Computer Science and Application 计算机科学与应用, 2019, 9(5), 912-920
Published Online May 2019 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2019.95103

Hans汉斯

Division Property Analysis of S-Boxes

Cuihua Nie, Hongru Wei

School of Mathematics and Physics, University of Science and Technology Beijing (USTB), Beijing
Email: 18813128404@163.com

Received: May 1st, 2019; accepted: May 14th, 2019; published: May 21st, 2019

Abstract

This paper uses two methods to analyze the division property of the S-boxes, mainly for the S-boxes of MISTY1, Camellia, AES, SMS4, DES, GIFT, Gost, KLEIN, LED, LBlock, MIBS, mCRYPTON, Midori64, RESENT, PRINCE, PRIDE, Piccolo, PUFFIN, RECTANGLE, SKINNY, SPONGENT, Serpent, TWINE, as well as 16 optimal S-boxes. The first method is based on algebraic degree, using the correspondence between Hamming weight and algebraic degree to find the division property of the S-boxes. The experimental results are obtained. According to the rules of division property propagation, we can obtain the theoretical derivation values of division property, compare and analyze experimental results with theoretical derivation values. It shows that there are a few differences. Since the same Hamming weight contains multiple cases, it may be hidden, so the second method, that is, the detailed division property, based on bit level, is adopted to analyze division property for each case. As a result, there are better results found than the first method. Lightweight 4-bit S-boxes have division property that can be utilized. 8-bit S-boxes based on finite field inverse have high security without balanced bits. This will facilitate the security analysis of block cipher algorithms and help to reduce time complexity and data complexity.

Keywords

S-Box, Division Property, Block Cipher, Division Trail

S盒的可分性质分析

聂翠华，卫宏儒

北京科技大学数理学院，北京
Email: 18813128404@163.com

收稿日期：2019年5月1日；录用日期：2019年5月14日；发布日期：2019年5月21日

摘　要

本文使用两种方法对S盒的可分性质进行了分析。主要针对MISTY1，Camellia，AES，SMS4，DES，GIFT，

文章引用: 聂翠华, 卫宏儒. S 盒的可分性质分析[J]. 计算机科学与应用, 2019, 9(5): 912-920.
DOI: 10.12677/csa.2019.95103

Gost，KLEIN，LED，LBlock，MISBS，mCRYPTON，Midori64，RESENT，PRINCE，PRIDE，Piccolo，PUFFIN，RECTANGLE，SKINNY，SPONGENT，Serpent，TWINE等分组密码算法中的S盒，以及16个最优S盒。第一种方法基于代数次数，利用汉明重量与代数次数的对应关系，分析S盒的可分性质，即得实验结果。根据可分性质传播规则，可得可分性质理论推导值。将实验结果与理论推导值相比较并进行分析，发现少部分有区别。由于第一种方法中同一个汉明重量对应多种情况，猜测有些可分性质可能被隐藏，于是采取第二种方法——基于比特级即细化的可分性质，针对每一种情况分析对应的可分性质，得到了比第一种方法更好的结果。轻量级4比特S盒具有可以使用的可分性质。基于有限域逆的8比特S盒没有平衡比特，具有高安全性。这将有助于分组密码算法的安全性分析，在降低时间复杂度与数据复杂度方面均有帮助。

关键词

S盒，可分性质，分组密码，可分迹

1. 引言

可分性质[1]是密码学者 Todo 在 2015 年欧密会上提出来的，它是积分性质的推广。即使分组密码算法具有非双射函数、比特导向结构和低次数函数，也可通过可分性质构造出有效的积分区分器，这弥补了积分性质的缺憾。

可分性质一经提出，引来无数学者的进一步研究。Bing SUN 等人[2]证明输入集中差分元素至少有个。Sun L 等人[3]对可分性质作了进一步的解释，引入了一种新的概念——奇偶性集合[4]，能够以简单的方式制定和表征任何秩序的可分性质。通过考虑奇偶校验集的更多属性来概括可分性质，从而在分组密码上构建区分器。还出现了另外一种新技术，即在独立考虑每个比特和整体考虑左半部和右半部之间实现折衷[5]，这实际上是时间存储复杂度和区分器精度之间的折衷。为了使用常量或子密钥传播 AND 和 OR 运算的基于比特的可分性质，Sun L 等人[6]证明了由输入可分性质推导出已知区域总是包含在从输出可分性质派生的已知区域中。尤瑞英[7]研究了已有的积分区分器在密钥恢复过程中的表现，进一步分析了 PRESENT、Serpent 和 Noekeon 在积分分析下的安全性。

