Alpha-DBL: A Reasonable High Secure Double-Block-Length Hash Function

Hoang Dinh Linh, Tran Hong Thai

***Abstract*—We propose a new double-block-length compression function which is called Alpha-DBL. This scheme uses two parallel secure single block length schemes based on a block cipher with $2n$-bit key and n-bit block size to compress a $3n$-bit string to a $2n$-bit one. We show that the Alpha-DBL scheme attains nearly optimal collision security and preimage security bounds (up to 2^n and 2^{2n} queries for finding a collision and a preimage, respectively). More precisely, for $n = 128$, no adversary making less than $2^{n-1.27} = 2^{126.73}$ queries can find a collision with probability greater than 1/2. To our knowledge, this collision security bound is nearly better than other such compression functions. In addition, we provide a preimage security analysis of Alpha-DBL that shows security bound of $2^{2n-5} = 2^{251}$ queries for $n = 128$. Using this scheme in the iterated hash function construction can preserve the collision resistance security and the preimage resistance security.**

***Tóm tắt*—Chúng tôi đề xuất một hàm nén độ dài khối kép mới được gọi là Alpha-DBL. Lược đồ này sử dụng hai lược đồ độ dài khối đơn an toàn song song dựa trên mã khối với khóa $2n$-bit và kích thước khối n-bit để nén chuỗi $3n$-bit thành chuỗi $2n$-bit. Chúng tôi đã chứng minh rằng, lược đồ Alpha-DBL đạt được cận an toàn kháng va chạm và kháng tiền ảnh gần như tối ưu (tối đa 2^n và 2^{2n} truy vấn tương ứng để tìm va chạm và tiền ảnh). Cụ thể với $n = 128$, một kẻ tấn công bất kỳ thực hiện ít hơn $2^{n-1.27} = 2^{126.73}$ truy vấn chỉ có thể tìm thấy một va chạm với xác suất nhỏ hơn 1/2. Theo hiểu biết của chúng tôi, cận an toàn kháng va chạm này tốt hơn so với các hàm nén khác. Ngoài ra, chúng tôi đã đưa ra phân tích độ an toàn kháng tiền ảnh của Alpha-DBL cho thấy cận an toàn là $2^{2^{2n-5}} = 2^{2^{251}}$ truy vấn cho $n = 128$. Sử dụng lược đồ này trong việc xây dựng hàm băm được lặp có thể bảo toàn độ an toàn kháng va chạm và an toàn kháng tiền ảnh.**

This manuscript is received on December 23, 2020. It is commented on December 24, 2020 and is accepted on December 24, 2020 by the first reviewer. It is commented on December 30, 2020 and is accepted on December 30, 2020 by the second reviewer.

***Keywords*—double-block-length compression function, collision security, preimage security, ideal cipher model.**

***Từ khóa*—hàm nén độ dài khối kép, độ an toàn kháng va chạm, độ an toàn kháng tiền ảnh, mô hình mã pháp lý tưởng.**

I. INTRODUCTION

A cryptographic hash function is a function which takes an input of arbitrary length and returns an output of fixed length. A general way of hashing messages of arbitrary length is to repeat a compression function using some general structures, e.g. Merkle-Damgard, HAIFA... A base compression function can be built from a mishmash of components or based on cryptographic primitives such as block ciphers.

Block cipher-based compression functions have been extensively studied. The most common approach is to build a $2n$-bit to n-bit compression function using a block cipher of n-bit block length, namely a single-block-length (SBL) compression function. However, such an SBL compression function may be susceptible to collision attacks because of its short output length. For example, we can successfully execute a birthday attack on an SBL compression function based on the AES-128 that only approximates 2^{64} queries. This prompted the study of double-block-length (DBL) compression functions which have the output length double the block length of the base block cipher.

DBL compression function can be classified into 2 classes: The first class are compression functions that use a block cipher with the key length of n-bit, i.e. $E: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$, denoted by DBL^n. The second class are compression functions that use a block cipher with the key length of $2n$-bit, i.e. $E: \{0,1\}^{2n} \times \{0,1\}^n \to \{0,1\}^n$, denoted by DBL^{2n}. This class consists of Tandem-DM [1] and Abreast-DM [1], Hirose's scheme [2], Stam's Type-I compression function [3] and general constructions of Hirose [4] and Ozen and Stam construction [5]. All the above compression functions provide optimal

collision security (up to 2^n queries), Tandem-DM, Abreast-DM and Hirose's scheme have also proven to be optimal preimage resistance (up to 2^{2n} queries).

