Grøstl-512 积分区分器的改进

毛明[1]，秦志光[1]，李艳俊[2]

（1. 电子科技大学 计算机科学与工程学院，四川 成都 610054；2. 北京电子科技学院 信息安全系，北京 100070）

摘 要：校正了 CANS2010 会议上 Minier 等人关于 Grøstl 区分器的分析结果，改进了 Grøstl 算法中压缩函数的积分区分器，充分利用渗透技术首次提出了关于 P 函数和 Q 函数的 11 轮积分区分器。虽然针对散列函数的分析是目前 SHA3 研究的主流，但是所提出的关于积分区分器的研究反映了压缩函数的随机性，对新的散列函数的设计具有重要意义。

关键词：SHA3；散列算法；积分分析；区分器

中图分类号：TN918.4 **文献标识码**：A **文章编号**：1000-436X(2012)07-0022-05

Improved integral distinguisher of Grøstl-512

MAO Ming[1], QIN Zhi-guang[1], LI Yan-jun[2]

(1. School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China;

2. Department of Information Security Beijing Electronic Science and Technology Institute, Beijing 100070, China)

Abstract: Firstly, the distinguisher of Grøstl-512 proposed by Minier in CANS 2010 was corrected. Then, the integral distinguisher of Grøstl-512 compression function was improved. By using the saturation technique new 11-round integral distinguishers of P function and Q function were proposed. Whereas the SHA-3 competition focuses the attacks of hash functions, the proposed analysis on integral distinguish reflect the randomness of the compression function, which is of great significance to design new hash function.

Key words: SHA3; hash function; integral cryptanalysis; distinguisher

1 引言

近年来，由于传统的散列函数如 MD4、MD5、SHA0、SHA1、RIPENMD 不断被成功攻击，美国国家标准技术研究所（NIST）在 2005 年、2006 年分别举行了两届散列研讨会以探讨更加安全的散列函数。2007 年，NIST 正式发起在全球范围内征集新的下一代散列算法的公告，新的散列算法将被称为 SHA-3，以作为新的安全散列标准，来增强现有的 FIPS 180-2 标准。算法征集工作已于 2008 年 10 月结束，NIST 分别于 2009 年和 2010 年举行了 2 次散列候选算法筛选会议，选出了进入最后一轮(final round)的 5 个算法，它们分别是：BLAKE、Grøstl、JH、Keccak 和 Skein，最终获胜的算法将于 2012 年通过散列候选算法会议确定。Grøstl 作为进入第 3 轮的 5 个候选算法之一，由丹麦的 Kundsen 和奥地利的 Mendel 等人设计提交[1]，自公布以来，Grøstl 即受到了大量的分析攻击，其中最好的分析方法是对 Grøstl-256 的 7 轮压缩函数进行的半自由起始碰撞攻击，该攻击需要的计算复杂度为 2^{120}，存储需求为 2^{64}[2]；而针对 Grøstl-512 压缩函数半自由起始碰撞攻击可以进行到 8 轮，其计算复杂度和存储需求分别为 2^{152} 和 2^{64}[3]。

从目前对 Grøstl 的分析状况来看，随着候选算

收稿日期：2011-07-16；**修回日期**：2012-02-10

基金项目：国家自然科学基金资助项目（60973161）

Foundation Item: The National Natural Science Foundation of China (60973161)

法之间的激烈竞争，更多的专家学者都将目光聚焦于散列算法所采用的模型及其置换运算等方面。显然，随着对散列函数以及分组密码算法攻击分析工作的不断深入，可能出现新的针对分组密码算法的攻击模型，因为散列函数与分组密码算法之间只是已知密钥和未知密钥的区别[4,5]。事实上，当利用分组结构来构建散列函数的压缩函数时，已不需要密钥的输入，这样构建的分组不仅是一个已知变换，同时也是可计算的。分析这样的分组结构有助于发现新的区分器构建技术，或找到在未知密钥的分组算法中恢复密钥的新技术，如对全轮 AES-192/256 的相关密钥攻击[6,7]。对于 Grøstl- 512，设计者给出了 9 轮积分区分器[1]。之后，在 CANS2010 会议上，Minier 等人对 Grøstl-512 的压缩函数进行了积分分析[8]，并对 P 函数和 Q 函数提出了 10 轮的积分区分器，计算复杂度为 2^{513}，需要的存储空间也很小，但是经验证后发现该结论存在错误。

