Cryptanalysis of Some Double-Block-Length Hash Modes of Block Ciphers with n-Bit Block and n-Bit Key

Deukjo Hong and Daesung Kwon

Abstract

In this paper, we make attacks on DBL (Double-Block-Length) hash modes of block ciphers with n-bit key and n-bit block. Our preimage attack on the hash function of MDC-4 scheme requires the time complexity $2^{3n/2}$, which is significantly improved compared to the previous results. Our collision attack on the hash function of MJH scheme has time complexity less than 2^{124} for $n = 128$. Our preimage attack on the compression function of MJH scheme find a preimage with time complexity of 2^n. It is converted to a preimage attack on the hash function with time complexity of $2^{3n/2+2}$. Our preimage attack on the compression function of Mennink's scheme find a preimage with time complexity of $2^{3n/2}$. It is converted to a preimage attack on the hash function with time complexity of $2^{7n/4+1}$.

These attacks are helpful for understanding the security of the hash modes together with their security proofs.

Key words: Hash Function, Hash Mode, Collision, Preimage

1 Introduction

Block ciphers and hash functions are the most widely used primitives for cryptographic applications. For secure building, hash functions are often designed based on block-cipher-like components. Preneel, Govaerts, and Vandewalle [19] analyzed the securities of 64 ways to make a compression function with $2n$-bit input and n-bit output from a single call of the underlying block cipher with n-bit block and n-bit key, under the assumption of the use of Merkle-Damgård domain extender, and suggested 12 of them as secure ones. Later, their objects are called PGV schemes. Black et al. [2] proved that the 12 schemes suggested in [19] make compression functions collision-resistant and preimage-resistant upto $O(2^{n/2})$ queries to the underlying block cipher in the ideal cipher model, and showed that for 8 of the remaining 52 schemes, the resulting compression functions are weak but the resulting hash functions by using Merkle-Damgård domain extender are collision-resistant and preimage-resistant upto $O(2^{n/2})$ queries to the underlying block cipher in the ideal cipher model. In [20], Stam extended Black et al.'s work to more generalized version of PGV schemes. Since [2], most researches for provable security of hash modes have used the assumptions of Merkle-Damgård domain extender and ideal cipher model.

Since the length of the hash value from PGV hash modes is the same as the block length of the underlying block cipher and block ciphers usually have too short block length to provide conventional securities for hash functions, double-block-length (DBL) hash modes have often been researched. Abreast-DM [11], Tandem-DM [11], and Hirose's [6] schemes make the compression functions with $3n$-bit input and $2n$-bit output from two calls of the underlying block cipher with n-bit block and $2n$-bit key. These DBL hash modes are proved to have the security bounds of $O(2^n)$ for collision resistance [6, 12, 13] and the security bounds close to 2^{2n} for preimage resistance [1], so they are very secure as long as the underlying block cipher is secure.

On the other hand, it seems to be more difficult to build secure hash modes from the block cipher with n-bit block and n-bit key. We can model MDC-2 and MDC-4 schemes [7, 17] as such cases, although the original ones used DES [4], which is the block cipher with 64-bit block and 56-bit key. MDC-2 scheme makes the compression function with $3n$-bit input and $2n$-bit output from two calls of the underlying block cipher with n-bit block and n-bit key, and MDC-4 scheme does the similar work from four calls of

the underlying block cipher. MDC-2 and MDC-4 schemes are proved to have the security bound $O(2^{3n/5})$ for collision resistance [21, 5]. MJH [14] and Mennink's [16] schemes are recently proposed DBL hash modes. The designers of MJH proved that it has security bound $O(2^{2n/3-\log n})$ for collision resistance. Mennink proved his scheme has security bound $O(2^n)$ for collision resistance and security bound $O(2^{3n/2})$ for preimage resistance.

Note that the except Mennink's scheme, the above mentioned DBL hash modes of the underlying block cipher with n-bit block and n-bit key have security bound not close to the conventional level n bits of collision resistance required for hash functions. Even Mennink's scheme does not reach the conventional level $2n$ bits of preimage resistance required for hash functions. Some one may use those hash modes by considering the block length suitable for certain security goals. However, what we want to address is that the proofs providing partial security do not give us any knowledge about the security for the adversary with much more query access than the bounds.

