第 34 卷 第 5 期 计 算 机 工 程 2008 年 3 月
Vol.34 No.5 Computer Engineering March 2008
·安全技术· 文章编号：1000—3428(2008)05—0158—03 文献标识码：A 中图分类号：TP301.6

SMS4 算法 S 盒的密码学性质

刘 佳[1,2]，韦宝典[1,2]，戴宪华[1]
（1. 中山大学电子与通信工程系，广州 510275；2. 广东省信息安全技术重点实验室，广州 510275）

摘 要：S 盒是分组密码的重要组成部分，在很大程度上决定了分组密码的安全性。该文研究了中国分组密码标准 SMS4 算法 S 盒的平衡性、差分性质、线性结构、非线性、Walsh 谱等性质，通过与美国高级加密标准、欧洲分组加密标准 Camellia 的 S 盒作比较，说明了 SMS4 算法 S 盒一些较好的安全特性。

关键词：SMS4 算法；高级加密标准；S 盒；布尔函数

Cryptographic Properties of S-box in SMS4

LIU Jia[1,2], WEI Bao-dian[1,2], DAI Xian-hua[1]
(1. Department of Electronic and Communication Engineering, Sun Yat-sen University, Guangzhou 510275;
2. Guangdong Provincial Key Laboratory of Information-security Technology, Guangzhou 510275)

【Abstract】S-boxes are quite important components of modern symmetric cryptosystems in the sense that S-boxes bring nonlinearity to block ciphers and strengthen their cryptographic security. This paper analyzes the cryptographic properties of the SMS4 S-box. Some algebraic properties of S-boxes such as Balanceness, differential characteristics, linear structures, nonlinearity and Walsh spectrums are investigated. It claims that SMS4 S-box possesses some better properties by being compared with those S-boxes used in AES and Camellia.

【Key words】SMS4 algorithm; AES; S-box; Boolean function

继美国将Rijndael算法[1]作为高级加密标准(AES)、欧洲将Camellia等算法[2]作为NESSIE分组加密标准之后，中国国家密码管理局于 2006 年 1 月 6 日发布第 7 号公告，将我国无线局域网产品的加密算法确定为SMS4 算法[3]。这是国内官方公布的第 1 个商用密码算法。S盒是大多数分组密码算法中唯一的非线性结构，其密码特性直接决定了密码算法的性能。设计优良的S盒保证密码算法能够较好地抵抗差分密码分析[4]和线性密码分析[5]等攻击，而S盒的任何缺陷都可能影响到整个算法的安全性。

1 SMS4 算法原理

SMS4 算法的分组长度为 128 b，密钥长度为 128 b，采用 32 轮非线性迭代结构。解密过程与加密过程的结构相似，但轮密钥的使用顺序相反。加/解密过程如图 1 所示。

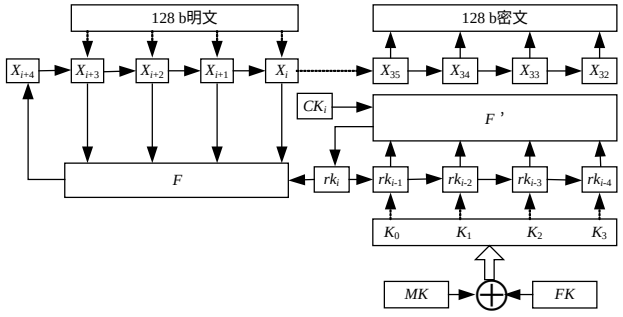

图 1 SMS4 算法流程

算法流程为：明文和密文均被看成 4 个 32 比特字，即明文 $(X_0, X_1, X_2, X_3) \in (Z_2^{32})^4$，密文 $(Y_0, Y_1, Y_2, Y_3) \in (Z_2^{32})^4$；各轮子密钥为 32 比特字：$rk_i \in Z_2^{32}$ $(i = 1, 2, \cdots, 31)$。轮函数 F 作用于前 4 轮结果 $X_i, X_{i+1}, X_{i+2}, X_{i+3}$ 和当前子密钥 rk_i，输出一个 32 比特字：

$$X_{i+4} = F(X_i, X_{i+1}, X_{i+2}, X_{i+3}, rk_i) = X_i \oplus T(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus rk_i)$$

