Security Analysis of Double Length Compression Function Based on Block Cipher

Pin Lin[1,2], Wen-Ling Wu[1], and Chuan-Kun Wu[1]
(Corresponding author: Pin Lin)

The State Key Laboratory Of Information Security[1]
Institute of Software, Chinese Academy of Sciences
Graduate School of Chinese Academy of Sciences[2]
Beijing 100080, P. R. China (Email: {linpin, wwl, ckwu}@is.iscas.ac.cn)

(Received Sept. 26, 2005; accepted Nov. 14, 2005)

Abstract

Recently Nandi etc. have proposed a 1/3-rate and a 2/3-rate double length compression functions and studied their security in the black-box model. They proved that to find a collision for the compression function, it requires $\Omega(2^{2n/3})$ queries, where n is the length of output size. In this paper, we show that not all hash functions based on block cipher constructed according to their model are of the same security .i.e., the complexity to find the collisions for these hash functions can be reduced to $O(2^{n/2})$.
Keywords: Hash function, block cipher, and PGV schemes

1 Introduction

A hash function is an easily implementable mapping from an arbitrary length of input to a fixed output. Hash functions are widely used in digital signatures, message authentication codes and simulating random oracle etc. There are many methods to construct hash functions. The most popular method is MD-method [4]. In this method, the first step is to design a compression function from a fixed length input to a fixed length output $f : \{0,1\}^{n+m} \to \{0,1\}^n$, and then extend the input length to an arbitrary length by iterating the compression function.

Before hashing in MD-method, the message may need to be padded with some binary string using an unambiguous padding rule. Usually the binary string represent the length of the message and the length of padded message is the multiple of some block length. This padding rule is called MD-Strengthening. By using padding, hash functions can avoid some trivial attacks. So with fixed initial value $h_0 \in \{0,1\}^n$ and a padded message $M = m_1 \| \cdots \| m_l \in (\{0,1\}^m)^*$, where the block length is $|m_i| = m$, the hash function H can be defined as follows:

Algorithm 1.

1 $H(h_0, m_1 \| \cdots \| m_l)$
2 *For* $i = 1$ *to* l
3 $h_i = f(h_{i-1}, m_i)$
4 *Return* h_l

Here $h_i (i < l)$ is called a *chain value* of the hash function, and h_l is the final hash output.

Many hash functions are based on MD-method mentioned above, such as MD4, MD5, SHA-0 and SHA-1. These hash functions are designed from scratch and the speed of these functions are very fast. Its disadvantage is that they have to be specifically designed and the security of these functions cannot be proved. Hash functions can also be constructed using the underlying block ciphers. Preneel etc. proposed 64 schemes based on block ciphers known as PGV schemes [12]. Black studied the security of PGV schemes in the black-box model [1] and classified them into three groups. The output length of PGV schemes is a single block length. To increase the security level, some double block length hash functions based on block ciphers such as MDC-2 [2] have been proposed. The hash rate of an iterated hash function is defined as the number of m-bit message blocks processed per encryption or decryption. The rate of MDC-2 is 1/2, so it is less efficient. Some hash functions based on block ciphers with rate 1 are designed for high efficiency [8, 11, 3]. Unfortunately Knudsen proved that they were not secure enough [9]. Hirose proposed some provably secure double-length hash functions with the black-box model, their key length is double block length [7]. The hash rate of these schemes is also 1/2. Nandi proposed a 1/3-rate scheme and constructed a 2/3-rate scheme based on it [10]. The complexity of free-start collision attack and pre-image attack for the two schemes is $O(2^{2n/3})$. The 2/3-rate scheme

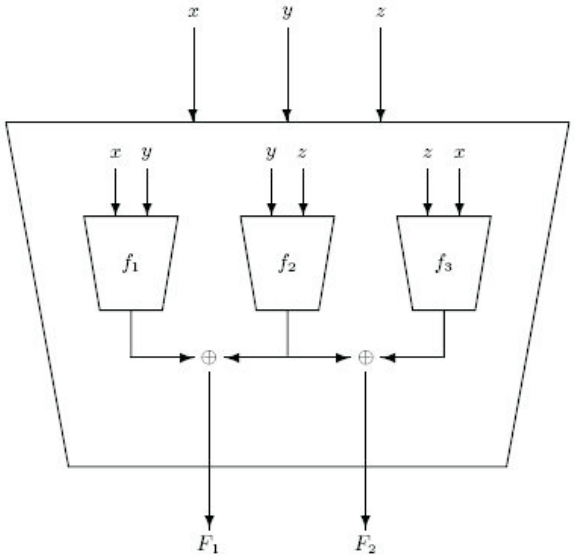

Figure 1: Nandi's model

is more secure than those 1-rate schemes and more efficient than these 1/2-rate schemes. We study the security of hash functions based on block ciphers constructed using Nandi's model.

