DES中S盒设计准则的分析*

梁 军

(国防科技大学)

【摘要】在被NSA称作S盒"设计准则"的一组已知的S盒特性中，有的比较浅显易懂，有的则不易理解。基于求本探源的思想，本文对这组特性进行了逐条分析，对其各自的理论依据和作用作了一些解释和推断，并且归纳了一种依据"设计准则"设计S盒的方法。

1 引言

作为美国联邦信息处理标准（FIPS）的DES，历经1977年、1983年和1987年的三次公开的论证和审查，将延期五年，使用到1992年。DES如此高的密码强度主要来源于其核心——S盒。长期以来，S盒特性及其设计方法的研究一直受到人们的重视。

被美国安全局(NSA)称作S盒"设计准则"的一组已知的S盒特性可作如下表述。

DES体制中有8个S盒，每个S盒可用一个4×16矩阵描述。

P_0：S盒的每一行是整数0至15的一个置换；

P_1：每个S盒的传输函数都不是线性的或仿射的；

P_2：对一个S盒而言，输入端每改变1位，至少引起输出端改变2位；

P_3：S(x)和S(x+0011 00)至少有2位不同；

P_4：对于任何e，f，$S(x) \neq S(x+11ef00)$；

P_5：当输入端任何一位保持不变时。S盒的设计应保证在其输出端0和1的个数之差达到最小[1]。

其中，六维矢量$x=(x_1\ x_2\ x_3\ x_4\ x_5\ x_6)=(a\ b\ c\ d\ e\ f)$表示一个S盒的六个输入端。S盒的四个输出端则由矢量$y=(y_1\ y_2\ y_3\ y_4)=(W\ X\ Y\ Z)$表示。

下面对上述特性进行逐条分析。

2 P_0的分析

设集合$A=\{0, 1, 2, \cdots, 15\}$，$P_0$表明了一种A到A自身的映射$S: A \to A$。若存在集合B和C，且满足

*本文于1989年12月18日收到。

$B \supset A$ （$B \fallingdotseq A$），

$C \subset A$ （$C \fallingdotseq A$），

则映射$S'：A \to B$和映射$S''：A \to C$将造成码位的多余或者信息的丢失。

虽然，S是A的一个变换，但其实现过程却不是简单进行的。设B＝｛0，1，2，…，63｝是一个含64个元素的集合，实际的映射过程如图1所示。设映射$\sigma：A \to B$和映射$\tau：B \to A$，则S为一个映射的乘积$S = \tau\sigma$。其中，σ由a、f和P_0控制，而τ仅由P_0决定。显然，τ不是σ的逆映射，S不是集合A的恒等映射，即$S \fallingdotseq 1_A$。

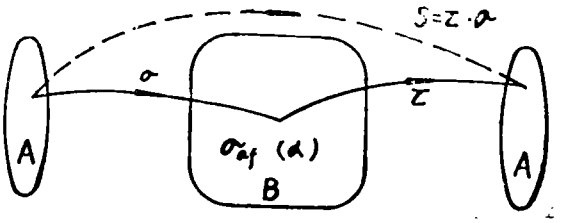


图 1

因为S的值域R（S）＝A，所以S是满射（映上的）。因为对于A中不同的元素i、j（i，$j \in A$，$i \neq j$），S（i）$\fallingdotseq$s（j）不恒成立，所以S不是单射（1—1的）。因此，在S盒的输入端和输出端之间不存在1—1对应关系。这一点正是S盒的特征，是S盒提供密码强度的基础。

3 P_1的必要性

可以证明（证明详见文献［3］），超加密算法不可能为违背P_1的设计带来多少密码强度。当在DES算法中违背P_1时，只要密码分析者掌握了一组明文——密文对，就可以由此求出密钥，达到破译的目的。

4 p_2是非线性强度的保证

P_2表明S盒本身就有完全的错误扩散特性，而这种特性却与函数S的非线性强度有很大的关系。在设计新S盒时，我们发现，因满足的条件不同，设计结果（指非线性强度）就相应不同（详见表1）。