Recently, there have been some proposed compression schemes such as Weimar-DM [6], and MR-MMO [7]. The Weimar-DM scheme uses two different keys for two block ciphers in a compression function call and is proven secure in the ideal cipher model (ICM). The MR-MMO scheme claimed by the author is more effective when using only one key for both block ciphers in a compression function call, but is proven in the weak cipher model (WCM). The MR-MMO is claimed that its collision resistance security bound is tighter than Weimar-DM's one but it is not valid in our understanding of the matter. The first reason is that Weimar-DM is considered in the ICM while MR-MMO is studied in the WCM. Secondly, there is an incorrect statement in the proof for MR-MMO's collision security bound. More precisely, in [7] the authors stated that:

$$\Pr[C_j] \le \frac{2(j-1)}{\left(2^n-(2j-1)\right)^2} \le \frac{2(j-1)}{2^{2n}},$$

for $j \le q$. It is clear that this statement is wrong which implies that the collision resistance security bound is incorrect. On the other hand, the two schemes are in the class of cyclic compression functions [8], which have been shown to be secure generally.

TABLE 1. THE ANALYSIS RESULTS OF SOME DOUBLE BLOCK LENGTH COMPRESSION FUNCTIONS [6], [8]-[10] THAT USES BLOCK CIPHER WITH THE KEY LENGTH OF 256 BITS AND THE BLOCK LENGTH OF 128 BITS

Scheme	Collision resistance	Preimage resistance	Key schedule
Alpha-DBL	$2^{126.73}$	2^{251}	2
Weimar-DM	$2^{126.23}$	2^{251}	2
Hirose-DM	$2^{124.55}$	2^{251}	1
Abreast-DM	$2^{124.42}$	2^{246}	2
Tandem-DM	$2^{120.87}$	2^{246}	2

In this paper, we propose a new compression scheme and demonstrate its security under the ICM. The rest of the paper is structured as follows: *Section II* presents some basic concepts about the iterated hash functions and the ideal cipher model. *Section III* presents the definition of the Alpha-DBL scheme. *Section IV* analyzes the collision resistance and preimage resistance of the proposed scheme. Finally, the conclusion is presented in *Section V*.

II. PRELIMINARIES

A. Iterated hash function

Let $H:\{0,1\}^* \to \{0,1\}^l$ be a hash function which often consists of a compression function $F:\{0,1\}^l \times \{0,1\}^{l'} \to \{0,1\}^l$ and a fixed initial value $H_0 \in \{0,1\}^l$. An input message M (after unambiguous padding) is divided into the l'-bit blocks $M_1, M_2, \cdots, M_l$. Then, $H_i = F(H_{i-1}, M_i)$ is computed successively for $1 \le i \le l$ and $H_l = H(M)$. H is called an iterated hash function.

B. Ideal cipher model

A (m,n) block cipher is a function $E:\{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$ such that $E(K,\cdot)$ is a permutation on $\{0,1\}^n$ for $K \in \{0,1\}^m$. Namely, m is the key length and and n is the block length of the block cipher E. Normally, we write $E_K(X)$ instead of $E(K,X)$ for $K \in \{0,1\}^m, X \in \{0,1\}^n$. Let $E_K^{-1}(\cdot)$ denotes the inverse of $E_K(\cdot)$.

Ideal cipher model. Let m,n be positive integers, denote

$$BC(m,n) = \left\{E:\{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n \middle| \forall K \in \{0,1\}^m, \atop E_K(\cdot) \text{ be a permutation on } \{0,1\}^n\right\}$$

In the ideal cipher model, a block cipher E is randomly chosen from $BC(m,n)$. It's allowed 2 types of query to $E_K(X)$ or $E_K^{-1}(Y)$ for $X,Y \in \{0,1\}^n, K \in \{0,1\}^m$. X, Y and K are plaintext, ciphertext and key, respectively. The answer of a backward query $E_K^{-1}(Y)$ is $X \in \{0,1\}^n$ such that $E_K(X) = Y$.

