The Security of Abreast-DM in the Ideal Cipher Model

Jooyoung Lee, Daesung Kwon

The Attached Institute of Electronics and Telecommunications Research Institute
Yuseong-gu, Daejeon, Korea 305-390
`jlee05@ensec.re.kr,ds_kwon@ensec.re.kr`

Abstract. In this paper, we give a security proof for Abreast-DM in terms of collision resistance, preimage resistance and adaptive preimage resistance. As old as Tandem-DM, the compression function Abreast-DM is one of the most well-known constructions for double block length compression functions. The bounds on the number of queries for collision resistance and preimage resistance are given by $O(2^n)$ and $O\left(2^{2n}\right)$, respectively. The adaptive preimage resistance is guaranteed up to $O(2^n)$ queries/commitments. Based on a novel technique using *query-response cycles*, our security proof is simpler than those for MDC-2 and Abreast-DM. We also present a wide range of Abreast-DM variants that enjoy a birthday-type security guarantee with a simple proof.

1 Introduction

A cryptographic hash function takes a message of arbitrary length, and returns a bit string of fixed length. The most common way of hashing variable length messages is to iterate a fixed-size compression function according to the Merkle-Damgård paradigm. The underlying compression function can either be constructed from scratch, or be built upon off-the-shelf cryptographic primitives such as blockciphers. Recently, the blockcipher-based construction is attracting renewed interest, as many dedicated hash functions, including those most common in practical applications, exhibit serious security weaknesses [1, 6, 16, 17, 22, 26–28]. Conveniently choosing an extensively studied blockcipher in the blockcipher-based construction, one can easily transfer the trust in the existing algorithm to the hash function. This approach is particularly useful in highly constrained environments such as RFID systems, since a single implementation of a blockcipher can be used for both a blockcipher and a hash function. Compared to blockciphers, the most dedicated hash functions require significant amounts of state and the operations in their designs are not hardware friendly [3].

Compression functions based on blockciphers have been widely studied [2, 4, 9–12, 14, 18–21, 23–25]. The most common approach is to construct a $2n$-to-n bit compression function using a single call to an n-bit blockcipher. However, such a function, called a *single block length* (SBL) compression function, might be vulnerable to collision attacks due to its short output length. For example, one could successfully mount a birthday attack on a compression function based on AES-128 using approximately 2^{64} queries. This observation motivated substantial research on *double block length* (DBL) compression functions, where the output length is twice the block length of the underlying blockciphers.

Unfortunately, it turned out that a wide class of DBL compression functions of rate 1 are not optimally secure in terms of collision resistance and preimage resistance [9, 10, 13]. The most classical DBL compression functions of rate less than 1 include MDC-2, MDC-4, TANDEM-DM and ABREAST-DM [5, 14]. In 2007, 20 years after its original proposal, Steinberger first proved the collision resistance of MDC-2 in the ideal cipher model [25]. The author showed that an adversary asking less than $2^{3n/5}$ queries has only a negligible chance of finding a collision. Motivated by this work, Fleischmann et. al. proved the security of TANDEM-DM [8]. Similar to MDC-2, the security of TANDEM-DM is estimated in terms of a parameter, say, α. Optimizing the parameter, they proved the collision resistance of TANDEM-DM up to the birthday bound. Currently, TANDEM-DM and the Hirose's scheme [12] are the only rate 1/2 DBL compression functions that are known to have a birthday-type security guarantee. The underlying blockciphers of these schemes use $2n$-bit keys, while MDC-2 accepts n-bit keys. For this reason, it seems to be natural that the security proof of MDC-2 is more challenging.

Results We give a security proof for ABREAST-DM in terms of collision resistance, preimage resistance and adaptive preimage resistance. As old as TANDEM-DM, the compression function ABREAST-DM is known to be more advantageous than TANDEM-DM in that two encryptions involved can be computed in parallel. The bounds on the number of queries for collision resistance and preimage resistance are given by $O\left(2^n\right)$ and $O\left(2^{2n}\right)$, respectively. The adaptive preimage resistance is guaranteed up to $O\left(2^n\right)$ queries/commitments.

The notion of adaptive preimage resistance is first introduced in [15]. A compression function that is collision resistant and adaptive preimage resistant can be composed with a public random function to yield a hash function that is indifferentiable from a random oracle. In addition, the Merkle-Damgård transform preserves adaptive preimage resistance as long as the underlying compression function is collision resistant. For this reason, we believe that adaptive preimage resistance would be one of the desirable properties of a secure compression function. We note that a similar security notion, called *preimage awareness*, was independently introduced in [7]. Since any compression function that is both collision resistant and adaptive preimage resistant is preimage aware, our result can be regarded as the proof of preimage awareness for ABREAST-DM.

Based on a novel technique using *query-response cycles*, our security proof is simpler than those for MDC-2 and ABREAST-DM. We also present a wide range of ABREAST-DM variants that enjoy a birthday-type security guarantee with a simple proof.