2 Preliminaries

In [9], Knudsen gave seven attacks on iterated hash functions which are described as follows:

1) Pre-image attack: Given h_0 and $Hash(h_0, M)$, to find $M^{'}$ such that $Hash(h_0, M) = Hash(h_0, M^{'})$ holds.

2) Second pre-image attack: Given h_0 and M, to find $M^{'}$ such that $M^{'} \neq M$ and $Hash(h_0, M) = Hash(h_0, M^{'})$ holds.

3) Free-start pre-image attack: Given h_0 and $Hash(h_0, M)$, to find $h_0^{'}$ and $M^{'}$ such that $Hash(h_0, M) = Hash(h_0^{'}, M^{'})$ holds.

4) Free-start second pre-image attack: Given h_0 and M, to find $h_0^{'}$ and $M^{'}$ such that both $(h_0, M) \neq (h_0^{'}, M^{'})$ and $Hash(h_0, M) = Hash(h_0^{'}, M^{'})$ holds.

5) Collision attack: Given h_0, to find M and $M^{'}$ such that $M \neq M^{'}$ and $Hash(h_0, M) = Hash(h_0, M^{'})$ holds.

6) Semi-free-start collision attack: To find h_0, M and $M^{'}$ such that $M \neq M^{'}$ and $Hash(h_0, M) = Hash(h_0, M^{'})$ holds.

7) Free-start collision attack: To find h_0, $h_0^{'}$, M and $M^{'}$ such that $(h_0^{'}, M^{'}) \neq (h_0, M))$ and $Hash(h_0, M) = Hash(h_0^{'}, M^{'})$ holds.

If a compression function is collision resistant, pre-image resistant and second pre-image resistant, the compression function is secure. There is a generalized collision attack on compression functions named birthday attack. Under birthday attack, if the output length of a compression function is n, the ideal complexity to find a collision for the compression function is $O(2^{n/2})$ and the complexity to find a pre-image for the compression function is $O(2^n)$. This result comes from the following lemma [5, 6].

Lemma 1. *When drawing a sample of size r from a set of N elements with replacements, where $r, N \to \infty$ and $r/N \to \lambda$, the probability that a given element is drawn converges to*

$$1 - exp(-\lambda).$$

The following theorem describes the connection between a hash function and its compression function and is proved in [4].

Theorem 1. *Free-start collision and free-start pre-image attacks against an iterated hash function with MD-strengthening have roughly the same complexities as free-start collision and free-start pre-image attacks against the hash compression function.*

This paper considers hash functions based on block ciphers constructed using Nandi's Model [10] with block length being equal to the key length.

A block cipher is a map $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$, where $k, n \geq 1$. For any key $K \in \{0,1\}^k$ of the block cipher, $E_K(.)$ is a permutation on $\{0,1\}^n$, and E_K^{-1} denotes the inverse of the permutation. A hash function based on block cipher can be defined in algorithm 1 by replacing f with a block cipher. In black-box model, block cipher is regarded as a random permutation to prove the security of hash functions. In Section 3, we use block ciphers to replace random functions.

3 Security Analysis of Hash Functions Based on Block Cipher Constructed Using Nandi's Model

In this section, we study the security analysis of hash functions based on block cipher constructed using Nandi's model [10]. In [10], Nandi proposed a 1/3-rate model and a 2/3-rate model, and we discusses the security of the two cases as follows.

3.1 Hash Functions with 1/3 Hash Rate

Nandi propose a new double length compression function with rate 1/3 [10]. The scheme is based on independent random functions. It is described as follows:

$$F(x, y, z) = (f_1(x, y) \oplus f_2(y, z)) \| (f_2(y, z) \oplus f_3(z, x)).$$

Here f_1, f_2 and f_3 are independent random functions. It can be seen from figure 1 that this scheme can be implemented in parallel, so it is highly efficient although its rate is only 1/3.

