

密码学中布尔函数的零化子

张文英¹, 武传坤¹, 于静之²

(1. 中国科学院软件研究所-中国科学院研究生院, 信息安全国家重点实验室, 北京 100080;

2. 山东医学高等专科学校, 山东济南 250002)

摘要: 布尔函数的零化子与代数攻击息息相关, 但是如何构造一个给定函数的低次零化子仍然是一个悬而未决的问题. 本文对此问题进行了研究, 研究结果表明, 如果布尔函数的零点集有一个 k 维子空间, 那么, 函数就会有代数次数为 $n-k$ 的零化子. 然而如何找到函数的具有最低代数次数的零化子仍然是一个亟待解决的难题.

关键词: 密码学; 代数攻击; 布尔函数的零化子

中图分类号: TN918.1 **文献标识码:** A **文章编号:** 0372-2112 (2006) 01-0051-04

On the Annihilators of Cryptographic Boolean Functions

ZHANG Wen-ying¹, WU Chuan-kun¹, YU Jing-zhi²

(1 State Key Lab of Information Security, Institute of Software - Graduate School of Chinese Academy of Sciences Beijing 100080 China; 2 Shandong Medical College, Jinan, Shandong 250002, China)

Abstract Annihilators of cryptographic Boolean functions have been shown to be closely related to algebraic attacks to stream ciphers. However how to construct annihilators of a given Boolean function remains a hard problem. As an approach to this problem, it gives an important characterization of annihilators with low degrees of a given Boolean function in terms of the set of x values on which $f(x) = 0$. This gives a way to find annihilators of a given Boolean function, however how to find an annihilator of the lowest algebraic degree still remains unsolved.

Key words cryptography; algebraic attack; annihilator of Boolean functions

1 引言

基于线性反馈移位寄存器的代数攻击是由法国密码学家 Nicolas T Courtois 和 Willi Meier 提出的^[1], 此方法一经提出就对现有流密码体制形成了巨大威胁. 运用代数攻击的方法人们成功的破译了运用具有良好性质的布尔函数所设计的、能够抵抗所有已知攻击的 Toyocrypt 和 LILI-128^[1-3]. Adi Shamir 在 2004 年的亚密会上作了题为 Stream Ciphers Dead or Alive^[4] (流密码是死还是活) 的报告, 使人们对流密码的前程倍感担忧. 由于代数攻击的提出, 使得一些新设计的流密码体制尽量避免使用基于线性移位寄存器的非线性组合函数, 而去仿效分组密码的块状结构模式, 例如最新提出的 Single-cycle F-function^[5] 和 Rabbit^[6] 流密码等. 在本文中我们假设读者已了解关于布尔函数的一些基本知识.

2 代数攻击的原理

定义 1 把 $f(x_1, \dots, x_n)$ 中每个单项式都看作一个新的变量后所得到的线性方程叫做多变量方程 (multivariate algebraic equations).

例如 5 元布尔函数

$$f(x) = x_1 + x_3 + x_5 + x_1x_2 + x_1x_5 + x_2x_4 + x_2x_5 + x_3x_4 + x_4x_5 + x_3x_4x_5 + x_1x_4x_5 + x_2x_3x_5.$$

是 12 个单项式的和, 从而 $f = 0$ 是有 12 个“变元”的多变量方程. 下文所说“变元”的个数均指方程中单项式的个数.

密码学中在设计同步流密码生成器时, 人们常将之分成两个部分——驱动部分和非线性组合部分. 它们的任务分别是: 驱动部分控制存储器的状态转移, 负责提供若干周期大、统计特性好的序列供组合部分使用, 而非线性部分则将由驱动部分提供的序列组合成满足要求的、性质好的密钥流序列. 驱动器部分采用 LFSR 序列, 非线性组合部分对线性部分产生的序列进行混淆和扩散. 这种流密码的工作流程方式可以用方程组如下表示:

$$\begin{cases} b_0 = f(s_0, s_1, \dots, s_{n-1}) \\ b_1 = f(L(s_0, s_1, \dots, s_{n-1})) \\ b_2 = f(L^2(s_0, s_1, \dots, s_{n-1})) \\ \dots \\ b_{m-1} = f(L^{m-1}(s_0, s_1, \dots, s_{n-1})) \end{cases} \quad (1)$$

