滤波生成器的旋转对称攻击

杨 笑[1,2],武传坤[1]
(1.中国科学院软件研究所信息安全国家重点实验室,北京 100190; 2.中国科学院研究生院,北京 100049)

摘 要: 滤波生成器的安全性主要由滤波函数提供.为抵抗代数攻击,通常选取代数免疫函数作为滤波函数.我们发现已知的几类代数免疫函数都具有较强的旋转对称性,并在此基础上给出了一种针对滤波函数的旋转对称性质的攻击方法.我们还讨论了布尔函数的旋转对称性质以及该性质对旋转对称攻击的影响,分析了最优代数免疫函数对旋转对称攻击的脆弱性,提出了选取滤波函数的一个新准则.

关键词: 密码学; 滤波生成器; 旋转对称函数
中图分类号: TN918.1 **文献标识码:** A **文章编号:** 0372-2112 (2011) 03-0494-06

Rotation-Symmetric Attack on Filter Generators

YANG Xiao[1,2], WU Chuan-kun[1]
(1. *The State Key Laboratory of Information Security, Institute of Software, Chinese Academy of Sciences, Beijing* 100190, *China*;
2. *Graduate University of the Chinese Academy of Sciences, Beijing* 100049, *China*)

Abstract: The security of filter generators is provided by the filter function. For the resistance to algebraic attack, functions with maximum algebraic immunity were used for designing filter functions. We find that the existing algebraic immune functions have a strong property of rotation symmetry and present a rotation-symmetric attack on the filter functions. We also discuss the rotation-symmetric property of filter functions and its influence on the rotation-symmetric attack. After the survey of the vulnerability of algebraic immunity function to the rotation-symmetric attack, we give a new criterion for the choice of filter function.

Key words: cryptography; filter generator; rotation-symmetric function

1 引言

流密码算法是一类重要的加密算法,对流密码算法的设计与分析是密码学界的热点研究问题[1,2].滤波生成器是一种重要的流密码体制,由一个线性反馈寄存器和一个非线性滤波函数组成.线性反馈寄存器的作用是产生长周期的伪随机序列,滤波生成器的安全性主要由一个我们称为滤波函数的非线性布尔函数来提供.为了保证密码算法的安全性,针对非线性布尔函数,密码学界先后提出了非线性度、相关免疫度等指标,要求滤波函数具备高代数次数、高非线性度、平衡性、相关免疫性等密码学性质[1],目的是抵抗相继出现的攻击.

近年来代数攻击引起了密码学家的关注,其基本思想是将恢复密码系统的密钥或内部状态的问题归约成多变量方程系统的求解问题.对于滤波生成器,一个代数次数较高的滤波函数如果有一个较低的零化子,人们就可以利用这个性质降低方程系统的次数,进而求解方程系统[3].为抵御这种攻击,密码学家提出了代数免疫度的概念[4],具有最优代数免疫阶的非线性滤波函数的零化子的代数次数不低于$\lceil n/2 \rceil$[5~7].滤波函数具有最优代数免疫阶成为抵御代数攻击的一个必要条件.目前发现的具有最优代数免疫阶的布尔函数大都具有很高的对称性质或者是纯粹的对称函数[8,9],或者是修改对称函数真值表得到的[10,11].

值得注意的是,如果滤波函数具有较好的对称性质,那么这个对称性质就成为了滤波生成器的一个严重的安全漏洞.本文主要给出了一个利用这个安全缺陷对滤波生成器的攻击:每个时钟滤波生成器输出一个密钥流比特同时更新 LFSR 的 n 个内部状态.对 LFSR 内部状态的更新是一个线性变换,可以由一个线性方程表示.当密钥流输出的连续两个比特不等时,由选择的滤波函数的特殊性我们可以得到一个额外的线性方程.用这种方式获得 n 个额外的线性方程,连同前面提到的 LFSR 线性更新方程就可以用 Gauss 消元法恢复 LFSR 的内部状态完成攻击.输出密钥流中出现一次连续两个比特不同的情况,就可获得 1 个额外的线性方程.而要获

收稿日期:2010-01-22;修回日期:2010-07-20
基金项目:国家自然科学基金(No.60833008,No.60902024)

得 n 个额外的线性方程,则只需在输出密钥流中连续两个比特不同的情况出现 n 次.假设滤波生成器的输出有较好的统计性质,即连续两个时刻的密钥流输出相等的概率接近 1/2,则 $z_t,\cdots,z_{t+2n+1}$ 连续 $2n+2$ 个密钥流中相邻两个时刻密钥比特对不等的个数平均为 n.因此,这个攻击平均只需要 $2n+2$ 个输出密钥流比特.

2　旋转对称函数的基本概念

记 $\boldsymbol{X}^{(0)}=(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)})\in\{0,1\}^n$ 是 n 维布尔向量,ρ_n 是循环移位变换.定义 $\{0,1\}^n$ 上的等价关系:n 维布尔向量 $\boldsymbol{X},\boldsymbol{Y}$ 在同一等价类中当且仅当存在整数 i 使得 $\boldsymbol{Y}=\rho_n^i(X)$.$\rho_n$ 将 $\{0,1\}^n$ 划分为若干个等价类 $\boldsymbol{A}_i$,称每个等价类为一个圈,等价类的个数为圈的个数,ρ_n 定义在 $\{0,1\}^n$ 上的圈的个数为 $Z_n=\frac{1}{n}\sum_{t\mid n}\phi(t)2^{n/t}$,其中 $\phi(t)$ 是欧拉函数[12].称等价类 $\boldsymbol{A}$ 中元素个数为圈 $\boldsymbol{A}$ 的长度.任选 $\boldsymbol{X}^{(0)}\in\boldsymbol{A}$ 为等价类 $\boldsymbol{A}$ 中代表元,令 $\boldsymbol{X}^{(i+1)}=\rho_n(\boldsymbol{X}^{(i)})=\rho_n^{i+1}(\boldsymbol{X}^{(0)})$,若圈 $\boldsymbol{A}$ 的长度为 k,则对于 $0\leqslant i<j<k$ 有 $\boldsymbol{X}^{(i)}\neq\boldsymbol{X}^{(j)}$,而且 $\boldsymbol{X}^{(k)}=\rho_n(\boldsymbol{X}^{(k-1)})=\rho_n^k(\boldsymbol{X}^{(0)})=\boldsymbol{X}^{(0)}$.