There exist some attacks on MDC-2 and MDC-4. [10] and [18] presented several attacks on MDC-2 and MDC-4 schemes with DES, which are translated for the underlying block cipher with n-bit block and n-bit key into, as follows: collision attacks on compression functions of MDC-2 and MDC-4 schemes with time complexities $2^{n/2}$ and $2^{3n/4}$, resp.; preimage attacks on compression functions of MDC-2 and MDC-4 schemes with 2^n and $2^{3n/2}$, resp.; preimage attacks on hash functions of MDC-2 and MDC-4 schemes with $2^{3n/2+1}$ and $2^{7n/4+1}$, respectively. Knudsen et al. [9] presented the collision attack on the hash function of MDC-2 scheme with the time complexity $2^{124.5}$ for $n = 128$, and the preimage attack on the hash function of MDC-2 scheme with the time complexity 2^n. These results are helpful for understanding the security of the hash modes together with the proofs. For example, for $n = 128$ Steinberger's proof of the collision resistance of MDC-2 implies any adversary making less than $2^{76.8}$ has only a negligible chance of finding a collision in the ideal cipher model, but Knudsen's attack shows an adversary computing $2^{124.5}$ compression functions can find a collision with high probability.

In this paper, we present some attacks on MDC-4, MJH, and Mennink's schemes as follows.

1. Preimage attack on MDC-4 scheme: It requires the time complexity $2^{3n/2}$, which is significantly improved compared to the previous result [10, 18].

2. Collision and preimage attacks on MJH scheme: Our collision attack on the hash function of MJH scheme has time complexity less than 2^{124} for $n = 128$. We show that a preimage is found for the compression function of MJH scheme with time complexity of 2^n. This preimage attack on the compression function is converted to a preimage attack on the hash function with time complexity of $2^{3n/2+2}$.

3. Preimage attack on Mennink's scheme: We make a preimage attack on the compression function in a different way from the previous one in [16], and convert it to a preimage attack on the hash function with time complexity of $2^{7n/4+1}$.

2 Description of DBL Hash Modes

In this section, we give a brief description of MDC-4, MJH and Mennink's schemes. We assume Merkle-Damgård domain extender is used for constructing the hash function from the compression functions of DBL schemes. 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.

For i-th call of the compression function $\mathsf{CF} : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n \times \{0,1\}^n$ in Merkle-Damgård domain extender, we consider the input chaining variable $(H_i, S_i) \in \{0,1\}^n \times \{0,1\}^n$, the message block $M_i \in \{0,1\}^n$, and the output chaining variable $(H_{i+1}, S_{i+1}) \in \{0,1\}^n \times \{0,1\}^n$.

Then, for the message blocks $M_0, M_1, ..., M_{\ell-1} \in \{0,1\}^n$ after padding the message and the initial value (H_0, S_0), Merkle-Damgård domain extender iterates the compression function CF to produce the

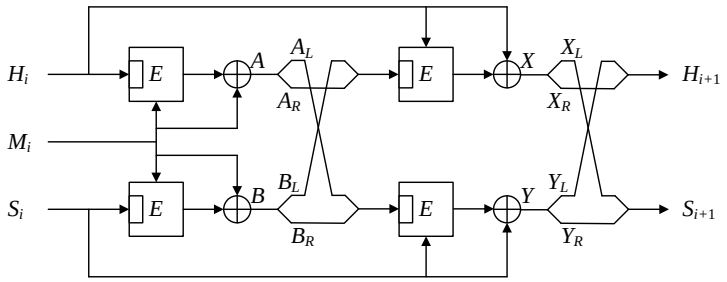

Figure 1: $\mathrm{CF}^{\text{MDC-4}}(H_i, S_i, M_i) = (H_{i+1}, S_{i+1})$

hash value (H_ℓ, S_ℓ) as follows:

$$(H_{i+1}, S_{i+1}) \leftarrow \mathsf{CF}(H_i, S_i, M_i) \quad 0 \le i \le \ell - 1.$$

We consider the padding rule following MD strengthen [15].

