

基于钟控非线性序列的 RFID 伪随机数发生器设计

秦雪丽,程 明,李 伟

(郑州大学 信息工程学院, 郑州 450006)

(mrheqf@163.com)

摘 要:以 RFID 加密系统的伪随机数发生器为研究对象,提出以线性反馈移位寄存器(LFSR)为基本部件的复合型钟控非线性伪随机数发生器的设计方法。通过 Matlab 和 QuartusII 对该设计的周期、线性复杂度、均匀性、功耗等特征参数进行分析,最后硬件电路采用 FPGA 产品中低成本、低功耗的 Cyclone II 实现。此设计既保持了基本钟控非线性序列循环周期长、线性复杂度高的特性,同时提高了输出序列取值分布的均匀性,电路结构简单,并行输出 16 位数据,能够满足 RFID 加密系统的要求。

关键词:线性反馈移位寄存器;钟控非线性序列;线性复杂度;均匀性;现场可编程门阵列

中图分类号: TP309.1 **文献标志码:** A

RFID pseudorandom number generator based on clock-controlled non-linear sequence

QIN Xue-li, CHENG Ming, LI Wei

(College of Information Engineer, Zhengzhou University, Zhengzhou Henan 450006, China)

Abstract: Focused on the study of Pseudo-Random Number Generator (PRNG) in RFID encryption system, a design of complex clock-controlled PRNG was proposed, which is constituted by the linear feedback shift register (LFSR). Some characteristic parameters of the design were analyzed, such as sequence cycle, linear-complexity, uniformity, power consumption etc, through the Matlab and Quartus II software. Hardware circuit used the low-cost and low-power Cyclone II of FPGA products. This method maintains these properties of long cycle and high linear complexity of the clock-controlled non-linear sequence, but also improves the uniformity of output sequence. In addition, this design of circuit structure is simple and can output 16 bit parallel data. Therefore, it can satisfy the requirements of RFID encryption system.

Key words: Linear Feedback Shift Register (LFSR); clock-controlled non-linear sequence; linear complexity; uniformity; Field Programmable Gate Array (FPGA)

0 引言

随机数在仿真、信息安全和软件测试等领域都起着重要的作用,特别是在加密系统中,随机数是整个加密系统的安全基石。真正的随机数可以通过物理噪声产生器实现,但是很难获得,且造价高,易受周围环境噪声的干扰。目前,信息网络中通常使用的是以递推算法为基础的伪随机数。

基于线性同余递推算法的伪随机数发生器是现在使用较多的一种伪随机数发生器,此种发生器实现简单,但是种子 $x(0)$ 一旦确定,以后的数就被确定性产生了,而且由于受指令周期限制,伪随机数的生成速度慢,便于攻击者连续监测获取随机数。在密码分析中,敌手如果获得数列的极少一部分 $x(i)$,就可以确定出算法参数,从而破解密码。改进的方法是利用系统时钟修改随机数序列,但是由于系统时钟本身具有规律性,所以该方法也容易破解。

也有一些使用由多级线性反馈移位寄存器(Linear Feedback Shift Register, LFSR)产生的伪随机数序列,为了达到较大的周期,LFSR 的级数 n 通常取得很大。当 n 很大时,寻找其本原多项式的时间将急剧增加,反馈链长度增长,增加了电路设计的复杂性和逻辑电路的延时,限制了 LFSR 的最高工作频率,因此循环周期和最高工作频率之间的矛盾又限

制了 n 的选择。另外,多级 LFSR 线性复杂度低,容易通过一段连续随机数推测出生成结构,预测以后的随机数。

由于无线射频识别(Radio Frequency Identification, RFID)是在无线的开放式环境中进行数据交换,所以为攻击者窃取数据提供了机会。为了避免“身份认证”随机数和加密密钥被破解,实现重放攻击,随机数必须具有强抗分析特性,只有这样才能保证整个加密系统的安全。

1 原理

1.1 伪随机序列定义与评判标准

伪随机序列是指由确定性过程或算法产生,其特性类似于白噪声的序列。为了达到密码体制要求的安全保密标准,避免某些攻击的威胁,文献[1]中提出伪随机序列应该满足:

