New Preimage Attack on MDC-4

Deukjo Hong and Daesung Kwon

Abstract

In this paper, we provide some cryptanalytic results for double-block-length (DBL) hash modes of block ciphers, MDC-4. Our preimage attacks follow the framework of Knudsen et al.'s time/memory trade-off preimage attack on MDC-2. We find how to apply it to our objects. When the block length of the underlying block cipher is n bits, the most efficient preimage attack on MDC-4 requires time and space about $2^{3n/2}$, which is to be compared to the previous best known preimage attack having time complexity of $2^{7n/4}$. Additionally, we propose an enhanced version of MDC-4, MDC-4* based on a simple idea. It is secure against our preimage attack and previous attacks and has the same efficiency as MDC-4.

Key Words. MDC4, Hash Function, Preimage

1 Introduction

Block ciphers and hash functions are widely used popular cryptographic primitives. Many hash functions are often designed based on block-cipher-like components. Some of them can be regarded as hash modes of block ciphers. PGV modes [12] are representative single-block-length (SBL) hash modes, where the length of the chaining and hash values is the same as the block length of the underlying block cipher. There are several double-block-length (DBL) hash modes such as MDC-2 , MDC-4 [5, 10], Hirose's scheme [4], Abreast-DM, and Tandem-DM schemes [8], where the length of the chaining and hash value is twice as long as the block length of the underlying block cipher.

MDC-4 is a security-enhanced version of MDC-2, which is a double-block-length (DBL) hash mode of block cipher. MDC-2 and MDC-4 have first described by Meyer and Schilling in 1988 [10], and have been patented by Brachtl et al. in 1990 [1]. Both of them have been standardized in ISO/IEC 10118-2 [5] and MDC-2 is used in practice (see [6, 13] for an exposition).

Steinberger showed that for MDC-2, any adversary making less than $2^{3n/5}$ queries has only a negligible chance of finding a collision in the ideal cipher model [13]. For MDC-2, Knudsen et al. provided the first collision attack with below-birthday-bound complexity, which is $2^{124.5}$ with $n = 128$, and the best known preimage attack with time and space complexity about 2^n [6]. In particular, the attacks in [6] appear to show that MDC-2 does not have enough security as a cryptographic hash function even if the underlying block cipher is secure.

We found previous cryptanalytic results for original MDC-4 based on DES [2] in [11, 7], where a collision attack on MDC-4 compression function, a preimage attack on MDC-4 compression function, and a preimage attack on MDC-4 hash function are provided with time complexities of 2^{41}, 2^{90}, and 2^{109}, respectively. They are often referred as the attacks with time complexities of about $2^{3n/4}$, $2^{3n/2}$, and $2^{7n/4}$, respectively, in the general setting with the underlying bock cipher with n-bit block and n-bit key. Fleischmann et al. proved that for MDC-4, any adversary making less than about $2^{3n/5}$ queries has only a negligible chance of finding a collision in the ideal cipher model [3]; precisely, their security bound is $2^{74.76}$ with $n = 128$. Mennink [9] also presented that his security proofs for MDC-4 in the ideal cipher model give the security bound of $2^{5n/8}$ in the aspect of collision resistance and the security bounds of $2^{5n/4}$ in two-independent-block-cipher setting and 2^n in single-block-cipher setting, respectively. In summary, as far as we know, no collision attack on MDC-4 hash function has been reported, and the only known preimage attack [7] on MDC-4 hash function has time complexity of $2^{7n/4}$ and a recently reported preimage attack [9] on MDC-4 compression function has time complexity of 2^n for a special case.

We provide a new preimage attack on MDC-4 hash function. It follows the framework of the time/memory trade-off preimage attack by Knudsen et al. [6], which was applied to MDC-2. We point out the compression function of MDC-4 can be divided into two parts with probability of $2^{-n/2}$, and find a pseudo-preimage for a target hash value faster than brute force attack. Based on this observation, we use Knudsen et al.'s approach to make a preimage attack on MDC-4 hash function. Our attack has time and memory complexities of about $2^{3n/2}$. It improves the best known preimage attack result. Additionally, we propose an enhanced version of MDC-4, MDC-4* based on a simple idea. It is secure against our preimage attack and previous attacks and has the same efficiency as MDC-4.

2 Description of MDC-2 and MDC-4 Hash Modes

MDC-2 and MDC-4 was originally defined using DES [2] as the underlying block cipher. However, for ease of presentation, we assume that the underlying block cipher E has n-bit block and n-bit key. Let $E_K(P)$ denote the encryption of a plaintext P using a key K with the block cipher E, which is assumed to be secure. If X is an n-bit string, then we let X_L denote the leftmost $n/2$ bits of X, and X_R denote the rightmost $n/2$ bits of X.

