Information Processing Letters 113 (2013) 764–770

Contents lists available at SciVerse ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Second order collision for the 42-step reduced DHA-256 hash function

Riham AlTawy, Aleksandar Kircanski, Amr Youssef*

Concordia Institute for Information Systems Engineering (CIISE), Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, H3G 1M8, Canada

ARTICLE INFO

Article history:
Received 29 April 2013
Received in revised form 16 July 2013
Accepted 16 July 2013
Available online 23 July 2013
Communicated by D. Pointcheval

Keywords:
Cryptography
Cryptanalysis
Hash functions
Second order collisions
Boomerang attack
DHA-256

ABSTRACT

At the Cryptographic Hash Workshop hosted by NIST in 2005, Lee et al. proposed the DHA-256 (Double Hash Algorithm-256) hash function. The design of DHA-256 builds upon the design of SHA-256, but introduces additional strengthening features such as optimizing the message expansion and step function against local collision attacks. Previously, DHA-256 was analyzed by J. Zhong and X. Lai, who presented a preimage attack on 35 steps of the compression function with complexity $2^{239.6}$. In addition, the IAIK Krypto Group provided evidence that there exists a 9-step local collision for the DHA-256 compression function with probability higher than previously predicted. In this paper, we analyze DHA-256 in the context of higher order differential attacks. In particular, we provide a practical distinguisher for 42 out of 64 steps and give an example of a colliding quartet to validate our results.

1. Introduction

In 2012, Keccak has been selected as the winner of the NIST hash function competition [5]. However, according to the eBASH project (ECRYPT Benchmarking of All Submitted Hashes),[1] SHA-2 outperforms Keccak on a significant number of widely used general purpose CPUs. Thus, the designs based on the SHA-2 hash may remain a compelling choice in the area of software oriented applications. On the side of attacks on the hash function following the SHA design approach, apart from the seminal attack [20] on SHA-1 by Wang, some progress has been reported in analysis of SHA-2: collisions and pseudo-collisions for variants of SHA-2 reduced to 32 and 38-step have been reported [15,14]. One of the early proposals that aim to strengthen the SHA-2 design is the DHA-256 hash function. The main innovation applied in the design of DHA-256 is tweaking the message expansion and the step function so that the possibility of repeating the patterns that potentially cause local collisions is minimal. Another change is increasing the number of modular additions in one step to 8. In this work, we revisit the security of DHA-256 hash function, assessing its security against second order differential distinguishers. Distinguishing attacks are important in the context of hash functions since compression functions are often modeled as random oracles and the potential existence of distinguishing properties can be used for disproving indifferentiability claims [6]. Furthermore, hash functions are modeled as random oracles in many schemes, e.g. the Optimal Asymmetric Encryption Padding and Probabilistic Signature Scheme [1] and the security of these schemes depends on the randomness of the underlying primitives [4]. Therefore, studying the distinguishing properties of compression functions is of importance from the perspective of proving the overall security of a particular system. Previous literature related to the higher order distinguishers on hash functions includes the pioneering works on higher order differentials by Lai [9] and the boomerang attacks by Wagner [19], both originally proposed for block ciphers. These two works have

* Corresponding author.
E-mail address: youssef@ciise.concordia.ca (A. Youssef).

[1] eBASH website: http://bench.cr.yp.to/ebash.html.

0020-0190/$ – see front matter
http://dx.doi.org/10.1016/j.ipl.2013.07.015

been adapted to hash functions independently by Biryukov et al. [3] and Lamberger and Mendel [10]. In particular, in [3], a distinguisher for the 7-round BLAKE-32 was provided, while in [10] a practical distinguisher for the 46-step SHA-2 compression function was given. The latter SHA-2 result was extended to 47 steps in [2]. Subsequently, boomerang distinguishers have been applied to other SHA-based functions. SIMD-512 compression function was analyzed by Mendel and Nad [13] and a boomerang distinguisher of theoretical complexity for the full function was presented. Sasaki et al. [17,18] provided a practical boomerang distinguisher on the full compression function of 5-pass HAVAL and also a distinguisher for the full HAS-160 compression function. Kircanski et al. [8] provided a practical zero sum for the 33-step SM3 compression function. In [12], Leurent and Roy showed that, under some conditions, three independent paths instead of two can be combined to achieve a better distinguishing complexity in a boomerang attack.