2.1 MDC-4 Scheme

MDC-4 scheme [7, 17] was originally defined using DES [4] as the underlying block cipher, but we assume that the underlying block cipher E has n-bit block and n-bit key. Given a block cipher E, 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_i, S_i) and the message block M_i, $(H_{i+1}, S_{i+1}) = \mathrm{CF}^{\text{MDC-4}}(H_i, S_i, M_i)$ is computed as follows:

$$\begin{aligned}
A &\leftarrow E_{H_i}(M_i) \oplus M_i;\\
B &\leftarrow E_{S_i}(M_i) \oplus M_i;\\
C &\leftarrow B_L \| A_R;\\
D &\leftarrow A_L \| B_R;\\
X &\leftarrow E_C(H_i) \oplus H_i;\\
Y &\leftarrow E_D(S_i) \oplus S_i;\\
H_{i+1} &\leftarrow Y_L \| X_R;\\
S_{i+1} &\leftarrow X_L \| Y_R.
\end{aligned}$$

2.2 MJH Scheme

MJH scheme [14] has two auxiliary components σ and $\cdot\theta$. σ is an involution on $\{0,1\}^n$ with no fixed point, and $\cdot\theta$ is a multiplication by a constant $\theta \neq 0, 1$ in $GF(2^n)$. The MJH compression function CF^{MJH} has $2n$-bit chaining variable and n-bit message block, based on the block cipher E with n-bit block and n-bit key. For the input chaining variable (H_i, S_i) and the message block M, $(H_{i+1}, S_{i+1}) = \mathrm{CF}^{\text{MJH}}(H_i, S_i, M_i)$ is computed as follows:

$$\begin{aligned}
X &\leftarrow H_i \oplus M_i;\\
H_{i+1} &\leftarrow E_{S_i}(X) \oplus X;\\
Y &\leftarrow E_{S_i}(\sigma(X)) \oplus \sigma(X);\\
S_{i+1} &\leftarrow (Y \cdot \theta) \oplus H_i.
\end{aligned}$$

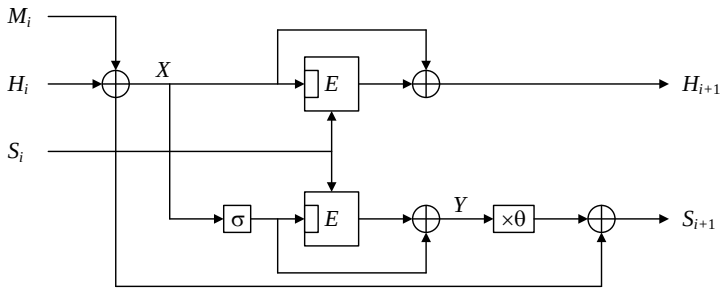

Figure 2: $\mathrm{CF}^{\mathrm{MJH}}(H_i, S_i, M_i) = (H_{i+1}, S_{i+1})$

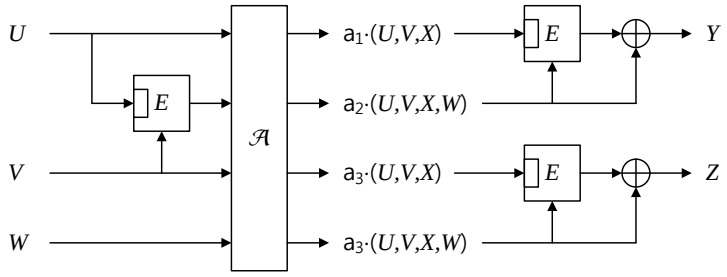

Figure 3: $\mathrm{CF}^{\mathrm{Mennink}}(U, V, W) = (Y, Z)$

2.3 Mennink's Scheme

Mennink's scheme [16] comprises three calls of the underlying block cipher E and a 4×4 matrix $\mathcal{A}$ over $GF(2^n)$, defined as follows:

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

Note that $\mathcal{A}$ is invertible and $a_{24}, a_{44} \neq 0$.