2 Preliminaries

General Notations For a positive integer n, we let $I_n = \{0,1\}^n$ denote the set of all bitstrings of length n. For two bitstrings A and B, $A||B$ and $\overline{A}$ denote the concatenation of A and B, and the bitwise complement of A, respectively. For a set U, we write $u \overset{\$}{\leftarrow} U$ to denote uniform random sampling from the set U and assignment to u.

Ideal Cipher Model For positive integers n and k, let

$$BC(n,k) = \{E : I_n \times I_k \to I_n : \forall K \in I_k,\ E(\cdot, K) \text{ is a permutation on } I_n\}.$$

In the ideal cipher model, an (n,k)-blockcipher E is chosen from $BC(n,k)$ uniformly at random. It allows for two types of oracle queries $E(X,K)$ and $E^{-1}(Y,K)$ for $X,Y \in I_n$ and $K \in I_k$. Here, X, Y and K are called a plaintext, a ciphertext and a key, respectively. The response to an inverse query $E^{-1}(Y,K)$ is $X \in I_n$ such that $E(X,K) = Y$.

The Abreast-DM Compression Function For positive integers m, t and r with $m > r$, let

$$\Phi = \{\phi_i^1 : I_n^{m+i-1} \to I_n \times I_k : i = 1,\ldots,t\} \cup \{\phi_i^2 : I_n^{m+t} \to I_n : i = 1,\ldots,r\}$$

be a set of arbitrary functions. Then Φ defines a *blockcipher-based compression function* F_{mtr}^{Φ} with oracle access to an ideal cipher $E \in BC(n,k)$ as follows.

$$\begin{aligned} F_{mtr}^{\Phi} : I_n^m &\longrightarrow I_n^r \\ (A_1,\ldots,A_m) &\longmapsto (B_1,\ldots,B_r), \end{aligned} \tag{1}$$

where $(B_1,\ldots,B_r)$ is computed by the algorithm described in Figure 1(a). The rate of F_{mtr}^{Φ} is defined as

$$\rho = \frac{m-r}{t}.$$

Now the compression function Abreast-DM F^{ABR} is defined by

$$\begin{aligned} \phi_1^1 :&(A_1,A_2,A_3) \longmapsto (A_1, A_2||A_3), \\ \phi_2^1 :&(A_1,A_2,A_3,Y_1) \longmapsto (\overline{A_2}, A_3||A_1), \\ \phi_1^2 :&(A_1,A_2,A_3,Y_1,Y_2) \longmapsto A_1 \oplus Y_1, \\ \phi_2^2 :&(A_1,A_2,A_3,Y_1,Y_2) \longmapsto A_2 \oplus Y_2, \end{aligned}$$

with $(m,t,r) = (3,2,2)$ and $k = 2n$. The algorithm of F^{ABR} is separately described in Figure 1(b).

```
**Algorithm** $F^{\Phi}_{mtr}(A_1,\dots,A_m)$
  **for** $i \leftarrow 1$ to $t$ **do**
    $(X_i,K_i) \leftarrow \phi^1_i(A_1,\dots,A_m,Y_1\dots,Y_{i-1})$
    $Y_i \leftarrow E(X_i,K_i)$
  **for** $i \leftarrow 1$ to $r$ **do**
    $B_i \leftarrow \phi^2_i(A_1,\dots,A_m,Y_1\dots,Y_t)$
  **return** $(B_1,\dots,B_r)$
```

(a) Compression function F^{Φ}_{mtr}

```
**Algorithm** $F^{ABR}(A_1,A_2,A_3)$
  $(X_1,K_1) \leftarrow (A_1, A_2||A_3)$
  $Y_1 \leftarrow E(X_1,K_1)$
  $(X_2,K_2) \leftarrow (\overline{A_2}, A_3||A_1)$
  $Y_2 \leftarrow E(X_2,K_2)$
  $B_1 \leftarrow A_1 \oplus Y_1$
  $B_2 \leftarrow A_2 \oplus Y_2$
  **return** $(B_1,B_2)$
```

(b) ABREAST-DM F^{ABR}

Fig. 1. Blockcipher-based compression functions

Collision Resistance and Preimage Resistance Given a blockcipher-based compression function $F := F^{\Phi}_{mtr}$ and an information-theoretic adversary $\mathcal{A}$ with oracle access to E and E^{-1}, we execute the experiment $\mathbf{Exp}^{\mathsf{coll}}_{\mathcal{A}}$ described in Figure 2(a) in order to quantify the collision resistance of F. The experiment records the queries that the adversary $\mathcal{A}$ makes into a *query history* $\mathcal{Q}$. A pair (X,K,Y) is in the query history if $\mathcal{A}$ asks $E(X,K)$ and gets back Y, or it asks $E^{-1}(Y,K)$ and gets back X. For $A = (A_1,\dots,A_m) \in I^m_n$ and $B = (B_1,\dots,B_r) \in I^r_n$, we write

$$A \vdash_{\mathcal{Q}} B,$$

if there exist query-response pairs $(X_i,K_i,Y_i) \in \mathcal{Q}$, $i = 1,\dots,t$, satisfying the following equations.