Nandi proved in black-box model that the complexity of free-start collision attack for this scheme is $\Omega(2^{2n/3})$. By replacing f_1, f_2 and f_3 with block cipher one can construct new hash functions based on block ciphers. We found after replacing f_1, f_2 and f_3 with a specified block cipher that the new constructed hash function is not as secure as Nandi claimed. We show that the complexity of free-start collision attack for the new hash function is $O(2^{n/2})$.

We first consider the scheme described as follows:

$$\begin{aligned} &F(M_i, H^1_{i-1}, H^2_{i-1}) = \\ &\qquad F_1(M_i, H^1_{i-1}, H^2_{i-1}) || F_2(M_i, H^1_{i-1}, H^2_{i-1}) \\ &F_1(M_i, H^1_{i-1}, H^2_{i-1}) = E^1_{H^1_{i-1}}(M_i) \oplus E^2_{H^1_{i-1}}(H^2_{i-1}) \\ &F_2(M_i, H^1_{i-1}, H^2_{i-1}) = E^2_{H^1_{i-1}}(H^2_{i-1}) \oplus E^3_{H^2_{i-1}}(M_i) \end{aligned} \quad (1)$$

In this scheme, there are three different block ciphers E^1, E^2 and E^3. So the three block ciphers can be regarded as three independent random permutations.

Proposition 1. *The complexity of a free-start collision attack on the compression function defined in Equation (1) is $O(2^{n/2})$*

Proof. First one randomly chooses $2^{n/2}$ values of H^1_{i-1} and H^2_{i-1}, then computes M_i by $(E^1_{H^1_{i-1}})^{-1}($

$E^2_{H^1_{i-1}}(H^2_{i-1}) \oplus v)$, where v is a constant string. For any two distinct 3-tuple $(H^1_{i-1}, H^2_{i-1}, M_i)$ and $(h^1_{i-1}, h^2_{i-1}, m_i)$, $E^1_{H^1_{i-1}}(M_i) \oplus E^2_{H^1_{i-1}}(H^2_{i-1}) = E^1_{h^1_{i-1}}(m_i) \oplus E^2_{h^1_{i-1}}(h^2_{i-1}) = v$ must hold. Therefore, one can get $2^{n/2}$ values of 3-tuple $(H^1_{i-1}, H^2_{i-1}, M_i)$. All these 3-tuples lead to collisions on $F_1(M_i, H^1_{i-1}, H^2_{i-1})$. Then one uses the $2^{n/2}$ 3-tuple to find a free-start collision on $F_2(M_i, H^1_{i-1}, H^2_{i-1})$. Concluded from lemma 1, the probability to find a free-start collision on F_2 is about 0.63. □

Let's see another scheme

$$\begin{aligned} &F(M_i, H^1_{i-1}, H^2_{i-1}) = \\ &\qquad F_1(M_i, H^1_{i-1}, H^2_{i-1}) || F_2(M_i, H^1_{i-1}, H^2_{i-1}) \\ &F_1(M_i, H^1_{i-1}, H^2_{i-1}) = E^1_{M_i}(M_i) \oplus H^1_{i-1} \oplus E^2_{H^1_{i-1}}(H^2_{i-1}) \\ &F_2(M_i, H^1_{i-1}, H^2_{i-1}) = E^2_{H^1_{i-1}}(H^2_{i-1}) \oplus E^3_{H^2_{i-1}}(M_i) \end{aligned} \quad (2)$$

Proposition 2. *The complexity of a free-start collision attack on the compression function defined in Equation (2) is $O(2^{n/2})$.*

Proof. The proof is essentially the same as Proposition 1. □

In the above, a 3-tuple (x, y, z) is used to construct double length hash functions, so the 3-tuple must contain two chain values of hash functions and the other must be a message block. The value of 3-tuple (x, y, z) can be described in Table 1.