Given a block cipher E, MDC-2 defines a $2n$-bit hash function. The MDC-2 compression function $\mathrm{CF}^{\text{MDC-2}}$ has $2n$-bit chaining variable and n-bit message block. For the input chaining variable H and the message block M, $V = \mathrm{CF}^{\text{MDC-2}}(H, M)$ is computed as follows:

$$\begin{aligned}
A &= E_{H_L}(M) \oplus M;\\
B &= E_{H_R}(M) \oplus M;\\
V_L &= A_L \| B_R;\\
V_R &= B_L \| A_R.
\end{aligned}$$

Given a block cipher E, MDC-4 defines a $2n$-bit hash function. The MDC-4 compression function $\mathrm{CF}^{\text{MDC-4}}$ has $2n$-bit chaining variable and n-bit message block. For the input chaining variable H and the message block M, $V = \mathrm{CF}^{\text{MDC-4}}(H, M)$ is computed as follows:

$$\begin{aligned}
A &= E_{H_L}(M) \oplus M;\\
B &= E_{H_R}(M) \oplus M;\\
C &= B_L \| A_R;\\
D &= A_L \| B_R;\\
X &= E_C(H_L) \oplus H_L;\\
Y &= E_D(H_R) \oplus H_R;\\
V_L &= Y_L \| X_R;\\
V_R &= X_L \| Y_R.
\end{aligned}$$

MDC-2 and MDC-4 take the Merkle-Damgård domain extender to hash arbitrary-length messages with the above compression functions. We assume they use a popular prefix-free padding, which embeds the message length information to the last message block.

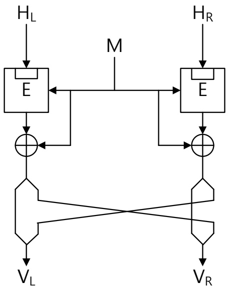

Figure 1: $\mathrm{CF}^{\text{MDC-2}}(H, M) = V$

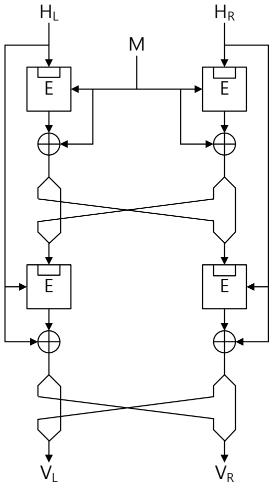

Figure 2: $\mathrm{CF}^{\text{MDC-4}}(H, M) = V$

3 Previous Attacks on MDC Hash Functions

In [11, 7], a collision attack on MDC-4 compression function, a preimage attack on MDC-4 compression function, and a preimage attack on MDC-4 hash function are presented. They have time complexities of $2^{3n/4}$, $2^{3n/2}$, and $2^{7n/4}$, respectively. So far, they have been the best attacks on MDC-4. Since the work, the only attack on MDC-4 reported is in [9]; it provided a simple preimage attack on MDC-4 compression function that finds a preimage with time complexity of 2^n when the left and right halves of the target hash value are equal.

[6] provided new collision and preimage attacks on MDC-2. Our approach is to apply the techniques from [6] to MDC-4. The collision attack on MDC-2 in [6] has the following framework.

1. Construct an r-collision on a half of the output chaining value of the first block.

2. Repeat the computation of the second block with randomly chosen message blocks until a collision on hash value is found.

The above attack has time complexity of $(r!2^{n(r-1)})^{1/r} + 2^n/(r-1)$. This approach is not applicable to MDC-4 because the attacker should expect a collision for at least $3n/2$ bits in the second block, while it is for n bits in MDC-2. So, the above attack has time complexity of $(r!2^{n(r-1)})^{1/r} + 2^{3n/2}/(r-1)$, which is not below 2^n.

[6] also provides preimage attacks on MDC-2. The most efficient one among them is as follows:

1. Choose two message blocks $M^{(0)}$ and $M^{(1)}$ arbitrarily, but with correct padding for a message of length $n+1$ blocks.
2. Compute $E_x(M^{(b)}) \oplus M^{(b)}$ and for every x from 0 to $2^n - 1$. Store the outputs in the tables U_b, sorted on the output.
3. Construct a binary tree with 2^n leaves having the target image V as root (with depth n): the nodes are labelled with intermediate hash values, and each edge is labelled with a message block value meaning that this message block maps from the intermediate hash value at the child node to the intermediate hash value at the parent. The two children of each node in the tree are found by lookups in U_0 and U_1, respectively.
4. Given 2^n new target images (namely the leaves in the tree), perform a brute force search starting from the initial value.