最初的积分攻击针对的是基于字节或字设计的分组密码。由于基于比特设计的分组密码在应用积分攻击时活跃字节的性质往往被线性变换毁坏，因此 Z'aba 等人在 FSE2008 上首次提出了基于比特的积分攻击方法[8]。可分性质也是在这样一个基础上推广来的。Todo 等人应用基于比特的可分性质，对全轮 MISTY1 进行了积分特征的分析[9]。虽然基于比特的可分性质能找到更准确的积分特征，但它的时间复杂度和数据复杂度都很高。因此，不能将其应用于分组长度超过 32 的算法。

S 盒是分组密码算法中的一个重要部件，本质上是一张替换表，对于给定的输入，通过查找该表能够得到相应的输出。S 盒对于整个密码算法的安全性十分重要，对 S 盒的可分性质研究也有许多。尤其是对轻量级分组密码算法[10]中的 S 盒，如 Simon 家族、PRESENT 等算法。可分性质的提出者 Todo 也做了对 S 盒性质的搜索，主要针对 MISTY1 算法，对其中的 S 盒及轮函数进行了可分性质分析，发现 S7 存在一种退化现象，这一现象有利于降低复杂度，因此对 S 盒的可分性质研究有实际意义。

本文针对现有分组密码算法中的 S 盒，首先使用基于代数次数的方法，得到 S 盒的一些可分性质；

之后使用基于比特的可分性质分析方法，得到了更好的可分性质结果。最后对两种方法的实验结果进行了分析比较，轻量级 4 比特 S 盒具有可以使用的可分性质。基于有限域逆的 8 比特 S 盒没有平衡比特因而具有高安全性。S 盒具有的不平衡比特越多，安全性就越高。

2. 相关知识

2.1. 符号说明

$\oplus$：表示 $\mathbb{F}_2^n$ 上求和运算。对于任意 $a \in \mathbb{F}_2^n$，第 i 个元素表示为 $a[i]$，汉明重量 w_a，计算公式 $w_a = \sum_{i=1}^{n} a[i]$。

子集 $\mathbb{S}_k^n$：对于任意的整数 $k \in \{0,1,\cdots,n\}$，设 $\mathbb{S}_k^n$ 是 $\mathbb{F}_2^n$ 的一个子集，它是满足 $k \le w_a$ 的所有 $a \in \mathbb{F}_2^n$ 的集合，定义如下：$\mathbb{S}_k^n := \{a \in \mathbb{F}_2^n \mid k \le w_a\}$。

比特产生函数 π_u：$\pi_u(x) := \prod_{i=1}^{n} x[i]^{u[i]}$，$\mathbb{F}_2^n \to \mathbb{F}_2$。

在计算可分性质时，为了区分输入数据集与输出数据集，其比特产生函数的自变量分别用 u 和 v 表示，即输出数据集比特产生函数记为 π_v。

2.2. 可分性质定义

设 X 是一个多重数据集，元素取值于 $\mathbb{F}_2^n$，k 取值范围是 0 到 n。当数据集 X 有可分性质 D_k^n 时，满足如下条件：当 $w_u < k$ 时，对于所有的 $x \in \mathrm{X}$，$\pi_u(x)$ 总是偶的；当 $w_u \ge k$ 时，$\pi_u(x)$ 的奇偶性则是未知的。

即当数据集 X 有可分性质 D_k^n 时，对于所有的 $u \in (\mathbb{F}_2^n \setminus \mathbb{S}_k^n)$，它满足 $\bigoplus_{x \in \mathrm{X}} \pi_u(x) = 0$。对任意的 $u \in \mathbb{S}_k^n$，对所有 $x \in \mathrm{X}$，$\pi_u(x)$ 奇偶性都是未知的。也就是说，在可分性质中，u 的集合被划分为两个子集，一个子集 $\oplus_{x \in \mathrm{X}} \pi_u(x)$ 奇偶性未知，另一个子集 $\oplus_{x \in \mathrm{X}} \pi_u(x)$ 是 0。

可分性质传播规则：设 S 是一个 n 比特到 n 比特的函数(S 盒)，其代数次数为 d。假设输入多重集 X 有可分性质 D_k^n，则输出多重集 Y 有可分性质 $D_{k'}^n$，$k' = \lceil k/d \rceil$。另外，假设 S 盒是排列，当输入集有可分性质 D_n^n 时，输出集的可分性质也是 D_n^n。