Table 1: The value of 3-tuple (x, y, z)

the value of 3-tuple (x,y,z)
$(M_i, H^1_{i-1}, H^2_{i-1})$
$(M_i, H^2_{i-1}, H^1_{i-1})$
$(H^1_{i-1}, H^2_{i-1}, M_i)$
$(H^1_{i-1}, M_i, H^2_{i-1})$
$(H^2_{i-1}, M_i, H^1_{i-1})$
$(H^2_{i-1}, H^1_{i-1}, M_i)$

Now let's review the three independent random functions defined in [10]. The generalized form is $f_i(a, b)$, where $i \in 1, 2, 3$ and (a, b) is a pair from the 3-tuple (x, y, z) defined in table 1. In [12], Preneel etc. considered 64 schemes of the form $E_a(b) \oplus c$, where $a, b, c \in \{h_{i-1}, m_i, h_{i-1} \oplus m_i, v\}$. We now replace f_1, f_2 and f_3 with three different specific block ciphers to get the hash functions we need. From table 1 it can be inferred that f_i has the similar form with the PGV schemes.

One can also be inferred from the values of 3-tuple (x, y, z) in table 1 that either f_1 or f_2 or f_3 has the form $E_a(b) \oplus c$, where $a, b, c \in \{H^1_{i-1}, H^2_{i-1}, H^1_{i-1} \oplus H^2_{i-1}, v\}$. Without loss of generality, we select the 3-tuple $(M_i, H^1_{i-1}, H^2_{i-1})$ from table 1 as an example and assume that f_2 has the form mentioned above. Then f_1 has the form $E_k(x) \oplus y$, where $k, x, y \in \{H^1_{i-1}, M_i, H^1_{i-1} \oplus M_i, v\}$, and f_3 has the similar form as f_1. All the forms of f_1 can be described in table 2 and all the forms of f_2 can be described in table 3. We can see that the two tables are similar to the tables in [1]. We divide the schemes in table 2 into three groups. The group marked with s means that the attack mentioned above can not be applied to the schemes $P \oplus Q$, where P is selected from this group and Q is selected from table 3. The group marked with h means the attack can be applied to some of the schemes $P \oplus Q$ where P is selected from this group and Q is selected from table 3. The group marked with u means that all the pairs from this group and table 3 can be attacked. We also find that all schemes marked with b in [1] are in group u and a majority of schemes marked with a are also in group u. All schemes marked with c and d in [1] are in group s. The rest of schemes in [1] are in group h. Similarly it is known that f_3 should have the similar form with f_1.

3.2 Hash Functions with 2/3 Hash Rate

In [10], Nandi also proposed a double length scheme with 2/3-rate. It can be described as

$$F(x, y, z, t) = (f_1(x, y, z) \oplus f_2(x, z, t)) || (f_2(x, z, t) \oplus f_3(x, y, t))$$

where f_1, f_2 and f_3 are independent random functions as mentioned above. The value of (x, y, z, t) are listed in