表1列出了五种不同的条件及设计结果。现以条件2、5为例，加以说明。

例1：将S_2盒按条件2修改后得S_2'（设$x_i = e$，$y_j = w$），如图2所示。

经化简，结果如下：

S_2'

15	7	8	0	6	14	3	11	9	1	2	10	12	4	5	13
3	11	4	12	15	7	8	0	1	9	2	10	6	14	13	5
0	8	7	15	10	2	13	5	3	11	12	4	9	1	6	14
13	5	10	2	3	11	4	12	14	6	7	15	0	8	1	9

图 2

表 1

序号	满 足 条 件	设 计 结 果
1	x每改变1位，引起y改变1位	线性函数 $s(x)=g(a+b+c+d+e+f)$
2	x某位x_i改变，仅引起y某位y_j改变，其余情况满足P_2。 ($i=1, 2, 3, 4, 5, 6$ $j=1, 2, 3, 4$)	部分线性 $y_j=x_i+g_{ij}(x_{k1},x_{k2},x_{k3},x_{k4},x_{k5})$ $y_t=g_t(x_{k1},x_{k2},x_{k3},x_{k4},x_{k5})$ ($k_s\neq i$ $s=1, 2, 3, 4, 5, t\neq j$)
3	x仅在某矢量x_0输入时，满足条件2，其余情况满足P_2	非线性函数
4	P_2	非线性函数
5	x每改变1位，引起y改变3位	线性函数 $s(x)=g(a+b+c+d+e+f)$

$$y_1=w=e+g_{51}(a, b, c, d, f)。$$

其中

$$g_{51}(a, b, c, d, f)=\bar{a}f+a\bar{f}+\bar{b}\bar{c}+b\bar{d}+ab\bar{c}+ab\bar{f}+ab\bar{d}f+b\bar{c}df,$$

$$y_2=X=g_2(a, b, c, d, f)=a+c+\bar{b}d+\bar{b}\,\bar{c}\,\bar{f}+ab\bar{c}\,\bar{f}+abd\,\bar{f},$$

$$y_3=Y=g_3(a, b, c, d, f)=d+\bar{b}c+b\bar{c}+\bar{a}\,\bar{c}+\bar{a}cdf+abdf,$$

$$y_4=Z=g_4(a, b, c, d, f)=\bar{d}+bc+a\bar{f}+\bar{a}\,\bar{b}c\,\bar{f}+ab\bar{c}。$$

例2：将S_1盒按条件5修改后得S_1'，如图3所示。

S_1'

14	5	3	8	9	2	4	15	0	11	13	6	7	12	10	1
0	11	13	6	7	12	10	1	14	5	3	8	9	2	4	15
9	2	4	15	14	5	3	8	7	12	10	1	0	11	13	6
7	12	10	1	0	11	13	6	9	2	4	15	14	5	3	8

图 3

经过化简，结果如下：

$$y_1=w=b+d+e+\bar{f},$$

$$y_2=X=a+b+c+d+\bar{f},$$

$$y_3=Y=a+b+c+e+\bar{f},$$

$$y_4=Z=a+c+d+e,$$

显然，表2中条件1、5背离了P_1（尽管条件5满足P_2）；条件2引起函数表达式简化而导致密码强度下降；而条件3中的x_0又必然是成对出现的，所以条件4（P_2）比条件3的相关特性更强，密码强度更高。

5 p_3的分析

因为在P_3涉及的矢量（0 0 1 1 0 0）中，a、f均为0，所以P_3刻划的是一个置换（4 × 16矩阵中某一行）内的特性。当初NSA透露的S盒的设计原则是P_1、P_2、P_6[2]。后来，E. F. Brickell 等人发现，在满足P_0、P_2、P_3的条件下，S盒便满足P_6[1]，P_6成为P_0、P_2、P_3的一个推论。由此可见，在条件P_0和P_2下，P_3是S盒满足P_6的充分条件。由于P_6的基础是统计学，所以P_3就显得很重要了。

