

文章编号:1001-9081(2010)08-2154-03

BLAKE-32 的自由起始原象攻击

贺 强¹,毛 明²,曾绍昆¹

(1. 西安电子科技大学 通信工程学院, 西安 710071; 2. 北京电子科技学院 信息安全系, 北京 100070)
(heqiang515@126.com)

摘要:SHA-3 第二轮候选算法 BLAKE 采用局部宽管道技术和改进的 MD 迭代结构, 其内核为 Chacha 密码算法的内核, 该算法的安全性还未得到证明。通过分析 BLAKE 算法的结构及其消息置换特征, 首次采用分段—连接技术对其进行了 3 轮的自由起始原象攻击。结果表明, 消息置换的设计存在缺陷, 而且这一设计缺陷影响了 BLAKE 算法的安全性。

关键词:消息摘要;安全性;分段—连接;自由起始原象攻击

中图分类号: TP309 文献标志码:A

Free-starting preimage attacks on BLAKE-32

HE Qiang¹, MAO Ming², ZENG Shao-kun¹

(1. College of Communication Engineering, Xidian University, Xi'an Shaanxi 710071, China;
2. Department of Information Security, Beijing Electronic Science and Technology Institute, Beijing 100070, China)

Abstract: As one of SHA-3 candidate algorithms for the second round competition, BLAKE adopts local wide-pipe technology and improved MD iteration structure. Its core is the core of Chacha cipher algorithm and its security has not been proved. By analyzing the structure and the characteristics of message permutation, three rounds free-starting preimage attack could be applied to BLAKE by using splice-and-subsection technology. The result shows that the designing deficiency of message permutation affects the security of BLAKE algorithm.

Key words: message digest; security; splice-and-subsection; free-starting preimage attack

0 引言

消息摘要函数 BLAKE^[1]作为消息摘要函数标准 SHA-3 的第二轮候选算法。BLAKE 算法继承了 LKE^[2]的算法思想和 HAIFA^[3]的迭代模式, 其内核采用 Chacha^[4]密码算法的内核, 算法内部采用宽通道模式, 是一种广义 MD 结构的哈希函数。根据消息摘要值长度的不同, 可将 BLAKE 算法系列分为 BLAKE-28、BLAKE-32、BLAKE-48、BLKAE-64, 分别对应的属性如表 1。

表 1 BLAKE 函数属性

算法	字长	消息长度	分组长度	摘要值长度	盐值长度
BLKAE-28	32	<2 ⁶⁴	512	224	128
BLAKE-32	32	<2 ⁶⁴	512	256	128
BLKAE-64	64	<2 ¹²⁸	1024	512	256

SHA-3 标准的第二轮候选算法共有 14 个, 根据各自结构大致可分为 3 类: 改进 MD 结构、流结构、海绵结构。其中有 9 个为改进 MD 结构或其变体, 由此可以看出, 消息摘要函数中广泛采用的 MD 结构对新标准哈希算法的设计产生了深远的影响, 仍然受到广泛认可。BLAKE 算法具有典型的扩展 MD 结构, 与传统哈希函数相比, 它在运算速度和安全性方面都具有很大优势:

1) 运算速度快。BLAKE 算法引入宽通道模式, 大大提高了算法执行速度, 虽然新标准算法安全性要求比以往要高, 而安全性高可能导致执行速度下降, 但 BLAKE 函数的运算速度

并没有明显降低, 相反, 与采用 MD 结构的 SHA-512 相比, 其运算速度比 SHA-512 要快得多。

2) 安全性高。BLAKE 算法中所采用的改进 MD 迭代结构抗第二原象攻击, 局部宽通道技术抗碰撞攻击。BLAKE 压缩函数的内核为 Chacha 密码内核, 而 Chacha 密码经过多次分析^[5]已确信具有很好的安全性。各轮参与运算的明文消息字顺序均不同且没有规律, 使得对其安全性的分析更为困难。因此, BLAKE 算法具有很高的安全性。

到目前为止, 还没有发现对 BLAKE 算法构成实质性威胁的攻击。Li 等人^[6]进行了 2.5 轮计算复杂度为 2²⁴ 的攻击, Guo 和 Matusiewicz^[7]得到了 4 轮的近似碰撞, 但并未公布其具体的分析方法和近似碰撞数据。他们的分析对 BLAKE 算法的安全性都不会产生实质的影响。

