一类基于时变逻辑的序列发生器

张焕国,孟庆树
(武汉大学计算机学院软件工程国家重点实验室,湖北武汉 430072)

摘 要: 基于带记忆组合逻辑的序列发生器虽然抗传统的相关攻击,但易受线性时序电路逼近攻击.结合表更新的思想,本文给出了一类基于时变逻辑的序列发生器模型,并分析了输入输出间的相关性等密码学性质.许多密码体制都可归于此种模型,该模型对设计序列发生器有借鉴意义.

关键词: 密码学; 组合逻辑; 时变逻辑; 序列
中图分类号: TN918.1 **文献标识码:** A **文章编号:** 0372-2112 (2004) 04-0651-03

Sequence Generator Based on Time-Varying Binary Combiner

ZHANG Huan-guo, MENG Qing-shu
(*School of Computer Science, State Key Lab of Software Engineering, Wuhan University, Wuhan, Hubei 430072, China*)

Abstract: The binary sequence generator with memory is vulnerable to "linear sequential circuit approximation" attack though it can resist traditional correlation attack. With the idea of table-shuffling, a kind of binary sequence generators based on time-varying binary combiner are given. Some cryptographic properties like correlation between input and output are discussed. Several known sequence generators are of this model. The model can be useful in designing sequence generators.

Key words: cryptography; binary combiner; time-varying binary combiner; sequence

1 引言

仙农证明了一次一密密码体制可实现完全保密.但如何传输和保存大量的密钥成为问题.为了使用的方便,同时满足保密性的要求,人们设计了各种密码体制,在抗穷尽攻击的条件下用尽量少的种子密钥生产大量的序列作为报文密钥来加密明文.

在流密码领域,提出过许多体制的序列发生器.Rueppel将二元序列发生器分成两部分:驱动部分(线性移位寄存器组Linear Feedback Shift Registers,记作LFSR)和非线性组合部分.这种体制易受各种相关攻击.因为根据能量守恒定理:输出函数和输入的各种线性组合函数的相关系数的平方和等于1.这样当组合函数是平衡布尔函数时,最大相关系数将大于$2^{(-n/2)}$,其中n是变元个数,而平衡性是一种基本要求.于是引进了相关免疫的概念.但由Xiao-Massey[1]定理推出:对于有n个变元的逻辑函数,其代数次数和相关免疫阶之和小于等于n.为克服相关免疫阶和代数次数间的制约,引进了带记忆组合逻辑的概念.

对于一个有m比特记忆,n比特输入的组合逻辑,它的一般描述为:

$s_{t+1}=F(x_t,s_t)$, $t\geq 0$,用作状态转移函数.

$y_t=f(x_t,s_t)$, $t\geq 0$,用作输出函数.

其中$s_t=(s_{1t},\cdots,s_{mt})$是$t$时刻的状态向量;$s_0$是初始向量;$x_t=(x_{1t},\cdots,x_{nt})$是$t$时刻的输入向量.

对于输出函数,如果能知道s_t,则可将y_t分解成2^m个无记忆逻辑函数.如果一个密码体制采用密文作为状态向量s_t,则可按密文值进行分类,将该带记忆组合函数分解成2^m个无记忆组合逻辑,对这2^m个组合逻辑可按通常的相关分析技术进行相关攻击.

虽然带记忆组合逻辑抗传统的相关攻击,然而文[2]用向量函数的相关性作为工具,指出这种带记忆的组合逻辑易受线性时序电路逼近(Linear sequential circuit approximation,记作LSCA)的攻击.

综上,对于无记忆组合逻辑,由于相关性的存在,当在时序上积累了足够的信息时,可进行相应的相关攻击.带m比特记忆的组合逻辑可认为是2^m个无记忆的逻辑函数的组合,比起无记忆逻辑组合函数,通过将信息分散到各个子函数上,增加了密码的抗相关攻击的能力,但依然受LSCA攻击的威胁.如能设计一个不受m大小限制,由尽量多的无记忆逻辑函数组成的组合逻辑,则基于此的密码体制的强度应更好.结合表更新的思想(Table-shuffling),令m趋于无穷大,便得到时变逻辑.本文给出了一类基于时变逻辑的序列发生器,分析了输入输出间的相关性,分析了该发生器与其他发生器比较的优缺点.