The DHA-256 hash function was first analyzed by the IAIK Krypto Group [7] where a local collision differential pattern was shown to exist with probability 2^{-63}, which is higher than it was anticipated by the designers. Later, Zhong and Lai [21] studied preimage resistance of DHA-256 using variants of meet in the middle attacks, reaching the complexity of $2^{239.6}$ function evaluations for a 35-step pseudo-preimage attack and $2^{248.8}$ for the preimage attack.

In this paper, we provide a practical distinguisher for the 42-step-reduced DHA-256 compression function. In particular, we present two independent differential characteristics and state the conditions imposed on each step during the quartet computation. The two provided paths consist of the internal and external parts: the approach used to satisfy the internal parts of the characteristics is to enforce one path by using message modification on the two faces of the boomerang, while simultaneously verifying whether the other path is satisfied on the other two faces of the boomerang. The external parts of the paths are satisfied probabilistically. Apart from satisfying the inner state differentials, by properly choosing the message words, we maximize the probability of expanding the messages following the message difference pattern, taking into account the non-linearity of the message expansion in DHA-256.

The paper is organized as follows. In the next section, the specification of the DHA-256 function along with the notation used throughout the paper is provided. A brief overview of second order differential attacks is provided in Section 3. Afterwards, we provide detailed description of our attack, differential characteristics, and the attack's complexity in Section 4. Finally, the paper is concluded in Section 5.

2. Specification of DHA-256

DHA-256 [11] is a Merkle–Damgård based hash function. Its compression function maintains a state of eight 32-bit words, produces 256-bit digests and takes 512-bit message blocks on the input. The input message is expanded according to the message expansion function. Each expanded message word is used twice in each step which increases the message diffusion rate. The compression function inner state consists of eight 32-bit words each step. See Fig. 1.

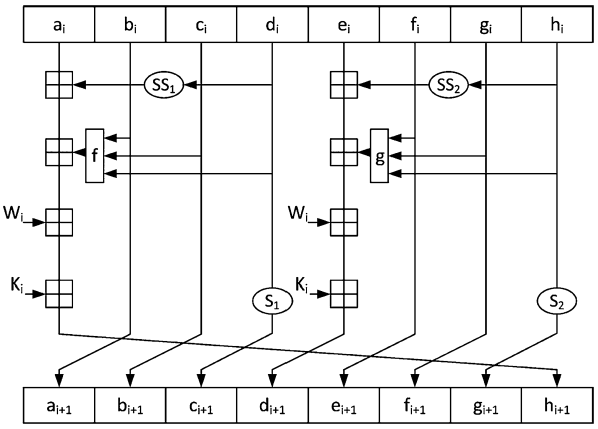

Fig. 1. The step function of DHA-256.

2.1. The state update function

The functions start by initializing the chaining variables $(a_0, b_0, c_0, \ldots, h_0)$ with eight 32-bit words initial values (IVs) and updates them iteratively for 64 steps using the following operations

$$h_{i+1} = a_i + SS_1(d_i) + f(b_i, c_i, d_i) + W_i + K_i$$
$$d_{i+1} = e_i + SS_2(d_i) + g(f_i, g_i, h_i) + W_i + K_i$$
$$b_{i+1} = S_1(c_i), \qquad f_{i+1} = S_2(g_i), \qquad a_{i+1} = b_i$$
$$c_{i+1} = d_i, \qquad e_{i+1} = f_i, \qquad g_{i+1} = h_i$$

The auxiliary functions, SS_1 and SS_2, both operating on 32-bit words, and are defined by

$$SS_1(X) = X \oplus (X \lll 11) \oplus (X \lll 25)$$
$$SS_2(X) = X \oplus (X \lll 19) \oplus (X \lll 29)$$

The rotations S_1 and S_2, are defined by

$$S_1(X) = X \lll 17, \qquad S_2(X) = X \lll 2$$

The Boolean functions f and g are defined by

$$f(X, Y, Z) = (X \wedge Y) \vee (\neg X \wedge Z)$$
$$g(X, Y, Z) = (X \wedge Y) \oplus (Y \wedge Z) \oplus (X \wedge Z)$$

The step constants K_i and the IV values are available in the full function specification [11].

2.2. The message expansion

Each 512-bit message block M is split into 16 words $M_0 \| M_1 \| \cdots \| M_{15}$ and then it is expanded to 64 32-bit words $(W_0, W_1, \ldots, W_{63})$ by

$$W_i = \begin{cases} M_i, & 0 \leqslant i \leqslant 15 \\ \sigma_1(W_{i-1}) + W_{i-9}, & \\ \quad + \sigma_2(W_{i-15}) + W_{i-16} & 16 \leqslant i \leqslant 63 \end{cases}$$