- 1) 随机序列的循环周期 T 要足够大;
- 2) 随机序列应具有良好的统计特性;
- 3) 随机序列其中的任何部分暴露时,要分析整个序列,提取产生它的电路结构或算法信息在计算上是不可行的。

以上这些条件是必要非充分的,但是要产生质量好的伪随机序列,必须满足这些条件。条件2)要求生成的随机序列能够满足或接近 Golomb 随机性公设,条件3)要求序列的线性复杂度,同时线性复杂度的稳定性要好。

收稿日期:2009-05-11;修回日期:2009-07-05。

作者简介:秦雪丽(1978-),女,河南郑州人,硕士,主要研究方向:RFID 中的信息安全、FPGA;程明(1949-),男,河南郑州人,教授,主要研究方向:电子技术、RFID;李伟(1981-),男,山东菏泽人,硕士,主要研究方向:数字信号处理、无线识别。

1.2 LFSR 的设计

如果移位寄存器的反馈函数 $f(a_1, a_2, \dots, a_n)$ 是 a_1, a_2, \dots, a_n 的线性函数,则称之为 LFSR。LFSR 的设计主要是找到一个最简特征多项式 (其反馈系数 c 尽可能多地为零),使其输出为 m 序列。

定义 1 LFSR 特征多项式:

$$p(x) = 1 + c_1x + \dots + c_{n-1}x^{n-1} + c_nx^n$$

使输出 x^n 为 m 序列的特征多项式称为本原多项式。为了简化硬件电路,可以通过 Matlab 编程找出最简单的本原多项式,但是随着级数 n 的增加,其最简本原多项式的寻找时间也急剧增加。通过编程计算发现:对于 n 级 LFSR,当 $p(x) = 1 + c_kx^k + c_nx^n$ 为本原多项式时,则 $p(x) = 1 + c_{n-k}x^{n-k} + c_nx^n$ 也为本原多项式,其中反馈链位置 c_k 和 c_{n-k} 关于 n 镜像对称,即 $c_k + c_{n-k} = n$,我们定义它们互为镜像本原多项式,因而,计算最简本原多项式只需计算 $n/2$ 次,缩短了实验时间,这对接下来的计算和选择非常方便。

LFSR 的输入输出反馈关系可以用下面的矩阵关系式表示:

$$\begin{bmatrix} x^n(T) \\ x^{n-1}(T) \\ \vdots \\ x^2(T) \\ x^1(T) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 1 \\ c_n & c_{n-1} & c_{n-2} & \dots & \dots & c_1 \end{bmatrix} \begin{bmatrix} x^n \\ x^{n-1} \\ \vdots \\ x^2 \\ x^1 \end{bmatrix}$$

即:

$$X(T) = CX \quad (1)$$

定义 2 矩阵 C 使 n 级 LFSR 输出为 m 序列的矩阵。其中 c_1, \dots, c_n 为该 LFSR 的本原多项式系数。

因为 LFSR 每个时钟周期仅输出一位 0/1 数据,为了提高带宽,将式(1)迭代 $w(w < n)$ 次得到:

$$X(wT) = C^w X \quad (2)$$

通过式(2)变换后 n 级移位寄存器以 C^w 矩阵为反馈链系数,构成可以在一个时钟周期内并行输出 w 位数据的 LFSR^[2],为了得到精简的电路结构,应使 C^w 矩阵内 c_i 为 1 的个数越少越好。根据上文所讲,每个 $p(x)$ 都有与之对称的镜像 $p(x)'$,因而也就有镜像 C' 矩阵存在,两个矩阵都可以使 n 级 LFSR 输出为 m 序列,所以可以选择最优矩阵 C 使 C^w 最简。研究发现: c_k 的位置离输出反馈 c_n 最近的矩阵 C 就是使 C^w 最简的反馈矩阵 (见表 1)。表 1 选择的是能找到最简本原多项式符合 $p(x) = 1 + c_kx^k + c_nx^n$ 形式的 LFSR,其中 $w = 16$ 。