收稿日期:2002-11-29;修回日期:2003-04-26
基金项目:国家自然科学基金(No.90104005;No.66973034;No.90204011,No.60373089).

2 基于时变逻辑的序列发生器

基于时变逻辑的序列发生器结构模型如图 1 所示. 时变逻辑控制部分负责控制时变逻辑的变化. 驱动部分作为输入, 经时变逻辑变换后输出作为生成器的最后输出, 用于加密明文.

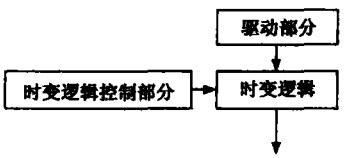


图 1 基于时变逻辑的序列发生器

有若干体制都可归于此种体制. Golic[3] 给出了一个体制并分析了其线性复杂度. 该称作 mem-bsg 的发生器是由三个 LFSR 和一个存储器组成的非线性二元生成器. 其中 LFSR1 生成读地址, LFSR2 生成写地址, LFSR3 根据写地址对存储器进行填充. 每一个时钟节拍, 根据读地址从存储器读出一项作为发生器的输出, 根据写地址将 LFSR3 生成的内容写入存储器. RC4[4] 也可认为是一种时变逻辑.

可以看出, 当我们用时变逻辑的角度来看时, 上述各发生器的关键在于如何设计时变逻辑. 我们设计的基于 LFSR 和时变逻辑的序列发生器如下:

(1)驱动部分 由能产生 序列的线性移位寄存器 LFSR1 组成.

(2)时变逻辑控制部分 由能产生 m 序列的线性移位寄存器 LFSR2、LFSR3 组成.

(3)时变逻辑 由取值为 0,1, 且 0,1 平衡的随机表 $s[0,1,\cdots,255]$ 组成.

(4)时变规则如下:

两个地址指针 a,b

a 为 LFSR2 的 8 个比特组成的整数

b 为 LFSR3 的 8 个比特组成的整数

交换 $s[a]$ 与 $s[b]$

(5)输出 由 LFSR1 的 8 个比特形成一个读地址 c, 输出为 $s[c]$, 用于加密明文. 称该发生器为时变逻辑序列发生器.

3 输入输出相关性

下面分析时变逻辑序列发生器的输入输出相关性.

定义 1 一个逻辑函数称作平衡的如果它的真值表中 0, 1 的个数相等.

定义 2 任意两个二元随机变量 x,y 的相关系数为: $c(x,y)=p\{x=y\}-p\{x\neq y\}$. 相应地, 两个逻辑函数的相关系数定义为 $c(f,g)=p\{f=g\}-p\{f\neq g\}$. 单个逻辑函数的相关系数 $c(f)$ 定义为 $c(f,0)$.

定义 3 称表 $s[0,1,\cdots,n-1]$ 是某个概率空间的一个随机表, 若其每一项都是取自该概率空间的随机变量, 且项与项之间相互独立.

定义 4 设 x,y 是某个概率空间的两个随机变量. 如果 $p(x|y)=p(x)$, 则称 x,y 是统计独立的.

定义 5 设 b 是某一概率空间的随机变量, 取值为 $\{0,1,\cdots,2^n-1\}$, $s[0,1,\cdots,2^n-1]$ 是另一概率空间的一个随机表. 如果 $p(s[i]|b)=p(s[i])$ 对任意 $i,b=0,1,\cdots,2^n-1$ 都成立. 则称随机变量 b 和表 $s[0,1,\cdots,2^n-1]$ 是统计独立的.

定义 6 称 T 是伪随机表 $s[0,1,\cdots,n-1]$ 的表周期, 如果存在一个常数 T_0, 当 $i>T_0$ 时, 有 $s[j]_i=s[j]_{i+T}, j=0,1,\cdots,n-1$ 成立.