积分分析是除了差分分析和线性分析之外的另一种重要的分析方法，它的思想起源于 Square 分析，是 Daemen 等针对 Square 算法提出的一种分析方法[8]，该分析方法更多的与算法的结构有关，而与算法部件的具体取值几乎无关[9~13]。在总结 Square 分析工作的基础上，Knudsen 等在 FSE2002 会议上将上述分析方法全部归入积分分析（integral）范畴，并给出了积分分析和高阶积分分析的一般原理和方法[14]。用积分性质构建散列函数中压缩函数的例子也有很多[15~17]。

本文首先纠正了 Minier 等人关于 Grøstl-512 的错误结果，进一步对 P 函数和 Q 函数进行了更为深入的研究，采用渗透技术，改进了他们的分析方法，首次提出 P 函数和 Q 函数的 11 轮积分区分器。主要内容安排如下：第 2 节简要介绍 Grøstl 算法及其压缩函数，第 3 节修正了 Minier 等人提出的 10 轮区分器及其相关数据，第 4 节构建了压缩函数的积分区分器，第 5 节进行了总结。

2　预备知识

2.1　Grøstl 算法介绍

Grøstl 采用宽管道 MD 迭代结构，能抵抗常见的攻击(generic attack)方法，它的压缩函数主要由 2 个排列组成，排列基于 AES 构造，采用 SPN 结构，S 盒与 AES 的相同，列混合使用 GF(2^8)上的 8×8 循环矩阵。由于其采用了 AES 的宽轨迹(wide trail)设计结构，大大提高了抵抗差分攻击的能力。Grøstl 有 2 个版本：Grøstl-256，输出摘要长度为 224bit 或 256bit；Grøstl-512，输出摘要长度为 384bit 或 512bit。相比之下，后者使用的压缩函数（包含 P 和 Q 函数）扩散性较差，因此更容易受到攻击。Grøstl 将分块为 t 的消息 M（padding 之后）经过压缩函数 $f(H_{i-1},M_i)$ 和输出变换 $g(H_t)$ 得到摘要值，2 个函数运算过程如下：

$$H_0 = IV$$
$$H_i = f(H_{i-1},M_i) = H_{i-1} \oplus P(H_{i-1} \oplus M_i) \oplus Q(M_i),\ 1 \leqslant i \leqslant t$$
$$h = g(H_t) = trunc(H_t \oplus P(H_t))$$

2 个置换 P 和 Q 是基于宽轨道设计，非常类似于已知密钥的 AES 加密算法。Grøstl-512 中压缩函数的 1 024bit 输入可以排成 8×16 的字节矩阵，然后加密 14 轮后得到输出。压缩函数中的轮变换包含以下 4 个部分。

1) 常数加（AC, addround constant）：对 P 和 Q 中 8×16 矩阵的每个字节异或一个常数。

2) 字节替代（SB，subbytes）：状态中每一个字节经过 AES 的 S 盒查询得到输出。

3) 行移位（ShB，shiftbytes）：第 j 行的向左移位方式有 2 种：第 1 种是 P 中的移位，$j=1,2,\cdots,7$ 时移动 $j-1$ 个字节，$j=8$ 时移动 11 个字节；第 2 种是 Q 中的移位，$j=1,2,3$ 时移动 1,3,5 个字节，$j=5,6,7,8$ 时移动 0,2,4,6 个字节，$j=4$ 时移动 11 个字节。如图 1 所示。