In this paper, we only study the case $m = 2n$ and denote $N = 2^n$.

Advantages of collision and preimage resistance. Let $F:\{0,1\}^{3n} \to \{0,1\}^{2n}$ be a compression function based on an ideal block cipher $E \in BC(2n,n)$, and let $\mathcal{A}$ be an information-theoretic adversary who has access to an oracle E or E^{-1}. Then, it carries out the experiment $\mathbf{Exp}_{\mathcal{A}}^{Coll}$ as illustrated in Table 2a, in order to quantify the collision resistance security of F. The experiment will record all the queries

made by the adversary $\mathcal{A}$ in a query history, denoted by Q. A tuple $(X, K, Y) \in Q$ if $\mathcal{A}$ ask $E_K(X)$ and receive an answer Y or ask $E_K^{-1}(Y)$ and receive an answer X. For $A \in \{0,1\}^{3n}, B \in \{0,1\}^{2n}$, we write $A \Rightarrow_Q B$ if there exists a query pair $(X_1, K_1, Y_1), (X_2, K_2, Y_2) \in Q$ such that $\mathcal{A}$ have the computation $F(A) = B$ using this query pair.

The advantage of $\mathcal{A}$ finding a collision is defined as

$$Adv_F^{Coll}(\mathcal{A}) = \Pr[\mathbf{Exp}_{\mathcal{A}}^{Coll} = 1].$$

The probability is taken over random block cipher E. For $q > 0$, we define $Adv_F^{Coll}(q)$ be the maximum of $Adv_F^{Coll}(\mathcal{A})$ over all adversaries that ask at most q oracle queries.

TABLE 2. EXPERIMENTS DETERMINE THE ADVANTAGES OF COLLISION AND PREIMAGE RESISTANCE

Experiment $\mathbf{Exp}_{\mathcal{A}}^{Coll}$	**Experiment $\mathbf{Exp}_{\mathcal{A}}^{Pre}$**
$E \xleftarrow{\$} BC(2n, n)$ $\mathcal{A}^{E,E^{-1}}$ update Q If $\exists A \neq A', B$ such that $A \Rightarrow_Q B$ and $A' \Rightarrow_Q B$ then return 1 else return 0	$E \xleftarrow{\$} BC(2n, n)$ $\mathcal{A}$ chooses $B \in \{0,1\}^{2n}$ $\mathcal{A}(B)^{E,E^{-1}}$ update Q If $\exists A$ such that $A \Rightarrow_Q B$ then return 1 else return 0
(a) Experiment for finding a collision	(b) Experiment for finding a preimage

The advantage of $\mathcal{A}$ finding a preimage is defined similarly using the experiment $\mathbf{Exp}_{\mathcal{A}}^{Pre}$ as in Table 2b. The adversary $\mathcal{A}$ chooses an image target $B \in \{0,1\}^{2n}$ before it asks queries. The advantage of $\mathcal{A}$ finding a preimage is defined as

$$Adv_F^{Pre}(\mathcal{A}) = \Pr[\mathbf{Exp}_{\mathcal{A}}^{Pre} = 1].$$

The probability is taken over random block cipher E. For $q > 0$, we define $Adv_F^{Pre}(q)$ be the maximum of $Adv_F^{Pre}(\mathcal{A})$ over all adversaries that ask at most q oracle queries.

The advantage of $\mathcal{A}$ finding a collision/preimage of an iterated hash function is defined similarly.

In this model, the experiments make a decision based on the history of the adversary's queries to encryption/decryption oracles. However, the adversary may, without asking anything from the oracles, try to construct a collision or a preimage, for example, to guess. But in this case, the complexity of constructing a collision is greater or equal than $O(2^n)$ (the optimal bound) and a preimage is greater or equal than $O(2^{2n})$ because we don't know anything about the structure of the compression function.