引理 1[5] 令 $x_i, 1\le i\le r$, 是 r 个独立均匀分布的 n 维二元随机变量. 令 f 为 n 个变元的均匀随机平衡逻辑函数, 令 g 为任意 n 变元的平衡布尔函数. 令 $u_r(2^{n-2})$ 是概率分布 $\left\{\binom{2^{n-2}}{k}\Big/\binom{2^n}{2^{n-1}}\right\}_{k=0}^{2^{n-1}}$ 的中心矩. 则 $\sum_{i=1}^{r}f(x_i)$ 和 $\sum_{i=1}^{r}g(x_i)$ 的相关系数为 $c=2^{-r(n-2)}u_r(2^{n-2})$.

若 r 为奇数, 则 $c=0$; 若 r 为偶数, 则当 $2n$ 很大时, $c=2^{-rn/2}(r-1)!!$, 其中 $(r-1)!!=1*3*\cdots*(r-1)$.

定理 1 起始随机表 $s[0,1,\cdots,2^n-1]$ 是一个取值为 0,1 的平衡的随机表, 序列 $a=\{a_0,a_1,\cdots\}$, 序列 $b=\{b_0,b_1,\cdots\}$, 序列 $c=\{c_0,c_1,\cdots\}$ 分别是三个 n 维二元均匀分布随机变量所生成的序列. $s[0,1,\cdots,2^n-1]$、序列 a,b,c 相互独立. 交换 $s[a_i]$ 和 $s[b_i]$ (将 n 维二元随机变量看成是整数的二进制表示), 输出 $s[c_i], i=0,1,\cdots$. 记输出序列为 d, 则(1)序列 d 与 c 的任意线性组合的相关系数为 0(将 c 看作 n 维二元均匀分布随机变量). (2)序列 d 与序列 a,b 分别统计独立.

证明 (1)由于 $s[0,1,\cdots,2^n-1]$、序列 a,b 相互独立, 序列 a,b 是均匀随机分布的, 所以在序列 a,b 对随机表 s 的作用下, s 是均匀随机分布的平衡布尔函数. 由引理 1 的 $r=1$ 的情形, 序列 d 和序列 c 的任意线性组合的相关系数为 0(将 c 看作 n 维二元均匀分布随机变量).

(2) $p(s[i]|b=k)$ 的意义为: 在序列 b 取 k 的时刻, $s[i]$ 取值的概率分布情况. 由于 $p(b=k)=1/2^n$, 所以当 $b=k$ 时, 在 a,b 的作用下 $s[0,1,\cdots,2^n-1]$ 至少应被随机更新一次, 所以 $(s[i]|b=k)=p(s[i])$. 由 i,k 的任意性, 则 b 和随机表 $s[0,1,\cdots,2^n-1]$ 统计独立, 即 $p(s[i]|b)=p(s[i])$ 对任意 $i, b=0,1,\cdots,2^n-1$ 成立. 又序列 d 是以 c 为下标从表 s 中取出的, 所以 $p(d|b)=p(s[c]|b)=p(s[c])=p(d)$. 所以序列 b,d 统计独立. 同理可证序列 a,d 统计独立.

对于本文设计的时变逻辑序列发生器, 由于序列 a,b,c 是由 LFSR 生成的, 具有代数结构, 并不完全是均匀随机序列, 但可近似符合. 实验数据和理论分析基本一致. 见附录 1.

4 结论

该发生器的优缺点:

(1)抗传统的相关攻击和 LSCA 攻击. 由定理 1 知, 输出序列和驱动序列的任意线性组合相关系数为 0, 时变控制序列与输出序列统计独立. 从 LSCA 方法的攻击过程知该发生器抗 LSCA 分析.

(2)好的统计特性.