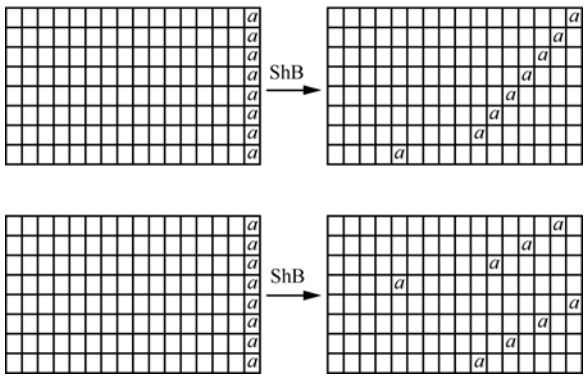

图 1　置换 P 和 Q 的行移位

4) 列混淆（MB, mixbytes）：对于状态中的每一列字节，用 GF(2^8)上的矩阵 $\boldsymbol{M}$ 进行乘法运算，矩阵 $\boldsymbol{M}$ 是 MDS 构成的，分支数为 9。

由于本文研究的积分性质不涉及具体的 S 盒

和矩阵 $\boldsymbol{M}$，所以详细的 Grøstl-512 描述可以参考文献[1]。

2.2　积分分析的几个概念

积分分析中主要包含以下几个概念。

活跃 A：如果某些比特/字节取遍所有可能值，而且出现次数相同，则称之为活跃比特/字节。

平衡 B：如果某些比特/字节的值之和为常数，则称之为平衡比特/字节。

常数 C：如果某些比特/字节出现的值始终不变，则称之为常数比特/字节。

本文中活跃字节 A 用阴影表示，平衡字节标识字母 B，空白表示常数字节 C。

3　文献[8]中的一个错误

在文献[8]中，图 7 的 2 轮变换是错误的，正确的扩散应该如图 2 所示。

如果解密方向的 5 轮区分器与加密方向的 5 轮区分器可以连接起来，那么中间状态的活跃字节集合就是这 2 个输入状态的并，如图 3 所示。因此，Minier 所提出的 10 轮区分器的时间复杂度应该是 2^{928}，而 P 和 Q 2 个函数的时间复杂度之和应为 2^{929}。相比之下，需要的存储空间仍然很小。