表 1 同级数 LFSR 中 K 值与 C^{16} 矩阵中 1 的个数关系

LFSR 的级数 n	K	C^{16} 中 1 的个数
17	14	35
	3	68
18	11	39
	7	45
21	19	37
	2	93
22	21	38
	1	158
23	18	39
	5	57

综上所述:LFSR 的设计应结合所需周期,根据矩阵 C 和

C^w 的构成,选定合适的级数 n ,以便得到最简的反馈链,从而设计出既能满足周期要求,又最精简的电路系统。

1.3 钟控序列发生器

钟控序列发生器的基本模型就是用一个 LFSR1 控制另一个 LFSR2 的移位时钟脉冲,如图 1 所示。当 LFSR1 输出 1 时,移位时钟脉冲通过与门使 LFSR2 进行一次移位,从而生成下一位。当 LFSR1 输出 0 时,移位时钟脉冲无法通过与门影响 LFSR2,此时 LFSR2 重复输出前一位。

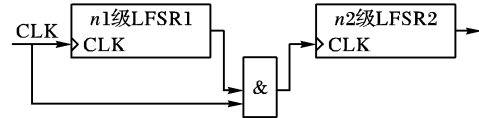


图 1 基本钟控序列生成器模型

假设 LFSR1 和 LFSR2 的级数分别为 $n1$ 和 $n2$,分别输出序列为 $\{a_i\}$ 和 $\{b_i\}$,记为 $G(i) = \sum a_i, (G(0) = 0)$ 其中 \sum 是在实数域上求和。那么钟控序列生成器的输出 $u(i) = b_{G(i)} (i \geq 0)$ 。

定理 1 假设图 2 钟控序列发生器的两个 LFSR 分别使用级数 $n1$ 和 $n2$,且 $n1$ 和 $n2$ 互素,则 $u(i)$ 的周期 $P = (2^{n1} - 1)(2^{n2} - 1)$,线性复杂度 $LC(u) = n2 \cdot (2^{n1} - 1)^{[3]}$ 。

通过定理 1 可以发现,钟控序列发生器与同样级数组成的 LFSR 和非线性组合发生器相比,其周期可以达到最大,线性复杂度是 $n2$ 和 $n1$ 指数级的乘积,这在其他 LFSR 组合发生器里是没有的。根据 Berlekamp-Massey 算法^[4],如果序列 u 的线性复杂度为 $LC(u)$,则只要知道该序列的任意 $2LC(u)$ 个连续比特,就可找到该序列所满足的齐次线性递归关系式,从而确定整个序列,因此密钥流序列的线性复杂度必须足够大。正是基于这种原因,我们选择钟控序列作为伪随机数发生器的核心部件。

2 随机数发生器设计

2.1 设计分析

虽然利用钟控序列发生器可以大大提升数据序列的线性复杂度,然而还不能直接用作伪随机数发生器。这是因为:

1) 根据上文所述,每个 LFSR 都可以通过 C^w 矩阵,形成新的反馈回路,使 LFSR 一次输出 w 位数据。然而,对于钟控序列发生器来讲,要在一个时钟周期内完成 w 位数据输出是个难题,这是因为 LFSR2 的时钟受 LFSR1 和 CLK 的双重控制。

2) 通过测试基本钟控序列发生器产生的数据组,发现数据分布的均匀性很差,不能满足随机数的均匀分布要求。

2.2 设计方案

综上所述,对基本钟控序列进行改进,使其和自相关函数为常数的 m 序列相加,如图 2 所示。

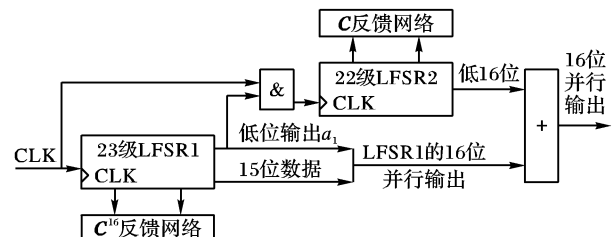


图 2 复合型钟控非线性序列伪随机数发生器设计框图

设计是由 LFSR1 和 LFSR2 组成一个钟控序列发生器, LFSR1 的反馈环路,即 C^{16} 矩阵,由 16 个或门组成,这样,

LFSR1 并行输出 16 位数据,其最低位 a_1 与时钟信号 CLK 相与后作为 LFSR2 的时钟控制信号,LFSR2 的低 16 位数据与 LFSR1 并行输出的 16 位数据异或后输出,输出波形(部分)见图 3。