$$(X_i,K_i) = \phi^1_i(A_1,\dots,A_m,Y_1\dots,Y_{i-1}), \qquad i = 1,\dots,t, \tag{2}$$

$$B_i = \phi^2_i(A_1,\dots,A_m,Y_1\dots,Y_t), \qquad i = 1,\dots,r. \tag{3}$$

Informally, $A \vdash_{\mathcal{Q}} B$ means that the query history $\mathcal{Q}$ determines the evaluation $F : A \mapsto B$. Now the *collision-finding advantage* of $\mathcal{A}$ is defined to be

$$\mathbf{Adv}^{\mathsf{coll}}_{F}(\mathcal{A}) = \mathsf{Pr}\left[\mathbf{Exp}^{\mathsf{coll}}_{\mathcal{A}} = 1\right]. \tag{4}$$

The probability is taken over the random blockcipher E and $\mathcal{A}$'s coins (if any). For $q > 0$, we define $\mathbf{Adv}^{\mathsf{coll}}_{F}(q)$ as the maximum of $\mathbf{Adv}^{\mathsf{coll}}_{F}(\mathcal{A})$ over all adversaries $\mathcal{A}$ making at most q queries.

The preimage resistance of F is quantified similarly using the experiment $\mathbf{Exp}^{\mathsf{pre}}_{\mathcal{A}}$ described in Figure 2(b). The adversary $\mathcal{A}$ takes as input a random $B \in I^r_n$ before it begins making queries to $E^{\pm 1}$. The *preimage-finding advantage* of $\mathcal{A}$ is defined to be

$$\mathbf{Adv}^{\mathsf{pre}}_{F}(\mathcal{A}) = \mathsf{Pr}\left[\mathbf{Exp}^{\mathsf{pre}}_{\mathcal{A}} = 1\right]. \tag{5}$$

For $q > 0$, $\mathbf{Adv}^{\mathsf{pre}}_{F}(q)$ is the maximum of $\mathbf{Adv}^{\mathsf{pre}}_{F}(\mathcal{A})$ over all adversaries $\mathcal{A}$ making at most q queries.

```
Experiment $\mathbf{Exp}^{\mathsf{coll}}_{\mathcal{A}}$
  $E \xleftarrow{\$} BC(n,k)$
  $\mathcal{A}^{E,E^{-1}}$ updates $\mathcal{Q}$
  if $\exists\ A \neq A'$, $B$ s.t. $A \vdash_{\mathcal{Q}} B$ and $A' \vdash_{\mathcal{Q}} B$ then
    output 1
  else
    output 0
```

(a) Collision resistance

```
Experiment $\mathbf{Exp}^{\mathsf{pre}}_{\mathcal{A}}$
  $E \xleftarrow{\$} BC(n,k)$
  $B \xleftarrow{\$} I_n^r$
  $\mathcal{A}(B)^{E,E^{-1}}$ updates $\mathcal{Q}$
  if $\exists\ A$ s.t. $A \vdash_{\mathcal{Q}} B$ then
    output 1
  else
    output 0
```

(b) Preimage resistance

Fig. 2. Experiments for quantification of collision resistance and preimage resistance

Adaptive Preimage Resistance The adaptive preimage resistance of F is quantified using the experiment $\mathbf{Exp}^{\mathsf{apre}}_{\mathcal{A}}$ described in Figure 3. At any point during the experiment, the adversary $\mathcal{A}$ can choose a "commitment" point $B \in I_n^r \backslash \mathsf{Range}_F(\mathcal{Q})$, where

$$\mathsf{Range}_F(\mathcal{Q}) = \{B \in I_n^r : A \vdash_{\mathcal{Q}} B \text{ for some } A \in I_n^m\}.$$

Then the experiment $\mathbf{Exp}^{\mathsf{apre}}_{\mathcal{A}}$ records the point B into a *commitment list* $\mathcal{L} \subset I_n^r$. At the end of the experiment, $\mathcal{A}$ would like to succeed in finding a preimage of some element in the commitment list. Now the *adaptive preimage-finding advantage* of $\mathcal{A}$ is defined to be

$$\mathbf{Adv}^{\mathsf{apre}}_F(\mathcal{A}) = \mathsf{Pr}\left[\mathbf{Exp}^{\mathsf{apre}}_{\mathcal{A}} = 1\right]. \tag{6}$$