III. DEFINITION OF ALPHA-DBL SCHEME

We propose a new DBL compression function that does not belong to the class of cyclic compression functions and demonstration its security in the ICM. The main idea of proof is according to in [6]. The proposed compression function uses two parallel secure single block length schemes, called Alpha-DBL (see Fig. 1, the name of the proposed scheme rises from its shape is like the symbol "α"), which are described as follows:

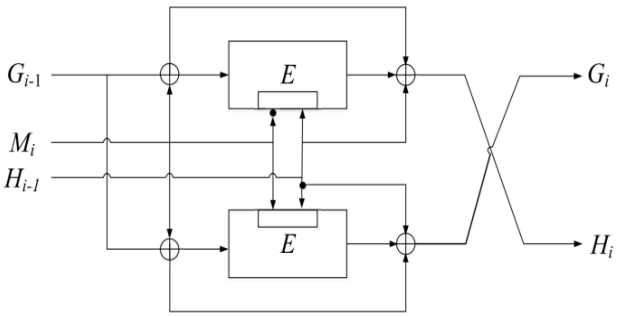

Fig. 1. The compression function Alpha-DBL. The black circle "•" denotes a bit complement.

Definition 1. *Let E be a block cipher which has an $2n$-bit key and an n-bit block size. Let $F^{Alpha}: \{0,1\}^{3n} \to \{0,1\}^{2n}$ be a compression function such that $(G_i, H_i) = F^{Alpha}(G_{i-1}, M_i, H_{i-1})$ where $G_i, H_i, M_i, G_{i-1}, H_{i-1} \in \{0,1\}^n$. F^{Alpha} is defined as follows:*

$$\begin{cases} G_i = E_{M_i||\bar{H}_{i-1}}(G_{i-1} \oplus M_i) \oplus G_{i-1} \oplus \bar{H}_{i-1} \oplus M_i \\ H_i = E_{\bar{M}_i||H_{i-1}}(G_{i-1} \oplus M_i) \oplus G_{i-1} \oplus H_{i-1} \oplus M_i \end{cases}$$

where $\bar{H}$ denotes the bit-by-bit complement of H.

IV. PROVABLE SECURITY OF ALPHA-DBL

A. Collision resistance security

Theorem 1. *Let $F^{Alpha}: \{0,1\}^{3n} \to \{0,1\}^{2n}$ be a compression function based on block cipher as defined in Definition 1. Then,*

$$Adv_{Alpha}^{Coll}(q) \leq \frac{q(q-1)}{(N-q)^2}.$$

Proof. Consider an arbitrary adversary $\mathcal{A}$ has made q queries to E or E^{-1} in order to attain a collision for the compression function F^{Alpha}. $\mathcal{A}$ will record a query history $Q = \{Q_i\}_{i=1}^{q}$, where $Q_i = (X_i, K_i, Y_i)$ such that $E_{K_i}(X_i) = Y_i$. Note that the adversary $\mathcal{A}$ never asks a query to which it already knows the answer. We build a more powerful adversary $\mathcal{A}'$ which copies $\mathcal{A}$ but it can ask an extra query to E in some cases. Therefore, we just need to find an upper bound of the advantage of $\mathcal{A}'$ finding a collision for F^{Alpha}.

The adversary $\mathcal{A}'$ maintains a list $\mathcal{L}$ (be null at the beginning) that represents any possible input/output of the compression function F^{Alpha} computed by adversary $\mathcal{A}$. An element $L \in \mathcal{L}$ is a quad-tuple $(X, K, Y, Y') \in \{0,1\}^{5n}$ where $X \in \{0,1\}^n, K \in \{0,1\}^{2n}$ is the $3n$-bit input to compression function such that $K = (\overline{M}_i, H_{i-1})$ and $X = G_{i-1} \oplus M_i$. The n-bit values Y, Y' can be computed by $Y = E_K(X)$ and $Y' = E_{\overline{K}}(X)$.

Let's define a collision in the list. Fix two integers a, b with $a \neq b$, such that $L_a = (X_a, K_a, Y_a, Y_a')$ represents the a-th element in $\mathcal{L}$ and $L_b = (X_b, K_b, Y_b, Y_b')$ is the b-th element in $\mathcal{L}$. We say that L_a and L_b "collide" if we can find a collision using the query results given in L_a and L_b. This event occurs if and only if one of the following two conditions is satisfied:

(i) $Y_a \oplus X_a \oplus Rightmost_n(K_a) = Y_b \oplus X_b \oplus Rightmost_n(K_b)$ and $Y_a' \oplus X_a \oplus Rightmost_n(\overline{K}_a) = Y_b' \oplus X_b \oplus Rightmost_n(\overline{K}_b)$,

(ii) $Y_a \oplus X_a \oplus Rightmost_n(K_a) = Y_b' \oplus X_b \oplus Rightmost_n(\overline{K}_b)$ and $Y_a' \oplus X_a \oplus Rightmost_n(\overline{K}_a) = Y_b \oplus X_b \oplus Rightmost_n(K_b)$,

where $Rightmost_n(K)$ and $Leftmost_n(K)$ are n bits farthest to the right and n bits farthest to the left of K, respectively.

