The Design and Analysis of a Hash Ring-iterative Structure

Shenghui Su[1], Yixian Yang[2], Bo Yang[2], and Shaolan Zhang[2]

[1]School of Information Engineering, University of Science & Technology Beijing
Beijing 100083, P. R. China
sheenway@126.com

[2]School of Information Engineering, Beijing University of Posts & Telecom
Beijing 100876, P. R. China

Abstract

This paper proposes a new type of hash iterative structure ─ the ring-iterative structure with feedback which is subdivided into the single feedback ring iteration and the multiple feedback ring iteration, namely SFRI and MFRI. The authors prove that if *f* is a secure one-way compression function, the security of the MFRI structure is greater than that of the classical MD iterative structure, analyze the resistance of MFRI, which is from the joint event on message modification, the endless loop on message modification and the incompatibility of the sufficient conditions, to the multi-block differential collision attack, and argue the ineffectiveness of the D-way second preimage attack on MFRI. The paper discusses the time and space expenses of MFRI, and points out the advantage of MFRI over the tree-iterative structure and the zipper-iterative structure.

Keywords: Digital signature, Hash function, Security, Ring Iteration, Compression function.

1 Introduction

It is well known that hash functions are predominantly employed for digital signature, data integrity or message authentication code. The security of hash functions is the foundation of security of digital signature.

At present, the hash functions used widely ─ MD5, SHA-0, and SHA-1 [1] for example all adopt the Merkle-Damgård (MD) iterative structure [2][3]. The design principle of this structure is that if there does not exist a computationally collision-resistant function *h* mapping a message of arbitrary polynomial length to a *k*-bit string, then there does not exist a computationally collision-resistant function *f* from *m* bits to *k* bits, where $k < m$ [2]. Thereby, it has been universally thought that the problem of designing a collision-resistant hash function may be reduced to the problem of designing a collision-resistant compression function, namely iterative function.

However, the multi-block differential collision attack on MD5、SHA-0 and SHA-1 [4][5][6] indicates that a collision-resistant compression function is not a sufficient condition of a collision- resistant hash function, but only a necessary condition[7]. It means that a secure and collision- resistant hash function will be based not only on a collision-resistant compression function, but also on a collision-resistant iteration structure.

Section 2 of this paper designs a new type of hash iteration structure, which is partitioned into the single feedback ring iteration and multiple feedback ring iteration, namely SFRI and MFRI. Section 3 proves that the MFRI structure is more secure than the MD iterative structure, analyzes the security of MFRI against the multi-block differential collision attack and the D-way second preimage attack. Section 4 makes the performance analyses of the MFRI structure in time and space expenses.

2 Design of Hash Ring-iterative Structures

2.1 Single Feedback Ring Iteration

Assume that a message to be hashed is X of l-bit length, and X is partitioned into n m-bit blocks $X_1, X_2, \ldots, X_n$, where $n = l / m$ and l is exactly divided by m, that is, the padding problem is neglected by us, which does not influence our discussion.

Let IV be the initial value of the chaining variable, f be a compression function, every iterative output be Y_i of k-bit length, where $k \le m$, $i = 1, 2, \ldots, n$, and D be the last iterative output, namely the message digest.

For the MD iterative structure, there are $Y_0 = IV$, $Y_i = f(Y_{i-1}, X_i)$, and $D = Y_n$.

The single feedback ring iteration, shortly SFRI, is a simple structure. It feeds back the reverse code of the MD iterative output Y_n into iterative box 1, sends the second output of iterative box 1 to iterative box n, and generates the message digest D last. See Figure 1.