其中，$i = 1, 2, \cdots, 31$。最后 4 轮的输出作为密文：

$$(Y_0, Y_1, Y_1, Y_2) = (X_{35}, X_{34}, X_{33}, X_{32})$$

轮函数中使用的子密钥由固定参数 $CK_i (i = 0, 1, \cdots, 31)$、系统参数 $FK = (FK_0, FK_1, FK_2, FK_3)$（$FK_i$，$CK_i$ 均为 32 比特字）和主密钥 $MK = (MK_0, MK_1, MK_2, MK_3)$ 按照密钥扩展算法生成：

$$(K_0, K_1, K_2, K_3) = (MK_0 \oplus FK_0, MK_1 \oplus FK_1, MK_2 \oplus FK_2, MK_3 \oplus FK_3)$$

$$rk_i = K_{i+4} = F'(K_i, K_{i+1}, K_{i+2}, K_{i+3}, CK_i) = K_i \oplus T'(K_{i+1} \oplus K_{i+2} \oplus K_{i+3} \oplus CK_i)$$

2 SMS4 算法 S 盒代数性质分析

2.1 S 盒代数表达式

如果密码中的元素有紧凑的代数表达式，并且这些元素能够被组合成具有可控制复杂性的表达式，那么插值攻击方法[6]对于该密码来说就是可行的，对变换代数次数和复杂度低的密码尤为奏效。为防止插入攻击，通常要求密码变换的代数式具有足够高的次数和复杂度。用拉格朗日插值多项式

$$f(x) = \sum_{i=0}^{255} y_i \prod_{j \neq i, j=0}^{255} \frac{x - x_j}{x_i - x_j}$$

基金项目：国家自然科学基金资助项目(60403007，60503006，60572059)；广东省自然科学基金资助项目(05300706)；中山大学青年教师科研启动基金资助项目(350405)

作者简介：刘 佳(1983－)，女，硕士研究生，主研方向：信息安全，密码学；韦宝典，讲师、博士；戴宪华，教授、博士

收稿日期：2007-05-09 **E-mail：**liujia_1116@163.com

可求得 SMS4 算法 S 盒的代数表达式。这是一个 254 次、255 项的多项式，具有最高的复杂程度。

2.2 差分特性

差分密码分析是一种选择明文攻击，其基本思想是通过分析特定明文差对相应密文差的影响来获得可能性最大的密钥，它是目前对分组密码最为有效的攻击方法之一。研究S 盒抗差分密码分析的能力主要从其差分均匀度和差分分布矩阵[7]着手：具有较小的差分均匀度是S盒抗击差分攻击的必要条件。

通过计算 SMS4 算法 S 盒 256 ×256 阶的差分矩阵不难发现，矩阵(首行首列除外)中只含有 3 种值：0, 2 和 4，且 4 在每行每列中仅出现一次。因此，SMS4 算法 S 盒的差分均匀度仅为 4。

2.3 线性特性

线性密码分析是一种已知明文攻击，目标是寻找到并利用明文P、密文C和密钥K的若干比特间的线性表达式。该表达式成立的概率与 1/2 的偏差大小是线性密码分析成功的一个重要衡量指标。线性密码分析思想归结到核心部件S盒就是要考察其输入输出比特间的相关关系，可用线性分布矩阵[8]来刻画。

计算 SMS4 算法 S 盒的线性分布矩阵，得到矩阵元素的最大绝对值为 16，这意味着 S 盒的线性逼近概率较低，即线性性较弱，符合分组密码强非线性性的要求。

3 SMS4 算法 S 盒布尔函数的性质

限于篇幅，本节仅给出 SMS4 算法 S 盒布尔函数 f_0 的性质，其他函数的性质类似。

3.1 布尔函数

布尔函数的输入输出关系可用多项式

$$f(x)=a_0+a_1x_0+\cdots+a_nx_{n-1}+a_{12}x_0x_1+\cdots+a_{(n-1)n}x_{n-2}x_{n-1}+\cdots+a_{12\cdots n}x_0x_1\cdots x_{n-1}=\sum_{u\in GF(2)^n}a_u x_0^{u_0}\cdots x_{n-1}^{u_{n-1}}$$