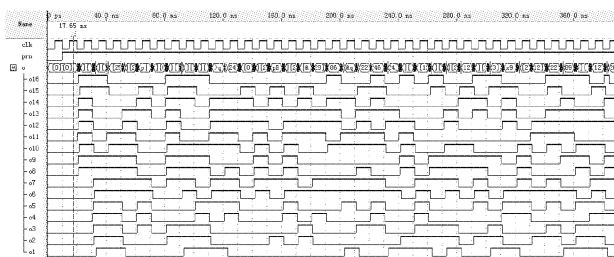


图 3 复合型钟控非线性序列伪随机数发生器部分波形

2.3 特征参数分析

在图 3 中,LFSR1 和 LFSR2 的级数分别为 $n_1 = 23, n_2 = 22$,周期分别为 $P1 = 2^{23} - 1, P2 = 2^{22} - 1$ 。钟控序列的输出为 $u(i), u(i)$ 周期 $P(u) = P1 \cdot P2, u(i)$ 的线性复杂度 $LC(u) = n_2 \cdot P1$ 。

下面对设计方案的随机性特征参数给予推导和证明。

1) 周期 $p(o) = \gcd(p(u), p1) = p(u) \approx 2^{45}$ 。

2) 线性复杂度分析:

定理 2 设伪随机序列 $\bar{a} = \{a_i, i \geq 0\}, \bar{b} = \{b_i, i \geq 0\}$ 的反馈函数分别是 $f_a(x), f_b(x)$, 那么 $L(\bar{a} \oplus \bar{b}) \leq L(a) + L(b)$ (这里 $\bar{a} \oplus \bar{b}$ 表示 $\{a_i \oplus b_i, i \geq 0\}$), 当且仅当 $(f_a(x), f_b(x)) = 1$ 时等号成立^[5]。

从输出关系式 $o(16 \times ((i-1), j) = \text{xor}(a(16 \times (i-1) + j), u(i+j-1))$ 可以看出, $u(i)$ 序列除了 $u(0)$ 之外,其余每一位都相当于重复了 16 次与 a_i 异或,可以推出与 a_i 异或的序列其实是 $u(i)$ 的 16 位扩展 $u'(i)$ 序列,即 $o(i) = a(i) \oplus u'(i)$,因而 $o(i)$ 是 $a(i)$ 和 $u'(i)$ 的并联生成序列, $o(i)$ 特征多项式 $f_o(x) = [f_{u'}(x), f_a(x)]$ 。

本设计中, $u(i)$ 和 $u'(i)$ 的特征多项式分别为:

$$f_u(x) = x^{LC(u)} + x^{2n_1} + 1$$

$$f_{u'}(x) = x^{2LC(u)} + x^{(2(n_1+1)-1)} + 1$$

因为 $f_a(x) = x^{23} + x^{18} + 1, f_{u'}(x)$ 与 $f_a(x)$ 互素,所以:

$$f_o(x) = x^{369098731} + x^{369098726} + x^{369098708} + x^{16777238} + x^{16777233} + x^{16777215} + x^{23} + x^{18} + 1$$

$$LC(o) = 369098731$$

3) 自相关函数分析:采用一个结构模型相同,级数缩减的伪随机数发生器进行举例模拟观测,其中 $n_1 = 7, n_2 = 6$,一次并行输出 2 位,通过 Matlab 对输出序列进行统计(见图 4)后发现,该设计输出序列虽然自相关值不具规律性,但具有良好的自相关特性。

2.4 采样测试与数据分析

1) o_{ij} 频数检验(采样 15000 点):1 的个数为 7477, 1 出现的概率为 $7477/15000 = 0.4985$,接近 0.5,可以认为通过检验。