For clarity, we use the notations $\mathsf{a}_1 = (a_{11}, a_{12}, a_{13}), \mathsf{a}_3 = (a_{31}, a_{32}, a_{33}) \in GF(2^n)^3$, $\mathsf{a}_2 = (a_{21}, a_{22}, a_{23}, a_{24}), \mathsf{a}_4 = (a_{41}, a_{42}, a_{43}, a_{44}) \in GF(2^n)^4$ in the description. The compression function $\mathrm{CF}^{\mathrm{Mennink}}$ of Mennink's scheme has $3n$-bit input and $2n$-bit output. For the input (U, V, W), $(Y, Z) = \mathrm{CF}^{\mathrm{Mennink}}(U, V, W)$ is computed as follows:

$$\begin{aligned} X &\leftarrow E_U(V); \\ K_1 &\leftarrow \mathsf{a}_1 \cdot (U, V, X); \\ P_1 &\leftarrow \mathsf{a}_2 \cdot (U, V, X, W); \\ Y &\leftarrow E_{K_1}(P_1) \oplus P_1; \\ K_2 &\leftarrow \mathsf{a}_3 \cdot (U, V, X); \\ P_2 &\leftarrow \mathsf{a}_4 \cdot (U, V, X, W); \\ Z &\leftarrow E_{K_2}(P_2) \oplus P_2. \end{aligned}$$

In [16], it was not specified how to correspond the variables (H_i, S_i, M_i) and (H_{i+1}, V_{i+1}) to (U, V, W) and (Y, Z), respectively. We will consider some cases for it in our attack.

3 Preimage Attack on MDC-4 Hash Function

Consider $\mathrm{CF}^{\text{MDC-4}}(H_i, S_i, M_i) = (H_{i+1}, S_{i+1})$. If the following equation

$$(E_{H_i}(M_i) \oplus M_i)_L = (E_{S_i}(M_i) \oplus M_i)_L \tag{1}$$

holds, then we can write

$$\begin{aligned} E_{E_{H_i}(M_i)\oplus M_i}(H_i) \oplus H_i &= X; \\ E_{E_{S_i}(M_i)\oplus M_i}(S_i) \oplus S_i &= Y, \end{aligned}$$

where A and B are n-bit values such that

$$\begin{aligned} X &= (S_{i+1})_L \| (H_{i+1})_R; \\ Y &= (H_{i+1})_L \| (S_{i+1})_R. \end{aligned}$$

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 construct a preimage attack on MDC-4 hash function with time complexity of $2^{\frac{3n}{2}}$ as follows, which is also based on the time-memory trade-off preimage attack which Knudsen et al. proposed for MDC-2 [9].

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 X and Y for the target hash value (H_ℓ, S_ℓ), and look up the tables to get the solution (a, b) to the following equations:

   $$\begin{aligned} E_{E_a(M)\oplus M}(a) \oplus a &= X; \\ E_{E_b(M)\oplus M}(b) \oplus b &= Y. \end{aligned}$$

   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

   $$(E_a(M) \oplus M)_L = (E_b(M) \oplus M)_L .$$

   With this solution, two child nodes of (H_ℓ, S_ℓ) are made by letting $H_{\ell-1} = a$ and $S_{\ell-1} = b$ 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. If the tree with 2^n leaves has the depth t, take $\ell = t + 1$, and perform a brute force search starting from the initial value (H_0, S_0) to find a match of (H_1, S_1) and the 2^n targets. If a match is found, then a preimage is comprised of the corresponding message blocks.

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.

4 Collision Attack on MJH

4.1 Collision Attack on MJH Compression Function

We find that we do not need to consider the right half of the hash value in the collision attack for the MJH compression function, because after finding a collision for the left half of the hash value, we can easily compute the left halves of the input chaining variable and the message blocks such that they give a same hash value. The following collision attack on the MJH compression function reflects on our observation.

1. Randomly choose S_i and S_{i+1}, and fix them.
2. Randomly choose r distinct $X = H_i \oplus M_i$: $X^{(1)}, X^{(2)}, ..., X^{(r)}$, compute $H_{i+1}^{(s)} = E_{S_i}(X^{(s)}) \oplus X^{(s)}$ for $s = 1, ..., r$, and then check whether there are at least one pair of (s, j) such that
$$X_s \neq X_j \text{ and } H_{i+1}^{(s)} = H_{i+1}^{(j)}. \tag{2}$$
3. If a pair (s_1, s_2) satisfying (2) is found, then compute $H_i^{(s_1)}$, $H_i^{(s_2)}$, $M_i^{(s_1)}$, and $M_i^{(s_2)}$ as follows:
$$\begin{aligned} H_i^{(s_j)} &= (E_{S_i}(\sigma(X^{(s_j)})) \oplus \sigma(X^{(s_j)})) \cdot \theta \oplus S_{i+1} \text{ for } j = 0, 1; \\ M_i^{(s_j)} &= H_i^{(s_j)} \oplus M_i^{(s_j)} \text{ for } j = 0, 1. \end{aligned}$$
4. Output $(H_i^{(s_1)}, S_i, M_i^{(s_1)})$ and $(H_i^{(s_2)}, S_i, M_i^{(s_2)})$ as a collision of the MJH compression function.