Table 2: The form of f_1

i	f_1	mark	i	f_1	mark
1	$E_{M_i}(M_i) \oplus v$	h	33	$E_{M_i}(H^1_{i-1} \oplus M_i) \oplus v$	s
2	$E_{H^1_{i-1}}(M_i) \oplus v$	u	34	$E_{H^1_{i-1}}(H^1_{i-1} \oplus M_i) \oplus v$	u
3	$E_{H^1_{i-1} \oplus M_i}(M_i) \oplus v$	s	35	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1} \oplus M_i) \oplus v$	h
4	$E_v(M_i) \oplus v$	u	36	$E_v(H^1_{i-1} \oplus M_i) \oplus v$	u
5	$E_{M_i}(M_i) \oplus M_i$	h	37	$E_{M_i}(H^1_{i-1} \oplus M_i) \oplus M_i$	s
6	$E_{H^1_{i-1}}(M_i) \oplus M_i$	s	38	$E_{H^1_{i-1}}(H^1_{i-1} \oplus M_i) \oplus M_i$	s
7	$E_{H^1_{i-1} \oplus M_i}(M_i) \oplus M_i$	s	39	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1} \oplus M_i) \oplus M_i$	h
8	$E_v(M_i) \oplus M_i$	h	40	$E_v(H^1_{i-1} \oplus M_i) \oplus M_i$	h
9	$E_{M_i}(M_i) \oplus H^1_{i-1}$	h	41	$E_{M_i}(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1}$	s
10	$E_{H^1_{i-1}}(M_i) \oplus H^1_{i-1}$	u	42	$E_{H^1_{i-1}}(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1}$	u
11	$E_{H^1_{i-1} \oplus M_i}(M_i) \oplus H^1_{i-1}$	s	43	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1}$	h
12	$E_v(M_i) \oplus H^1_{i-1}$	u	44	$E_v(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1}$	u
13	$E_{M_i}(M_i) \oplus H^1_{i-1} \oplus M_i$	h	45	$E_{M_i}(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1} \oplus M_i$	s
14	$E_{H^1_{i-1}}(M_i) \oplus H^1_{i-1} \oplus M_i$	s	46	$E_{H^1_{i-1}}(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1} \oplus M_i$	s
15	$E_{H^1_{i-1} \oplus M_i}(M_i) \oplus H^1_{i-1} \oplus M_i$	s	47	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1} \oplus M_i$	h
16	$E_v(M_i) \oplus H^1_{i-1} \oplus M_i$	h	48	$E_v(H^1_{i-1} \oplus M_i) \oplus H^1_{i-1} \oplus M_i$	h
17	$E_{M_i}(H^1_{i-1}) \oplus v$	s	49	$E_{M_i}(v) \oplus v$	h
18	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus v$	u	50	$E_{H^1_{i-1}}(v) \oplus v$	h
19	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1}) \oplus v$	s	51	$E_{H^1_{i-1} \oplus M_i}(v) \oplus v$	h
20	$E_v(H^1_{i-1}) \oplus v$	u	52	$E_v(v) \oplus v$	u
21	$E_{M_i}(H^1_{i-1}) \oplus M_i$	s	53	$E_{M_i}(v) \oplus M_i$	h
22	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus M_i$	u	54	$E_{H^1_{i-1}}(v) \oplus M_i$	u
23	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1}) \oplus M_i$	s	55	$E_{H^1_{i-1} \oplus M_i}(v) \oplus M_i$	h
24	$E_v(H^1_{i-1}) \oplus M_i$	u	56	$E_v(v) \oplus M_i$	u
25	$E_{M_i}(H^1_{i-1}) \oplus H^1_{i-1}$	s	57	$E_{M_i}(v) \oplus H^1_{i-1}$	h
26	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1}$	u	58	$E_{H^1_{i-1}}(v) \oplus H^1_{i-1}$	u
27	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1}) \oplus H^1_{i-1}$	s	59	$E_{H^1_{i-1} \oplus M_i}(v) \oplus H^1_{i-1}$	h
28	$E_v(H^1_{i-1}) \oplus H^1_{i-1}$	u	60	$E_v(v) \oplus H^1_{i-1}$	u
29	$E_{M_i}(H^1_{i-1}) \oplus H^1_{i-1} \oplus M_i$	s	61	$E_{M_i}(v) \oplus H^1_{i-1} \oplus M_i$	h
30	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1} \oplus M_i$	u	62	$E_{H^1_{i-1}}(v) \oplus H^1_{i-1} \oplus M_i$	u
31	$E_{H^1_{i-1} \oplus M_i}(H^1_{i-1}) \oplus H^1_{i-1} \oplus M_i$	s	63	$E_{H^1_{i-1} \oplus M_i}(v) \oplus H^1_{i-1} \oplus M_i$	h
32	$E_v(H^1_{i-1} \oplus H^1_{i-1} \oplus M_i)$	u	64	$E_v(v) \oplus H^1_{i-1} \oplus M_i$	u