收稿日期: 2005-01-20 修回日期: 2005-09-06

基金项目: 国家自然科学基金 (No. 90304007); 国家 973 重点研究发展规划 (No. 2004CB318004); 中国博士后科学基金

其中 $(s_0, s_1, \dots, s_{n-1})$ 是初态. 假设状态转移函数 L 及非线性组合函数 f 都是事先给定的, 只有初态及初态的平移等价类未知. 攻击者的目的是通过分析观察到的输出密钥流序列 (b_0, b_1, \dots) 来恢复初态.

设 f 的代数次数为 k . 由于线性变换不改变函数的代数次数, 所以对任意的 $t, f(s_0, s_1, \dots, s_{n-1})$ 和 $f(L^t(s_0, s_1, \dots, s_{n-1}))$ 的代数次数都相同, 因此方程组 (1) 中“变元”的个数至多为 $C_n^1 + C_n^2 + \dots + C_n^k$ 个. 若式 (1) 中方程的个数 m 大于“变元”的个数, 则称方程组 (1) 为超定系统 (over defined system).

设 $f(x)$ 是 n 元布尔函数, 满足 $f(x)g(x) = 0$ 的非 0 函数 $g(x)$ 称为 $f(x)$ 的零化子^[7]. 利用布尔函数的零化子, 攻击者可以得到一个关于初态和输出密钥流序列的代数次数较低的方程, 从而减少多变量方程中变元的个数, 此方法被称为代数攻击. 代数攻击的数学描述为: 若存在函数 g , 使得 g 的代数次数 d 远远低于 f 的代数次数 k , 那么在满足 $f(x) = 1$ 的点 $L^j(s_0, s_1, \dots, s_{n-1}), j = 1, 2, \dots$ 有

$$\begin{cases} g(L^1(s_0, s_1, \dots, s_{n-1})) = 0 \\ g(L^2(s_0, s_1, \dots, s_{n-1})) = 0 \\ \dots \\ g(L^j(s_0, s_1, \dots, s_{n-1})) = 0 \end{cases} \quad (2)$$

且方程组 (2) 中变元个数至多为 $C_n^1 + C_n^2 + \dots + C_n^d$ 这个数目也远远小于方程组 (1) 中变元个数. 于是只要 (2) 中线性无关的方程的个数大于变元的个数 $C_n^1 + C_n^2 + \dots + C_n^d$ 就可以唯一的恢复密钥, 这比直接解方程组 (1) 要快得多. 由于在实际应用中, 都要求非线性组合函数 f 是平衡函数, 若观察到密钥流序列的长度是 m , 则 (1) 中将有 $m/2$ 个 $b_i = 1 (0 \leq i \leq m)$, 从而方程组 (2) 中方程个数的期望值为 $m/2$.

3 布尔函数零化子的代数次数

由前面的介绍可知, 对基于线性移位寄存器的流密码体制代数攻击的成效取决于该体制中所使用的非线性组合函数的零化子的代数次数, 零化子的代数次数越低, 方程组 (2) 所含变元的个数越少, 代数攻击的复杂度越低. 于是根据函数的结构形式寻找函数的具有较低代数次数 (尤其是最低次数) 的零化子将对代数攻击具有重要的指导意义. 本节的主要工作是给出了函数有代数次数为 $n-k$ 次零化子的一个判定定理, 另外还研究了布尔函数的代数次数最低的零化子集合的结构特征, 给出了关于布尔函数代数免疫阶的一个紧的上界.

首先介绍一下 Meier 等的工作. 若以 R_n 记所有 n 元布尔函数所做成的代数系统, Meier 等在文献 [7] 中给出了关于布尔函数的零化子集合的一个简洁、完美的刻画:

定理 1^[7] 设 f 是 n 元布尔函数. 则 f 的所有零化子的集合 $An(f)$ 是 R_n 中由 $1+f$ 生成的主理想, 即: $An(f) = \{ (1+f)r \mid r \in R_n \} = \langle 1+f \rangle$.