The time complexity of the above attack is dominated by r encryptions of the block cipher E. With $r = 2^{n/2}$, we expect at least one collision for MJH compression function, because the right half of the hash value is fixed.

4.2 Collision Attack on MJH Hash Function

We provide a 2-block collision attack on MJH hash function. The first step to find a collision for MJH hash function is similar to Knudsen et al.'s attack for MDC-2 [9]. We make a multi-collision for the right half S_1 of the $2n$-bit output chaining value in the first block. The multi-collision from the first block fix the key inputs to the block ciphers in the second step. In the second step, we take a different approach of choosing $X = M_1 \oplus H_1$ at random, instead of M_1. Due to the fixed key input, the computations in the second block are almost independent of the first block. With this observation, we make a collision attack on MJH hash function as follows.

1. Choose sufficiently many message blocks in the first block, and obtain an r-collision for S_1. Denote the corresponding left halves of the output chaining variable and the message blocks by $H_1^{(1)}, H_1^{(2)}, ..., H_1^{(r)}$, and $M_0^{(1)}, M_0^{(2)}, ..., M_0^{(r)}$, respectively.
2. Choose randomly q distinct $X = H_1 \oplus M_1$: $X^{(1)}, X^{(2)}, ..., X^{(q)}$, compute $H_2^{(i)} = E_{S_1}(X^{(i)}) \oplus X^{(i)}$ for $i = 1, 2, ..., q$, and collect the pairs of (i, j) such that $i \neq j$ and $H_2^{(i)} = H_2^{(j)}$ for $1 \leq i, j \leq q$.
3. For the pairs of (i, j) colliding on H_2, compute $Y^{(k)}$ for $k = i$ and j as follows:
$$Y^{(k)} = (E_{S_1}(\sigma(X^{(k)})) \oplus \sigma(X^{(k)})) \cdot \theta,$$
and check whether there is at least one pair of (u, v) for $1 \leq u, v \leq r$ such that $u \neq v$ and $H_1^{(u)} \oplus Y^{(i)} = H_1^{(v)} \oplus Y^{(j)}$. For such a tuple (i, j, u, v), the same S_2 is produced from $H_1^{(u)} \oplus Y^{(i)}$ and $H_1^{(v)} \oplus Y^{(j)}$ or from $H_1^{(v)} \oplus Y^{(i)}$ and $H_1^{(u)} \oplus Y^{(j)}$. If such a tuple is found, output $M_0^{(u)} \| (H_1^{(u)} \oplus X^{(i)})$ and $M_0^{(v)} \| (H_1^{(v)} \oplus X^{(j)})$, or $M_0^{(v)} \| (H_1^{(v)} \oplus X^{(i)})$ and $M_0^{(u)} \| (H_1^{(u)} \oplus X^{(j)})$ as a collision.

Table 1: Time complexity of the collision attack on MJH with an n-bit block cipher, compared to birthday complexity, where $T_E \cong T_K$.

n	r	Collision Attack	Birthday Attack
64	8	$2^{60.58}$	2^{64}
128	14	$2^{123.81}$	2^{128}
256	23	$2^{251.03}$	2^{256}

Table 2: Time complexity of the collision attack on MJH with an n-bit block cipher, compared to birthday complexity, where $T_E \gg T_K$.

n	r	Collision Attack	Birthday Attack
64	9	$2^{61.01}$	2^{64}
128	15	$2^{124.27}$	2^{128}
256	25	$2^{251.50}$	2^{256}

For more precise estimation of time complexity, we separate the costs of encryption and key schedule in a call of block cipher. We consider that the compression function of MJH requires two encryptions and one key schedule operations. Let T_E, T_K, T_{EK}, and T_{CF} be the time costs wasted in one encryption, one key schedule operation, one block cipher call and one compression function operation, respectively. We estimate the complexity for the case that T_E is almost equal to T_K (so, $T_{EK} \cong 2T_E$ and $T_{CF} \cong 3 \cdot T_E$), and the case that T_E is much larger than T_K (so, $T_{EK} \cong T_E$ and $T_{CF} \cong 2 \cdot T_E$).