Table 3: The form of f_2

i	f_2	i	f_2
1	$E_{H^2_{i-1}}(H^2_{i-1}) \oplus v$	33	$E_{H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus v$
2	$E_{H^1_{i-1}}(H^2_{i-1}) \oplus v$	34	$E_{H^1_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus v$
3	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^2_{i-1}) \oplus v$	35	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus v$
4	$E_v(H^2_{i-1}) \oplus v$	36	$E_v(H^1_{i-1} \oplus H^2_{i-1}) \oplus v$
5	$E_{H^2_{i-1}}(H^2_{i-1}) \oplus H^2_{i-1}$	37	$E_{H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^2_{i-1}$
6	$E_{H^1_{i-1}}(H^2_{i-1}) \oplus H^2_{i-1}$	38	$E_{H^1_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^2_{i-1}$
7	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^2_{i-1}) \oplus H^2_{i-1}$	39	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^2_{i-1}$
8	$E_v(H^2_{i-1}) \oplus H^2_{i-1}$	40	$E_v(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^2_{i-1}$
9	$E_{H^2_{i-1}}(H^2_{i-1}) \oplus H^1_{i-1}$	41	$E_{H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1}$
10	$E_{H^1_{i-1}}(H^2_{i-1}) \oplus H^1_{i-1}$	42	$E_{H^1_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1}$
11	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^2_{i-1}) \oplus H^1_{i-1}$	43	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1}$
12	$E_v(H^2_{i-1}) \oplus H^1_{i-1}$	44	$E_v(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1}$
13	$E_{H^2_{i-1}}(H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	45	$E_{H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$
14	$E_{H^1_{i-1}}(H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	46	$E_{H^1_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$
15	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	47	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$
16	$E_v(H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	48	$E_v(H^1_{i-1} \oplus H^2_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$
17	$E_{H^2_{i-1}}(H^1_{i-1}) \oplus v$	49	$E_{H^2_{i-1}}(v) \oplus v$
18	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus v$	50	$E_{H^1_{i-1}}(v) \oplus v$
19	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1}) \oplus v$	51	$E_{H^1_{i-1} \oplus H^2_{i-1}}(v) \oplus v$
20	$E_v(H^1_{i-1}) \oplus v$	52	$E_v(v) \oplus v$
21	$E_{H^2_{i-1}}(H^1_{i-1}) \oplus H^2_{i-1}$	53	$E_{H^2_{i-1}}(v) \oplus H^2_{i-1}$
22	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus H^2_{i-1}$	54	$E_{H^1_{i-1}}(v) \oplus H^2_{i-1}$
23	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1}) \oplus H^2_{i-1}$	55	$E_{H^1_{i-1} \oplus H^2_{i-1}}(v) \oplus H^2_{i-1}$
24	$E_v(H^1_{i-1}) \oplus H^2_{i-1}$	56	$E_v(v) \oplus H^2_{i-1}$
25	$E_{H^2_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1}$	57	$E_{H^2_{i-1}}(v) \oplus H^1_{i-1}$
26	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1}$	58	$E_{H^1_{i-1}}(v) \oplus H^1_{i-1}$
27	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1}$	59	$E_{H^1_{i-1} \oplus H^2_{i-1}}(v) \oplus H^1_{i-1}$
28	$E_v(H^1_{i-1}) \oplus H^1_{i-1}$	60	$E_v(v) \oplus H^1_{i-1}$
29	$E_{H^2_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	61	$E_{H^2_{i-1}}(v) \oplus H^1_{i-1} \oplus H^2_{i-1}$
30	$E_{H^1_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	62	$E_{H^1_{i-1}}(v) \oplus H^1_{i-1} \oplus H^2_{i-1}$
31	$E_{H^1_{i-1} \oplus H^2_{i-1}}(H^1_{i-1}) \oplus H^1_{i-1} \oplus H^2_{i-1}$	63	$E_{H^1_{i-1} \oplus H^2_{i-1}}(v) \oplus H^1_{i-1} \oplus H^2_{i-1}$
32	$E_v(H^1_{i-1} \oplus H^1_{i-1} \oplus H^2_{i-1})$	64	$E_v(v) \oplus H^1_{i-1} \oplus H^2_{i-1}$

Table 4: The value of 4-tuple (x, y, z, t)

