binary Symmetric Channel ,BSC)后的加噪声接收序列[9]，也就是

Z=($z_1, z_2, ..., z_N$) =$(s_1^i + e_1, s_2^i + e_2, ..., s_N^i + e_N)$

其中e_1，e_2 …，e_N是独立同分布的随机变量，而且P(e_k=1)=P_i =0.5一ε $_i$，P(e_k =1)=1/2+ε $_i$这样就可使用纠错码的技术恢复[9]。

3 小结

本文介绍了序列密码中密钥生成器工作原理——LFSR 的工作机制，分析了密钥生成器两大组成部分 LFSR 和 Geffe、钟控收缩生成器等非线性组合生成器的安全性，在此基础上提出了一般情况下非线性组合函数的重构条件。文中密钥流序列安全性分析对设计基于 LFSR 的序列生成器有一定的指导意义。

参考文献

[1] RUEPPEL R. A. Stream Ciphers [M].New York: IEEE Press,1992:65-134.

[2] COURTOIS N. T. Higher order correlation attacks, XL algorithm and cryptanalysis of Toyocrypt [A]. Information Security and Cryptology 2002 [C].Berlin: Springer-Verlag,2003.182-199.

[3] COURTOIS N. T. MEIER W. Algebraic attacks on stream ciphers with linear feedback [A]. Advances in Cryptology-Eurocrypt 2003 [C].Berlin: Springer-Verlag,2003.345-359.

[4] TEZUKA S. Lattice Structure of Pseudorandom Sequences from Shift Register Generators [A]. Proceedings of the 1990 Winter Simulation Conference[C].IEEE Press,1990.

[5] SCHNEIER B. 应用密码学[M].吴世忠，祝世雄，张文政，等译.北京:机械工业出版社 2000.

[6] MEIER W., STAFFELBACH O. The Self-Shrinking Generator [A]. Advances in Cryptology-Eurocrypt'94 [C].1994,205-214.

[7] Blackburn,S R. The linear complexity of the self-shinking generator [J].Information Theory IEEE Transactions,1999,45(6):2073-2077.

[8] 武传坤. 流密码的比特安全性[J]. 通信学报：1994, l5(1):73-77

[9] 杨礼珍, 傅晓彤, 肖国镇, 陈克非. 对非线性组合生成器的相关攻击[J]. 电子科技大学学报(自然科学版)：2001, 28(5): 566-568