例　对于 $n=4$ 的情况,ρ_4 将 $\{0,1\}^4$ 分划为 6 个等价类分别为

$\boldsymbol{A}_1=\{(0,0,0,0)\}$

$\boldsymbol{A}_2=\{(0,0,0,1),(0,0,1,0),(0,1,0,0),(1,0,0,0)\}$

$\boldsymbol{A}_3=\{(0,0,1,1),(0,1,1,0),(1,1,0,0),(1,0,0,1)\}$

$\boldsymbol{A}_4=\{(0,1,0,1),(1,0,1,0)\}$

$\boldsymbol{A}_5=\{(0,1,1,1),(1,1,1,0),(1,1,0,1),(1,0,1,1)\}$

$\boldsymbol{A}_6=\{(1,1,1,1)\}$

下面给出旋转对称布尔函数的定义.

定义 1　称布尔函数 f 为旋转对称函数,如果满足对任意 $(x_1,\cdots,x_n)\in\{0,1\}^n$,$1\leqslant k\leqslant n$ 等式

$$f(\rho_n^k(x_1,\cdots,x_n))=f(x_1,\cdots,x_n)$$

成立.这里 ρ_n^k 表示循环移位运算 $\rho_n^k(x_1,\cdots,x_n)=(x_{1+k \bmod n},\cdots,x_{n+k \bmod n})$,记 $\mathrm{Rot}S_n$ 是由所有 n 元旋转对称布尔函数组成的集合.

从旋转对称函数的定义不难发现 n 元布尔函数 f 是旋转对称函数,当且仅当 f 在 ρ_n 定义的任一圈上的不同元素的取值相同.

3　攻击旋转对称滤波函数

滤波生成器是一种密钥流生成器,包含线性反馈寄存器 LFSR 和非线性滤波函数两部分.记 $\boldsymbol{X}_t$ 为 LFSR 在 t 时刻的内部状态:$\boldsymbol{X}_t=(x_{t+n-1},\cdots,x_{t+1},x_t)$,其中 $x_{t+i}\in\boldsymbol{F}_2,0\leqslant i\leqslant n-1$.LFSR 的反馈多项式是 $L(x)$,滤波函数 f 是 $\boldsymbol{F}_2^n\to\boldsymbol{F}_2$ 的映射.每个节拍滤波生成器输出密钥流比特 $z_t=f(\boldsymbol{X}_t)$,并更新内部状态:$\boldsymbol{X}_{t+1}=(x_{t+n},\cdots,x_{t+2},x_{t+1})$,其中 $x_{t+n}=L(\boldsymbol{X}_t)$.当滤波函数 f 是旋转对称函数时,存在一种有效的代数攻击恢复滤波生成器 t 时刻的内部状态 $(x_{t+n-1},\cdots,x_{t+1},x_t)$.下面我们叙述这种针对旋转对称滤波函数的攻击方法:

根据旋转对称函数定义:对于 $\forall\,\boldsymbol{X},\boldsymbol{Y}\in\boldsymbol{F}_2^n$,如果 $f(\boldsymbol{X})\neq f(\boldsymbol{Y})$,则 $\rho_n(\boldsymbol{X})\neq\boldsymbol{Y}$ 成立.观察滤波生成器连续两个时刻的输出 z_t 和 z_{t+1},如果 $z_t\neq z_{t+1}$,即 $f(\boldsymbol{X}_t)\neq f(\boldsymbol{X}_{t+1})$,那么 $\rho_n(\boldsymbol{X}_t)\neq\boldsymbol{X}_{t+1}$.不等式左边部分 $\rho_n(\boldsymbol{X}_t)=\rho_n(x_{t+n-1},\cdots,x_{t+1},x_t)=(x_t,x_{t+n-1},\cdots,x_{t+1})$,不等式右边部分 $\boldsymbol{X}_{t+1}=(x_{t+n},x_{t+n-1},\cdots,x_{t+1})$,二者只有第一个分量元素不同,因此 $\rho_n(\boldsymbol{X}_t)\neq\boldsymbol{X}_{t+1}$ 推出 $x_t=x_{t+n}\oplus 1$.这样,就得到了一个关于滤波生成器内部状态的线性方程.观察 N 个连续的密钥流输出,如果存在 n 个不同的时刻 $\{t+i_k\mid 1\leqslant k\leqslant n\}$ 都有 $z_{t+i_k+1}\neq z_{t+i_k}$,那么就可以得到一个含 n 个方程等式的线性方程系统:

$$\begin{cases}x_{t+i_1}\oplus x_{t+i_1+n}=1\\ x_{t+i_2}\oplus x_{t+i_2+n}=1\\ \quad\vdots\\ x_{t+i_n}\oplus x_{t+i_n+n}=1\end{cases}\tag{1}$$

这 n 个线性方程连同 LFSR 自身 i_n+1 次状态更新产生的含 i_n+1 个线性方程等式的方程系统:

$$\begin{cases}x_{t+n}\oplus L(x_{t+n-1},\cdots,x_{t+1},x_t)=0\\ x_{t+n+1}\oplus L(x_{t+n},\cdots,x_{t+2},x_{t+1})=0\\ \quad\vdots\\ x_{t+i_n+n}\oplus L(x_{t+i_n+n-1},\cdots,x_{t+i_n+1},x_{t+i_n})=0\end{cases}\tag{2}$$

组成了一个变元数和方程个数都是 i_n+n+1 的线性方程系统,$\{x_i\mid t\leqslant i\leqslant t+i_n+n\}$ 是方程系统未知数的集合.只要这些方程相互独立,就可以使用 Gauss 消元法解这个线性方程系统求得 t 时刻内部状态,完成状态恢复攻击.