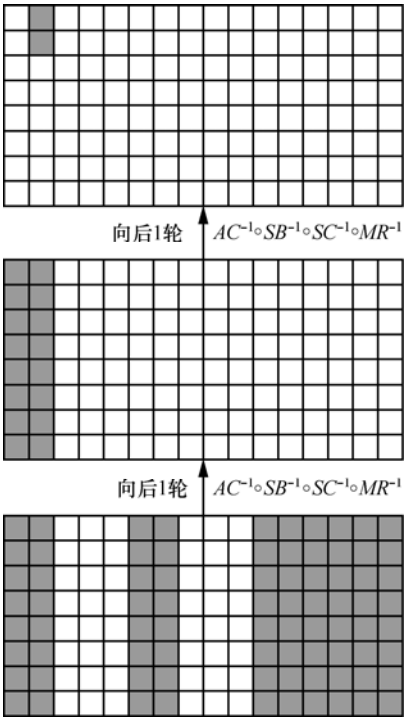

图 2　向后（解密方向）2 轮变换

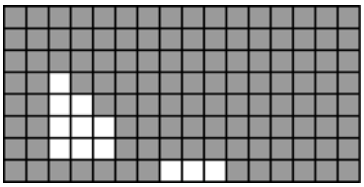

图 3　10 轮区分器的中间状态

4　关于 P（或 Q）函数的 11 轮积分区分器

构建 P 函数的 11 轮区分器分 2 个步骤完成，包括向后加密方向的 6 轮区分器和向前解密方向的 5 轮区分器。

4.1　加密方向的 6 轮区分器

选择一个字节活跃时，可以由加密过程得到。经过第 3 轮运算后，输出中每个字节都平衡，如图 4 所示，其中 a 表示一个活跃字节的扩散情况。

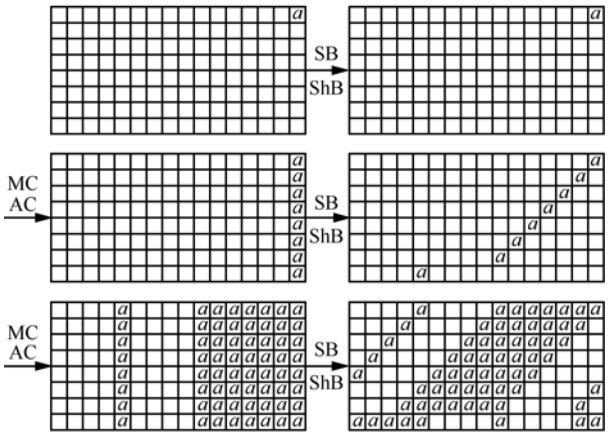

图 4　活跃字节 a 经过 2.5 轮加密

这里用 $b_{i,j}^{r}$ 表示第 r 轮 ShB 变换后输出第 i 行第 j 列的字节。仔细观察第 3 轮 ShB 的输出 $b_{i,j}^{3}$，发现每一列的活跃字节个数不同，最少为 2。如果第一轮 ShB 的最后一列有 2 个活跃字节，即 $b_{i,15}^{1}, 0\leqslant i\leqslant 7$ 中任意 2 个字节活跃，则第 3 轮 ShB 的输出状态中每一列的活跃字节两两独立。因此容易得出，如果输入的 2 个活跃字节位置合适，则经过第 3 轮 MC 之后，可以继续得到第 4 轮 ShB 之后的状态（如图 5 所示）。

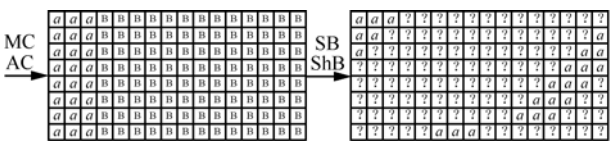

图 5　第 4 轮 ShB 之后状态

如果要使经过第 4 轮 MC 后某些字节平衡，那么就需要采取渗透技术[10]。假设要使得第 4 轮 ShB 后的第一列字节全部活跃，则第一轮 ShB 后的第 16 列、第 3 列、第 4（或 5 或 6）列，第 9（或 10 或 11）列，这 4 列的每列至少 4 个字节活跃。显然这样的组合会有很多，用组合知识计算共有 $(C_8^2)^2\times(C_3^1\times C_8^2)^2\approx 2^{22.4}$ 种。图 6 是其中的一种输入状态，选择这 8 个字节活跃，可以构成 8 阶 4 轮区分器。

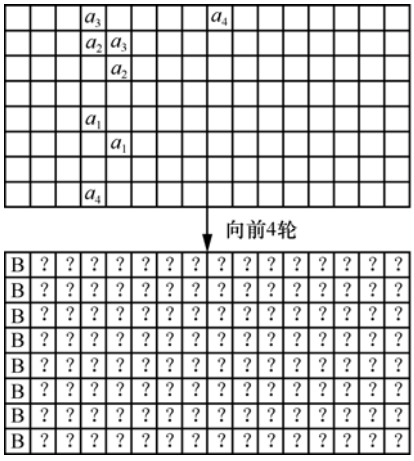

图 6　向前（加密方向）4 轮积分区分器

由此可以继续往前（解密方向）扩展一轮，得到 24 阶 5 轮积分区分器。由于输入的活跃字节在经过一轮加密之后，3 列中的每一列活跃字节已远远大于 4 轮区分器所需要的活跃字节，所以在第 5 轮输出中的平衡字节 B 也同样不止一列。根据 5 轮区分器的输入状态可以看出，积分区分器可以继续向前推进一轮，即如图 7 所示的 104 阶 6 轮积分区分器。