For $q_1, q_2 > 0$, we define $\mathbf{Adv}^{\mathsf{apre}}_F(q_1, q_2)$ as the maximum of $\mathbf{Adv}^{\mathsf{apre}}_F(\mathcal{A})$ over all adversaries $\mathcal{A}$ that make at most q_1 queries to E and E^{-1} and make at most q_2 commitments.

```
Experiment $\mathbf{Exp}^{\mathsf{apre}}_{\mathcal{A}}$
  $E \xleftarrow{\$} BC(n,k)$
  $\mathcal{A}^{E,E^{-1}}$ updates $\mathcal{Q}$ and $\mathcal{L}$ (in an arbitrarily interleaved order)
  if $\exists\ A$ such that $A \vdash_{\mathcal{Q}} B$ for some $B \in \mathcal{L}$ then
    output 1
  else
    output 0
```

Fig. 3. Experiment for quantification of adaptive preimage resistance

Here we state two important properties of adaptive preimage resistance without proof. We refer to [15] for details.

Theorem 1. *Let F be a compression function, and let MD^F be the Merkle-Damgård transform based on the function F. Then it holds that*

$$\mathbf{Adv}^{\mathsf{apre}}_{MD^F}(q) \leq \mathbf{Adv}^{\mathsf{coll}}_{F}(q) + \mathbf{Adv}^{\mathsf{apre}}_{F}(q).$$

Theorem 2. *For $V \subset I^*$, let $H : V \to I_r$ be a hash/compression function that uses at most L calls to the underlying ideal primitives. If $G : I_r \to I_s$ is a public random function, then the composite function $G \circ F$ is $(q/L, \epsilon)$-indifferentiable from a public random function $H : V \to I_s$, where*

$$\epsilon = \mathbf{Adv}^{\mathsf{coll}}_{F}(q) + \mathbf{Adv}^{\mathsf{apre}}_{F}(q).$$

3 Security of Abreast-DM

3.1 Query-response Cycle and Modified Adversary

Let $F := F^{ABR}$ be the compression function Abreast-DM based on a blockcipher $E \in BC(n, 2n)$, and let $Q_1, \ldots, Q_6$ be query-response pairs obtained by oracle access to E and E^{-1}. If the 6-tuple $\Delta = (Q_1, \ldots, Q_6) \in \mathcal{Q}^6$ satisfies

$$\begin{aligned} Q_1 &= (A_1, A_2||A_3, Y_1), & Q_2 &= (\overline{A_2}, A_3||A_1, Y_2), & Q_3 &= (\overline{A_3}, A_1||\overline{A_2}, Y_3), \\ Q_4 &= (\overline{A_1}, \overline{A_2}||\overline{A_3}, Y_4), & Q_5 &= (A_2, \overline{A_3}||\overline{A_1}, Y_5), & Q_6 &= (A_3, \overline{A_1}||A_2, Y_6), \end{aligned}$$

for some A_i's and Y_i's, then it is called a *query-response cycle*(or simply a cycle). Observe that the first three blocks of the query-response pairs are moving cyclically under the permutation

$$\begin{aligned} \pi : I_n^3 &\longrightarrow I_n^3 \\ (A_1, A_2, A_3) &\longmapsto (\overline{A_2}, A_3, A_1). \end{aligned}$$

We state some useful properties of query-response cycles as follows.

Property 1. For query-response cycles Δ and Δ', either $\Delta = \Delta'$ or $\Delta \cap \Delta' = \emptyset$.

Property 2. For a query-response cycle $\Delta = (Q_1, \ldots, Q_6)$, either

- Q_i's are all distinct, or
- $Q_1 = Q_3 = Q_5 = (A_1, A_1, \overline{A_1})$ and $Q_2 = Q_4 = Q_6 = (\overline{A_1}, A_1, A_1)$.

Property 3. If Q_i is used as the first blockcipher call in an evaluation of F, then the second query-response pair should be Q_{i+1}. If Q_i is used as the second blockcipher call, then the first query-response pair should be Q_{i-1}. Moreover, Q_i and Q_{i+1} are always distinct. Here, the subscripts are interpreted up to modulo 6. The evaluations of F determined by Q_i and Q_{i+1}, $i = 1, \ldots, 6$, are as follows.

$$\begin{aligned} &(A_1, A_2, A_3) \vdash_{Q_1,Q_2} (A_1 \oplus Y_1, A_2 \oplus Y_2), && (\overline{A_2}, A_3, A_1) \vdash_{Q_2,Q_3} (\overline{A_2} \oplus Y_2, A_3 \oplus Y_3), \\ &(\overline{A_3}, A_1, \overline{A_2}) \vdash_{Q_3,Q_4} (\overline{A_3} \oplus Y_3, A_1 \oplus Y_4), && (\overline{A_1}, \overline{A_2}, \overline{A_3}) \vdash_{Q_4,Q_5} (\overline{A_1} \oplus Y_4, \overline{A_2} \oplus Y_5), \\ &(A_2, \overline{A_3}, \overline{A_1}) \vdash_{Q_5,Q_6} (A_2 \oplus Y_5, \overline{A_3} \oplus Y_6), && (A_3, \overline{A_1}, A_2) \vdash_{Q_6,Q_1} (A_3 \oplus Y_6, \overline{A_1} \oplus Y_1). \end{aligned}$$

Given an adversary $\mathcal{A}$ with oracle access to E and E^{-1}, one can transform $\mathcal{A}$ into an adversary $\mathcal{B}$ that records its query history in terms of query-response cycles. The modified adversary $\mathcal{B}$ is described in Figure 4. We can easily check the following properties of $\mathcal{B}$.