3. 基于代数次数的 S 盒可分性质分析

3.1. 基于代数次数的分析方法定义

Todo 在对 MISYT1 进行积分分析时，对其中的 S_7 和 S_9 进行了可分性质搜索。以 S_7 为例，其代数次数[11]是 3，已有 S 盒的布尔函数 S_7 代数标准型 ANF(Algebraic Normal Form)。计算复合布尔函数 $(\pi_v \circ S_7)$ 的可分性质从而得到该 S 盒的可分性质传播特征。$\pi_v \circ S_7$ 的代数次数 deg 与 V 的汉明重量 w_v 对应关系如表 1。

Table 1. w_v-deg relationship of the7-bit S-box in MISTY1
表 1. MISTY1 中 7 比特 S 盒 w_v-deg 对应关系

w_v	0	1	2	3	4	5	6	7
deg	0	3	5	5	6	6	6	7

该 S 盒可分性质传播特征如表 2。

这是实际传播特征。值得注意的是，根据可分性质传播规则，即 $k' = \lceil k/d \rceil$，可以得到理论值传播情况如下表 3 所示。

Table 2. The 7 bit S-box division property propagation characteristics
表 2. S7 可分性质传播特征

X _ D_k^7	D_0^7	D_1^7	D_2^7	D_3^7	D_4^7	D_5^7	D_6^7	D_7^7
Y _ $D_{k'}^7$	D_0^7	D_1^7	D_1^7	D_1^7	D_2^7	D_2^7	D_4^7	D_7^7

Table 3. The 7 bit S-box theoretical division property propagation characteristics
表 3. S7 理论可分性质传播特征

X _ D_k^7	D_0^7	D_1^7	D_2^7	D_3^7	D_4^7	D_5^7	D_6^7	D_7^7
Y _ $D_{k'}^7$	D_0^7	D_1^7	D_1^7	D_1^7	D_2^7	D_2^7	D_2^7	D_7^7

可发现，实际值与理论值有不同。即 D_6^7 理论上传播到 D_2^7，但实际上传播到 D_4^7，这是由于最高次项被抵消造成的。这就是此方法分析 S 盒可分性质的关键之处。

3.2. 基于代数次数的分析方法原理和过程

根据 $\pi_v \circ S_n$ 的代数次数 deg 与 v 的汉明重量 w_v 之间的关系，可以推出 S 盒的可分性质传播特征。由可分性质的定义可知，假设输入数据集的可分性质满足 D_k^n，即当 $w_u < k$ 时，$\oplus_{x\in \mathrm{X}} \pi_u(x)$ 是 0，经复合布尔函数作用之后的输出数据集的可分性质记为 $D_{k'}^n$，即当 $w_v < k'$ 时，$\oplus_{y\in \mathrm{Y}} \pi_v(y)$ 是 0。其中 k 与 k' 的关系相对应于布尔函数 $\pi_v \circ S_n$ 的代数次数 deg 与变量 v 的汉明重量 w_v 的对应关系。

研究可分性质传播时，猜测满足可分性质为 D_k^n 的最大集合，设其元素长度为 n 比特，令其中 k 个比特的值遍历(每个位置取 0 或 1，共 2^k 种情况)，其余$(n-k)$个比特的值固定。经验证，固定值是 0 或者是 1 对结果没有影响。取出这样的数据集 X 之后，根据可分性质的定义，计算该数据集的可分性质，结果证明该数据集满足可分性质，但并不能证明这是满足可分性质是 D_k^n 的最大数据集。

在 S 盒的可分性质传播中，S 盒本身是一个布尔函数，可分性质计算时使用的比特产生函数 $\pi_v(x)$ 也是布尔函数，因此可将这两个布尔函数复合，即 $\pi_v \circ S_n$。分析过程如下：

1) 该复合布尔函数 $\pi_v \circ S_n$ 的汉明重量与代数次数对应关系，共 $n+1$ 种情况，这里的汉明重量指的是进行比特产生函数计算时，需要遍历的参数 v 的汉明重量。

2) 接下来根据此关系推导可分性质传播情况：已知输入集 X 有可分性质 D_k^n，记输出集 Y 的可分性质为 $D_{k'}^n$。由搜索原理可分别求出与 k 对应的 k' 的值，即可得到可分性质传播的实际值。