In the first step, we need $((r!) \cdot 2^{(r-1)n})^{1/r}$ block cipher encryptions with key schedule operations to get an r-collision for H_R. The second step requires q block cipher encryptions. The time complexity of the last step is negligible compared to other steps. Since there are $\binom{r}{2}\binom{q}{2}$ possibilities for pairing $(X^{(i)}, X^{(j)})$'s and $(H_1^{(u)}, H_1^{(v)})$'s, we expect a collision for V with $\binom{r}{2}\binom{q}{2} = 2^{2n}$, where $\binom{r}{2}\binom{q}{2} \cong \frac{(rq)^2}{4}$ and $q \cong 2^{n+1}/r$. Overall, the time complexity of the above attack is estimated as

$$((r!) \cdot 2^{(r-1)n})^{1/r} T_{EK} + 2^{n+1}/r T_E. \tag{3}$$

(3) is approximated to

$$((r!) \cdot 2^{(r-1)n})^{1/r} \cdot \frac{2}{3} + 2^{n+1}/r \cdot \frac{1}{3} \tag{4}$$

for the case of $T_E \cong T_K$, and

$$(((r!) \cdot 2^{(r-1)n})^{1/r} + 2^{n+1}/r) \cdot \frac{1}{2} \tag{5}$$

for the case of $T_E \gg T_K$., respectively. For $n = 128$, we get the most efficient complexities of $2^{123.81}$ with $r = 14$ for (4) and $2^{124.27}$ with $r = 15$ for (5).

5 Preimage Attack on MJH Scheme

5.1 Preimage Attack on MJH Compression Function

It is easy to find a preimage of MJH compression function with time complexity of about $2^n \cdot T_E$. The preimage attack on MJH compression function is as follows.

1. Given a $2n$-bit target hash value (H_{i+1}, S_{i+1}), choose randomly $S_i \in \{0,1\}^n$ and fix it.

2. Find $X \in \{0,1\}^n$ such that $E_{S_i}(X) \oplus X = H_{i+1}$ with brute force attack.

3. If such X is found, compute $Y = (E_{S_i}(\sigma(X)) \oplus \sigma(X)) \cdot \theta$, $H_i = Y \oplus S_{i+1}$, and $M_i = H_i \oplus X$. Then, (H_i, S_i, M_i) is a preimage.

5.2 Preimage Attack on MJH Hash Function

We have to consider the padding rule for constructing a preimage attack on a hash function from a preimage attack on a compression function. We assume that the last message block contains a length information of the message. In the attack described in Section , the attacker does not have a control on the message block unlike the preimage attack on MDC-4 compression function described in Section . So, the attacker can not intend to embed a predetermined length information to the preimage, and we can not use Knudsen et al.'s time-memory trade off technique to make a preimage attack for MJH hash function from the preimage attack for MJH compression function. Alternatively, we make it using the meet-in-the-middle technique [15, Fact 9.99] and expandable messages with fixed-points [3, 8].

A fixed-point for a compression function CF is defined as (H, S, M) such that $\mathrm{CF}(H, S, M) = (H, S)$. We can find fixed-points for MJH compression function as follows.

1. Choose randomly $M \in \{0,1\}^n$ and $S \in \{0,1\}^n$, fix them.

2. Compute $X = E_S^{-1}(M)$, $H = M \oplus X$, $Y = (E_S(\sigma(X)) \oplus \sigma(X)) \cdot \theta$, and $H \oplus Y = S'$.

3. If $S' = S$, then output (H, S, M) as a fixed-point; else, repeat the computations in step 2 with new random choices of M and S.

On average, we expect to find a fixed-point with time complexity of 2^n.

An expandable message is constructed as follows.

1. Collect $2^{n/2}$ fixed-points $(H^{(1)}, S^{(1)}, M_2^{(1)})$, ..., $(H^{(2^{n/2})}, S^{(2^{n/2})}, M_2^{(2^{n/2})})$ by repeating the above search.