在设计新S盒时，我们发现以下两点仍然值得注意。

（1）从几何对称性看，与P_3对等的准则可能还有三个：

P_{31}：s(x)与s(x+ 0 0 1 0 1 0)至少有两位不同；

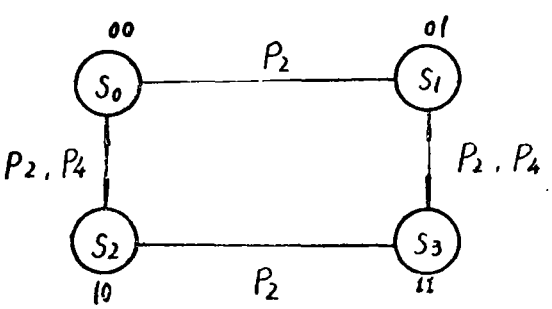
P_{32}：s(x)与s(x+ 0 1 0 0 1 0)至少有两位不同；

P_{33}：s(x)与s(x+ 0 1 0 1 0 0)至少有两位不同。

（2）P_3和P_2之间有某种联系。在S盒中取出一些置换，当有意改变其中若干位以背离P_3时，发现结果违背3 P_2。当利用P_2加P_3设计一个置换时，很容易设计出满足于P_2的置换。因此，P_3对P_2有一定的保障作用。但是在用P_2加P_3的方法设计一个置换时，仍然不可避免误入歧途的可能性，因而不能保证一次设计成功。

6 p_4的分析

由P_4涉及的矢量（1 1 e f 0 0）中的a = 1，f = 0，可知P_4是用来刻划一个S盒内两对不同置换之间关系的。这里，对的划分是依据f进行的。f相同的两个置换组成一对，如图4所示。图中还以准则形式标出了各置换之间的约束关系。



4

NSA声称P_4也是S盒“设计准则”P_1、P_2、P_3的一个推论[1]。我们发现，在一个S盒的4 × 16矩阵中调换某些数值的位置以违背P_4时，结果都导致了新设计的失败（违背P_2）。由此可见，要在一个4 × 16 S盒矩阵中实现P_2，就必须满足P_4条件。

7 关于p_5

P_5的基础是统计学。P_5的作用比较明显，在此不作赘述。

8 根据“设计准则”设计S盒的方法

由S盒的“设计准则”很容易构成一套设计S盒的方法。由于这种方法是按4 × 16

矩阵分行进行设计的，所以只需两张表（表3和表4）和四个四元卡诺框就够了，设计方法简述如下。

首先，我们按P_2和P_3制出表2，按P_4制出表3。

表2

准则	α（或β）	0	1	2	3	4	⑤	6	7	8	9	10	11	12	13	⑭	15
P_2	w_1	8	9	10	11	12	⑬	14	15	0	1	2	3	4	5	⑥	7
	w_2	4	5	6	7	0	①	2	3	12	13	14	15	8	9	⑩	11
	w_3	2	3	0	1	6	⑦	4	5	10	11	8	9	14	15	⑫	13
	w_4	1	0	3	2	5	④	7	6	9	8	11	10	13	12	⑮	14
P_3	α′	6	7	4	5	2	3	0	1	14	15	12	13	10	11	⑧	9

表2的第一行是矢量α（或β）行。其中，α=（b c d e），β=y=（W X Y Z）。注意，这里是用十进制数表示的。表2中，同一列填写的是矢量 wi（i=1，2，3，4）的十进制表示数。wi与α（或β）有且仅有一位(第i位)不同。表中α′=α+（0 1 1 0）。如果按表2取6×16矩阵P，则P中元素满足

$$\begin{cases} P_{1j}=P_{i1}+P_{ij} & (i=1,2,3,4,5) \\ P_{0j}=P_0+P_{1j} \end{cases}$$