刻画，称为$f(x)$的多项式表达式，系数 a_u 可用Mobius变换求得[9]：

$$a_u=\oplus_{x\prec u}f(x)$$

其中，$\prec$ 代表偏序关系，即若对于所有 $0\le i\le n-1$ 都有 $x_i\le u_i$，则 $x\prec u$。由此求得 SMS4 算法 S 盒布尔函数：

$$f_0=x_2+x_1x_2+x_1x_3+\cdots+x_2x_3x_4x_5x_6x_7x_8$$

其项数为 134，次数为最大可能值 7，保证了 S 盒的代数复杂度和安全强度。

3.2 布尔函数代数性质

SMS4 算法 S 盒在整个算法中共用到 256 次，其任何缺陷都将影响到整个算法的安全性。

定义 1[8] 若n元布尔函数 $f:GF(2)^n\to GF(2)$ 真值表中“1”和“0”的个数均为 2^{n-1}，则称f满足平衡性。

SMS4 算法S盒是$GF(2)^8$上的一一映射，当输入遍历 0~255 时，输出也遍历 0~255，每一个输出比特位置上都含有 128 个“1”，128 个“0”，因此，每一比特都是平衡的。

定理 SMS4算法S盒各输出比特布尔函数的代数次数上界为 7。

证明 SMS4 算法 8×8 的 S 盒 $f=(f_7,f_6,f_5,f_4,f_3,f_2,f_1,f_0)$ 是$GF(2)^8$上的置换，当输入遍历 0~255 时，其输出也遍历 0~255。因此，$f_i(i=0,1,\cdots,7)$ 为平衡函数，其小项表达式中含有 128 项，而每一项展开均含有 $x_7x_6x_5x_4x_3x_2x_1x_0$ 项，化简后的 128 个 $x_7x_6x_5x_4x_3x_2x_1x_0$ 相互抵消，最终表达式不含该项，即无 8 次项，因此，上界只能为 7。

代数次数在一定程度上表明 S 盒的线性复杂度：代数次数越高，线性复杂度越高，越难以被线性表达式所逼近。SMS4 算法 S 盒各布尔函数代数次数均为 7，达到最大可能值，说明此 S 盒具有高线性复杂度。

定义 2[8] 设$f(x)$是一个n元布尔函数，记 $L_n[x]$ 为所有n元线性函数(包括仿射函数)之集。$f(x)$的非线性度记为N_f，其定义为

$$\min_{l\in L_n[x]}d(f,l)=\min_{l\in L_n[x]}w(f+l)$$

其中，$\max_{l\in L_n[x]}d(f,l)$ 为 $f(x)$的线性度。若 $f(x)$的非线性度满足

$$N_f=2^{n-1}-2^{\frac{1}{2}n-1}$$

则称 $f(x)$为 Bent 函数。

布尔函数$f(x)$的非线性度N_f是用来衡量抵抗“线性攻击”能力的一个非线性准则，N_f越大，则布尔函数$f(x)$抵抗“线性攻击”的能力越强；反之，N_f越小，则布尔函数$f(x)$抵抗“线性攻击”的能力越弱。对于SMS4 算法S盒来说，计算结果显示各布尔函数的非线性度均为 112。而Bent函数的非线性度为最大值 $2^{n-1}-2^{\frac{1}{2}n-1}$ (n=8 时为 120)，因此，SMS4 算法S盒各布尔函数均非Bent函数，但是它们的非线性程度已经非常接近Bent函数。

定义 3[8] 若对于任意汉明重量为 1 的向量

$$\boldsymbol{c}=(c_1,c_2,\cdots,c_n)\in GF(2)^n$$

有 $w(f(x)+f(x+\boldsymbol{c}))=2^{n-1}$，即 $f(x)+f(x+\boldsymbol{c})$ 是平衡函数，则称布尔函数 $f(x)$满足严格雪崩准则。

严格雪崩准则用于考察 $f(x)$输入变化对输出变化的影响：当任一输入比特改变(从 x 变为 $x+\boldsymbol{c}$)时，输出发生变化($f(x)\ne f(x+\boldsymbol{c})$)的个数与不发生变化($f(x)=f(x+\boldsymbol{c})$)的个数：如果相等，则 $f(x)$满足严格雪崩准则；否则不满足严格雪崩准则。满足严格雪崩准则的函数，其轻微的输入变化将导致输出的巨大改变。