3) 利用传播规则，理论上可分性质传播得到的值为 $k' = \lceil k/d \rceil$，与实际值作比较。

另外，尝试遍历所有 16！个 4 比特 S 盒。遍历过程如下：

1) 用一个递归函数 *function* 来求取 0 到 15 这 16 个数的排列即一个 4 比特 S 盒，结果用数组 *result* 表示；

2) 初始化一个索引与数值对应的链表 *list* 值为 0 到 15，给定初始索引位置 *index* 为 0 ，这两个赋值变量作为递归函数 *function* 的两个参数；

3) 取第一个变量 *list* 的长度 *length* ，当 $length > 0$ 时，执行下一步，否则执行步骤 5)；

4) 更新索引位置 $index1 = index + 1$，定义一个循环，当 $i < length$ 时，将 $list[i]$ 赋值给 $result[index]$，为了保证 *list* 不变，将其赋值给一个新的链表 *newlist* ，这时移去 $newlist[i]$，调用函数本身进行递归，*newlist* 和 *index*1 作为参数；

5) 当 $length = 0$ 时，16 个数的排列已完成，调用求 S 盒代数次数的函数，这样输出 S 盒及其代数次数。

3.3. 实验结果

主要针对已有轻量级分组密码算法中的 S 盒，基于代数次数做了可分性质分析。基于篇幅限制，本

文举例进行结果分析。

Camellia、AES、SMS4 算法中的 8 比特 S 盒，代数次数 7，其结果如下表 4 所示。

Table 4. The 8-bit S-box division property propagation characteristics
表 4. 8 比特 S 盒可分性质传播特征

$X: D_k^n$	D_0^8	D_1^8	D_2^8	D_3^8	D_4^8	D_5^8	D_6^8	D_7^8	D_8^8
$Y: D_k^n$	D_0^8	D_1^8	D_1^8	D_1^8	D_1^8	D_1^8	D_1^8	D_1^8	D_8^8

实际值与理论值一致。

研究的 4 比特 S 盒有：302 个等价类，16 个最优 S 盒 G0-G15 [12]，DES 中 4 个，GIFT，Gost 中 8 个，KLEIN，LED，LBlock 中 10 个，MIBS，mCRYPTON 中 4 个，Midori64 中 8 个，RESENT，PRINCE，PRIDE，Piccolo，PUFFIN，RECTANGLE，SKINNY，SPONGENT，Serpent 中 8 个，TWINE。

302 个等价类中第 14 个：12，9，1，2，3，5，4，7，6，0，10，11，8，13，14，15，代数次数为 2，其可分性质传播特征如下表 5 所示。

Table 5. The 14th of the 302 equivalence classes division property
表 5. 302 个等价类中第 14 类 S 盒的可分性质

$X: D_k^n$	D_0^4	D_1^4	D_2^4	D_3^4	D_4^4
理论值	D_0^4	D_1^4	D_1^4	D_2^4	D_4^4
实际值	D_0^4	D_1^4	D_1^4	D_1^4	D_4^4

其他的 4 比特 S 盒代数次数均为 3(除了 Midori64 算法中有一个 S 盒代数次数为 4)，可分性质传播的实际值均与理论值一致，具体如下表 6。

Table 6. Other 4-bit S-boxes division property
表 6. 其余 4 比特 S 盒的可分性质

$X: D_k^n$	D_0^4	D_1^4	D_2^4	D_3^4	D_4^4
理论值	D_0^4	D_1^4	D_1^4	D_1^4	D_4^4
实际值	D_0^4	D_1^4	D_1^4	D_1^4	D_4^4

结果分析：以 302 个等价类中第 14 类 S 盒为例，由基于代数次数的可分性质分析结果可知，实际值与理论值有不同，理论上可分性质 D_3^4 经过该 S 盒传播到 D_2^4，但实际上传播后的可分性质是 D_1^4，即除了汉明重量是 0 的情况是平衡状态，其余均为非平衡状态。这样就会比理论扩散的慢，所以这些可分性质应用到整体算法的安全性分析中可能会相应地降低复杂度。整体来看，除了 D_0^4、D_4^4 两种情况传播到其本身，其余均传播到 D_1^4。也就是说只有 0 这一个平衡状态，说明该 S 盒的安全性是极好的。其余大多数 S 盒的可分性质实际值与理论值一致。