Indeed, the condition (i) implies a collision pair $(G_a, H_a, M_a), (G_b, H_b, M_b)$ with

$H_a = Rightmost_n(K_a)$,
$M_a = Leftmost_n(\overline{K}_a), G_a = X_a \oplus M_a$,
$H_b = Rightmost_n(K_b)$,
$M_b = Leftmost_n(\overline{K}_b), G_b = X_b \oplus M_b$.

The condition (ii) implies a collision pair $(G_a, H_a, M_a), (G_b, H_b, M_b)$ with

$H_a = Rightmost_n(K_a)$,
$M_a = Leftmost_n(\overline{K}_a), G_a = X_a \oplus M_a$,
$H_b = Rightmost_n(\overline{K}_b)$,
$M_b = Leftmost_n(K_b), G_b = X_b \oplus M_b$.

Construction of the list: The adversary $\mathcal{A}$ will make a query number i to E or E^{-1} for $1 \leq i \leq q$. Then the adversary gets a triple-tuple (X_i, K_i, Y_i) such that $E_{K_i}(X_i) = Y_i$ in case of a forward query and $E_{K_i}^{-1}(Y_i) = X_i$ in case of a backward query. In either case, the value $X_i \oplus Y_i \oplus Rightmost_n(K_i)$ is randomly determined by the output of the query.

Now, $\mathcal{A}'$ checks if an entry $L = (X_i, K_i, *, *)$ or $L' = (X_i, \overline{K}_i, *, *)$ belongs to the recent list $\mathcal{L}$, where "$*$" is an arbitrary value. Obviously, there are 2 scenarios: both L, L' are not in $\mathcal{L}$, or both of them are already in $\mathcal{L}$. Indeed, if $L_i := (X_i, K_i, Y_i, Y_i') \in \mathcal{L}$ then we also have $L_i := (X_i, \overline{K}_i, Y_i', Y_i) \in \mathcal{L}$.

Scenario 1: If L or L' are not in $\mathcal{L}$. Then $\mathcal{A}'$ will make an additional forward query $Y_i' = E_{\overline{K}_i}(X_i)$. Since $\overline{K}_i \neq K_i$ for every $K_i \in \{0,1\}^n$ then the value of Y_i' is independently and randomly distributed with Y_i. Then, the adversary sets

$$L_i := (X_i, K_i, Y_i, Y_i')$$

and appends to the list $\mathcal{L}$.

Let $Success^i$, for $1 \leq i \leq q$, be the event that the i^{th} success, i.e. there exists $j < i$ such that L_i collide with L_j. For $1 \leq j < i$, we have:

$$\Pr\begin{bmatrix} X_i \oplus Y_i \oplus Rightmost_n(K_i) \\ = X_j \oplus Y_j \oplus Rightmost_n(K_j) \end{bmatrix} \leq \frac{1}{N-q}$$

and

$$\Pr\begin{bmatrix} X_i \oplus Y_i' \oplus Rightmost_n(\overline{K}_i) \\ = X_j \oplus Y_j' \oplus Rightmost_n(\overline{K}_j) \end{bmatrix} \leq \frac{1}{N-q}.$$

Since these above events are independent then the probability of condition (i) occurring is at most $\frac{1}{(N-q)^2}$. Similarly, the probability of condition (ii) occurring is at most $\frac{1}{(N-q)^2}$.

Therefore, the probability of success of the i^{th} query is

$$\Pr[Success^i] \leq \sum_{j=1}^{i-1} \frac{2}{(N-q)^2} = \frac{2(i-1)}{(N-q)^2}.$$

Thus, the total probability of success for q queries is

$$\Pr[Success(q)] \leq \sum_{i=1}^{q} \Pr\left[Success^i\right] \leq \frac{q(q-1)}{(N-q)^2}.$$

Scenario 2: Both L and L' are in $\mathcal{L}$. Therefore, $\mathcal{A}'$ ignores this query and we know that $\mathcal{A}$ has no chance of winning.