Property 4. If $\mathcal{A}$ makes at most q queries, then the corresponding adversary $\mathcal{B}$ makes at most $6q$ queries, and records at most q query-response cycles.

Property 5. $\mathbf{Adv}_F^{\mathsf{sec}}(\mathcal{A}) \leq \mathbf{Adv}_F^{\mathsf{sec}}(\mathcal{B})$ for $\mathsf{sec} \in \{\mathsf{coll}, \mathsf{pre}, \mathsf{apre}\}$.

```
Algorithm $\mathcal{B}^{E,E^{-1}}$
  $\mathcal{Q}_\Delta \leftarrow \emptyset$
  Run $\mathcal{A}$
  if $\mathcal{A}$ makes a fresh query $E(A_1, A_2||A_3)$ then
      Make queries

      $Y_1 = E(A_1, A_2||A_3)$,   $Y_2 = E(\overline{A_2}, A_3||A_1)$,   $Y_3 = E(\overline{A_3}, A_1||\overline{A_2})$,
      $Y_4 = E(\overline{A_1}, \overline{A_2}||\overline{A_3})$,   $Y_5 = E(A_2, \overline{A_3}||\overline{A_1})$,   $Y_6 = E(A_3, \overline{A_1}||A_2)$,

      $\mathcal{Q}_\Delta \leftarrow \mathcal{Q}_\Delta \cup \{\Delta\}$ ($\Delta$=the cycle defined by the above six queries)
      Return $Y_1$ to $\mathcal{A}$
  else if $\mathcal{A}$ makes a fresh query $E^{-1}(Y_1, A_2||A_3)$ then
      Make queries

      $A_1 = E^{-1}(Y_1, A_2||A_3)$,   $Y_2 = E(\overline{A_2}, A_3||A_1)$,   $Y_3 = E(\overline{A_3}, A_1||\overline{A_2})$,
      $Y_4 = E(\overline{A_1}, \overline{A_2}||\overline{A_3})$,   $Y_5 = E(A_2, \overline{A_3}||\overline{A_1})$,   $Y_6 = E(A_3, \overline{A_1}||A_2)$,

      $\mathcal{Q}_\Delta \leftarrow \mathcal{Q}_\Delta \cup \{\Delta\}$
      Return $A_1$ to $\mathcal{A}$
  else
      Return the response using query history $\mathcal{Q}_\Delta$
```

Fig. 4. Modified algorithm $\mathcal{B}$. A query is called "fresh" if its response is not obtained from the query history of $\mathcal{B}$

3.2 Security Results

Given Property 5, we will analyze the security of the compression function Abreast-DM with respect to the modified adversary $\mathcal{B}$. Without loss of generality, we might assume that $\mathcal{B}$ makes exactly $6q$ queries (including redundant queries in a same cycle), and records q query-response cycles. The query history of $\mathcal{B}$ is denoted

$$\mathcal{Q}_\Delta = \{\Delta^i : i = 1, \ldots, q\},$$

where $\Delta^i = (Q_1^i, Q_2^i, Q_3^i, Q_4^i, Q_5^i, Q_6^i)$ and Q_j^i is the $(6(i-1)+j)$-th query-response pair for $1 \leq i \leq q$ and $1 \leq j \leq 6$.

Collision Resistance Let $\mathcal{E}$ denote the event that $\mathcal{B}$ makes a collision of F. Then, by definition, $\mathbf{Adv}_F^{\mathsf{coll}}(\mathcal{B}) = \mathsf{Pr}\left[\mathcal{E}\right]$. In order to estimate $\mathsf{Pr}\left[\mathcal{E}\right]$, we decompose $\mathcal{E}$ as follows.

$$\mathcal{E} = \bigcup_{i=1}^{q} \left(\mathcal{E}^i \cup \bigcup_{j=1}^{i-1} \mathcal{E}^{i,j} \right), \tag{7}$$

where

$$\mathcal{E}^i \Leftrightarrow \text{two evaluations from a single cycle } \Delta^i \text{ determines a collision,} \tag{8}$$

$$\mathcal{E}^{i,j} \Leftrightarrow \text{two evaluations from } \Delta^i \text{ and } \Delta^j \text{ determine a collision.} \tag{9}$$

Then it follows that

$$\mathsf{Pr}\left[\mathcal{E}\right] = \sum_{i=1}^{q} \left(\mathsf{Pr}\left[\mathcal{E}^i\right] + \sum_{j=1}^{i-1} \mathsf{Pr}\left[\mathcal{E}^{i,j}\right] \right). \tag{10}$$

Lemma 1. *Let $N' = 2^n - q$ and $1 \leq j < i \leq q$. Then,*

1. $\mathsf{Pr}\left[\mathcal{E}^i\right] \leq 1/N'$,

2. $\mathsf{Pr}\left[\mathcal{E}^{i,j}\right] \leq 36/\left(N'\right)^2$.