4　对一般滤波函数的攻击

上一节描述的攻击方法主要得利于密钥流生成器所使用的滤波函数 f 的旋转对称性质,但是当 f 不是旋转对称函数时,线性方程系统(1)不再成立,那么也就不能直接求解等式(1)和式(2)组成的线性方程系统完成攻击了.下面我们说明当滤波函数 f 不是旋转对称函数时,只要 f 具有较强的旋转对称性质,就可以推广前面针对旋转对称滤波函数的攻击来恢复滤波生成器的内部状态.

我们用上一节相同的方法写出方程系统(1),因为滤波函数不再是旋转对称函数,因此方程中每个等式都有一定的失败概率,记为 p'.这个失败概率 p' 就是在 $z_{t+i_t}\neq z_{t+i_t+1}$ 的条件下 $x_{t+i}\oplus x_{t+i_t+n}=0$ 的概率,即 $p'=$

$\Pr(x_{t+n}=x_t \mid z_t \neq z_{t+1})$. 记方程系统(1)中出错的方程个数为 r,则其平均值(期望)为 $E(r)=p'n$.

如果线性方程(1)中每个等式都成立,我们就可以通过求解等式(1)和式(2)联立的线性方程系统恢复 LFSR 的初始状态. 假设方程系统(1)中有 k 个等式是错误的,那么首先修改其中 k 个等式($k \leqslant r$),即选定其中的第 $j_1, j_2, \cdots, j_k$ 个等式,将 $x_{t+i_{j_s}} \oplus x_{t+i_{j_s}+n}=1$ 修改为 $x_{t+i_{j_s}} \oplus x_{t+i_{j_s}+n}=0$,这里 $1 \leqslant s \leqslant k$;然后尝试用 Gauss 消元法求解这个新的线性方程系统. 如果方程无解,则尝试对方程(1)的其他修改;如果有解,则验证求得的 $x_t, x_{t+1}, \cdots, x_{t+n}$的正确性. 验证的方法是:以 $x_t, x_{t+1}, \cdots, x_{t+n-1}$作为 LFSR 的内部状态,运行滤波生成器产生连续 n 个比特的密钥流输出 $z'_t, z'_{t+1}, \cdots, z'_{t+n-1}$. 然后和滤波生成器的实际输出 $z_t, z_{t+1}, \cdots, z_{t+n-1}$对比,如果$(z'_t, z'_{t+1}, \cdots, z'_{t+n})=(z_t, z_{t+1}, \cdots, z_{t+n})$则断言攻击成功,$x_t, x_{t+1}, \cdots, x_{t+n-1}$为滤波生成器 t 时刻内部状态,否则尝试对系统(1)的其他修改并重复上面步骤. 通过不断的尝试修改系统(1),最终可以成功恢复 t 时刻内部状态.

引理 1[13] $\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{r} \leqslant 2^{H(\frac{r}{n})n}$,其中 r, n 是正整数且 $r \leqslant n$, $H(x)=-(x\log x+(1-x)\log(1-x))$是熵函数.

定理 1 f 是图 1 中的滤波函数,记 $p=\Pr(f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t)))$,则上述对滤波生成器攻击的时间复杂度为 $O(2^{H(p)n} \cdot n^{2.37})$,平均需要 $2n+2$ 个输出密钥流比特,这里的 n 是线性反馈寄存器的级数.

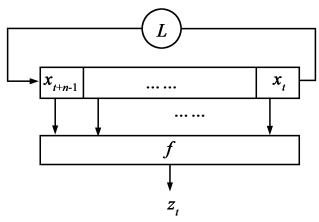

图1 滤波生成器

证明 恢复 t 时刻内部状态需要 $z_t, \cdots, z_{t+N}$连续 $N+1$ 个密钥流比特,其中 N 是使得集合 $S=\{i \mid 0 \leqslant i < N, z_{t+i} \neq z_{t+i+1}\}$恰好有 n 个元素的最小自然数. 假设滤波生成器的输出有较好的统计性质 $\Pr(z_{t+i} \neq z_{t+i+1}) \approx 1/2$,即连续两个时刻的密钥流输出相等的概率接近 1/2,则连续 $2n+2$ 个比特密钥流 $z_t, \cdots, z_{t+2n+1}$中相邻两个时刻密钥比特对不等的个数平均为 n. 因此上述攻击平均需要 $2n+2$ 个输出密钥流比特.

首先假设方程系统(1)中每个等式失败的概率为 p',那么系统中有平均 $p'n$ 个方程出错. 考虑最坏的情况,需要修改$\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{r}$次方程,每次修改的方程系统求解验证过程的时间复杂度为 $n^{2.37}$(主要是求解线性方程组的复杂度, 2.37 是 Gauss 消元法复杂度指数的近似值[14]). 根据引理 1,本攻击最坏情况的攻击时间复杂度为 $O(2^{H(p')n}n^{2.37})$.