本文将研究 BLAKE 算法家族中具有代表性的 BLAKE-32^[1]的安全性。针对消息置换的设计缺陷, 利用分段—连接技术对算法进行了 3 轮的自由起始原象攻击。分析表明, 此攻击方法理论上可以降低其计算复杂度。

1 BLAKE-32 算法原理及置换特征

1.1 算法原理

BLAKE-32 有 4 个输入参数: 初始变量 $h = h_0^0, \dots, h_7^0$, 消息字 $m = m_0, \dots, m_{15}$, 盐值 $s = s_0, \dots, s_3$ 以及计数值 $t = t_0, t_1$ 。压缩函数记为 $h' = \text{compress}(h^0, m, s, t)$ 。此外, 还有一组常数参与轮函数运算, 分别用 c_0, \dots, c_{15} 表示。

BLAKE-32 算法主要分 3 步: 变量初始化、变量运算、终值

收稿日期:2010-02-01;修回日期:2010-03-02。

作者简介: 贺强(1982-), 男, 贵州思南人, 硕士研究生, 主要研究方向: 信息安全; 毛明(1963-), 男, 山西稷山人, 教授, 主要研究方向: 信息安全、密码学; 曾绍昆(1986-), 男, 江西吉安人, 硕士研究生, 主要研究方向: 信息安全。

压缩。

1) 初始化。

变量初始值 v_i ($0 \leq i \leq 15$) 通过下列方程:

$$\begin{bmatrix} v_0 & v_1 & v_2 & v_3 \\ v_4 & v_5 & v_6 & v_7 \\ v_8 & v_9 & v_{10} & v_{11} \\ v_{12} & v_{13} & v_{14} & v_{15} \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 & h_7 \\ s_0 \oplus c_0 & s_1 \oplus c_1 & s_2 \oplus c_2 & s_3 \oplus c_3 \\ t_0 \oplus c_4 & t_1 \oplus c_5 & t_0 \oplus c_6 & t_1 \oplus c_7 \end{bmatrix}$$

得到,其中等号右边的参数值为已知常数值或随机值。

2) 变量运算。

经过初始化后的 16 个变量作为 G 函数的输入参数,每个 G 函数有 4 个变量参与运算,前 4 个 G 函数各自的输入变量互不相同,因此可并行计算,其输出值作为后 4 个同样可以并行运算的 G 函数的输入变量。 G 函数及其参数分别表示为:

$$\begin{aligned} G_0(v_0, v_4, v_8, v_{12}), G_1(v_1, v_5, v_9, v_{13}), G_2(v_2, v_6, v_{10}, v_{14}), \\ G_3(v_3, v_7, v_{11}, v_{15}), G_4(v_0, v_5, v_{10}, v_{15}), G_5(v_1, v_6, v_{11}, v_{12}), \\ G_6(v_2, v_7, v_8, v_{13}), G_7(v_3, v_4, v_9, v_{14}) \end{aligned}$$

$G_i(a, b, c, d)$ 函数的运算规则如下:

$$\begin{aligned} a \leftarrow a + b + (m_s \oplus c_j) \\ d \leftarrow (d \oplus a) \gg 16 \end{aligned}$$

```

c ← c + d
b ← (b ⊕ c) >>> 12
a ← a + b + (mj ⊕ cs)
d ← (d ⊕ a) >>> 8
c ← c + d
b ← (b ⊕ c) >>> 7

```

其中: $s = \sigma_r(2i)$, $j = \sigma_r(2i+1)$ ($0 \leq i \leq 7$) 分别表示 16 个明文或常量中的第 s 、 j 个字。 r 表示第 r ($0 \leq r \leq 9$) 轮, 具体置换顺序见表 2 所示。

3) 终值压缩

G 函数经过各轮变换运算后,得到 16 个变量 v_0, \dots, v_{15} , 将变量值与盐值 s 以及初始值 h 按如下方式作异或运算,即得到最终消息摘要值:

$$h'_i = h_i^0 \oplus v_i \oplus s_j \oplus v_{i+8}$$

其中: $0 \leq i \leq 7, j = i \pmod{4}$, h'_i 表示第 r 轮摘要值的第 i 个字。

1.2 消息置换特征

BLAKE-32 共执行 10 轮的消息置换,在 1.1 节中,消息字 m_s 及常量 c_j 在不同的轮运算中分别运用不同的置换 $s = \sigma_r(2i), s = \sigma_r(2i+1)$, 置换规则见表 2 所示。