2. Repeat to compute $\mathrm{CF}^{\mathrm{MJH}}(\mathrm{CF}^{\mathrm{MJH}}(H_0, S_0, M_0), M_1)$ for a randomly chosen two-block message (M_0, M_1) until the result is matched with any $(H^{(i)}, S^{(i)})$ for $i = 1, ..., 2^{n/2}$.

3. If a match is found, then output the corresponding $(M_0, M_1, M_2^{(i)})$ as a $(2, \infty)$-expandable message.

On average, the number of repetition in the step 2 should be $2^{3n/2}$ to expect a match. So, the time complexity for the above construction of an expandable message is about $2^{3n/2+1}$.

Assume that we are given a $(2, \infty)$-expandable message (M_0, M_1, M_2) made from a fixed-point (H, M_2). Let $\mathrm{len}(M)$ be the length information of the hashed message M contained in the last message block. With this expandable message, we can construct a preimage attack on MJH hash function as follows.

1. Given a target hash value (H, S), collect $2^{n/2}$ preimages $(U^{(1)}, V^{(1)}, M_L^{(1)})$, $(U^{(2)}, V^{(2)}, M_L^{(2)})$, ..., $(U^{(2^{n/2})}, V^{(2^{n/2})}, M_L^{(2^{n/2})})$ for the last compression function.

2. Repeat to compute $\mathrm{CF}^{\mathrm{MJH}}(H, S, M_{L-1})$ for a randomly chosen one-block message M_{L-1} until the result is matched with any $U^{(i)}$ for $i = 1, ..., 2^{n/2}$.

3. If a match is found, then output the corresponding $(M_0, M_1, M_2, ..., M_2, M_{L-1}, M_L)$ as a preimage for V, where the repetition number of M_2 depends on $\mathrm{len}(M)$ contained in M_L.

On average, the number of repetition in the step 2 should be $2^{3n/2}$ to expect a match. So, the time complexity for the above construction of an expandable message is about $2^{3n/2+1}$.

Finally, a preimage attack on MJH hash function is made from the preimage attack on MJH compression function and the expandable message by the meet-in-the-middle technique in [15, Fact 9.99]. Overall time complexity is about $2^{3n/2+2}$.

6 Preimage Attack on Mennink's Scheme

In [16], it was not specified how to correspond the variables (H_i, S_i, M_i) and (H_{i+1}, V_{i+1}) to (U, V, W) and (Y, Z), respectively. In the cases of $M_i = U$ or $M_i = V$, we can make a preimage attack on the hash function of Mennink's scheme. Without loss of generality, we assume $M_i = U$.

6.1 Preimage Attack on Mennink's Compression function

For a given target image $(Y, Z) \in GF(2^n)^2$, our preimage attack works as follows.

1. Choose $2^{n/2}$ P_1 randomly. They are denoted by $P_1^{(0)}, ..., P_q^{(2^{n/2}-1)}$.
2. For $0 \leq i \leq 2^{n/2} - 1$, find $(U^{(i)}, V^{(i)}, X^{(i)})$'s such that $E_{U^{(i)}}(V^{(i)}) = X^{(i)}$ and $P_1^{(i)} = \mathbf{a}_1 \cdot (U^{(i)}, V^{(i)}, X^{(i)})$ in the following way:
   (a) Choose $U^{(i)}$.
   (b) Find $V^{(i)}$ such that $P_1^{(i)} = \mathbf{a}_1 \cdot (U^{(i)}, V^{(i)}, E_{U^{(i)}}(V^{(i)}))$ with a brute-force method.
   (c) Compute $X^{(i)} = E_{U^{(i)}}(V^{(i)})$.
3. For $0 \leq j \leq 2^{n/2} - 1$, find $K_1^{(j)}$ such that $E_{K_1^{(j)}}(P_1^{(j)}) \oplus P_1^{(j)} = Y$.
4. Make 2^n tuples $(U^{(i)}, V^{(i)}, X^{(i)}, W^{(\rho(i,j))})$ such that $K_1^{(j)} = \mathbf{a}_2 \cdot (U^{(i)}, V^{(i)}, X^{(i)}, W^{(\rho(i,j))})$ for $0 \leq i, j \leq 2^{n/2}-1$, where $\rho(i, j)$ is an index dependent on i and j. Renumber the tuples as $(U^{(i)}, V^{(i)}, X^{(i)}, W^{(i)})$ for $0 \leq i \leq 2^n - 1$.
5. For $0 \leq i \leq 2^n - 1$, compute $P_2^{(i)} = \mathbf{a}_3 \cdot (U^{(i)}, V^{(i)}, X^{(i)})$ and $K_2^{(i)} = \mathbf{a}_4 \cdot (U^{(i)}, V^{(i)}, X^{(i)}, W^{(i)})$, and check whether the relation $E_{K_2^{(i)}}(P_2^{(i)}) \oplus P_2^{(i)} = Z$. One $(P_2^{(i)}, K_2^{(i)})$ solution is expected for the relation. Then, the corresponding tuple $(U^{(i)}, V^{(i)}, W^{(i)})$ is a preimage of the compression function.