Therefore, the probability of the adversary $\mathcal{A}'$ success is

$$Adv_{F^{Alpha}}^{Coll}(\mathcal{A}') \leq \frac{q(q-1)}{(N-q)^2}.$$

Now, we return to evaluate the advantage of $\mathcal{A}$. We have

$$Adv_{F^{Alpha}}^{Coll}(\mathcal{A}) \leq Adv_{F^{Alpha}}^{Coll}(\mathcal{A}') \leq \frac{q(q-1)}{(N-q)^2}.$$

Since $\mathcal{A}$ is an arbitrary q-query adversary then

$$Adv_{Alpha}^{Coll}(q) \leq \frac{q(q-1)}{(N-q)^2}.$$

We can easily get the following corollary:

Corollary 1. *Let $F^{Alpha}: \{0,1\}^{3n} \to \{0,1\}^{2n}$ be a compression function based on block cipher as defined in Definition 1. Then for $q \leq 2^{n-1.27}$ we have*

$$Adv_{Alpha}^{Coll}(q) \leq \frac{1}{2} + o(1)$$

where $o(1)$ tends to 0 when n tends to infinity.

Proof. Firstly, it can be seen that the right hand side of Theorem 1 is an increasing function in q for $q < N$. Consider

$$\frac{q(q-1)}{(N-q)^2} = \frac{1}{2}.$$

We get

$$q \approx N\left(\sqrt{2}-1\right) = 2^{n-1.27}.$$

Applying Theorem 1, we have the proof.

For example, for $n = 128$ Corollary 1 implies that any adversary making less than $2^{126.73}$ queries cannot find a collision with probability greater than 1/2.

The MD-strengthening design preserves collision resistance (see Theorem 2.4.1, [11]). Combining this with Theorem 1, we get the following theorem:

Theorem 2. *Let H be an iterated hash function built on the compression function F defined in Definition 1. Then*

$$Adv_H^{Coll}(q) \leq \frac{q(q-1)}{(N-q)^2}, \text{ for every } 1 \leq q < N.$$

B. Preimage resistance security

Theorem 3. *Let $F^{Alpha}: \{0,1\}^{3n} \to \{0,1\}^{2n}$ be a compression function based on block cipher as defined in Definition 1. Then*

$$Adv_{Alpha}^{Pre}(q) \leq \frac{16q}{N^2}.$$

Proof. The idea of the proof follows the proofs of Theorems 1 and 2 in [9]. Let $U||V \in \{0,1\}^{2n}$ be the preimage target (chosen by the adversary before he mounts any query to E). We need to upper bound the probability that the adversary finds a point $A||L||M \in \{0,1\}^{3n}$ such that $F^{Alpha}(A,L,M) = (U,V)$ using q queries.

We also reuse the notion of *free queries* and *super queries* [9]:

- After the adversary asks a forward query $E_{\bar{L}||M}(A \oplus L)$, it is given the answer of the query $E_{L||\bar{M}}(A \oplus L)$ for free. Similarly, if the adversary makes a backward query $E_{\bar{L}||M}^{-1}(R)$, and receives an answer $A \oplus L = E_{\bar{L}||M}^{-1}(R)$ then the answer of the forward query $E_{L||\bar{M}}(A \oplus L)$ is given for free. Therefore, the entries of the adversary's query history Q can be grouped into adjacent pairs of the form $(A \oplus L, \bar{L}||M, R), (A \oplus L, L||\bar{M}, S)$, namely "adjacent query pair".

- After completing each adjacent query pair, we check whether the key $K \in \{0,1\}^{2n}$ used for the latest query satisfies the query history contains exactly $N/2$ queries with this key. If this occurs, all remaining queries under the key K and the remaining queries under key $\bar{K}$ will be given to the adversary for free. We add these $N/2$ free query pairs to the query history. Since the adversary is assumed never to make a query to which it knows the answer, then the adversary cannot make any more queries under this key after these free queries have been added into the query history. We say that a *super query* occurs if and only if $N/2$ free query pairs are given to the adversary. Note that a *super query*

is the set of $N/2$ free query pairs that returned to the adversary.