Proof. Inequality 1: First, assume that Δ^i consists of two distinct query-response pairs. A collision within this cycle implies that $Q_1^i = (A_1, A_1||\overline{A_1}, Y_1)$, $Q_2^i = (\overline{A_1}, A_1||A_1, Y_2)$ and $(A_1 \oplus Y_1, A_1 \oplus Y_2) = (\overline{A_1} + Y_2, \overline{A_1} + Y_1)$. Since the second query-response pair Q_2^i is obtained by a forward query and Y_2 should be equal to $\overline{Y_1}$, the probability that this case occurs is not greater than $1/N'$.

Next, assume that Δ^i consists of six distinct query-response pairs. Suppose that (Q_1^i, Q_2^i) and (Q_2^i, Q_3^i) makes a collision. As seen in Property 3, it should be the case that $(A_1 \oplus Y_1, A_2 \oplus Y_2) = (\overline{A_2} \oplus Y_2, A_3 \oplus Y_3)$. In this case, we have $Y_2 = A_1 \oplus Y_1 \oplus \overline{A_2}$ and $Y_3 = A_2 \oplus Y_2 \oplus A_3$. The probability that Y_2 and Y_3 satisfy these equations is not greater than $\left(1/N'\right)^2$. The same argument applies to every pair of (Q_h^i, Q_{h+1}^i) and $(Q_{h'}^i, Q_{h'+1}^i)$. Since the number of such pairs is $\binom{6}{2} = 15$ and $15/\left(N'\right)^2 \leq 1/N'$ for a sufficiently large N', the first inequality is proved.

Inequality 2: Cycle Δ^j determines at most six evaluations of F. For a fixed $1 \leq h' \leq 6$, let

$$(A'_1, A'_2, A'_3) \vdash_{Q^j_{h'}, Q^j_{h'+1}} (B_1, B_2).$$

The probability that

$$(A_1, A_2, A_3) \vdash_{Q_1^i, Q_2^i} (A_1 \oplus Y_1, A_2 \oplus Y_2) = (B_1, B_2)$$

is not greater than $\left(1/N'\right)^2$. The same argument applies to (Q_h^i, Q_{h+1}^i) for $h = 2, \ldots, 6$. It completes the proof of the second inequality. $\square$

By Lemma 1, equality (10) and Property 5, we obtain the following theorem.

Theorem 3. *Let F^{ABR} be the compression function* Abreast-*DM. Then,*

$$\mathbf{Adv}_{F^{ABR}}^{\mathsf{coll}}(q) \leq \frac{q}{(2^n - q)} + \frac{18q^2}{(2^n - q)^2}.$$

Preimage Resistance Suppose that a modified adversary $\mathcal{B}$ is given an image point $B = (B_1, B_2)$. Let $\mathcal{E}$ denote the event that $\mathcal{B}$ makes an evaluation $F(A_1, A_2, A_3) = (B_1, B_2)$ for some A_i's. Then, by definition, $\mathbf{Adv}_F^{\mathsf{pre}}(\mathcal{B}) = \mathsf{Pr}[\mathcal{E}]$. Define

$$\mathcal{E}^i \Leftrightarrow \Delta^i \text{ determines a preimage of } B. \tag{11}$$

Then it follows that

$$\mathsf{Pr}[\mathcal{E}] = \sum_{i=1}^{q} \mathsf{Pr}\left[\mathcal{E}^i\right]. \tag{12}$$

Consider the case where $Q_1^i = (A_1, A_2||A_3, Y_1)$ and $Q_2^i = (\overline{A_2}, A_3||A_1, Y_2)$ determine

$$F(A_1, A_2, A_3) = (A_1 \oplus Y_1, A_2 \oplus Y_2) = (B_1, B_2).$$

This event occurs with probability at most $(1/N')^2$ for $N' = 2^n - q$. Since each cycle determines at most six evaluations of F, we obtain $\mathsf{Pr}\left[\mathcal{E}^i\right] \le 6/(N')^2$ for $1 \le i \le q$, and the following theorem.