The time complexity of the above attack is $2^{3n/2}$ because steps 2 and 3 require more dominant cost than the others.

6.2 Preimage Attack on Mennink's Hash Function

Our preimage attack on the compression function explained in Section 6.1 is converted to a preimage attack in the case of $M_i = U$ or $M_i = V$ because the attacker has the control on the message length in the last block, while the previous attack in [16] cannot lead to a preimage attack on the hash function.

Again, a preimage attack on the hash function of Mennnink's scheme is made from the preimage attack on the compression function by the meet-in-the-middle technique in [15, Fact 9.99]. Overall time complexity is about $2^{7n/4+1}$.

7 Conclusion

We presented several attacks on MDC-4, MJH, and Mennink's schemes which are DBL hash modes. The preimage attacks on MDC-4 and Mennink's significantly improves the previous results, and the attacks on MJH scheme are the first cryptanalytic results. Our attacks provide deeper understanding about the securities of our target hash modes. Some of them show that there are gaps between the complexities of ours and the bounds stated in the security proofs. Studying how to reduce the gaps is interesting; whether the attacks are more developed or the proofs are more improved.

References

[1] F. Armknecht, E. Fleischmann, M. Krause, J. Lee, M. Stam, and J. Steinberger, "The Preimage Security of Double-Block-Length Compression Functions," *ASIACRYPT 2011*, LNCS 7073, pp. 233–251, Springer, 2011.

[2] J. Black, P. Rogaway, and T. Shrimpton, "Black-Box Analysis of the Block-Cipher-Based Hash-Function Constructions from PGV," *CRYPTO 2002*, LNCS 2442, pp. 320–335, Springer, 2002.

[3] R. D. Dean, *Formal Aspects of Mobile Code Security*, Ph. D Dissertion, Princeton University, January 1999.

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

[5] E. Fleischmann, C. Forler, and S. Lucks, "The Collision Security of MDC-4," *AFRICACRYPT 2012*, LNCS 7374, pp. 252–269, 2012.

[6] 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, 2006.

[7] 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.

[8] J. Kelsey and B. Schneier, "Second Preimages on n-Bit Hash Functions for Much Less than 2^n Work," *EUROCRYPT 2005*, LNCS 3494, pp. 474–490, Springer, 2005.

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

[10] L. R. Knudsen and B. Preneel, "Fast and Secure Hashing Based on Codes," *CRYPTO'97*, LNCS 1294, pp. 485–498, Springer, 1997.

[11] 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, 1993.

[12] J. Lee and D. Kwon, "The Security of Abreast-DM in the Ideal Cipher Model," *IEICE Transactions on Fundamentals*, Vol. 95-A, No. 1, pp. 104–109, 2011.

[13] J. Lee, M. Stam, and J. Steinberger, "The Collision Security of Tandem-DM in the Ideal Cipher Model," *CRYPTO 2011*, LNCS 6841, pp. 561–577, Springer, 2011.

[14] J. Lee and M. Stam, "MJH: A Faster Alternative to MDC-2," In A. Kiayias (Ed.), *CT-RSA 2011*, LNCS 6558, pp. 213–236, Springer, 2011.

[15] A. J. Menezes, P. C. Oorschot and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1997.

[16] B. Mennink, "Optimal Collision Security in Double Block Length Hashing with Single Length Key," *ASIACRYPT 2012*, LNCS 7658, pp. 330–347, Springer, 2012.

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

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

[19] 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, 1994.

[20] M. Stam, "Blockcipher-Based Hashinf Revisited," *FSE 2009*, LNCS 5665, pp. 67–83, Springer, 2009.

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