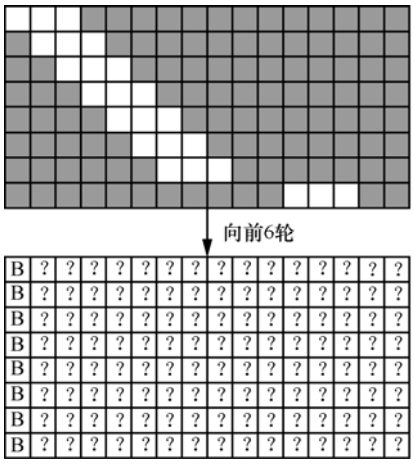

图 7　6 轮积分区分器

同样，这种 6 轮区分器还有很多，这里不再一一例举。

4.2　解密方向的 5 轮区分器

解密方向的 5 轮区分器仍然采用文献[8]提出的 5 轮区分器，但是对于活跃字节集合的选取需要慎重。本文将图 2 中的活跃字节按列移动位置，得到一个新的输入状态，使得其中的活跃字节与图 7 中输入活跃字节求并集后，活跃字节个数最少，如图 8 所示。

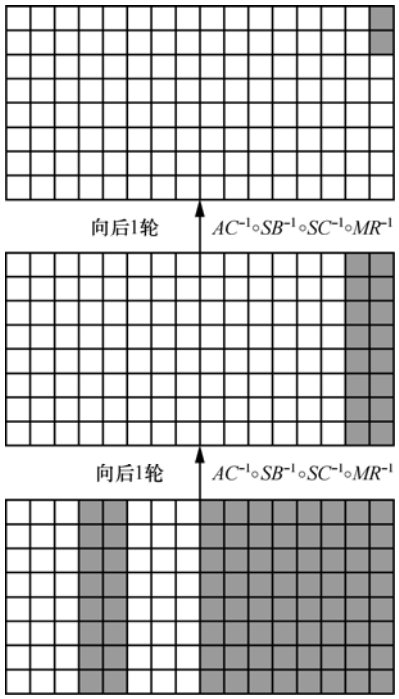

图 8　向后（解密方向）2 轮变换

这样，中间状态需要的活跃字节集合就为图 9 中的 114 个字节。

通过这个中间状态，可以连接 11 轮积分区分器，如图 9 所示。

5　结束语

本文分析了 SHA-3 候选算法 Grøstl-512 压缩函数的积分性质，校正了文献[8]中 10 轮积分区分器的时间复杂度；进一步针对轮函数扩散性较差的特点，采用渗透技术构建了 Grøstl-512 的 11 轮积分区分器，该积分区分器的时间复杂度小于文献[8]中 10 轮区分器的时间复杂度，其结果比较如表 1 所示。显然，新的 11 轮积分区分器的构建更加体现了渗透技术的充分利用。

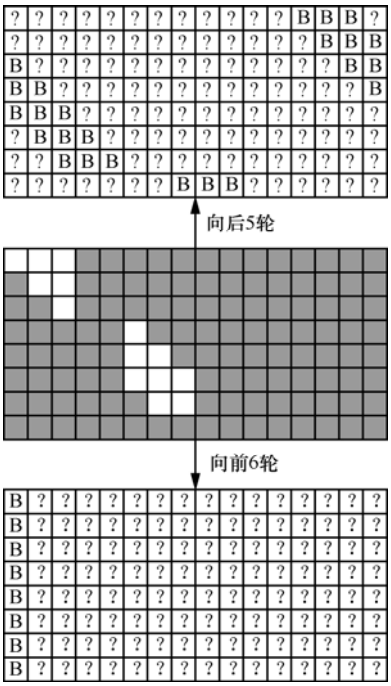

图 9 新的 11 轮积分区分器

表 1　关于 Grøstl-512 积分区分器的总结

出处	区分器轮数	时间复杂度	存储
文献[1]	9	2^{704}	小
文献[8]*	10	2^{929}	小
本文	11	2^{913}	小

*表示改正后的结果。

参考文献：

[1] GAURAVARAM P, KNUDSEN L R, MATUSIEWICZ K, *et al*. Grøstl-a sha-3 candidate[EB/OL]. http://ehash.iaik.tugraz.at/wiki/The SHA-3 Zoo.