基于代数次数的 S 盒可分性质分析结果显示，在算法的实际安全性分析中能够真正降低复杂度的情况较少，考虑是汉明重量掩盖造成的，因为同一个汉明重量对应多个具体的取值情况，因此，下一章在基于比特的基础上，继续挖掘更多的可分性质。

4. 基于比特的 S 盒可分性质分析

4.1. 基于比特的分析方法定义

定义 1 基于比特的可分性质[13]

当分析分组长度为 n 比特的密码算法时，传统的可分性质使用 $D_{\mathrm{K}}^{l_1,l_2,\cdots,l_m}$，其中 l_i 和 m 是由攻击者在 $n=\sum_{i=1}^{m} l_i$ 范围内选定。这里考虑传统的基于比特的可分性质即 $D_{\mathrm{K}}^{1^n}$。

定义 2 可分迹[14]

设 f_r 是分组密码算法的轮函数。假设分组密码算法的输入多重数据集有初始可分性质 $D_{\boldsymbol{k}}^{n,m}$，记经过 f_r 函数 i 轮传播之后的可分性质为 $D_{\mathrm{K}_i}^{n,m}$。因此，有可分性质传播链：$\{\boldsymbol{k}\} \overset{def}{=} \mathrm{K}_0 \xrightarrow{f_r} \mathrm{K}_1 \xrightarrow{f_r} \mathrm{K}_2 \xrightarrow{f_r} \cdots$。另外，对于 K_i 中任意向量 $\boldsymbol{k}_i^* (i \geq 1)$，一定存在一个 K_{i-1} 中的向量 $\boldsymbol{k}_{i-1}^*$，通过可分性质传播规则能够传播到 $\boldsymbol{k}_i^*$。

对于 $(\boldsymbol{k}_0, \boldsymbol{k}_1, \cdots, \boldsymbol{k}_r) \in (\mathrm{K}_0 \times \mathrm{K}_1 \times \cdots \times \mathrm{K}_r)$，如果 $\boldsymbol{k}_{i-1}$ 能够传播到 $\boldsymbol{k}_i$，则称 $(\boldsymbol{k}_0, \boldsymbol{k}_1, \cdots, \boldsymbol{k}_r)$ 为一个 r 轮的可分迹。

可分迹便是基于比特的一种概念，接下来使用它对 S 盒的可分性质进一步分析。可分迹是用来描述可分性质的传播特征的，通过检查所有可分迹的最后一个向量可估计是否存在有用的区分器。方法一中同一个汉明重量包括多种情况，因此，这里将其细化到每一种情况，期待发现比之前更好的结果。

4.2. 分析过程

符号说明：

n：S 盒比特长度

k：输入值比特级表示，形如 $(k_{n-1}, \cdots, k_1, k_0)$，$k_i \in \{0,1\}$

$Unbal$：k 的非平衡项集合

K_out：k 值对应的输出集

$\succeq : \boldsymbol{k} \succeq \boldsymbol{k}' : \text{if } k_i \geq k_i' \text{ for all } i$

(x_3, x_2, x_1, x_0)：K 中元素

k 的非平衡项是指，u 遍历所有可能的 2^n 个值，若某个 u 值满足 $u \succeq k$，那么此 u 就是 k 的一个非平衡项，遍历完成后即可得到 k 的非平衡项集合 $Unbal$。

取定一个 k 值，当 u 遍历的过程中，对于每一个 u 值，判断相应的 $\pi_u(y)$ 中是否存在非平衡项集合 K_out 中的项，先计算 $\pi_u(y) = \pi_{(u_3,u_2,u_1,u_0)}((y_3, y_2, y_1, y_0)) = y_3^{u_3} \oplus y_2^{u_2} \oplus y_1^{u_1} \oplus y_0^{u_0}$。

其中 $y_m \oplus y_n$ 计算过程如下：

1) 遍历 y_m，非零位置记为 1，零位置记 0，存为 $list1$，同理遍历 y_n 可得 $list2$。这一步是因为 0 的存在，本可以直接对原数组 y_m 进行遍历，索引值与数值一致。但是，若“0(0000)”这一项存在，索引却为 0，检索不到，无法与后一个数组 y_n 进行运算，因此建立新数组 $list$，来区别“0”项的存在。

2) 遍历 $list1$，对于 y_m 的每个非零项，与 y_n 中所有项做逻辑或运算，以结果为索引，值为 1 存入新的数组 $result$ 中，注意数值采取二进制加法(如果该项出现过两次，则抵消，因为实质上项与项之间是异或运算)。