下面说明 $p' \approx \Pr(f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t)))$, p'就是 $z_{t+1} \neq z_t$ 时等式 $x_t=x_{t+n}$成立的概率,即 $p'=\Pr(x_{t+n}=x_t \mid z_t \neq z_{t+1})$,由条件概率公式得到

$$\Pr(x_{t+n}=x_t \mid z_t \neq z_{t+1})=\frac{\Pr(z_t \neq z_{t+1} \mid x_{t+n}=x_t)\Pr(x_{t+n}=x_t)}{\Pr(z_t \neq z_{t+1})} \tag{3}$$

设 $\boldsymbol{a}=(x_0, x_1, x_2, \cdots)$是二元 m 序列,定义作用在$\underline{\boldsymbol{a}}$ 上的左移变换 $L: L(\underline{\boldsymbol{a}})=(x_1, x_2, \cdots)$,并定义 $L^0(\underline{\boldsymbol{a}})=\underline{\boldsymbol{a}}$, $L^t(\underline{\boldsymbol{a}})=L(L^{t-1}(\underline{\boldsymbol{a}}))$. 记$\underline{\boldsymbol{a}}=\{x_t\}_{t=0}^{\infty}$是 LFSR 的输出序列,则序列$\underline{\boldsymbol{b}}=\{x_{t+n}\}_{t=0}^{\infty}$是序列$\underline{\boldsymbol{a}}$ 的移位序列,$\underline{\boldsymbol{b}}=L^n(\underline{\boldsymbol{a}})$. 由文献[15]中推论 6.2.13 知$\underline{\boldsymbol{a}}+\underline{\boldsymbol{b}}=\{x_{n+t}+x_t\}_{t=0}^{\infty}$也是$\underline{\boldsymbol{a}}$ 的移位序列,因此$\{x_{n+t}+x_t\}_{t=0}^{\infty}$是 n 级线性反馈移位寄存器生成的 m 序列,进一步 $\Pr(x_t \oplus x_{n+t}=0)=(2^{n-1}-1)/(2^n-1) \approx 1/2$. 在流密码算法设计良好的情况下,$\Pr(z_t \neq z_{t+1}) \approx 1/2$,这样的假设是合理的,否则存在对该算法的区分攻击. 这样等式(3)可以化简为

$$\Pr(x_{t+n}=x_t \mid z_t \neq z_{t+1}) \approx \Pr(z_t \neq z_{t+1} \mid x_{t+n}=x_t) \tag{4}$$

计算 $\Pr(z_t \neq z_{t+1} \mid x_{t+n}=x_t)$:

$$\begin{aligned}
&\Pr(z_t \neq z_{t+1} \mid x_{t+n}=x_t)\\
&=\Pr(f(x_{t+n-1}, \cdots, x_{t+1}, x_t) \neq f(x_{t+n}, \cdots, x_{t+1}) \mid x_{t+n}=x_t)\\
&=\Pr(f(x_{t+n-1}, \cdots, x_{t+1}, x_t) \neq f(x_t, \cdots, x_{t+1}) \mid x_{t+n}=x_t)\\
&=\Pr(f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t)) \mid x_{t+n}=x_t)
\end{aligned} \tag{5}$$

在假设事件 $x_{t+n}=x_t$ 和事件$f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t))$独立的条件下,有

$$\Pr(z_t \neq z_{t+1} \mid x_{t+n}=x_t)=\Pr(f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t))) \tag{6}$$

由等式(4)、(6)就证明了 $p' \approx \Pr(f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t)))$.

因此,上面描述的攻击的时间复杂度为 $O(2^{H(p)n}n^{2.37})$.

我们看到这个攻击的时间复杂度是关于 n 和 p 的函数, n 是线性反馈寄存器长度, $p=\Pr(f(\boldsymbol{X}_t) \neq f(\rho(\boldsymbol{X}_t)))$是滤波函数自变量旋转移位后函数值产生变化的概率,简称为函数的旋转变化率. 在线性反馈寄存器长度不变的情况下, p 越接近 1/2,滤波生成器抵抗攻击的能力越强. 但是从应用的角度看, p 在很小的情况下,就可以使得攻击复杂度足够大. 下面给出 $n=100$ 时,不同的 p 对应的攻击复杂度以 2 为底的对数值(见表 1). 我们看到当 $p>0.17$ 时,攻击时间复杂度高于 2^{80}. 因此在实际的密码应用中,给定线性反馈寄存器级数 n 的大小,只要所选取得滤波函数的旋转变化率 p 适当大,就可抵御这种针对旋转对称性质的攻击.

表 1 $n=100$ 时攻击复杂度的对数值与滤波函数的旋转变化率 p 的关系

p	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
Log(T)	29.9	40.0	48.5	56.0	62.6	68.7	74.2	79.2	83.8	88.0

注:本节提出的旋转对称攻击的有效性主要取决于方程系统(1)中出错等式的比率大小.旋转对称变化率有效地刻画了这个比率.上面我们说明了当滤波函数具有较小(<10%)的旋转对称变化率时,旋转对称攻击有不错的攻击效果.事实上,当旋转变化率很大时(>90%)用类似的攻击方法也能够有效恢复 LFSR 的内部状态,不同只有方程系统(1)的生成规则:观察 N 个连续的密钥流输出,如果存在 n 个不同的时刻 $\{t+i_k \mid 1\leqslant k\leqslant n\}$ 都有 $z_{t+i_k+1}=z_{t+i_k}$,那么就可以得到一个含 n 个方程等式的线性方程系统(1).在这种情况下,方程系统(1)中等式失败的概率 $p'=\Pr(x_{t+n}=x_t \mid z_t=z_{t+1})\approx\Pr(z_t=z_{t+1} \mid x_{t+n}=x_t)=\Pr(f(\boldsymbol{X}_t)=f(\rho(\boldsymbol{X}_t)))=1-p$,攻击的时间复杂度为 $O(2^{H(1-p)n}n^{2.37})$.

5　布尔函数的旋转对称性质

布尔函数 f 到旋转对称函数类的距离是布尔函数的旋转对称性质的直观的度量指标,我们简称之为旋转对称距离.本节讨论旋转对称距离和上节提到的旋转对称变化率的计算方法以及二者的相互关系.

定义 2　n 元布尔函数 f 的旋转对称距离是指函数 f 到 n 元旋转对称函数集合 $\mathrm{Rot}S_n$ 的汉明距离,记为 $D^n_{\mathrm{Rots}}(f)$,即 $D^n_{\mathrm{Rots}}(f)=\min_{g\in\mathrm{Rots}(n)}d_H(f,g)$,其中 $d_H(f,g)$ 表示 n 元布尔函数 f 和 g 的汉明距离,$d_H(f,g)=wt(f\oplus g)$.