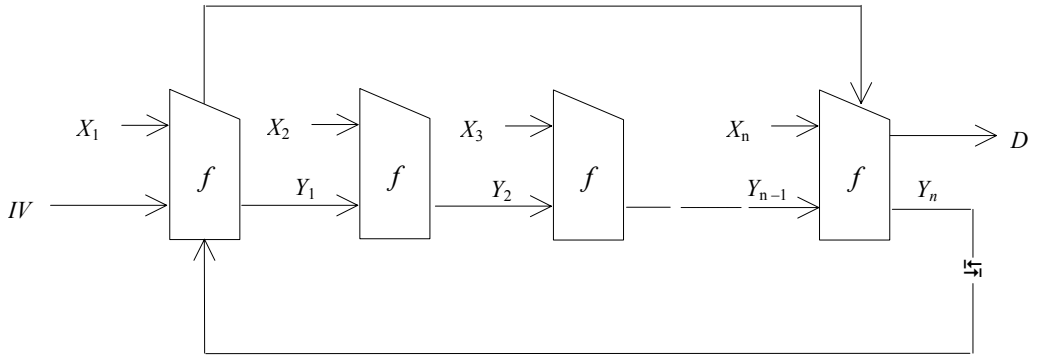

Figure 1: The Single Feedback Ring-iterative Structure

In Figure 1, we define $X_{n+1} = X_1$, $X_{n+2} = X_n$, and the message digest $D = f(X_{n+2}, f(X_{n+1}, \leftrightarrows Y_n))$.

Although the inputs X_1 and X_n of box 1 and box n are employed twice, and they have respectively two outputs, it is not incompatible in logicality according to the above definitions.

Notice that sign '⇆' denotes reversal operation, that is, the bits of a variable are rearranged in reverse order. For example, the reverse code of '100110' is '011001'.

2.2 Multiple Feedback Ring Iteration

The multiple feedback ring iteration, shortly MFRI, is a comparatively complex structure. It feeds back the modular sum of reverse codes of all iterative outputs into iterative box 1, sends the second output of iterative box 1 to iterative box n, and generates the message digest D last. See the following Figure 2.

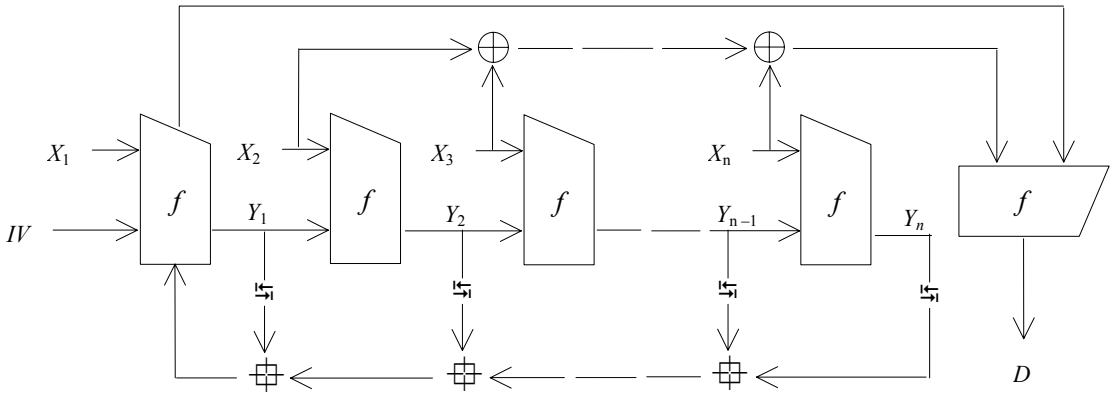

Figure 2: The Multiple Feedback Ring-iterative Structure

According to the above diagram, we define $X_{n+1} = X_1$, $X_{n+2} = X_2 \oplus X_3 \oplus \ldots \oplus X_n$, and the message digest $D = f(X_{n+2}, f(X_{n+1}, \lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_n))$, where sign '$\boxplus$' denotes modular addition operation.

Herein, we substitute $X_{n+2} = X_2 \oplus X_3 \oplus \ldots \oplus X_n$ for $X_{n+2} = X_n$ in the SFRI structure to make the last D depend relatively uniformly on $X_1, X_2, \ldots$, and X_n. $(X_2 \oplus X_3 \oplus \ldots \oplus X_n)$ may be regarded as a feedforward.