表 2 置换规则表

置换	G_0		G_1		G_2		G_3		G_4		G_5		G_6		G_7	
σ_0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
σ_1	14	10	4	8	9	15	13	6	1	12	0	2	11	7	5	3
σ_2	11	8	12	0	5	2	15	13	10	14	3	6	7	1	9	4
σ_3	7	9	3	1	13	12	11	14	2	6	5	10	4	0	15	8
σ_4	9	0	5	7	2	4	10	15	14	1	11	12	6	7	3	13
σ_5	2	12	6	10	0	11	8	3	4	13	7	5	15	14	1	9
σ_6	12	5	1	15	14	13	4	10	0	7	6	3	9	2	8	11
σ_7	13	11	7	14	12	1	3	9	5	0	15	4	8	6	2	10
σ_8	6	15	14	9	11	3	0	8	12	2	13	7	1	4	10	5
σ_9	10	2	8	4	7	6	1	5	15	11	9	14	3	12	13	0

分析 BLAKE 算法原理并观察各轮运算的明文置换规则,不难发现:在任意连续的两轮运算中,如果同时修改前一轮的某个输入变量 v_i ($0 \leq i \leq 3$) 及该步运算中相应的明文 $m_{\sigma_r(2i)}$,那么在该步运算结束后,可以控制其临时结果变量没有差分,并且该轮运算结束后,16 个结果变量不会产生差分。而在下一轮运算中,存在差分的该明文只出现在后半轮 G_{i+4} 函数的运算中,并且只能使函数 G_{i+4} 中的所有变量产生差分,而其余 G 函数生成的变量值不存在差分。因此,如果不考虑盐值,我们可以根据这一消息置换特征,通过不固定初始状态值,控制输出差分变量,找到生成摘要值的原象并降低 BLAKE 算法的计算复杂度。

2 三轮 BLAKE-32 的自由起始原象攻击

2.1 攻击思想

针对 BLAKE 函数的安全性特点,它有抗第二原象攻击和抗碰撞攻击的能力。本文针对 1.2 节中消息置换的特点,尝试对该消息摘要函数采用一种新的分析方法——自由起始原象攻击。

自由起始原象攻击是在不固定初始状态值的前提下,对于任意给定的摘要值 y , 寻找生成该摘要值 y 的消息摘要函数 h 的消息 x 的一种攻击方法。

为了便于分析,在不固定初始状态值的情况下,我们仅研究只有 3 轮运算的 BLAKE-32,在已知其消息摘要值 y 的情况下如何寻找消息 x 的情况,即寻找自由起始状态下的原象。

在 3 轮的 BLAKE-32 算法中,把 3 轮运算分割成两部分,第一部分为第一轮,第二部分为后两轮。给定消息摘要值 h_i^t ($0 \leq i \leq 7$), 并随机生成一组明文, 计算得到一个新的摘要值 h_i^3 , 并存储第一轮运算结果值 v_k ($0 \leq k \leq 15$); 然后从第二部分开始, 把 v_k 作为其初始状态值, 根据 1.1 节中 BLAKE 终值计算规则的特点和 1.2 节中的消息特征分析, 利用方程 $h_i^t = h_i^3 \oplus v_k \oplus v'_{k'} (k = k')$, 计算得到一个有差分的初始值与给定摘要值和新摘要值以及初始状态值之间的等式关系; 在保持该步运算结果不变的情况下, 根据 $v'_{k'}$ 修改明文并重新计算得到一个修改明文后的摘要值 $h_i^{3'}$, 控制第三轮运算结束后的变量 v_i^3, v_{i+8}^3 不产生差分, 则 $h_i^t = h_i^{3'}$, 并找到自由起始的原象消息。对于第一部分即第一轮, 根据前面修改后的变量值 v' 和明文值, 通过修改某些特定变量值进行反推, 得到初始状态变量值 $h_i^{0'}$, 并使得 $h_i^{0'} = v'_{i'}$ 。最后把两部分连接起来, 完成一次完整分析过程。

我们对上述两部分分别进行分析,然后把两部分连接起来从而完成对整体攻击的这种技术称为分段—连接技术。根据这个技术,可以提高找到对应原象的概率。对于 512 比特

的明文,假设自由起始原象攻击的计算复杂度为 $2^x (x < 254)$, 利用分段—连接技术进行原象攻击的计算复杂度^[8]为 $2^{\lceil \frac{256-x}{2} \rceil + 1}$, 大大低于直接进行原象攻击的计算复杂度。