3) 遍历完成后，得到的是 $y_m \oplus y_n$ 结果的索引数组 $list$，这样直接利用该 $list$ 继续下一组运算。

这样 $\pi_u(y) = \pi_{(u_3,u_2,u_1,u_0)}((y_3, y_2, y_1, y_0)) = y_3^{u_3} \oplus y_2^{u_2} \oplus y_1^{u_1} \oplus y_0^{u_0}$ 计算完毕，得到的是索引数组，要转变为值数组，主要判断“0”项是否存在。接下来，判断结果中是否包含非平衡项集合 $Unbal$ 中的项。

若 $\pi_u(y)$ 中存在非平衡项集合 $Unbal$ 中的项，此 u 便是输出集 K_out 中的一个元素，遍历完成后，

可初步确定 k 值对应的输出集 K_out 。

最后需要使用 SizeReduce()函数去掉冗余：对于输出集 K_out 中任一项 $k^{(i)}$，若存在一个 j，使得 $k^{(i)} \succeq k^{(j)}$，则 $k^{(i)}$ 是冗余的，删去。这样就可得到该 S 盒输入 k 值时的输出集合 K_out 。

4.3. 实验结果

由于篇幅限制，本文未列出所有实验结果，下面以 SKINNY 算法中的 S 盒为例，举例进行结果分析，结果如下表 7。

Table 7. Division trails of the S-box in the SKINNY
表 7. SKINNY 算法中 S 盒的可分迹

Input $D_k^{1,4}$	Output $D_K^{1,4}$
(0, 0, 0, 0)	(0, 0, 0, 0)
(0, 0, 0, 1)	(0, 0, 0, 1) (0, 0, 1, 0) (1, 0, 0, 0)
(0, 0, 1, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (0, 1, 0, 0)
(0, 0, 1, 1)	(0, 0, 0, 1) (0, 0, 1, 0) (1, 1, 0, 0)
(0, 1, 0, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (0, 1, 0, 0) (1, 0, 0, 0)
(0, 1, 0, 1)	(0, 0, 0, 1) (0, 1, 1, 0) (1, 0, 1, 0) (1, 1, 0, 0)
(0, 1, 1, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (0, 1, 0, 0)
(0, 1, 1, 1)	(0, 0, 0, 1) (0, 1, 1, 0) (1, 1, 0, 0)
(1, 0, 0, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (0, 1, 0, 0) (1, 0, 0, 0)
(1, 0, 0, 1)	(0, 0, 0, 1) (0, 1, 1, 0) (1, 1, 0, 0)
(1, 0, 1, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (1, 1, 0, 0)
(1, 0, 1, 1)	(0, 0, 1, 1) (0, 1, 1, 0) (1, 1, 0, 1)
(1, 1, 0, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (1, 0, 0, 0)
(1, 1, 0, 1)	(0, 0, 1, 1) (1, 1, 1, 0)
(1, 1, 1, 0)	(0, 0, 0, 1) (0, 0, 1, 0) (1, 1, 0, 0)
(1, 1, 1, 1)	(1, 1, 1, 1)

(0, 0, 0, 1)对应的结果集有 3 个单位向量(0, 0, 0, 1) (0, 0, 1, 0) (1, 0, 0, 0)，也就是说，x_0, x_1, x_3 这 3 个比特是非平衡的，而 x_2 这 1 个比特是平衡的；(0, 0, 1, 0)对应的结果集有 3 个单位向量，即第 x_0, x_1, x_2 这 3 个比特是非平衡的，而 x_3 这 1 个比特是平衡的；(0, 0, 1, 1)对应的结果集(0, 0, 0, 1) (0, 0, 1, 0) (1, 1, 0, 0)，x_0, x_1 这 2 个比特和 x_2x_3 乘积项是非平衡的；(0, 1, 0, 0)对应的结果集是 4 个单位向量，即 4 个比特都是非平衡的；(1, 0, 1, 1)对应的结果集(0, 0, 1, 1) (0, 1, 1, 0) (1, 1, 0, 1)，没有单位向量，即 4 个比特都是平衡的，这一点在 S 盒安全性分析时可以降低扩散的速度，更好地分析其安全性。