其中，P_0=（0 1 1 0）。显然，在S盒输入端每改变一位（第i位）（暂不考虑a、f位）时，α就变为wi。下面举例说明表2的用法。

定义映射 S_i：α→β（i=0，1，2，3 ——由a，f决定）。设α=14，且在S_0下，β=5。即，S_0（14）=5。

因此，应该在第i（i=0）个四元卡诺框的第α（α=14）格内，填入值β（β=5）。并且根据P_2，在α的四个邻格6，10，12，15中，根据P_3在格8中一律禁用值5，13，1，7，4（参见表2），结果如图5（a）所示。

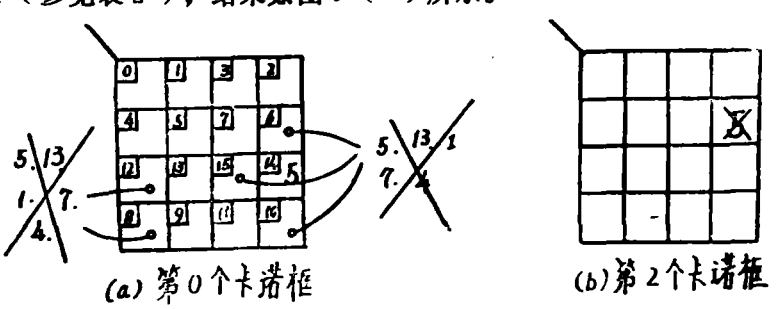


图 5

表3中，第一列为a f值；第二列为置换代号；第三列中α″=α+（1 e f 0）。

表3的用法：表3将四个置换分为两组。现合在一起叙述如下：

表3

0 0	S_0	α	0	1	2	3	4	5	6	7	8	9	10	11	12	13	⑭	15
1 0	S_2	α″	8	13	10	15	12	9	⑭	11	0	5	2	7	4	1	6	3
0 1	S_1	α	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 1	S_3	α″	10	15	8	13	14	11	12	9	2	7	0	5	6	3	4	1

在s_0（s_1）求得后，须根据P_4考虑其对s_2（s_3）设计的影响。具体地说，就是在第2（3）个卡诺框的第α格内禁用第0（1）个卡诺框的第α″格内的β值。例如，在第2个卡诺框的第6格内禁用5。因为5是第α个卡诺框的第14格内的β取值（参见表3），如图5所示。

注意到P_0，P_1，P_5在下述设计步骤中已能得到满足，参考图4关系，现归纳一个S盒的设计步骤如下。

（1）按P_2加P_3求s_0；

（2）按P_2加P_3求s_1（包括按P_2考虑s_0的影响）；

（3）按P_2加P_3求s_2（包括按P_2和P_4考虑s_0的影响）；

（4）按P_2加P_3求s_3（包括按P_2考虑s_1、s_2的影响，按P_4考虑s_1的影响）。

9 结论

由NSA透露的DES的S盒“设计准则”P_0～P_5是目前设计新S盒的主要依据。其中，P_0，P_1，P_2和P_5是建立在密码学理论基础之上的，P_3和P_4为设计者的工作带来了方便。但注意到由P_0～P_5构成的设计方法仍可能误入设计歧途，所以可以断言，就设计方法而言P_0～P_5尚未达到完备的地步。关于P_3和P_4与P_1，P_2和P_5的关系至今也尚无数学上的表述。所以，要做的工作还很多。

10 鸣谢

本文在写作过程中得到汪漱玉副教授、李情与副教授和唐朝京同志的热情帮助，谨在此表示感谢！

参考文献

〔1〕E.F.Brickell, J.H.Moore, M. R. Purtill: "Structure in the S—Boxes of the DES", Advances in Cryptology, Proceedings of CRYPTO86.

〔2〕A·G. Konheim: "Cryptography, A Primer", John Wiley & Sons, Inc. 1981.

〔3〕梁军: "DES中S盒设计准则的分析"，全国信息论与通信理论学术会论文资料，1989年10月。