Theorem 4. *Let F^{ABR} be the compression function* Abreast-*DM. Then,*

$$\mathbf{Adv}_{F^{ABR}}^{\mathsf{pre}}(q) \le \frac{6q}{(2^n - q)^2}.$$

Adaptive Preimage Resistance For $1 \le i \le q_1$ and $1 \le j \le 6$, consider the event $\mathcal{E}^{i,j}$ that the j-th query-response pair of the i-th cycle determines a preimage of a commitment point $B = (B_1, B_2) \in \mathcal{L}$. If $j = 1$, then $\mathsf{Pr}\left[\mathcal{E}^{i,j}\right] = 0$ since the first query of a cycle does not determine any evaluation of F. Next, assume that $j = 2$. In order for Q_1^i and Q_2^i to determine an evaluation $F(A_1, A_2, A_3) = (A_1 \oplus Y_1, A_2 \oplus Y_2) = (B_1, B_2)$, Y_2 should be equal to $A_2 \oplus B_2$. This event occurs with probability at most $1/N'$ for $N' = 2^n - q_1$. Note that the adversary $\mathcal{B}$ might choose $B_1 = A_1 \oplus Y_1$ after the first query. The same argument applies to $j = 3, 4, 5$, and a similar argument shows that $\mathsf{Pr}\left[\mathcal{E}^{i,j}\right] \le 2/N'$ for $j = 6$. Since $|\mathcal{L}| \le q_2$, we obtain the following theorem.

Theorem 5. *Let F^{ABR} be the compression function* Abreast-*DM. Then,*

$$\mathbf{Adv}_{F^{ABR}}^{\mathsf{apre}}(q_1, q_2) \le \frac{6q_1q_2}{(2^n - q_1)^2}.$$

4 Abreast-DM Variants

In this section, we present a wide range of Abreast-DM variants that enjoy a birthday-type security guarantee. Let π be a permutation on $I_n^3 (\equiv I_n \times I_n^2)$ such that every cycle in π is of length $2 \le l \le L$ for a positive integer L. Then we define

$$\begin{aligned} F_\pi^{ABR} : I_n^3 &\longrightarrow I_n^2 \\ (A_1, A_2, A_3) &\longmapsto (E(X_1, K_1) \oplus X_1, E(X_2, K_2) \oplus X_2), \end{aligned} \tag{13}$$

where $(X_1, K_1) = (A_1, A_2||A_3)$ and $(X_2, K_2) = \pi(A_1, A_2, A_3)$. By the same argument as the previous section, we can prove the following theorem.

Theorem 6. *Let* F_π^{ABR} *be the compression function defined in (13), and let* $2^n \geq q + \binom{L}{2}$*. Then,*

$$\mathbf{Adv}^{\mathsf{coll}}_{F_\pi^{ABR}}(q) \leq \frac{q}{(2^n - q)} + \frac{L^2 q^2}{2(2^n - q)^2},$$

$$\mathbf{Adv}^{\mathsf{pre}}_{F_\pi^{ABR}}(q) \leq \frac{Lq}{(2^n - q)^2},$$

$$\mathbf{Adv}^{\mathsf{apre}}_{F_\pi^{ABR}}(q_1, q_2) \leq \frac{Lq_1 q_2}{(2^n - q_1)^2}.$$

If π *contains no cycle of length* 2*, then*

$$\mathbf{Adv}^{\mathsf{coll}}_{F_\pi^{ABR}}(q) \leq \frac{L^2(q + q^2)}{2(2^n - q)^2}.$$

Example 1. Let $\pi : (A_1, A_2, A_3) \mapsto (A_1 \oplus C, A_2, A_3)$ for a constant $C \in I_n$. Then F_π^{ABR} is reduced to the Hirose's scheme [12].

Example 2. Let $\pi : (A_1, A_2, A_3) \mapsto (\overline{A_1}, A_3, \overline{A_2})$. Then every cycle in π is of length 4. By Theorem 6, we have

$$\mathbf{Adv}^{\mathsf{coll}}_{F_\pi^{ABR}}(q) \leq \frac{8(q + q^2)}{(2^n - q)^2}.$$

In numerical terms with $n = 128$, any adversary asking less than $2^{125.67}$ queries cannot find a collision with probability greater than 1/2.

5 Conclusion

In this paper, we have analyzed the security of Abreast-DM in terms of collision resistance, preimage resistance and adaptive preimage resistance. The bounds on the number of queries for collision resistance and preimage resistance are given by $O(2^n)$ and $O(2^{2n})$, respectively. The adaptive preimage resistance is guaranteed up to $O(2^n)$ queries/commitments. We presented a wide range of Abreast-DM variants that enjoy a birthday-type security guarantee. The variants include the Hirose's scheme as a special case. It would be an interesting open problem whether our approach could apply to more complicated constructions such as MDC-2 and MDC-4.

References

1. E. Biham, R. Chen, A. Joux, P. Carribault, C. Lemuet and W. Jalby. Collisions of SHA-0 and reduced SHA-1. Eurocrypt 2005, LNCS 3494, pp. 36–57, Springer-Verlag, 2005.
2. J. Black, M. Cochran and T. Shrimpton. On the impossibility of highly-efficient blockcipher-based hash functions. Eurocrypt 2005, LNCS 3494, pp. 526–541, Springer-Verlag, 2005.

