第 35 卷 第 6 期 计 算 机 工 程 2009 年 3 月
Vol.35 No.6 Computer Engineering March 2009
·安全技术· 文章编号：1000—3428(2009)06—0192—02 文献标识码：A 中图分类号：TP309.2

# 基于 LFSR 的演化随机序列发生器

王玉华[1]，管爱红[1]，侯志强[1]，詹 静[2]，张焕国[2]
(1. 河南工业大学信息科学与工程学院，郑州 450001；2. 武汉大学计算机学院，武汉 430079)

**摘 要：**针对基于线性反馈移位寄存器的随机序列发生器产生的随机数线性复杂度低的问题，设计一个新的随机序列发生器，使用遗传算法演化线性反馈移位寄存器产生的随机序列，新产生的序列可以通过 SP800-22 的测试。测试结果表明，生成的序列周期大、线性复杂度高，能够满足安全协议和密码算法的安全强度要求。

**关键词：**随机序列；安全；遗传算法；线性反馈移位寄存器

# Evolutionary Random Sequence Generator Based on LFSR

WANG Yu-hua[1], GUAN Ai-hong[1], HOU Zhi-qiang[1], ZHAN Jing[2], ZHANG Huan-guo[2]
(1. School of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001;
2. School of Computer, Wuhan University, Wuhan 430079)

**【Abstract】**The random number generated by random sequence generator based on Linear Feedback Shift Register(LFSR) has low linear complexity. This paper proposes a new random sequencer generator to solve this problem. Genetic algorithm is used to evolve the sequences produced by LFSR to improve the linear complexity of the random number generator based on LFSR. The new sequence can pass the statistical test suit SP800-22. The result of tests shows the new sequence owns longer period and higher linear complexity, meeting the requirements of security protocols and encryption.

**【Key words】**random sequence; security; genetic algorithm; Linear Feedback Shift Register(LFSR)

## 1 概述

随着密码学的发展，随机数质量的要求也越来越高。在非对称算法中，公/私密钥对由随机位流提供[1]；在认证协议对称算法中，验证时生成密钥所需的填充字节和填充值要用到随机数；在智能卡的应用中，为了对抗旁路攻击而采取的对抗措施也要求有高质量的随机数[2]。密码系统的安全性主要依赖于随机序列的不可预测性[3]，这成为许多数学库中随机数发生器设计的准则。

线性反馈移位寄存器(Linear Feedback Shift Register, LFSR)是一类随机位序列生成器，它可以应用于流密码，也能满足低功耗或高速度的要求[4]。LFSR 的数学意义很容易理解。选择一个反馈向量 $(a_0a_1\cdots a_n)$，这个向量相当于一个本原多项式，发生器将产生一个随机位序列 $(s_0s_1\cdots s_m)$，用于重复所有可能的 $2^n-1$ 线性状态。由于输出位和内部状态的相互依赖性可以通过一个线性方程式构建，自身不能提供一个随机强度足够大的位流序列，因此即使是最大序列的 LFSR，在流密码的设计中也不予以考虑。LFSR 的一些非随机特性制约了它在密码学中的应用：(1)由长度为 $n$ 的 LFSR 生成序列的前 $n$ 位是 LFSR 的内部初始状态。只要能得到连续的 $2n$ 位，就能够破解 LFSR 的反馈函数。(2)LFSR 随机序列的线性特征使它不能直接用于加密。而且，LFSR 随机序列具有较高的相关性，这也限制了它在一些领域的应用。然而，增强它的复杂性可以提高 LFSR 抗线性攻击的能力，因此，LFSR 在伪随机数的生成中是一个有用的模块。从硬件实现的角度，LFSR 可以用触发器链实现，这些触发器不仅能提供一个寄存位，而且能在不增加触发器的情况下提供一个置反位，从而节省了硬件成本，提高了硬件集成度。因此，许多非线性的方法被用来改善 LFSR 的随机特性。

遗传算法具有很好的函数逼近特性，这使它在许多科学领域成为一个有用的工具。本文利用遗传算法设计了一个基于 LFSR 的新型随机序列发生器。

## 2 线性反馈移位寄存器

LFSR 是用来生成二进制位序列的一种机制。寄存器由许多基本存储单元构成，这些基本存储单元由一个初始化向量来填充，即由密钥种子来填充。寄存器由时钟控制，在每一个时钟节拍基本存储单元的值右移一位，基本存储单元的值的一个子集异或后放置到最左边的单元，在这个更新的过程中，可以获得一个位的输出。