[2] GILBERT H, PEYRIN T. Super-sbox cryptanalysis: improved attacks for aes-like permutations[EB/OL]. http://eprint.iacr.org/,2009.

[3] MENDEL F, RECHBERGER C, SCHLAFFER M, *et al*. Rebound attacks on the reduced Grøstl hash function[A]. CT-RSA 2010[C]. Springer, Heidelberg,2010. 350-365.

[4] KNUDSEN L R, RIJMEN V. Known-key distinguishers for some block ciphers[A]. ASIACRYPT 2007[C]. Springer, Heidelberg,2007. 315-324.

[5] MINIER M, PHAN R C W, POUSSE B. Distinguishers for ciphers and known key attack against Rijndael with large blocks[A]. AFRICACRYPT 2009[C]. Springer, Heidelberg, 2009. 60-76.

[6] BIRYUKOV A, KHOVRATOVICH D. Related-key cryptanalysis of the full AES-192 and AES-256[A]. ASIACRYPT 2009[C]. Springer, Heidelberg, 2009.1-8.

[7] BIRYUKOV A, KHOVRATOVICH D, NIKOLIC I. Distinguisher and related-key attack on the full AES-256[A]. CRYPTO 2009[C]. Springer, Heidelberg, 2009. 231-249.

[8] MINIER M, PHAN R C W, POUSSE B. Integral distinguishers of some SHA-3 candidates[A]. CANS 2010[C]. Springer, Heidelberg,2010. 106-123.

[9] DAEMEN J, KNUDSEN L, RIJMEN V. The block cipher Square[A]. FSE 1997[C]. Springer, Heidelberg, 1997. 149-165.

[10] GALICE S, MINIER M. Improving integral attacks against Rijndael-256 up to 9 rounds[A]. AFRICACRYPT 2008[C]. Springer, Heidelberg,2008. 1-15.

[11] COLLARD B, STANDAERT F X. A statistical saturation attack against the block cipher present[A]. CT-RSA 2009[C]. Springer, Heidelberg, 2009. 195-210.

[12] 王薇, 王小云. 对 CLEFIA 算法的饱和度分析[J]. 通信学报, 2008, 29(10): 88-92.
WANG W, WANG X Y. Saturation cryptanalysis of CLEFIA[J]. Journal on Communications, 2008, 29(10): 88-92.

[13] 孙兵, 李瑞林, 屈龙江等. 对低代数次数分组密码的 SQUARE 攻击[J]. 中国科学: 信息科学, 2010, 40(6): 777-785.
SUN B, LI R L, QU L J, *et al*. SQUARE attack on block ciphers with low algebraic degree[J]. Science China: Information Science, 2010, 40(6): 777-785.

[14] KNUDSEN L, WAGNER D. Integral cryptanalysis[A]. FSE 2002[C]. Springer, Heidelberg, 2002.112-127.

[15] AUMASSON J P, KASPER E, KNUDSEN L R, *et al*. Distinguishers for the compression function and output transformation of hamsi-256[EB/OL]. http://ehash.iaik.tugraz.at/wiki/The SHA-3 Zoo.

[16] BOURA C, CANTEAUT A. A zero-sum property for the keccak-f permutation with 18 rounds[EB/OL]. http://ehash.iaik.tugraz.at/wiki/The SHA-3 Zoo.

[17] CALIK C, TURAN M S. Message Recovery and Pseudo-preimage Attacks on the Compression Function of Hamsi-256[R]. Cryptology ePrint Archive, 2010.

作者简介：

毛明（1963-），男，山西稷山人，电子科技大学博士生，主要研究方向为密码学和信息安全等。

秦志光（1956-），男，四川隆昌人，博士，电子科技大学教授、博士生导师，主要研究方向为网络安全性、移动社会网等。

李艳俊（1979-），女，山西阳城人，博士，北京电子科技学院讲师，主要研究方向为密码学。