计算表明 SMS4 算法 S 盒各布尔函数任何一行数值都不全为 0，因此，SMS4 算法 S 盒各布尔函数均不满足严格雪崩准则。但这些数值都很接近 0，说明当某一输入比特取反时，每个输出比特均会以较大概率取反，因此，SMS4 算法 S 盒仍然具有很好的 SAC 性质。

定义 4[8] 设 $\alpha\in GF(2)^n,\alpha\ne 0$，若

$$w(f(x)+f(x+\alpha))=2^{n-1}$$

即 $f(x)+f(x+\alpha)$ 是平衡函数，则称 $f(x)$关于 α 满足扩散准则。若对任意满足 $1\le w(\alpha)\le k$ 的 α，$f(x)$关于 α 满足扩散准则，则称 $f(x)$ 满足 k 次扩散准则，记为 $PC(k)$。

k 次扩散准则用于进一步考察 $f(x)$输入变化对输出变化的影响：当任何几个输入比特改变(从 x 变为 $x+\alpha$)时，输出发生变化($f(x)\ne f(x+\alpha)$)的个数与不发生变化($f(x)=f(x+\alpha)$)的个数：如果相等，则 $f(x)$满足 k 次扩散准则；否则不满足 k 次扩散准则。

计算表明，SMS4 算法S盒的 8 个布尔函数均不满足任何次数的扩散准则。但是，布尔函数f_0关于 32 个 α 满足扩散准则。

定义 5[8] 对n元布尔函数$f(x)$和 $\alpha\in GF(2)^n$，如果对于任意 $x\in GF(2)^n$ 都有 $f(x+\alpha)+f(x)=m$ (常数)，则称 α 为$f(x)$的

一个线性结构。若m=0，则称 α 为$f(x)$的不变线性结构；若m=1，则称 α 为$f(x)$的恒变线性结构。

S盒中任何非零线性结构的存在，都会是密码算法的一个缺陷，可能导致算法的全面失败。计算表明，SMS4 算法S盒没有非零线性结构。对f_0，存在 32 个 α 值使

$$f(x+\alpha)+f(x)=0$$

以 1/2 的概率成立，即有 128 个x使 $f(x+\alpha)+f(x)=0$ 成立，即这些 α 值的线性结构特性最差。其他 α 值也以接近 1/2 (1/2-16/256~1/2+16/256)的概率使 $f(x+\alpha)+f(x)=0$ 成立，再次说明f_0线性性确实不明显。

通过对 SMS4 算法 S 盒以上性质的分析，可以相信 SMS4 算法 S 盒具有良好的非线性特性，在理论上保证算法能够很好地抵抗线性密码分析。

3.3 布尔函数 Walsh 谱

定义 6[8] n元布尔函数循环Walsh谱变换定义为

$$S_{(f)}(w)=2^{-n}\sum_{x=0}^{255}(-1)^{f(x)+x\cdot w}$$

其中，$w\in GF(2)^n$；$x\in GF(2)^n$；$w\cdot x=w_0x_0+\cdots+w_{n-1}x_{n-1}$。

布尔函数 Walsh 循环谱值越大，说明布尔函数 Walsh 环谱与线性函数越接近，对应的 S 盒就容易被攻击，即性能越差。

计算SMS4 算法S盒布尔函数f_0的 256 倍循环谱发现：各点谱值的绝对值主要在 4~30 之内，最大不超过 32，处于较小的状态。这再次表明其与线性函数相似程度很小。

需要注意的是，零谱值个数和非线性度之间存在着制约关系[10]：零谱值个数过大将导致非线性度的下降，反之，非线性度上升。计算发现，SMS4 算法S盒布尔函数循环谱的零谱值个数较小，仅为 17。

4 密码学性质比较

表 1、表 2 是 SMS4 算法、AES 算法和 Camellia 算法 S 盒各项数据指标的计算结果，本文对其表达式、非线性性和抗攻击能力等进行了全面的比较。

表 1 S 盒代数性质的比较

	SMS4	AES	Camellia			
			S1	S2	S3	S4
次数	254	254	254	254	254	254
项数	255	9	254	253	254	255
差分矩阵最大值	4	4	4	4	4	4
线性矩阵最大绝值	16	16	16	16	16	16