2) o_{ij} 游程检验(采样 15000 点):令 $\pi = (\sum o_{ij})/n$,则理想游程期望值为 $2n\pi(1-\pi) = 7499$,实际游程数 $v_n = 7231$,接近理想游程数,可以认为通过检验。

3) 均匀分布检验:对 16000 点进行卡方分布拟合检验,在 $\alpha = 0.05$ 的显著水平上假设该伪随机数发生器生成数值满足 $(0, 65535)$ 上的均匀分布。

得出: $\sum (f(i) - np_i)^2 / np_i = 2.2535; i = 1, 2, \dots, 7$

因为 $2.2535 < \chi_{0.05(7-1)}^2 = \chi_{0.05(6)}^2 = 12.592$,所以假设成立。

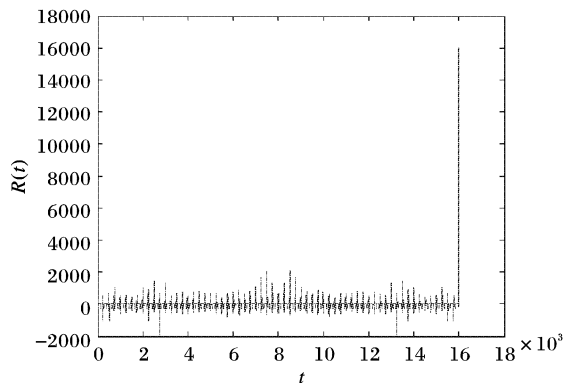


图 4 自相关函数统计

表 2 χ^2 检验计算表

$A(i)$	$L(i) \times 10^4$	$f(i)$	np_i	$(f(i) - np_i)^2 / np_i$
A(1)	[0, 1)	2418	2441.4	0.2243
A(2)	[1, 2)	2438	2441.4	0.0047
A(3)	[2, 3)	2469	2441.4	0.3120
A(4)	[3, 4)	2434	2441.4	0.0224
A(5)	[4, 5)	2494	2441.4	1.1333
A(6)	[5, 6)	2406	2441.4	0.5133
A(7)	[6, 6.5535)	1341	1351.3	0.0435

2.5 功耗分析

由于 RFID 系统对读卡器和标签的功耗、体积、成本都有特定的要求,所以设计方案的硬件采用低功耗、低成本的 Cyclone II EP2C5T144C6 芯片来实现。Cyclone II 采用全铜层、低 K 值、1.2 V SRAM 工艺设计,以 TSMC 成功的 90 nm 工艺技术为基础,功耗只有竞争低成本 90 nm FPGA 的一半,大大降低了静态和动态功耗。在 25℃ 温度下,通过 Quartus II PowerPlay Power Analysis & Optimization Technology 测试,设计方案的静态功耗是 18.04 mW,总功耗是 44.86 mW。UHF-RFID 应答器可以获得的典型功率在一百毫瓦数量级,因此,该随机数发生器的设计可以满足 UHF-RFID (915 MHz) 系统的要求。

3 结语

通过上面的分析和验证,本文设计的伪随机数发生器与多级 LFSR 相比,在级数相同的情况下,具有相近的周期长度,较高的线性复杂度,因而具有强抗分析特性,而且输出数据分布均匀,自相关特性良好。其硬件电路设计简单,便于实现,相对于软件递推算法,伪随机数的生成周期短,一个时钟周期可以并行输出 16 位数据,因此适用于体积小、功耗低、速度快的 RFID 系统。

参考文献:

- [1] 王相生. 序列密码设计与实现的研究[D]. 上海: 中国科学院上海冶金研究所, 2001.
- [2] 张文豪, 王春梅, 姚秀娟. 可配置 m 序列并行生成算法研究[J]. 微计算机信息, 2008, 24(4): 248-249.
- [3] 杨波. 现代密码学[M]. 2 版. 北京: 清华大学出版社, 2007: 30-32.
- [4] de VISME G H. Binary sequences[M]. [S. l.]: Hodder & Stoughton Ltd, 1971.
- [5] 杨义先, 林须端. 编码密码学[M]. 北京: 人民邮电出版社, 1992.