3. A. Bogdanov, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw and Y. Seurin. Hash functions and RFID tags: mind the gap. CHES 2008, LNCS 5154, pp. 283–299, Springer-Verlag, 2008.
4. J. Black, P. Rogaway and T. Shrimpton. Black-box analysis of the block-cipher-based hash-function construction from PGV. Crypto 2002, LNCS 2442, pp. 320–325, Springer-Verlag, 2002.
5. B. Brachtl, D. Coppersmith, M. Heyden, 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. US Patent #4,908,861. Awarded March 13, 1990 (filed August 28, 1987).
6. De. Canniere and C. Rechberger. Preimages for reduced SHA-0 and SHA-1. Crypto 2008, LNCS 5157, pp. 179–202, Springer-Verlag, 2008.
7. Y. Dodis, T. Ristenpart and T. Shrimpton. Salvaging Merkle-Damgård for practical applications. Eurocrypt 2009, To appear. Available at http://www.cs.nyu/ dodis.
8. E. Fleischmann, M. Gorski and S. Lucks. On the security of TANDEM-DM. Preproceedings of FSE 2009, pp. 85–105, 2009.
9. M. Hattori, S. Hirose and S. Yoshida. Analysis of double block length hash functions. IMA 2003, LNCS 2898, pp. 290–302, Springer-Verlag, 2003.
10. S. Hirose. A security analysis of double-block-length hash functions with the rate 1. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E89-A, NO. 10, pp. 2575–2582, 2006.
11. S. Hirose. Provably secure double-block-length hash functions in a black-box model. ICISC 2004, LNCS 3506, pp. 330–342, Springer-Verlag, 2005.
12. S. Hirose. Some plausible construction of double-block-length hash functions. FSE 2006, LNCS 4047, pp. 210–225, Springer-Verlag, 2006.
13. L. R. Knudsen, J. L. Massey and B. Preneel. Attacks on fast double block length hash functions. Journal of Cryptology, Vol. 11, NO. 1, pp. 59–72, 1998.
14. X. Lai and J. L. Massey. Hash function based on block ciphers. Eurocrypt 1992, LNCS 658, pp. 55–70, Springer-Verlag, 1993.
15. J. Lee and J. H. Park. Adaptive preimage resistance and permutation-based hash functions. Available at http://eprint.iacr.org/2009/066.
16. G. Leurent. MD4 is not one-way. FSE 2008, LNCS 5086, pp. 412–428, Springer-Verlag, 2008.
17. F. Mendel, N. Pramstaller, C. Rechberger and V. Rijmen. Analysis of step-reduced SHA-256. FSE 2006, LNCS 4047, pp. 126–143, Springer-Verlag, 2006.
18. B. Preneel, R. Govaerts and J. Vandewalle. Hash functions based on block ciphers: A synthetic approach. Crypto 1993, LNCS 773, pp. 368–378, Springer-Verlag, 1994.
19. T. Ristenpart and T. Shrimpton. How to build a hash function from any collision-resistant function. Asiacrypt 2007, LNCS 4833, pp. 147–163, Springer-Verlag, 2007.
20. P. Rogaway and J. Steinberger. Constructing cryptographic hash functions from fixed-key blockciphers. Crypto 2008, LNCS 5157, pp. 433–450, Springer-Verlag, 2008.
21. P. Rogaway and J. Steinberger. Security/efficiency tradeoffs for permuation-based hashing. Eurocrypt 2008, LNCS 4965, pp. 220–236, Springer-Verlag, 2008.
22. Y. Sasaki and K. Aoki. Finding preimages in full MD5 faster than exhaustive search. Eurocrypt 2009, LNCS 5479, pp. 134–152, Springer-Verlag, 2008.
23. T. Shrimpton and M. Stam. Building a collision-resistant function from non-compressing primitives. ICALP 2008, LNCS 5126, pp. 643–654, Springer-Verlag, 2008.
24. M. Stam. Beyond uniformity: Better security/efficiency tradeoffs for compression functions. Crypto 2008, LNCS 5157, pp. 397–412, Springer-Verlag, 2008.

25. J. Steinberger. The collision intractability of MDC-2 in the ideal-cipher model. Eurocrypt 2007, LNCS 4515, pp. 34–51, Springer-Verlag, 2008.
26. X. Wang, X. Lai, D. Feng, H. Chen and X. Yu. Cryptanalysis of the hash functions MD4 and RIPEMD. Eurocrypt 2005, LNCS 3494, pp. 1–18, Springer-Verlag, 2005.
27. X. Wang, X. Lai and H. Yu. Finding collisions in the full SHA-1. Crypt0 2005, LNCS 3621, pp. 17–36, Springer-Verlag, 2005.
28. X. Wang and H. Yu. How to break MD5 and other hash functions. Eurocrypt 2005, LNCS 3494, pp. 19–35, Springer-Verlag, 2005.