对于 MISTY1 中的 7 比特 S 盒 S7，(0, 0, 0, 1, 1, 1)和(0, 0, 1, 0, 1, 1, 0) 这两个向量作为初始数据集时，结果集中不存在单位向量，即 7 个比特都是平衡的，而其中一些乘积项如 x_0x_1 、 x_0x_2 、 x_1x_2 等等是非平衡的。

Camellia、AES、SMS4 中的 8 比特 S 盒结果都是 8 个单位向量，即这 8 个比特都是非平衡的。

5. 两种分析结果对比

下面仍以 SKINNY 算法中的 S 盒为例，按第一种方法，可分性质 D_3^4 经过该 S 盒传播到 D_1^4，即 4 个比特都是非平衡的，应包含 4 个单位向量；按第二种方法，汉明重量为 3 的共 4 种情况(0, 1, 1, 1)，(1, 0,

1, 1)，(1, 1, 0, 1)，(1, 1, 1, 0)。(0, 1, 1, 1)作为初始向量，经 S 盒传播后得到集合(0, 0, 0, 1) (0, 1, 1, 0)(1, 1, 0, 0)，即 x_0 这 1 个比特、x_1x_2 和 x_2x_3 这两个乘积项是非平衡的；(1, 0, 1, 1)作为初始集，结果集合为(0, 0, 1, 1)(0, 1, 1, 0)(1, 1, 0, 1)，不包含单位向量，4 个比特都是平衡的，这称之为“退化”；(1, 1, 0, 1)作为初始集，结果集合为(0, 0, 1, 1) (1, 1, 1, 0)，也是“退化”；(1, 1, 1, 0)作为初始集，经 S 盒传播后得到集合(0, 0, 0, 1)(0, 1, 1, 0)(1, 1, 0, 0)，这种变化与第一种情况相似。又比如(0, 1, 0, 0)作为初始集，结果集是 4 个单位向量，这时与第一种方法相比就是没有变化的。

下表 8 对实验中的所有 S 盒进行了统计。

Table 8. Compare the results of S-boxes division property
表 8. S 盒可分性质分析结果对比

算法	变化率%	退化细节
MISTY1-S7	93	0001011,0010110
Camellia-S8		
AES-S8		
SMS4-S8		
等价类 14	75	1101->(0011)(0110)(1101),1110->(0011)(1110)
最优 G1	50	1011->(0101)(0110)(1011),1110->(0011)(1001)(1010)
最优 G2	62.5	1011->(0101)(0110)(1011),1110->(0011)(1001)(1010)
最优 G9	56.25	1011->(0101)(0110)(1011)
DES-4	43.75	
GIFT	62.5	
Gost_6	43.75	1011->(0111)(1010)(1101)
KLEIN	37.5	
LED	62.5	1110->(0101)(1011)(1110)
LBlock-10	62.5	0111->(0111)(1100),1011->(0101)(1011) and so on
MIBS	31.25	
mCRYPTON	31.25	
Midori64	43.75	
PRESENT	75	1110->(0101)(1011)(1110)
PRINCE	43.75	
PRIDE	75	0111->(0011)(1101),1011->(1001)(1110)
Piccolo	75	1011->(0011)(0110)(1101),1101->(0011)(1110)
PUFFIN	37.5	
RECTANGLE	62.5	1011->(0110)(1011)(1101),1101->(0110)(1010)(1101)
SKINNY	81.25	1011->(0011)(0110)(1101),1101->(0011)(1110)
SPONGENT	62.5	0111->(0111)(1001)(1010)(1100)
Serpent_0	62.5	1011->(0111)(1011)(1100)
Serpent_1	62.5	0111->(0101)(1011)(1110)
Serpent_2	25	1110->(0101)(1001)(1110)
Serpent_6	62.5	1101->(0111)(1010)(1101)
TWINE	50	

从 S 盒整体分析，比如 Camellia 中的 8 比特 S 盒，除了 D_0^8 和 D_8^8，其余都是传播到 D_1^8，也就是说所有比特都是非平衡的，这样的 S 盒安全性是比较高的。再比如 302 个等价类中的第 14 个，理论上有一个平衡比特，但是实际上却没有，此 S 盒的安全性也很好。但是对于 MISTY1 中的 S7 来说，D_4^7、D_5^7、D_6^7 这三种情况都是传播到 D_2^7，即均有一个平衡比特，那么在攻击该 S 盒时便可利用此平衡比特，因此该 S 盒的安全性便不是最高的。