the value of (x, y, z, t)	the value of (x, y, z, t)
$(M_i^1, M_i^2, H_{i-1}^1, H_{i-1}^2)$	$(M_i^2, M_i^1, H_{i-1}^1, H_{i-1}^2)$
$(M_i^1, M_i^2, H_{i-1}^2, H_{i-1}^1)$	$(M_i^2, M_i^1, H_{i-1}^2, H_{i-1}^1)$
$(M_i^1, H_{i-1}^1, M_i^2, H_{i-1}^2)$	$(H_{i-1}^1, M_i^1, H_{i-1}^2, M_i^2)$
$(M_i^1, H_{i-1}^1, H_{i-1}^2, M_i^2)$	$(H_{i-1}^1, M_i^1, H_{i-1}^2, M_i^2)$
$(M_i^1, H_{i-1}^2, M_i^2, H_{i-1}^1)$	$(H_{i-1}^2, M_i^1, H_{i-1}^1, M_i^2)$
$(M_i^1, H_{i-1}^2, H_{i-1}^1, M_i^2)$	$(H_{i-1}^2, M_i^1, H_{i-1}^1, M_i^2)$
$(M_i^2, H_{i-1}^2, M_i^1, H_{i-1}^1)$	$(H_{i-1}^2, M_i^2, H_{i-1}^1, M_i^1)$
$(M_i^2, H_{i-1}^2, H_{i-1}^1, M_i^1)$	$(H_{i-1}^2, M_i^2, H_{i-1}^1, M_i^1)$
$(M_i^2, H_{i-1}^1, M_i^1, H_{i-1}^2)$	$(H_{i-1}^1, M_i^2, H_{i-1}^2, M_i^1)$
$(M_i^2, H_{i-1}^1, H_{i-1}^2, M_i^1)$	$(H_{i-1}^1, M_i^2, H_{i-1}^2, M_i^1)$
$(H_{i-1}^1, H_{i-1}^2, M_i^1, M_i^2)$	$(H_{i-1}^2, H_{i-1}^1, M_i^1, M_i^2)$
$(H_{i-1}^1, H_{i-1}^2, M_i^2, M_i^1)$	$(H_{i-1}^2, H_{i-1}^1, M_i^2, M_i^1)$

table 4. We now use $(M_i^1, H_{i-1}^1, M_i^2, H_{i-1}^2)$ to design an example. The scheme is described as follows.

$$
\begin{aligned}
&F(M_i^1, M_i^2, H_{i-1}^1, H_{i-1}^2) = \\
&\qquad F_1(M_i^1, M_i^2, H_{i-1}^1, H_{i-1}^2) \| F_2(M_i^1, M_i^2, H_{i-1}^1, H_{i-1}^2) \\
&F_1(M_i^1, M_i^2, H_{i-1}^1, H_{i-1}^2) = \\
&\qquad E^1_{M_i^1 \oplus M_i^2}(H_{i-1}^1) \oplus E^2_{M_i^1 \oplus M_i^2}(H_{i-1}^2) \\
&F_2(M_i^1, M_i^2, H_{i-1}^1, H_{i-1}^2) = \\
&\qquad E^2_{M_i^1 \oplus M_i^2}(H_{i-1}^2) \oplus E^3_{H_{i-1}^2}(M_i^1 \oplus H_{i-1}^1) \qquad (3)
\end{aligned}
$$

Proposition 3. *The complexity of a collision attack on the compression function defined in Equation (3) is* $O(2^{n/2})$.

Proof. First one randomly chooses $2^{n/2}$ message pairs (M_i^1, M_i^2) satisfying $M_i^1 \oplus M_i^2 = v$, where v is a constant string. For any two distinct message pairs (M_i^1, M_i^2) and (m_i^1, m_i^2), $E^1_{M_i^1 \oplus M_i^2}(H_{i-1}^1) \oplus E^2_{M_i^1 \oplus M_i^2}(H_{i-1}^2) = E^1_{m_i^1 \oplus m_i^2}(H_{i-1}^1) \oplus E^2_{m_i^1 \oplus m_i^2}(H_{i-1}^2)$ must hold. In this way one has $2^{n/2}$ collisions on F_1. Then use these message pairs to find the collision on F_2. From lemma 1, the probability to find a collision on F_2 is about 0.63. □

There are many combinations of the 4-tuple (x, y, z, t) and the forms of f_i are different from those PGV schemes. We cannot analyze all the combinations and classify them as we do in Section 3.1. But if it is very easy to find the collision or free-start collision for F_1 or F_2 we can find collision or free-start collision for F with complexity $O(2^{n/2})$.

4 Conclusion

This paper has analyzed the security of some specified schemes of hash function using Nandi's model, and has shown that the complexity to get a free-start collision is $O(2^{n/2})$, lower than in the general case as claimed by Nandi. We also found that although many PGV schemes are not secure in [1], they can be used in Nandi's model. The result for double length hash function is similar to this. The result does not contradict with the result in [10]. In [10], the security of Nandi's model was proved in black-box model, and three independent random functions are used in this model. One does not know how these independent random functions are constructed. In this paper,three different block ciphers are used to replace these random functions and because hash functions have no secret information, the keys of block ciphers are public to everyone and the three block ciphers can be easily reversed. So Nandi's model should be carefully implemented.