如果旋转对称函数 g 与 f 的汉明距离等于 $D^n_{\mathrm{Rots}}(f)$,则称 g 是 f 的一个旋转对称最佳逼近函数.易见 f 的旋转对称最佳逼近函数不唯一.

引理 2　若 n 元旋转对称布尔函数 g 是 f 的旋转逼近函数,圈 $\boldsymbol{A}$ 是 ρ_n 定义的任一等价类集合,则 g 和 f 在集合 $\boldsymbol{A}$ 上取值不同的点的个数 $\leqslant\lfloor|\boldsymbol{A}|/2\rfloor$.

证明　用反证法证明,假设 g 和 f 在集合 $\boldsymbol{A}$ 上取值不同的点的个数 $r>\lfloor|\boldsymbol{A}|/2\rfloor$,易验证 $|\boldsymbol{A}|-2r<0$.我们定义 n 元布尔函数

$$g'(\boldsymbol{X}):=\begin{cases}g(\boldsymbol{X}) & ,\text{如果 } \boldsymbol{X}\notin\boldsymbol{A}\\ g(\boldsymbol{X})\oplus 1 & ,\text{否则}\end{cases}$$

根据汉明距离的定义 $d_H(f,g')=d_H(f,g)-r+(|\boldsymbol{A}|-r)=d_H(f,g)+|\boldsymbol{A}|-2r$.因为 $|\boldsymbol{A}|-2r<0$,所以 $d_H(f,g')<d_H(f,g)$,这与 g 是 f 的一个旋转对称最佳逼近函数的条件矛盾.

下面我们研究如何计算布尔函数到旋转对称函数集合的距离.受引理 2 证明过程启发,我们统计布尔函数输入的每个圈结构上的函数值:令 $w_{j,1}$ 为 f 在圈 j 上取值为 1 的计数,令 $w_{j,0}$ 为 f 在圈 j 上取值为 0 的计数,则 $D^n_{\mathrm{Rots}}(f)=\sum_j\min\{w_{j,1},w_{j,0}\}$.算法的工作量主要是扫描布尔函数的真值表,时间复杂度和空间复杂度都为 $O(2^n)$.

上面算法中一个重要环节是要依次遍历布尔函数输入空间 $\{0,1\}^n$ 的圈结构,我们给出一个方法:将二元向量空间 $\{0,1\}^n$ 中元素看成二进制数,这样我们可以用 $\mathbb{Z}/2^n\mathbb{Z}$ 中元素表示空间 $\{0,1\}^n$.用长度为 2^n 的数组 B 标识空间 $\mathbb{Z}/2^n\mathbb{Z}$ 中元素的处理状态,数组元素下标依次为 $0,\cdots,2^n-1$,初始化 $B[i]$ 为 0,其中 $0\leqslant i\leqslant 2^n$.集合 $\boldsymbol{C}[1],\boldsymbol{C}[2],\cdots,\boldsymbol{C}[\boldsymbol{Z}_n]$ 分别存储 ρ_n 定义的 Z_n 个圈的元素,初始化为空.然后运行下面步骤:

(1) $i=0,j=1$;

(2)如果 $i<2^n$ 则把 i 添入集合 $\boldsymbol{C}[j]$,置 $B[i]$ 为 1;否则程序结束,并输出 $\boldsymbol{C}[1],\boldsymbol{C}[2],\cdots,\boldsymbol{C}[\boldsymbol{Z}_n]$;

(3)将 i 看成 n 维二元向量,更新 $i=\rho_n(i)$;

(4)如果 $B[i]=0$,则跳转到步骤(2);否则更新 $i=i+1,j=j+1$ 进入步骤(5);

(5)如果 $B[i]=0$,则跳转到步骤(2);否则更新 $i=i+1$,重复步骤(5).

通过上面的方法可以有效的遍历空间 $\{0,1\}^n$ 的圈结构,注意到旋转变化率 $p=\Pr(f(\boldsymbol{X}_t)\neq f(\rho(\boldsymbol{X}_t)))=\sum_i|\boldsymbol{C}^i_\rho|/2^n$,其中 $\boldsymbol{C}^i_\rho=\{\boldsymbol{X}\in\boldsymbol{C}[i]\mid f(\boldsymbol{X})\neq f(\rho(\boldsymbol{X}))\}$,因此在上面遍历圈结构的同时统计 $|\boldsymbol{C}^i_\rho|$ 的值,就可以计算出旋转变化率.易见,这个计算旋转变化率的算法的时间、空间复杂度都是 $O(2^n)$.

旋转对称函数的距离 $D^n_{\mathrm{Rots}}(f)$ 与旋转对称变化率 $p=\Pr(f(\boldsymbol{X}_t)\neq f(\rho(\boldsymbol{X}_t)))$ 有如下关系:

定理 2　f 是 n 元布尔函数,记 f 的旋转对称距离 $D^n_{\mathrm{Rots}}(f)$ 为 d,则 $p\leqslant 2d/2^n$.