The above preimage attack has time and memory complexities of about 2^{n+1}. The following theorem and corollary guarantee that the above attack makes a 2^n-leaf tree with high probability.

Theorem 1 ([6]). *Given a target hash value $V = V_L \| V_R$, a pseudo-preimage can be found in time at most 2^{n-1} with probability about $(1-1/e)^2$. By a pseudo-preimage we mean a chaining variable $H = H_L \| H_R$ and a message block M such that $\mathrm{CF}^{MDC\text{-}2}(H, M) = V$.*

Corollary 1 ([6]). *Given t target hash values, in time 2^{n-1} one pseudo-preimage (on average) can be found for each target hash value. Here, t can be any number between 1 and 2^n.*

The above preimage attack is not directly applicable to MDC-4 because of the interruption of the intermediate swap between upper block ciphers and lower block ciphers. However, in the next section, we give an observation about MDC-4 and based on it, make an application of Knudsen et al.'s approach for a preimage attack on MDC-4.

4 Preimage Attack on MDC-4 Hash Function

Consider $\mathrm{CF}^{\mathrm{MDC\text{-}4}}(H, M) = V$. If the following equation

$$\left(E_{H_L}(M) \oplus M\right)_L = \left(E_{H_R}(M) \oplus M\right)_L \qquad (1)$$

holds, then we can write

$$\begin{aligned} E_{E_{H_L}(M)\oplus M}(H_L) \oplus H_L &= A; \\ E_{E_{H_R}(M)\oplus M}(H_R) \oplus H_R &= B, \end{aligned}$$

where A and B are n-bit values such that

$$\begin{aligned} A &= (V_R)_L \| (V_L)_R; \\ B &= (V_L)_L \| (V_R)_R. \end{aligned}$$

Note that the match in (1) is for $\frac{n}{2}$ bits. That is, the event occurs with the probability of $2^{-n/2}$.

Using the above observation, we can construct a preimage attack on MDC-4 hash function with time complexity of $2^{\frac{3n}{2}}$. It is also based on the time-memory trade-off preimage attack which Knudsen et al. proposed for MDC-2 [6].

1. Choose $2^{\frac{n}{2}+1}$ n-bit message blocks M. For each M, compute

   $$E_{E_x(M)\oplus M}(x) \oplus x \quad \text{for all } x \in \{0,1\}^n$$

   and store the results in the table T_M. After this precomputation, we get $2^{\frac{n}{2}+1}$ tables $\{T_M\}$, where each has 2^n entries.

2. Compute A and B for the target hash value V, and look up the tables to get the solution (x, y) to the following equations:

   $$\begin{aligned} E_{E_x(M)\oplus M}(x) \oplus x &= A; \\ E_{E_y(M)\oplus M}(y) \oplus y &= B. \end{aligned}$$

   On average, it is expected that at least $2^{\frac{n}{2}+1}$ solutions are found. Finally, it is expected that there are two solutions satisfying

   $$\left(E_x(M) \oplus M\right)_L = \left(E_y(M) \oplus M\right)_L.$$

   With this solution, two child nodes of V are made by letting $H_L = x$ and $H_R = y$ and labeling the each edge with the corresponding message block M.

3. For the new nodes, make their child nodes with the same tables. Repeat this procedure until 2^n leaves are produced.

4. Given 2^n new target images (namely the leaves in the tree), perform a brute force search starting from the initial value.

The time complexity of the above attack is $2^{3n/2}$ and the memory requirement is about $2^{3n/2}$ cells for $2n$-bit values.

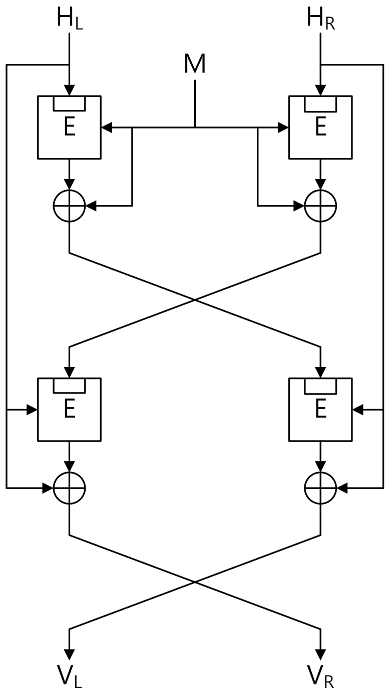

Figure 3: $\mathrm{CF}^{\text{MDC-4}^*}(H, M) = V$

5 Enhanced Version of MDC-4

Considering our attack, we propose an enhanced version of MDC-4, MDC-4*. It is based on a very simple idea. The main reason that MDC-4 is attacked with our work is the intermediate $n/2$-bit-wise swapping. Our attack shows that swapping only partial bits at the middle of the compression function can lead to a separation of the whole structure into left and right pieces with a probability. So, we recommend the use of n-bit swapping in our proposal. For the input chaining variable $H \in \{0,1\}^{2n}$ and the message block $M \in \{0,1\}^n$, the MDC-4* produces the output chaining variable $V \in \{0,1\}^{2n}$ as follows.