(3)和 men-bsg 的比较. 对于 mem-bsg, 当把 mem-bsg 发生器的存储器看作一张表时, 则表的周期为 LFSR2 和 LFSR3 的周期的最小公倍数. 另外文[6]指出由于一个 LFSR3 的内容直

接存入存储器内,这样在一段时间内(这段时间与表的大小有关系),该内容将被读出作为发生器的输出.所以从发生器的输出序列可推得 LFSR3 的一些信息,即此种生成器有熵漏.对于我们给出的发生器,当 LFSR1、LFSR2、LFSR3 相同,表的大小相同时,由于 LFSR2 和 LFSR3 形成的地址对表内容的交换,当运行的拍数为 LFSR2 和 LFSR3 的周期的最小公倍数时,表一般不会回到初始状态.所以我们给出的发生器具有更大的表周期.由于我们给出的时变逻辑控制序列不直接参与生成输出序列,有效避免了类似的熵漏.

该发生器的缺点是:由于线性移位寄存器的使用及对表内容的交换,软件实现速度受影响.构造高效率的时变逻辑有待研究.除 Golic[5] 用随机函数的方法对时变逻辑进行过分析外,类似文献较少.

从本文的分析看,基于时变逻辑的序列发生器从设计思想上能抗相关攻击和 LSCA 攻击.许多密码体制都可归于本文所给的模型,本文所给的模型对设计新的序列发生器有一定的借鉴作用.

附录 1

取三个本原线性移位寄存器:LFSR1 为 $x^{11}+x^2+1$,初态为 0X2*aa*;LFSR2 为 $x^{15}+x+1$,初态为 0X2*aaa*;LFSR3 为 $x^{17}+x^3+1$,初态为 0X*aaaa*.产生 $N=800000$ 比特的输出.

输出序列 d 和输入序列 c 的任意线性组合的相关性,其中 $T=(n_0-N/2)/\sqrt{N/4}$.

表 1 给出的是 c 序列(8bit)的所有线性组合中 $T>2.5$ 的结果.

表 1

线性组合系数	n_0(d 等于 c 组合值个数)	n_1(d 不等于 c 组合值个数)	T 值
18	401258	398742	2.812973
27	401294	398706	2.893472
61	401139	398861	2.546881
73	401383	398617	3.092482
120	401258	398742	2.812973

d 和 a 的统计独立性,其中 $T=(n_0-N/512)/\sqrt{N/512*(1-1/512)}$.表 2 给出了 a 取所有值,d 取 1 时相应的 $T>2.5$ 的结果.

表 2

a 的值	n_0($d=1$ 的个数)	T 值
19	1662	2.532701
191	1662	2.532701

参考文献:

[1] Xiao G Z, Massey J L. A spectral characterization of correlation-immune combining functions [J]. IEEE Trans. Inform. Theory, 1988, 34: 569 - 571.

[2] J D Golic. Correlation properties of a general binary combiner with memory [J]. Journal of Cryptology, 1996, 9: 111 - 126.

[3] J D Golic, M J Mihaljevic. Minimal linear equivalent analysis of a variable-memory binary sequence generator [J]. IEEE Trans. Inform. Theory, 1990, 36(1): 190 - 192.

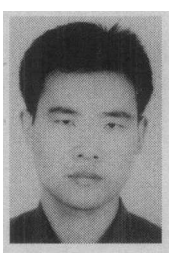
[4] Bruce Schneier. Applied Cryptography. John, second edition [M]. Wiley & Sons, 1996. 282.

[5] J D Golic. Linear models for a time-variant permutation generator [J]. IEEE Trans. Inform. Theory, 1999, 45(7): 2374 - 2382.

[6] 吕述望,范修斌. 序列密码的分析与设计 [M]. 北京:中软电子出版社, 2003. 6(2).

作者简介:

张焕国　男,1945 年出生于河北省元氏县,武汉大学教授,担任中国密码学会理事,中国计算机学会容错专业委员会委员,创建了全国第一个信息安全本科专业,发表论文 60 多篇,出版著作 6 部,主要研究领域:演化密码,纠错编码,安全计算机,智能卡,网络安全等.



孟庆树　男,1972 年出生于江苏省赣榆县,武汉大学博士研究生,主要研究领域:演化密码,编码密码学等.