证明　设圈 $\boldsymbol{A}$ 是 ρ_n 定义的某一个等价类,我们首先考虑 $|\boldsymbol{A}|=k>1$ 的情况:

记圈 $\boldsymbol{A}=\{\boldsymbol{X}_0,\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{k-1}\}$ 其中 $\boldsymbol{X}_i\in\{0,1\}^n$ 且 $\boldsymbol{X}_{i+1 \bmod k}=\rho_n(\boldsymbol{X}_i),i<k$.令 g 是 f 的某一旋转对称最佳逼近函数,不妨设 g 在圈 $\boldsymbol{A}$ 上的取值为 0,记 $\boldsymbol{A}_f=\{\boldsymbol{X}_{i_1},\cdots,\boldsymbol{X}_{i_s}\}$ 是 f 在圈 $\boldsymbol{A}$ 上取值为 1 的点的集合,根据引理 2,$s\leqslant\lfloor k/2\rfloor$.记 $\boldsymbol{A}_\rho=\{\boldsymbol{X}\in\boldsymbol{A}\mid f(\boldsymbol{X})\neq f(\rho_n(\boldsymbol{X}))\}$ 是圈 $\boldsymbol{A}$ 中循环移位函数值发生变化的点的集合.针对 $\boldsymbol{A}_f$ 在 $\boldsymbol{A}$ 中不同的分布情况,我们分别讨论 $|\boldsymbol{A}_\rho|$ 的取值.当 $\boldsymbol{A}_f$ 满足条件 $\boldsymbol{X}_{i_r}\neq\rho(\boldsymbol{X}_{i_s}),\forall(\boldsymbol{X}_{i_r},\boldsymbol{X}_{i_s})\in\boldsymbol{A}_f\times\boldsymbol{A}_f$,即 $\boldsymbol{A}_f$ 中任意两个不同的元素在圈 $\boldsymbol{A}$ 中都不相邻时,$|\boldsymbol{A}_\rho|=2|\boldsymbol{A}_f|$.考虑另一种极端情况:当 $\boldsymbol{A}_f$ 满足条件 $\forall\boldsymbol{X}_{i_s}\in\boldsymbol{A}_f,\exists\boldsymbol{X}_{i_r}$ s.t. $\boldsymbol{X}_{i_k}=\rho_n(\boldsymbol{X}_{i_r})$ 时,即 $\boldsymbol{A}_f$ 中元素在 $\boldsymbol{A}$ 连续分布,此时易见 $|\boldsymbol{A}_\rho|\leqslant 2$.当 $\boldsymbol{A}_f=\varnothing$ 时,$|\boldsymbol{A}_\rho|=0$.因此 $0\leqslant|\boldsymbol{A}_\rho|\leqslant 2|\boldsymbol{A}_f|$.

而当 $|\boldsymbol{A}|=1$ 时,$\boldsymbol{A}=\{2^n-1\}$ 或者 $\{0\}$,此时 $|\boldsymbol{A}_\rho|=0\leqslant 2|\boldsymbol{A}_f|$.综合上面的讨论,我们得到:$|\boldsymbol{A}_\rho|\leqslant 2|\boldsymbol{A}_f|$.

ρ_n 将 $\{0,1\}^n$ 划分为若干个圈 $\boldsymbol{A}^i$,$\{0,1\}^n=\cup_i\boldsymbol{A}^i$,计算

$$\Pr(f(\boldsymbol{X}_t)\neq f(\rho(\boldsymbol{X}_t)))=\frac{|\{\boldsymbol{X}_t\in\{0,1\}^n\mid f(\boldsymbol{X}_t)\neq f(\rho(\boldsymbol{X}_t))\}|}{2^n}$$

$$=\frac{\sum_i|\{\boldsymbol{X}_t\in\boldsymbol{A}^i|f(\boldsymbol{X}_t)\neq f(\rho(\boldsymbol{X}_t))\}|}{2^n}$$
$$=\frac{\sum_i|\boldsymbol{A}_\rho^i|}{2^n}$$
$$\leqslant\frac{2\sum_i|\boldsymbol{A}_f^i|}{2^n}\qquad(7)$$

由对称最佳逼近函数的定义可知道 $d=\sum_i|\boldsymbol{A}_f^i|$，因此 $p=\Pr(f(\boldsymbol{X}_t)\neq f(\rho(\boldsymbol{X}_t)))\leqslant 2d/2^n$.

从这个结论我们看到，如果滤波函数是修改对称函数的真值表的得到的代数免疫函数，q 是修改比率($q=d/2^n$)，线性反馈寄存器级数为 n，那么滤波函数的旋转变化率就不会超过 $2q$，第四节介绍的攻击方法就可以以不大于 $O(2^{H(2q)n}n^{2.37})$的复杂度进行状态恢复攻击.以 $n=100$ 为例，为了使我们的攻击的复杂度提升到 2^{62}以上，要求滤波函数至少要修改了对称函数的真值表的 5%.

6 最优代数免疫函数对旋转逼近攻击的脆弱性

近年来代数攻击引起了密码学家的关注，其基本思想是将恢复密码系统的密钥或内部状态的问题归约成多变量方程系统的求解问题.对于滤波生成器，一个代数次数较高的滤波函数如果有一个较低的零化子，人们就可以利用这个性质降低方程系统的次数，进而求解方程系统.例如流密码算法 E_0, Toyocrypt, LILI-128 等就因为设计时没考虑到大规模超定方程系统的可解性情况，结果不能抵抗代数攻击.

为抵御代数攻击，密码学家提出了代数免疫度的概念，滤波函数具有最优代数免疫阶成为抵御代数攻击的一个必要条件.目前发现的具有最优代数免疫阶的布尔函数大都具有很高的对称性质或者是纯粹的对称函数，或者是修改对称函数真值表得到的.下面我们简单地介绍这两类主要的构造方法.

引理 3[8] 设正整数 $n=2t+1$，若 n 元布尔函数 f 定义如下：

$$f(\boldsymbol{X})=\begin{cases}a & ,\text{如果 } wt(\boldsymbol{X})\leqslant t\\ a\oplus 1 & ,\text{否则}\end{cases}$$

其中 $a\in\boldsymbol{F}_2$，则 $AI(f)=t+1$.