2.2 攻击步骤

如果不计盐值和计数值的影响,利用 1.2 节分析得到的消息置换特征和 2.1 节的分析思路,对于明文分组长度为 512 比特的 BLAKE-32 算法,对三轮运算进行自由起始原象攻击,步骤如下:

- 1) 随机设置三轮的 hash 值 $h^t = h_0^t, \dots, h_7^t$ 。
- 2) 随机产生消息字 $m = m_0, m_1, \dots, m_{15}$ 。
- 3) 计算三轮运算的摘要值 $h^3 = h_0^3, \dots, h_7^3$, 并存储该运算过程中在第一轮运算后生成的链路变量 $v^1 = v_0^1, v_1^1, \dots, v_{15}^1$ 及明文消息 m_0, m_1, \dots, m_{15} 。
- 4) 根据公式 $v_2^{1'} = h_2^t \oplus h_2^3 \oplus h_2^0$ 计算 $v_2^{1'}$ 。
- 5) 根据第二轮 G_0 函数算法及方程 $v_2^1 + v_6^1 + (m_9 \oplus c_{15}) = v_2^{1'} + v_6^1 + (m_9' \oplus c_{15})$, 反推得到 $m_9' = (v_2^1 + (m_9 \oplus c_{15}) - v_2^{1'}) \oplus c_{15}$, 将 m_9 修改为 m_9' 。
- 6) 固定变量值 $v_0^1, v_1^1, v_2^{1'}, \dots, v_{15}^1$ 和消息字 $m_0, m_1, \dots, m_8, m_9', m_{10}, \dots, m_{15}$, 设置 $h_2^{0'} = v_2^{1'}$, 倒推第一轮得到初始变量值。

由于要保证 m_9 修改为 m_9' , 其余消息字不变, 第一轮运算后, $v_2^{1'}$ 需满足条件 $v_2^{1'} = h_2^{0'} = h_2^t \oplus h_2^3 \oplus h_2^0$, 并且生成的其余 15 个变量没有差分, 必须修改初始状态值: $v^0 = v_0^0, \dots, v_{16}^0$ 。通过对前 0.5 轮初始值的修改, 消除由 m_9' 引起的差分。

a) 观察第一轮 G_4 函数内部运算法则, 利用方程 $v_0 + v_5 + (m_9 \oplus c_8) = v_0' + v_5 + (m_9' \oplus c_8)$ 得到 $v_0' = v_0 - (m_9 \oplus c_8) - (m_9' \oplus c_8)$, 消除该步链路变量差分;

b) 而 $v_0' = v_0 + v_5 + (m_8 \oplus c_9)$, 为了得到 v_0' , 修改 v_0 , 根据前一步的公式, 前 0.5 轮中 G_0 函数的输出为: $v_0 \leftarrow v_0 - v_3 - (m_9 \oplus c_8) - (m_9' \oplus c_8) - (m_8 \oplus c_9)$, 其中 v_5 为 G_1 函数的输出, 该临时变量值可通过第一轮后的结果变量逆向推导得到, 此处略, 同时产生有差分的 v_{15}' ;

c) 根据 $G_i(a, b, c, d)$ 算法进行逆向运算, 对于 G_0 函数: $v_0 \leftarrow v_0 - v_4 - (m_1 \oplus c_0)$, 而 $v_0 = h_0^0 + h_4^0 + (m_0 + c_1)$, 因此修改 h_0^0 , 使 $h_0^{0'} = v_0 - v_4 - (m_1 \oplus c_0) - h_4^0 - (m_0 \oplus c_1)$;

d) v_{15}' 由 G_3 函数产生的 v_{15} 与 v_0' 异或产生, 而 v_0' 不能再更改, 因此, 可通过方程 $v_0' \oplus v_{15}' = v_0 \oplus v_{15}$ 得: $v_{15}' = v_0' \oplus v_0 \oplus v_{15}$, 依此进行逆推运算, 可以通过修改初始值 v_3^0 和 v_{15}^0 来满足要求;

e) 把 G_2 函数中的初始状态 h_2^0 修改为 $h_2^{0'}$, 使 $h_2^{0'} = v_2^{1'}$ 。根据变量 $v_2^{1'}$ 的值进行逆向推导, 最后得到一个关于 G_2 函数第一个变量的中间临时状态值, 记为 t , 保存该变量; 设置 $h_2^{0'}$, 计算得到一个中间临时状态值, 记为 t' , 根据两个临时状态差分值修改初始状态值 h_6^0, v_{10}, v_{14} 。利用类似的方法进行逆向推导可设置其他初始状态值。