6. 结论

本文使用基于代数次数、基于比特即可分迹这两种方法，对目前轻量级分组密码算法中的 S 盒，进行了可分性质分析。第一种方法实验结果中，可利用的可分性质不多，考虑在分析可分性质时，是用汉明重量来区分的，但是同一种汉明重量包括多种情况，很可能存在覆盖隐藏，因此将可分性质传播细化到比特级，即将长度为 n 的 S 盒的 2^n 种情况分别计算，针对每一个值，研究可能得到的可分性质。发现了确实存在覆盖现象，挖掘出更多隐藏的可利用的 S 盒可分性质，在对算法做整体安全性分析时，可以降低复杂度或者拓展攻击轮数。

基金项目

本文获得国家自然科学基金项目(No. 61672509、U1603116)和内蒙古自治区科技创新引导奖励资金资助项目的资助。

参考文献

[1] Todo, Y. (2015) Structural Evaluation by Generalized Integral Property. *Advances in Cryptology—EUROCRYPT* 2015, Springer, Berlin, Heidelberg, 287-314. https://doi.org/10.1007/978-3-662-46800-5_12

[2] Sun, B., Hai, X., Zhang, W.Y., Cheng, L. and Yang, Z.C. (2017) New Observation on Division Property. *Science China* (*Information Sciences*), **60**, 274-276. https://doi.org/10.1007/s11432-015-0376-x

[3] Sun, L. and Wang, M.Q. (2017) Toward a Further Understanding of Bit-Based Division Property. *Science China* (*Information Sciences*), **60**, 277-279. https://doi.org/10.1007/s11432-016-9170-y

[4] Boura, C. and Canteaut, A. (2016) Another View of the Division Property. *Advances in Cryptology—CRYPTO* 2016, Springer, Berlin, Heidelberg, 654-682. https://doi.org/10.1007/978-3-662-53018-4_24

[5] Xiang, Z., Zhang, W. and Lin, D. (2016) On the Division Property of Simon 48 and Simon 64. *Advances in Information and Computer Security*, Springer International Publishing, 147-163. https://doi.org/10.1007/978-3-319-44524-3_9

[6] Sun, L., Wang, W., Liu, R., *et al*. (2018) MILP-Aided Bit-Based Division Property for ARX Ciphers. *Science China Information Sciences*, **61**, 1-31. https://doi.org/10.1007/s11432-017-9321-7

[7] 尤瑞英. 应用 MILP 方法搜索基于分离特性的算法积分区分器[D]: [硕士学位论文]. 济南: 山东大学, 2017.

[8] Z'Aba, M.R., Henricksen, M. and Dawson, E. (2008) Bit-Pattern Based Integral Attack. *Fast Software Encryption*, Springer-Verlag, 363-381. https://doi.org/10.1007/978-3-540-71039-4_23

[9] Todo, Y. (2015) Integral Cryptanalysis on Full MISTY1. *Advances in Cryptology—CRYPTO* 2015, Springer, Berlin, Heidelberg, 413-432. https://doi.org/10.1007/978-3-662-47989-6_20

[10] 贾平, 徐洪, 戚文峰. 轻量 S 盒密码性质研究[J]. 密码学报, 2015, 2(6): 497-504.

[11] 杨默涵, 来学嘉. 布尔函数代数次数的计算方法[C]//中国密码学会 2009 年会. 2009: 35-42.

[12] Leander, G. and Poschmann, A. (2007) On the Classification of 4 Bit S-Boxes. *Arithmetic of Finite Fields*, Springer Berlin Heidelberg, 115-118.

[13] Todo, Y. and Morii, M. (2016) Bit-Based Division Property and Application to Simon, Family. *Fast Software Encryption*, Springer, Berlin, Heidelberg, 357-377. https://doi.org/10.1007/978-3-662-52993-5_18

[14] Xiang, Z., Zhang, W., Bao, Z., *et al*. (2016) Applying MILP Method to Searching Integral Distinguishers Based on Division Property for 6 Lightweight Block Ciphers. *International Conference on the Theory and Application of Cryptology and Information Security*, Springer, Berlin, Heidelberg, 648-678. https://doi.org/10.1007/978-3-662-53887-6_24

知网检索的两种方式：

1. 打开知网页面 http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD
   下拉列表框选择：[ISSN]，输入期刊 ISSN：2161-8801，即可查询
2. 打开知网首页 http://cnki.net/
   左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：csa@hanspub.org