LFSR 的结构如图 1 所示。

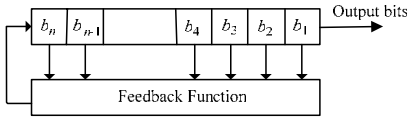


**图 1 LFSR 的结构**

一个长度为 $n$ 的 LFSR 具有的最大内部状态为 $2^n$，由于“0”状态是全封闭的，因此它的周期最大为 $2^{n-1}$。当抽头序列加 1 构成的多项式是本原多项式时，LFSR 具有最大的周期为 $2^{n-1}$，$n$ 位 LFSR 必须在有限域 $GF(2)$ 上找到指数为 $n$ 的

**基金项目：**国家自然科学基金资助项目(60673071, 60373087, 90104005, 60473023)；河南工业大学博士基金资助项目(2007BS043)
**作者简介：**王玉华(1973－)，女，讲师、博士，主研方向：人工智能，信息安全；管爱红，讲师、博士；侯志强，硕士；詹 静，博士；张焕国，教授、博士生导师
**收稿日期：**2008-09-20 **E-mail：**yuhua.w@tom.com

本原多项式。从密码学安全的角度考虑，周期越长，安全性越高。然而随着 $n$ 的增大，指数为 $n$ 的本原多项式生成也越来越困难。通常，通过 $2^{n-1}$ 次因式分解，可以确定指数为 $n$ 的多项式为本原多项式。而且，指数为 $n$ 的本原多项式很难攻破。算法越复杂，攻破的难度也越大。本原多项式的指数越大，LFSR 就越长，本原多项式的项数越多，LFSR 的抽头序列就越长，实现也就越困难。

## 3 遗传算法

遗传算法被广泛应用于解决复杂非线性问题。演化算法源自自然界中的进化过程。遗传算法通过计算机来模仿自然界中的生物遗传和自然选择。在遗传算法中，问题的参数由长度有限的位串进行编码，像人类一样，由核酸链构成的染色体(DNA 链)来编码生命形式。一个串可以是任何形式的符号序列，本文的演化随机序列发生器用二进制编码每个 LFSR 的位比例。

遗传算法不需要描述问题的所有特征，却可以根据自然原则产生更好的新解。它用简单的编码技术描述各种复杂的问题结构，确定搜索方向并指导学习。因为使用了群体搜索的策略，遗传算法可以在解空间的多个区间上同时进行搜索；而且很适合进行大规模并行处理过程。除了通过自组织、自学习、自适应来进行演化计算外，优胜劣汰的自然选择法则和简单的遗传操作使遗传算法不受解空间和其他辅助信息的限制。使用遗传算法时必须用一个基因组描述问题的解。遗传算法产生一个解的群体，应用变异和交叉等遗传操作演化解结构以获取最好的解。遗传算法有 3 个重要的因素：(1)必须定义目标函数；(2)必须定义和实现基因描述；(3)必须定义和实现遗传操作。许多不同的变化可以用来提高算法的性能，寻找多重优化或并行算法。

在随机位序列发生器的设计中，根据位比例，遗传算法被用来演化 LFSR 的组合结构。演化随机序列发生器是多个 LFSR 复杂的非线性组合。遗传算法扩展了序列的周期，简化了电路的设计。

## 4 演化随机序列发生器

实验中使用了标准的遗传算法，没有选择操作，群体比例变化后，整个群体作为新的一代。设定交叉概率为 $p_c$，变异概率为 $p_m$。适应度函数是遗传算法的关键，因此，使用 FIPS 140-1[4]测试组构建适应度函数。FIPS 140-1 包含 4 个测试，它们可以确定一个二进制的序列 $\{a_n\}$ 是否具有随机序列所要求的随机属性。本文对 FIPS 140-1 进行了 2 处修改：(1)增加了测试的强度。 FIPS 140-1 中的 Poker 测试 4 位的非重叠序列，本文使用 8 位的非重叠序列。(2)对这些测试使用了一个公式以确保 LFSR 的比例之和为 1。

(1)Monobit Test。这个测试表明序列中 0 和 1 的个数是否与预期的大致相同。用 $n_0$，$n_1$ 分别表示 0 和 1 的个数，则第 1 个统计量为

$$X_1 = \frac{(n_0 - n_1)^2}{n} \tag{1}$$

若 $n > 30$，则 $X_1$ 服从自由度为 1 的卡方分布。

(2)Max Run Test。这个测试确定序列中 0 和 1 的最大游程。在本文的 20 000 位测试中，最大游程不超过 34，第 2 个测试统计量 $X_2 \le 34$。