受定理 1 的启发, 我们得到了关于布尔函数的代数次数最低的零化子集合的结构特征:

定理 2 设 f 是 n 元布尔函数, 且其零化子的代数次数最低为 d , 则其所有 d 次零化子做成 R_n 的一个子空间.

证明 设 d 是所有零化子代数次数的最小值, N 是 f 的所有代数次数为 d 的零化子所做成的集合. 若 $g_1, g_2 \in N$, 则 $g_1 + g_2 \in N$ 且 $\deg(g_1 + g_2) \leq d$, 因 d 是所有零化子代数次数的最小值, 故一定有 $\deg(g_1 + g_2) = d$, 所以 N 对加法运算封闭, 因此 N 是 R_n 的一个子空间.

一般的来讲, 函数 f 自身不一定有低次零化子, 但是 $f+1$ 可能有低次零化子. 例如函数 $f = 1 + x_1x_2 \dots x_n$ 仅在 $(1, 1, \dots, 1)$ 处函数值为 0 其余点处函数值都是 1, 它只有一个不平凡零化子, 是 n 次布尔函数 $x_1x_2 \dots x_n$. 容易验证任意线性函数 $w \cdot x, w \in GF^n(2)$ 都是 $f+1$ 的零化子. 如果 f 没有低次零化子, 但 $f+1$ 有低次零化子 g , 则在 $f+1=0$ (即 $f+1=1$) 的点有 $g(x) = 0$ 攻击者便可以在所观察到的满足 $f=0$ 的点建立方程组 (2). 因此函数 f 抗代数攻击的能力取决于 f 和 $f+1$ 的代数次数的最小值, 称这个最小值为布尔函数的代数免疫阶, 记作 $AI(f)$ (Algebraic Attack Immune). 代数免疫阶愈大, 函数抗代数攻击的能力愈强, 反之就弱. 下面我们给出关于布尔函数的代数免疫阶上限的一个定理.

引理 1^[7] 设 f 是 n 元布尔函数. 那么存在代数次数至高为 $\lceil n/2 \rceil$ 的非 0 布尔函数 g 使得 fg 的代数次数不超过 $\lceil n/2 \rceil$.

定理 3 设 f 是 n 元布尔函数. 则 $AI(f) \leq \lceil n/2 \rceil$.

证明 由引理 1, 存在代数次数至多为 $\lceil n/2 \rceil$ 的非 0 布尔函数 $g \neq 0$ 使得 $fg = h$ 的代数次数至多为 $\lceil n/2 \rceil$. 因 $fg = fh$ 若 $g \neq h$, 则 $f(g+h) = 0$ 从而 $g+h$ 是 f 的代数次数不大于 $\lceil n/2 \rceil$ 的零化子. 若 $g = h$, 则有 $fh = h, (f+1)h = 0$ 则 h 是 $f+1$ 的代数次数不大于 $\lceil n/2 \rceil$ 的零化子.

令 $offset(f) = \{x \mid f(x) = 0\}$, $supp(f) = \{x \mid f(x) = 1\}$, 并分别称它们为 $f(x)$ 的零点集和支撑. 易知 g 是 f 的零化子的充要条件是 g 的支撑是 f 的零点集的子集. 即 $g = 1$ 时 f 一定等于 0 于是以 f 的零点集的子集为支撑的所有函数作成 f 的零化子的全体. 于是欲构造 f 的零化子, 只须从 $offset(f)$ 中取一子集, 以此集合为支撑的函数便是 f 的一个零化子, 零化子的代数次数完全取决于所取集合的结构特点. 基于此我们给出了布尔函数有 $n-k$ 次零化子的一个充分条件, 首先来看一个引理:

引理 2 设 S 是 $GF^n(2)$ 的一个 k 维子空间, M 是以 S 中所有元素为行作成的矩阵. 若 M 的第 j 列非零, 那么该列中 0, 1 个数各半.