当 $a=1$ 时，引理 3 中的函数 f 是择多函数，记为 G_n.这类函数是密码学界最早发现的最优代数免疫函数.记 $\mathbf{1}_{Gn}$ 为 G_n 的支撑集，即 $\mathbf{1}_{Gn}=\{\boldsymbol{X}\in\{0,1\}^n:G_n(\boldsymbol{X})=1\}$；记 $\mathbf{0}_{Gn}$ 为 G_n 的零点集，即 $\mathbf{0}_{Gn}=\{\boldsymbol{X}\in\{0,1\}^n:G_n(\boldsymbol{X})=0\}$.文献[10]给出了一种方法，通过分别修改 G_n 的支撑集 $\mathbf{1}_{Gn}$ 和零点集 $\mathbf{0}_{Gn}$ 中 k 个元素的函数值来构造代数免疫函数，并证明了所有的代数免疫函数理论上都可以通过该方法构造.构造方法叙述如下：

任意一个 n 元布尔函数可以唯一表示成 $\boldsymbol{F}_2[x_1,\cdots,x_n]/(x_1^2-x_1,\cdots,x_n^2-x_n)$中的一个多项式.设 $<_v$ 是定义在 n 维二元空间 $\{0,1\}^n$ 上的一个全序，则可以由 $<_v$ 定义 $\boldsymbol{F}_2[x_1,\cdots,x_n]/(x_1^2-x_1,\cdots,x_n^2-x_n)$中的单项式序 $<_m$，$x_1^{a_1}\cdots x_n^{a_n}<_m x_1^{b_1}\cdots x_n^{b_n}$ 当且仅当 $(a_1,\cdots,a_n)<_v(b_1,\cdots,b_n)$.记 $\boldsymbol{V}(\boldsymbol{X})=(1,x_1,\cdots,x_n,x_1x_2,\cdots,x_{n-1}x_n,\cdots,x_1,\cdots,x_t,\cdots,x_{t+2},\cdots,x_n)$，向量中的单项式按照序 $<_m$ 排列.将 $\mathbf{1}_{Gn}$ 中向量按照 $<_v$ 排列，定义 2^{n-1}阶方阵 $\boldsymbol{V}(\mathbf{1}_{Gn})$，$\boldsymbol{V}(\mathbf{1}_{Gn})$的第 i 行为向量 $\boldsymbol{V}(\boldsymbol{X}_i)$，其中 $\boldsymbol{X}_i$ 是 $\mathbf{1}_{Gn}$ 中第 i 个向量.类似地，定义 $\boldsymbol{V}(\mathbf{0}_{Gn})$，即 $\boldsymbol{V}(\mathbf{0}_{Gn})$的第 i 行是向量 $\boldsymbol{V}(\boldsymbol{Y}_i)$，其中 $\boldsymbol{Y}_i$ 是 $\mathbf{0}_{Gn}$ 中第 i 个向量.

引理 4[10,11] 设正整数 $n=2t+1$，选取正整数 $1\leqslant k\leqslant 2^{n-2}$，任意选取 k 个整数 $1\leqslant i_1<\cdots<i_k\leqslant 2^{n-1}$，找 k 个整数 $1\leqslant j_1<\cdots<j_k\leqslant 2^{n-1}$，使得 $\boldsymbol{M}(n)_{(i_1,\cdots,i_k;j_1,\cdots,j_k)}$ 可逆，构造布尔函数

$$G_{i_1,\cdots,i_k;j_1,\cdots,j_k}(\boldsymbol{X})=\begin{cases}G_n(\boldsymbol{X})\oplus 1 & ,\text{如果 } \boldsymbol{X}\in\{\boldsymbol{X}_{j_1},\cdots,\boldsymbol{X}_{j_k},\boldsymbol{Y}_{i_1},\cdots,\boldsymbol{Y}_{i_k}\}\\ G_n(\boldsymbol{X}) & ,\text{否则}\end{cases}\qquad(8)$$

那么函数 $G_{(i_1,\cdots,i_k;j_1,\cdots,j_k)}(\boldsymbol{X})$就是最优代数免疫函数.

引理 4 中的 $\boldsymbol{M}(n)_{(i_1,\cdots,i_k;j_1,\cdots,j_k)}$ 是 2^{n-1} 阶方阵 $\boldsymbol{M}(n)$的 k 阶子阵，取自 $\boldsymbol{M}(n)$的 $i_1,\cdots,i_k$ 行 $j_1,\cdots,j_k$ 列元素.方阵 $\boldsymbol{M}(n)=\boldsymbol{V}(\mathbf{0}_{Gn})\boldsymbol{V}(\mathbf{1}_{Gn})^{-1}$.

这个构造方法的主要策略是通过寻找矩阵 $\boldsymbol{M}(n)$的 k 阶可逆子阵来确定择多函数 G_n 的支撑集中 k 个元素和零点集中 k 个元素，对这 $2k$ 个位置的函数值取反，其它位置函数值保持不变就可以得到新的最优代数免疫函数.而判断一个 k 阶子方阵的可逆性至少需要 k^2 的运算量(方阵含 k^2 个元素).当 $k=2^{30}$时，用该方法构造代数免疫函数的运算量就超过 2^{60}，这种规模的运算在现实中已不可能实现.也就是说应用此构造方法在现实中不可能修改择多函数超过 $2k=2^{31}$个位置的真值表的值来构造新的最优代数免疫函数.在这个条件下，$n>40$ 时，用这种方法得到的最优代数免疫函数 f'的旋转变化率 $\leqslant 2D^n_{\mathrm{Rots}}(f')/2^n\leqslant 2^{32}/2^{41}<0.2\%$.以 f' 为滤波函数构造滤波生成器且令滤波生成器的 LFSR 的长度 $n\geqslant 40$，那么旋转对称攻击对这样的滤波生成器的攻击复杂度的一个上界是 $O(2^{H(2^{31}/2^n)n}n^{0.37})$.图 2 给出了攻击复杂度的以 2 为底的对数值与滤波函数的长度 n 的变化关系.从图中可以看到，滤波函数抵抗旋转对称攻击的能力随 LFSR 的长度增加而增强的幅度较小：$n=40$ 时旋转对称攻击复杂度小于 $2^{13.15}$，n 增加到 100 时，旋转对称攻击复杂度小于 $2^{15.7}$.事实上，如果以 f' 为滤波函数构造滤波生成器，即使使用 1000 级的

LFSR,旋转对称攻击复杂度也不超过 $2^{23.6}$.