表 2 S 盒布尔函数性质的比较

	SMS4	AES	Camellia			
			S1	S2	S3	S4
平衡性	平衡	平衡	平衡	平衡	平衡	平衡
次数	7	7	7	7	7	7
项数	123~139	110~145	126~135	126~135	126~135	126~135
非线性度	112	112	112	112	112	112
循环谱(最大绝对值)	32	32	32	32	32	32
循环谱(零谱值个数)	17	17	17	17	17	17

虽然 SMS4 算法、AES 算法和 Camellia 算法 S 盒的代数表达式的次数均为 254，但 SMS4 算法和 Camellia 算法的项数多达 255，而 AES 算法仅为 9。可见，SMS4 算法和 Camellia 算法 S 盒的代数表达式比 AES 的更复杂，在一定程度上更好地保证了算法的安全性。

3 种算法 S 盒的差分矩阵的最大值均为 4，且最大值在每一行每一列(首行首列除外)中仅出现一次。因此，可以初步认为 SMS4 算法、AES 算法和 Camellia 算法 S 盒的差分特性相当。此外，3 种算法 S 盒的线性矩阵中最大绝对值均为 16，且最大值在每一行每一列(首行首列除外)中仅出现 5 次。因此，它们的线性特性也是相当的。

由表 2 可知，3 种算法 S 盒的所有布尔函数都是平衡函数，次数都达到最大上限 7，非线性度与完全非线性函数都有(120-112)/120=6.67%的距离，循环谱的最大值和零谱值个数也都一致。仅有的区别在于 SMS4 算法 S 盒布尔函数的最小项数为 123，Camellia 算法为 126，比 AES 的最小项数 110 更多，从这个方面来说，SMS4 算法和 Camellia 算法 S 盒布尔函数的复杂度略优于 AES 算法。

5 结束语

本文对 SMS4 算法 S 盒的密码性进行了深入分析，讨论了算法种 S 盒的代数次数、平衡性、非线性度、雪崩特性、扩散特性及其不变线性结构与恒变线性结构等性质。通过与 Rijndael, Camellia 等算法的 S 盒比较发现，SMS4 算法 S 盒的设计已经达到欧美分组密码标准算法 S 盒的设计水准。但算法的整体安全特性还有待进一步的研究。

参考文献

[1] Daemen J, Rijmen V. AES Proposal: Rijndael(Version 2)[EB/OL]. [2006-04-12]. http://www.east.kuleuven.ac.be/~rijmen/rijndael.

[2] Aoki K, Ichikawa T, Kanda M, et al. Camellia: A 128-bit Block Cipher Suitable for Multiple Platforms[EB/OL]. [2006-09-12]. http://info.isl.ntt.co.jp/camellia. 2000.

[3] 国家密码管理办公室. 无线局域网产品使用的SMS4 密码算法[EB/OL]. [2006-04-20]. http://www.oscca.gov.cn/Doc/6/News_1106.htm.

[4] Biham E, Shamir A. Differential Cryptanalysis of DES-like Cryptosystems[C]//Proc. of CRYPTO'90. Berlin: Springer-Verlag, 1991.

[5] Matsui M. Linear Cryptanalysis Method for DES Cipher[C]//Proc. of EUROCRYPT'93. Berlin: Springer-Verlag, 1994.

[6] Jakobsen T, Kundsen L R. The Interpolation Attack on Block Ciphers[C]//Proc. of the 4th Fast Software Encryption Workshop. Berlin: Springer-Velag, 1997.

[7] Nyberg K. Differentially Uniform Mappings for Cryptography[C]// Proc. of EUROCRYPT'93. Berlin: Springer-Verlag, 1994.

[8] 温巧燕, 钮心忻, 杨义先. 现代密码学中的布尔函数[M]. 北京: 科学出版社, 2000.

[9] Filiol E. A New Statistical Testing for Symmetric Ciphers and Hash Functions[EB/OL]. [2006-08-05]. http://eprint.iacr.org/2002/.

[10] 冯登国. 频谱理论及其在密码学中的应用[M]. 北京: 科学出版社, 2000.