An adjacent query pair $(A \oplus L, \bar{L}||M, R)$, $(A \oplus L, L||\bar{M}, S)$ is called "winning" if

$A \oplus L \oplus R \oplus M = U$ and $A \oplus L \oplus S \oplus \bar{M} = V$,

or if

$A \oplus L \oplus R \oplus M = V$ and $A \oplus L \oplus S \oplus \bar{M} = U$.

Therefore, if the adversary obtains a winning adjacent query pair then it obtains a preimage of $U||V$. In addition, the winning query pair is part of a super query or not. Let $SuperQueryWin(Q)$ and $NormalQueryWin(Q)$ be the event that the adversary obtains a winning query pair that is part of a super query and normal queries, respectively. Therefore, we need to upper bound

$$\Pr[SuperQueryWin(Q)] + \Pr[NormalQueryWin(Q)].$$

When the event $NormalQueryWin(Q)$ occurs. Assume that the adversary asks a forward query $E_{\bar{L}||M}(A \oplus L)$, then at most $(N/2 - 2)$ queries (including free queries) have been previously answered with the key $\bar{L}||M$. It's implies that,

$$\Pr[A \oplus L \oplus R \oplus M = U] \leq \frac{2}{N}.$$

If $A \oplus L \oplus R \oplus M = U$ then the probability of the free query $E_{L||\bar{M}}(A \oplus L)$ returns $A \oplus L \oplus V \oplus M$ that is at most $1/(N/2) = 2/N$, since the answer to the free query comes uniformly at random from a set of size at least $N/2 + 2 > N/2$. Therefore, we have

$$\Pr\begin{bmatrix}(A \oplus L \oplus R \oplus M = U) \wedge \\ (A \oplus L \oplus S \oplus \bar{M} = V)\end{bmatrix} \leq \frac{4}{N^2}.$$

Similarly,

$$\Pr\begin{bmatrix}(A \oplus L \oplus R \oplus M = V) \wedge \\ (A \oplus L \oplus S \oplus \bar{M} = U)\end{bmatrix} \leq \frac{4}{N^2}.$$

Moreover, since the adversary makes q queries total then we have

$$\Pr[NormalQueryWin(Q)] \leq \frac{8q}{N^2}. \quad (1)$$

In case the event $SuperQueryWin(Q)$ occurs. Assume that a super query occur on keys $\bar{L}||M$ and $L||\bar{M}$, then the value of $E_{\bar{L}||M}(.)$ and $E_{L||\bar{M}}(.)$ is already known on exactly $N/2$ points. Let $\mathcal{D}$ and $\mathcal{R}$ be the domain and the range of that super query, respectively. If $A \oplus L \in \mathcal{D}$ then the probability that $E_{\bar{L}||M}(A \oplus L) \oplus A \oplus L \oplus M = U$ is either 0 if $A \oplus L \oplus M \oplus U \notin \mathcal{R}$, or is exactly $2/N$ if $A \oplus L \oplus M \oplus U \in \mathcal{R}$. Thus, the probability that $E_{\bar{L}||M}(A \oplus L) \in \{U \oplus A \oplus L \oplus M, V \oplus A \oplus L \oplus M\}$ is at most $4/N$.

If $E_{\bar{L}||M}(A \oplus L) \in \{U \oplus A \oplus L \oplus M,\ V \oplus A \oplus L \oplus M\}$, then the probability that $E_{L||\bar{M}}(A \oplus L) \in \{U \oplus A \oplus L \oplus \bar{M}, V \oplus A \oplus L \oplus \bar{M}\}$ is at most $1/(N/2)$. Therefore, the probability that the super query produces a winning pair of adjacent queries is at most $\frac{N}{2} \times \frac{8}{N^2} = \frac{4}{N}$. Since there are at most $q/(N/2)$ super queries, we have

$$\Pr[SuperQueryWin(Q)] \leq \frac{8q}{N^2}. \quad (2)$$

Combining (1) with (2) completes the proof.