证明 设 $\alpha_1, \dots, \alpha_k$ 是 S 的一组基. 当 $\lambda = (\lambda_1, \dots, \lambda_k)$ 取遍 $GF^k(2)$ 时, $\lambda(\alpha_1, \dots, \alpha_k)^T$ 取遍 M 的所有行向量, 此处“ T ”表示矩阵或向量的转置. 以 a_1, \dots, a_k 记 $\alpha_1, \dots, \alpha_k$ 的

$$\sum_{i=1}^r a_{j_{i-1}} a_{j_{i-2}} \cdots a_{j_i} = 1$$

因此, h 的代数次数为 $n-k$, 从而 g 的代数次数也为 $n-k$.

$$\text{例 令 } f(x) = x_1 + x_3 + x_5 + x_1x_2 + x_1x_4 + x_1x_5 + x_2x_4 + x_2x_5 + x_3x_4 + x_1x_2x_3 + x_2x_3x_4x_5$$

则 $f(x)$ 是 F_2^5 上的平衡函数, 在 $f(x)$ 的零点集中有一个 3 维子空间

$$S = \{ (00000), (11100), (11010), (00110), (01001), (10101), (10011), (01111) \}.$$

$$\text{若定义 } g(x) = \begin{cases} 1 & x \in S \\ 0 & \text{其它} \end{cases}$$

则 g 是 f 的 $5-3=2$ 次零化子, g 的 ANF 是

$$g(x) = 1 + x_1 + x_2 + x_3 + x_4 + x_5 + x_1x_2 + x_1x_3 + x_1x_4 + x_1x_5 + x_2x_3 + x_2x_4 + x_3x_5 + x_4x_5.$$

4 结论

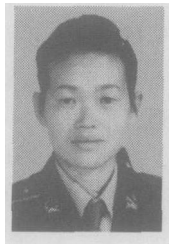
代数攻击是分析现有密码体制的一种新的强有力的分析方法. 如果 f 的零点集包含一个维数较高的子空间或仿射子空间, 将导致 f 存在低次零化子, 进而攻击者就可以建立一个关于密钥和输出序列 (明文) 的低次方程, 于是得以快速恢复密钥. 因此, 为了抵抗代数攻击, 设计者应尽量避免使用零点集中包含一个维数较高的子空间或仿射子空间的密码函数.

参考文献:

- [1] Nicolas T Courtois, Willi Meier Algebraic attacks on stream ciphers with linear feedback [A]. Advances in Cryptology-EUROCRYPT 2003 [C]. LNCS 2656, Berlin: Springer-Verlag 2003. 346- 359.
- [2] M Mihaljević, H Inai Cryptanalysis of ToyocryptHSI stream cipher IEICE Transactions on Fundamentals vol E85-A, 66-73 [OL]. <http://www.csl.esat.sony.co.jp/atlpapers/IEICEJan02.pdf>

- [3] Steve Babbage Cryptanalysis of LILI-128, Nessie project internal report [OL]. <http://www.cosic.esat.kuleuven.ac.be/nessie/reports/>, January 2001. 22.
- [4] Adi Shamir Stream ciphers dead or alive [A]. Advances in Cryptology-ASIACRYPT 2004 [C]. LNCS 3329 Berlin: Springer-Verlag 2004. 78.
- [5] Jin Hong, Dong Hoon Lee, Yongjin Yoon, Daewan Han A new class of single cycle T-functions [A]. Fast Software Encryption 2005 [C]. LNCS 3557, Berlin: Springer-Verlag 2005.
- [6] Martin Boesgaard, Mette Vestergaard, Thomas Pedersen, Jesper Christiansen, Ove Cavenius Rabbit A new high-performance stream cipher [A]. Fast Software Encryption 2003 [C]. LNCS 2887, Berlin: Springer-Verlag 2003. 307- 329.
- [7] Willi Meier, Enes Pasalic, Claude Carlet Algebraic attacks and decomposition of Boolean functions [A]. In Advances in Cryptology-EUROCRYPT 2004 [C]. LNCS 3027, Berlin: Springer-Verlag 2004. 474- 491.

作者简介:



张文英 女, 1970年6月出生于山东省鄄城县, 中科院软件所博士后, 副教授, 研究方向为密码学. E-mail: wenyngzh@iscas.ac.cn

武传坤 男, 1964年7月出生于山东沂水县, 中科院软件所研究员, 博导, 中科院百人计划引进人才, 主要研究方向为密码学, 网络安全和电子商务等. E-mail: ckwu@iscas.ac.cn