Acknowledgement

Research supported by National Natural Science Foundation of China(No. 60373047, 90304007) and the National Grand Fundamental Research 973 Program of China(No. 2004CB318004).

References

[1] J. Black, P. Rogaway, and T.Shrimpton, "Black-box analysis of the block-cipher based hash function constructions from PGV," *Advances in Cryptology-Crypto'02*, LNCS 2442, pp. 320-335, Springer-Verlag, 2002.

[2] B. Brachtl, D. Coppersmith, M. Hyden, S. Matys, 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 4,908,861, Mar 13, 1990.

[3] L. Brown, J. Pieprzyk, and J. Seberry, "LOKI-a cryptographic primitive for authentication and secrecy applications," *Advances in Cryptology-AusCrypt'90*, LNCS 453, pp. 229-236, Springer-Verlag, 1990.

[4] I. B. Damg rd, "A design principle for hash functions," *Advances in Cryptology-Crypto'89*, LNCS 435, pp. 416-427, Springer-Verlag, 1989.

[5] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, Wiley, New York, 1968.

[6] M. Girault, R. Cohen, and M. Campana, "A generalized birthday attack," *Advances in Cryptology-Eurocrypt'88*, LNCS 330, pp. 129-156, Springer-Verlag, 1988.

[7] S. Hirose, "Provably secure double-block-length hash functions in a black-box model," *ICISC 2004*, LNCS 3506, pp. 330-342, Springer-Verlag, 2005.

[8] W. Hohl, X. Lai, T. Meier, and C. Waldvogel, "Security of iterated hash function based on block ciphers," *Advances in Cryptology-Crypto'93*, LNCS 773, pp. 379-390, Springer-Verlag, 1993.

[9] L. Knudsen, X. Lai, and B. Preneel, "Attacks on fast double block length hash functions," *Journal of Cryptology*, vol. 11, no. 1, pp. 59-72, winter 1998.

[10] M. Nandi, W. Lee, K. Sakurai and S. Lee, "Security ananlysis of a 2/3-rate double length compression function in the black-box model," *FSE2005*, LNCS 3557, pp. 243-254, Springer-Verlag, ENSTA, 2005.

[11] B. Preneel, A. Bosselaers, R. Govaerts, and J.Vandewalle, "Collsion-free hash functions based on block cipher algorithms," in *Proceedings of 1989 International Carnahan Conference on Security Technology*, pp. 203-210, 1989.

[12] B. Preneel, R. Govaerts, and J. Vandewalle, "Hash functions based on block ciphers: A synthetic approach," *Advances in Cryptology-Crypto'93*, LNCS 773, pp. 368-378, Springer-Verlag, 1994.

Pin Lin is now a phd candidate at the State Key Laboratory of Information Security, Graduate University of Chinese Academy of Sciences. He received his bachelor degree in computer science from ShanDong University in 2001.He received his Master degree in network information security from ShangDong University in 2004. His current research interest is the design and analysis of hash functions.

Wenling Wu is now a professor at the State Key Laboratory of Information Security, Institute of Software,Chinese Academy of Sciences. She received her B.S. degree and M.S. degree in Maths from Northwest University in 1987 and 1990, respectively. She received her Ph.D degree in Cryptography from Xidian University in 1997. From 1998 to 1999 she was a postdoctoral fellow in the Institute of Software, Chinese Academy of Science. Her current research interests include theory of cryptography, mode of operation, block cipher, stream cipher and hash function.

Chuan-Kun Wu was teaching at Xidian University since January 1988. He was promoted by Xidian University as a Lecturer in 1990, an Associate professor in 1992, and a full professor in 1995. In September 1995, he became a postdoctoral fellow in Australia, then from 1997 a research fellow, and from 2000 a Lecturer in the Department of Computer Science, Australian National University. Since 2003, He has joined the Institute of Software, Chinese Academy of Sciences. He has got many awards while he was in China, including China Government Special Subsidy awarded in 1993. He founded and served as a program co-chair for 2001, 2002 and 2003 International Workshop on Cryptology and Network Security (CANS) which has become one of the influential international conferences since 2005, and has served as a program committee member for many international conferences. He is an associate editor of IEEE.