从上面的分析我们看出,用文献[8~11]中的方法构造出的 n 元最优代数免疫函数,在 n 较大的情况下($n>40$)得到的布尔函数具有很强的旋转对称性,如果将其用作滤波函数则极易遭受第四节提出的旋转对称攻击.

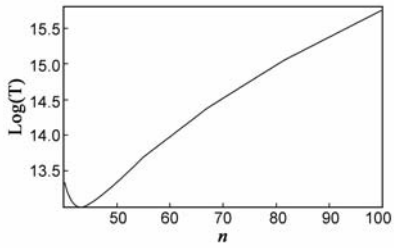

图2　旋转对称攻击对以 f' 为滤波函数的滤波生成器的攻击复杂度与滤波函数的长度 n 的变化关系,横坐标表示 n 的值,纵坐标表示攻击复杂度的以2为底的对数值

7　结论

我们利用滤波函数的旋转对称性质给出了对滤波生成器的一种新的攻击,时间复杂度是 $O(2^{H(p)n}n^{2.37})$.我们讨论了滤波函数的旋转对称距离和旋转对称变化率 p 的关系以及计算这两个指标的算法;指出为保证密码体制的安全性,所选择的滤波函数要具备较大的旋转对称距离.文献[8~11]提出的构造最优代数免疫函数的方法是目前构造最优代数免疫函数的主要方法,通过分析我们发现这些方法构造的最优代数免疫函数具有很强的旋转对称性,易于遭受旋转对称攻击,不适于直接选作滤波生成器的滤波函数.

参考文献

[1] 张木想,肖国镇.流密码中非线性组合函数的分析与设计[J].电子学报,1996,24(01):48-52.
ZHANG Mu-xiang, XIAO Guo-zhen. Analysis and design of nonlinear combining functions in stream ciphers[J]. Acta Electronica Sinica, 1996, 24(01): 48-52. (in Chinese)

[2] 张斌,金晨辉.对迭代型混沌密码的逆推压缩攻击[J].电子学报,2010,38(01):129-134.
ZHANG Bin, JIN Chen-hui. Inversion and compression attacks to iterative chaotic ciphers[J]. Acta Electronica Sinica, 2010, 38(01): 129-134. (in Chinese)

[3] Nicolas T Courtois, Wili Meier. Algebraic attacks on stream ciphers with linear feedback[A]. Advances in Cryptology-EUROCRYPT 2003[C]. LNCS 2656, Berlin: Springer-Verlag, 2003. 346-359.

[4] Willi Meier, Enes Pasalic, Claude Carlet. Algebraic attacks and decomposition of boolean functions[A]. Advances in Cryptology-EUROCRYPT 2004[C]. Berlin, Germany: Springer-Verlag, 2004. 474-491.

[5] 张文英,武传坤,等.密码学中布尔函数的零化子[J].电子学报,2006,34(01):51-54.
ZHANG Wen-ying, WU Chuan-kun, et al. On the annihilators of cryptographic boolean functions[J]. Acta Electronica Sinica, 2006, 34(01): 51-54. (in Chinese)

[6] D K Dalai, K C Gupta, S Maitra. Results on algebraic immunity for cryptographically significant Boolean functions[A]. INDOCRYPT 2004 (Lecture Notes in Computer Science)[C]. Berlin, Germany: Springer-Verlag, 2004. 92-106.

[7] D K Dalai, K C Gupta, S Maitra. Cryptographically significant Boolean functions: construction and analysis in terms of algebraic immunity[A]. FSE 2005 (Lecture Notes in Computer Science)[C]. Berlin, Germany: Springer-Verlag, 2005. 98-111.

[8] A Braeken, B Preneel. On the algebraic immunity of symmetric Boolean functions[A]. INDOCRYPT 2005 (Lecture Notes in Computer Science)[C]. Berlin, Germany: Springer-Verlag, 2005. 35-48.

[9] Na Li, Wen-Feng Qi. Symmetric Boolean functions depending on an odd number of variables with maximum algebraic immunity[J]. IEEE Transactions on Information Theory, 2006, 52(5): 2271-2273.

[10] Na Li, Wen-Feng Qi. Construction and analysis of boolean functions of 2t + 1 variables with maximum algebraic immunity[A]. Advances in Cryptology-ASIACRYPT 2006[C]. Berlin, Germany: Springer-Verlag, 2006. 84-98.

[11] Na Li, Wen-Feng Qi. Boolean functions of an odd number of variables with maximum algebraic immunity[J]. Science in China Series F: Information Sciences, 2007, 50(3): 307-317.

[12] P Stanica, S Maitra. Rotation symmetric boolean functions-count and cryptographic properties[J]. Discrete Applied Mathematics, 2008, 156(10): 1567-1580.

[13] Meier W, Staffelbach. Fast correlation attacks on stream ciphers[A]. Advances in Cryptology-EUROCRYPT88[C]. Berlin: Springer-Verlag, 1989. 301-314.

[14] D Coppersmith, S Winograd. Matrix multiplication via arithmetic progressions[J]. Symbolic Computation, 1990, 9(3): 251-280.

[15] 林东岱.代数学基础与有限域[M].北京:高等教育出版社,2006.

作者简介

杨　笑　男,1982 年 9 月生于北京,现为中国科学院软件研究所信息安全国家重点实验室博士研究生,研究方向为密码学.
E-mail: yangxiao@is.iscas.ac.cn

武传坤　男,1964 年生于山东省沂水县,现为中国科学院软件研究所研究员、博士生导师,研究方向为密码学和网络安全.
E-mail: ckwu@is.iscas.ac.cn