Corollary 2. *Let $F^{Alpha}: \{0,1\}^{3n} \to \{0,1\}^{2n}$ be a compression function based on block cipher as defined in Definition 1. Then*

$$Adv^{Pre}_{Alpha}(2^{2n-5}) \leq \frac{1}{2}.$$

Proof. Considering $q \leq \frac{1}{32}N^2$. Applying Theorem 3, we have

$$Adv^{Pre}_{Alpha}(2^{2n-5}) \leq \frac{1}{2}.$$

For example, for $n = 128$ Corollary 2 implies that any adversary making less than 2^{251} queries cannot find a preimage with a considerable probability.

The Merkle-Damgard design also preserves preimage resistance (see Theorem 2.4.2, [11]). Combining Theorem 3 with Theorem 2.4.2 [11], we get the preimage resistance of a hash function composed of F in Definition 1.

Theorem 4. *Let H be an iterated hash function built on the compression function F specified in Definition 1. Then*

$$Adv^{Pre}_{H}(q) \leq \frac{16q}{N^2}.$$

V. CONCLUSION

In this paper, we have proposed a double block length compression function called Alpha-DBL. We have shown very tight collision security bound for the proposed scheme. To our knowledge, the collision security bound of the proposed scheme is nearly better than other double block length schemes. On the other hand, the proposed scheme also achieves the same preimage security bound as the Weimar-DM scheme, which is nearly the best second preimage security bound. Using our compression function in the iterated hash function construction can preserve the collision resistance and preimage resistance security. Moreover, it is shown in [12] that under certain conditions, collision resistance implies second preimage resistance. Thus, we conclude that our proposed hash function is second preimage resistance as well.

REFERENCES

[1] Lai, X. and Massey, J.L. "Hash functions based on block ciphers". Workshop on the Theory and Application of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1992.

[2] Hirose, S. "Some plausible constructions of double-block-length hash functions". International Workshop on Fast Software Encryption. Springer, Berlin, Heidelberg, 2006.

[3] Stam, M. "Blockcipher-based hashing revisited". Fast Software Encryption. Springer, Berlin, Heidelberg, 2009.

[4] Hirose, S. "Provably secure double-block-length hash functions in a black-box model. International Conference on Information Security and Cryptology. Springer, Berlin, Heidelberg, 2004.

[5] Özen, O. and Stam, M. "Another glance at double-length hashing". IMA International Conference on Cryptography and Coding. Springer, Berlin, Heidelberg, 2009.

[6] Fleischmann, Ewan, et al. "Weimar-DM: a highly secure double-length compression function". Australasian Conference on Information Security and Privacy. Springer, Berlin, Heidelberg, 2012.

[7] Miyaji, A. and Rashed, M. "A new (n, 2n) double block length hash function based on single key scheduling". The 29th International Conference on Advanced Information Networking and Applications (AINA), IEEE, 2015.

[8] Fleischmann, E., Gorski, M., and Lucks, S. "Security of cyclic double block length hash functions". IMA International Conference on Cryptography and Coding. Springer, Berlin, Heidelberg, 2009.

[9] Armknecht, F., et al. "The preimage security of double-block-length compression functions". International Conference on the Theory and Application of Cryptology and Information Security. Springer, Berlin, Heidelberg, 2011.

[10] Lee, J., Stam, M., and Steinberger, J. "The security of Tandem-DM in the ideal cipher model". Journal of Cryptology, 2017. 30(2): p. 495-518

[11] Mennink, B., "Provable security of cryptographic hash functions". University of Bristol, UK, 2013.

[12] Rogaway, P., Shrimpton, T. "Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance". In International workshop on fast software encryption. Springer, Berlin, Heidelberg, 2004.

ABOUT THE AUTHOR

Hoang Dinh Linh

Workplace: Institute of Cryptographic Science and Technology

Email: hoangdinhlinh@bcy.gov.vn

Education: Received a Bachelor's degree in Mathematics in Hanoi University of Science in 2014.

Current research direction: symmetric cryptography algorithm, random number generation, randomness tests.

Tran Hong Thai

Workplace: Institute of Cryptographic Science and Technology

Email: ththai@bcy.gov.vn

Education: Received an Engineer's degree in 2000 and Master's degree in Cryptographic Engineering in 2007, Academy of Cryptography Techniques.

Current research direction: research and evaluate the security of cryptographic hashes and block ciphers.