(上接第 2100 页)

- [8] WEN C H, ZHANG X Y, TAN T S. Generating an ω -tile set for texture synthesis[C]// Proceedings of the 23rd Computer Graphics International. New York: ACM, 2005: 177–184.
- [9] XUE F, ZHANG Y S, JIANG J L, et al. Real-time texture synthesis using s-tile set[J]. Journal of Computer Science and Technology, 2007, 22(4): 590–596.
- [10] WEI L Y, HAN J Y, ZHOU K, et al. Inverse texture synthesis

7) 使用连接技术, 把第一部分与第二部分在 $v_2^{1'}$ 满足 $v_2^{1'} = h_2^0$ 时连接起来, 执行一次完整的运算, 则得到 $h_2^t = h_2^{3'}$ 。

重复尝试以上步骤, 修改初始状态值, 直到找到所有自由起始状态下的原象, 其计算复杂度为 2^{224} 。本文只作了 3 轮的分析, 利用类似分析思想, 同样能够实现对 BLAKE-32 进行 10 轮的攻击。

3 结语

文章针对三轮 BLAKE-32 消息摘要函数为研究对象从第二轮开始作自由起始原象攻击并获得一个随机明文, 在不考虑盐值和计数值影响的情况下, 根据 BLAKE 算法结构的特点, 通过任意连续的两轮运算得到数据并能够进行逆向运算, 得到自由起始状态下的原象。虽然文中的分析结果不是真正的原象, 但根据前面的分析过程, 能够以 2^{241} 的计算复杂度找到给定摘要值的原象。此攻击过程的重点在于第二部分初始状态差分的设定, 以及使算法在第一部分运算后的结果状态值与其对应的初始状态值相等。虽然三轮的自由起始原象攻击还不足以威胁该算法的安全, 但进一步将详细对完整的 10 轮运算进行研究, 降低其计算复杂度, 那么 BLAKE 算法在理论上就不是一个安全的算法。

参考文献:

- [1] AUMASSON J P, HENZEN L, MEIER W, et al. SHA-3 proposal blake[EB/OL]. [2009-08-11]. <http://131002.net/blake/blake.pdf>.
- [2] AUMASSON J P, MEIER W, RAPHAEL C W. The hash function family LAKE[C]// FSE 2008: Proceedings of 15th Fellow of the Society of Engineers, LNCS 5086. Berlin: Springer-Verlag, 2008: 36–53.
- [3] BIHAM E, DUNKELMAN O. A framework for iterative hash functions — HAIFA[EB/OL]. [2009-08-25]. http://csrc.nist.gov/groups/ST/hash/documents/DUNKELMAN_NIST3.pdf.
- [4] BERNSTEIN D J. Chacha, a variant of salsa20[EB/OL]. [2009-09-06]. <http://cr.yp.to/chacha.html>.
- [5] HENZEN L, CARBOGNANI F, FELBER N, et al. VLSI hardware evaluation of the stream ciphers Salsa20 and ChaCha, and the compression function Rumba[C]// SCS' 08: 2th IEEE International Conference on Signals, Circuits and Systems. Washington, DC: IEEE, 2008: 7–9.
- [6] LI J, XU L Y. Attacks on round-reduced BLAKE[EB/OL]. [2009-11-08]. http://ehash.iaik.tugraz.at/wiki/The_SHA-3_Zoo.
- [7] GUO J, MATUSIEWICZ K. Round-reduced near-collisions of BLAKE-32[EB/OL]. [2009-11-05]. http://ehash.iaik.tugraz.at/wiki/The_SHA-3_Zoo.
- [8] SASAKI Y, WANG L, AOKI K. Preimage attacks on 41-step SHA-256 and 46-step SHA-512[EB/OL]. [2009-12-15]. <http://eprint.iacr.org/2009/479.pdf>.

[J]. ACM Transactions on Graphics, 2008, 27(3): 1–9.

[11] 林定, 陈崇成, 唐丽玉, 等. 基于 Image Quilting 算法的纹理合成[J]. 系统仿真学报, 2008, 20(Z1): 381–384.

[12] SHEVELEV I A, KAMENKOVICH V M, SHARAEV G A. The role of lines and corners of geometric figures in recognition performance [J]. Acta Neurobiologiae Experimentalis, 2005, 63(4): 361–368.