(3)Blocks and Gaps Test。这个测试表明 0 和 1 各种游程在序列中出现的频率是否与预期的频率相同。在长度为 $n$ 的序列中，长度为 $l$ 的 0 或 1 游程的期望值为 $e_l = (n-l+3)/2^{l+2}$。设 $k$ 为使 $e_l$ 成立的最大正整数 $l$，$B_l$，$G_l$ 分别表示序列中长度为 $l$ 的 0 和 1 的个数，$1 \le l \le k$。则第 3 个统计量为

$$X_3 = \sum_{l=1}^{k} \frac{(B_l - e_l)^2}{e_l} + \sum \frac{(G_l - e_l)^2}{e_l} \tag{2}$$

其中，$X_3$ 为近似服从自由度为 $2k-2$ 的卡方分布。

(4)Poker Test。设 $k \le n/m$，$n/m \ge 5 \times 2^m$，将长度为 $n$ 的序列分割成不重叠的 $k$ 个长度为 $m$ 的块。设 $n_i$ 是第 $i$ 个长度为 $m$ 的子序列出现的次数。Poker 测试表明每个长度为 $m$ 的分段是否以预期的比率出现。第 4 个统计量为

$$X_4 = \frac{2^m}{k}\left(\sum_{i=1}^{2^m} n_i^2\right) - k \tag{3}$$

其中，$X_4$ 为服从自由度为 $2^m - 1$ 的卡方分布。Poker 测试通常使用 $m = 4$ 位非重叠位，而本文使用 $m = 5$ 位非重叠位。

FIP 140-1 是一个通过/不通过的测试组，为了用 FIP 140-1 作为适应度函数，必须对上面的统计量进行公式化。本文对上述 4 个项测试统计量，构建了一个简单的适应度函数，设 $V$ 为 *Fitness* 的下界：

$$Fitness = X_1 + 0.11X_2 + X_3 + X_4 \tag{4}$$

为了实现随机位序列发生器的演化结构，选取 $m$ 个长度 $n$ 不同的 LFSR，要求 LFSR 反馈函数为 $GF(2)$上的本原多项式。新的随机位序列发生器构建如下：

(1)用 *Fitness* 对各个 LFSR 生成的序列进行评估。

(2)构建一个 LFSR 表，根据 *Fitness* 大小对所有的 LFSR 排序。

(3)根据轮盘赌的方法，确定各个 LFSR 的取位比例 $p_i$ $(1 \le i \le m)$。

(4)将各个比例值用二进制表示。

(5)以 4 000 位为一个单元，按比例依次从各个 LFSR 中取得 4 000 $p_i$ 位，按顺序将其构成一个子序列。

(6)重复(5)，直到取得足够位数的序列 $s$。

(7)对 $s$ 进行上述的 4 个测试，计算序列的 *Fitness*。

(8)如果 $Fitness > V$，进行(9)~(12)，否则结束。

(9)使用单点交叉，随机选取 2 对 LFSR，根据 $p_c$ 进行单点交叉。

(10)使用基本变异，随机选取 4 个，根据 $p_m$ 实行单点变异。

(11)计算 $\sum_{i=1}^{4\,000} p_i$：如果 $\sum_{i=1}^{4\,000} p_i > 1$，令 $\Delta = \sum_{i=1}^{4\,000} p_i - 1$；如果 $\sum_{i=1}^{4\,000} p_i < 1$，$\Delta = 1 - \sum_{i=1}^{4\,000} p_i$，随机选取 4 个比例值 $p_i$，将每个减去 $\Delta/4$。

(12)返回(5)。

基于 LFSR，演化随机位序列发生发生器能提供密码强度所要求的非线性以增加输出序列的非线性。由它生成的随机序列长度可以满足抽样的要求。而且位序列的周期大于单个 LFSR 的周期。不同的 LFSR 集合可以构建一个新的随机序列发生器。

## 5 实验结果

目前已有许多测试用于检验生成的随机序列的质量。在测试随机序列时，由于 SP800-22 是针对已经存储的序列进行的，因此可能有较明显的偏差。*P-value* 确定序列能否通过检测。序列通过检测的比例应该在正态分布的置信区间内。如果通过比例在可信区间内，随机序列发生器就是符合要求的。为了达到检测的目的，本文设定置信度为 99%，抽样尺寸为 $10^6$ 二进制位。表 1 给出了使用 SP800-22 测试组对演化随机

(下转第 196 页)