$$\begin{aligned} A &= E_{H_L}(M) \oplus M; \\ B &= E_{H_R}(M) \oplus M; \\ X &= E_B(H_L) \oplus H_L; \\ Y &= E_A(H_R) \oplus H_R; \\ V_L &= Y; \\ V_R &= X. \end{aligned}$$

MDC-4* is still secure against Knudsen et al.'s collision attack [6] as well as secure against our preimage attack.

6 Conclusion

In this paper we presented a new preimage attack on MDC-4. We found a property for MDC-4 compression function, which holds with $2^{-n/2}$. Based on it, we use Knudsen et al.'s approach to make a preimage attack on MDC-4 whose time and memory complexities are about $2^{3n/2}$. Our attack improves previous preimage attack results and shows that MDC-4 has the security level of preimage resistance much less than 2^{2n}, popularly expected. Additionally, we propose an enhanced version of MDC-4, MDC-4* based on a simple idea. It is secure against our preimage attack and previous attacks and has the same efficiency as MDC-4.

References

[1] B. Brachtl, D. Coppersmith, M. Hyden, S. Matyas, C. Meyer, J. Oseas, S. Pilpel, and M. Schilling, "Data Authentication Using Modification

Detection Codes Based on a Public One Way Encryption Function," U.S. Patent Number 4,908,601, March 13, 1990.

[2] U.S.Department of Commerce/National Institute of Standards and Technology, "Data Encryption Standard (DES)," FIPS PUB 46-3, Reaffirmed October 25th 1999.

[3] E. Fleischmann, C. Forler, and S. Lucks, "The Collision Security of MDC-4," In A. Mitrokotsa and S. Vaudeay (Eds.), *AFRICACRYPT 2012*, LNCS 7374, pp. 252–269, Springer-Verlag, 2012.

[4] S. Hirose, "Some Plausible Constructions of Double-Block-Length Hash Functions," In M. J. B. Robshaw (Ed.), *FSE 2006*, LNCS 4047, pp. 231–246, Springer-Verlag, 2006.

[5] ISO/IEC 10118-2:2010, "Information technology — Security techniques — Hash-functions — Part 2: Hash-functions using an n-bit block cipher," 1994, revised in 2010.

[6] L. R. Knudsen, F. Mendel, C. Rechberger, and S. Thomsen, "Cryptanalysis of MDC-2," *EUROCRYPT 2009*, LNCS 5479, pp. 106–120, Springer-Verlag, 2009.

[7] L. R. Knudsen and B. Preneel, "Fast and Secure Hashing Based on Codes," In B. S. Kaliski Jr. (Ed.), *CRYPTO'97*, LNCS 1294, pp. 485–498, Springer-Verlag, 1997.

[8] X. Lai and J. L. Massey, "Hash function based on block ciphers," In R. A. Rueppel (Ed.), *EUROCRYPT'92*, LNCS 658, pp. 55–70, Springer-Verlag, 1993.

[9] B. Mennink, "On the Collision and Preimage Security of MDC-4 in the Ideal Cipher Model," IACR ePrint Archive 2012/113. FSE 2012 Rump Session (Slide).

[10] C. Meyer and M. Schilling, "Secure Program Load with Manipulation Detection Code," *Proc. Securicom*, pp. 111–130, 1988.

[11] B. Preneel, *Analysis and Design of Cryptographic Hash Functions*, Doctorial Dissertation, Katholieke Universiteit Leuven, 1993.

[12] B. Preneel, R. Govaerts and J. Vandewalle, "Hash Functions Based on Block Ciphers: A Synthetic Approach," In D. R. Stinson (Ed.), CRYPTO 1993, LNCS 773, pp. 363–378, Springer-Verlag, 1994.

[13] J. Steinberger, "The Collision Intractability of MDC-2 in the Ideal-Cipher Model," *EUROCRYPT 2007*, LNCS 4515, pp. 34–51, Springer-Verlag, 2007.