Remark: the single feedback ring-iterative structure is dominantly used to assist in proving the following theorem 1. In practical applications, we should employ the multiple feedback ring-iterative structure.

3 Security Analysis of Ring-iterative Structures

3.1 More Secure than the MD structure

Assume that f is a compression function. It combines with any iteration structure to construct a one-way hash function H.

[***Definition* 1**] For a given message M, if we can not find a message $M' \neq M$ such that $H(M') = H(M)$ in polynomial time, then H is called weakly collision-resistant.

[***Definition* 2**] If we can not find any two messages M and M' satisfying $M' \neq M$ and $H(M') = H(M)$ in polynomial time, then H is called strongly collision-resistant.

Obviously, if a hash function is strongly collision-resistant, then it must be weakly collision-resistant.

[***Lemma* 1**] The SFRI structure is equivalent to the MD structure in security.

Remark: In cryptology, security is measured with time complexity. Therefore, that the SFRI structure is equivalent to the MD structure in security means that the time complexity of breaking SFRI is equivalent to the time complexity of breaking MD.

Proof:

Suppose that f combines with the MD structure to construct the hash function H_1, and f does with the SFRI structure to construct the hash function H_2.

First, suppose that H_1 is strongly collision-resistant, and we need to infer that H_2 is also strongly collision-resistant.

Presume that H_2 is not strongly collision-resistant. Namely we can find the two messages M and M' satisfying $M \neq M'$ and $H_2(M) = H_2(M')$ in polynomial time.

Suppose that M is exactly partitioned into n m-bit blocks $X_1, X_2, \ldots, X_n$, and M' is exactly partitioned into n' m-bit blocks $X'_1, X'_2, \ldots, X'_{n'}$.

Then, in terms of the SFRI structure, there are

$$H_2(M) = f(X_n, f(X_1, \leftrightarrows Y_n)), \text{ and } H_2(M') = f(X'_{n'}, f(X'_1, \leftrightarrows Y'_{n'})),$$

where $Y_n = H_1(M)$, and $Y_{n'} = H_1(M')$.

Let $M_1 = M \,\$\, X_1 \,\$\, X_n$, and $M'_1 = M' \,\$\, X'_1 \,\$\, X'_{n'}$, where '$' represents the concatenation of strings. Then

$$H_2(M) = H_1(M_1), \text{ and } H_2(M') = H_1(M'_1).$$

Hence, we can find the two messages M_1 and M'_1 which meet $M_1 \neq M'_1$ and $H_1(M_1) = H_1(M'_1)$ in polynomial time. It is contrary to the presupposition, so H_2 is also strongly collision-resistant, which illustrates that the security of H_2 is not less than the security of H_1.

Secondly, suppose that H_2 is strongly collision-resistant, and we need to infer that H_1 is also strongly collision-resistant.

According to reference [5] and [6], define the message block differential as $\Delta X_i = X'_i - X_i$, and the iterative output differential as $\Delta Y_i = Y'_i - Y_i$.

In the SFRI structure, the collision differential characteristics will be $\Delta X_1, \Delta X_2, \ldots, \Delta X_n, \Delta X_{n+1}, \Delta X_{n+2}$ and $0 = \Delta Y_0, \Delta Y_1, \Delta Y_2, \ldots, \Delta Y_n, \Delta Y_{n+1}, \Delta Y_{n+2} = 0$.

Notice that $\Delta X_{n+1} = X'_1 - X_1$ and $\Delta X_{n+2} = X'_n - X_n$. Thus, $\Delta X_1 = \Delta X_{n+1}$ and $\Delta X_n = \Delta X_{n+2}$.

If the attackers set $\Delta X_1 = \Delta X_n = \Delta Y_n = 0$, then the attack on SFRI degenerates to the attack on the MD structure.

Thereby, if H_1 is not strongly collision-resistant, H_2 is not strongly collision-resistant. This is contrary to the presupposition, and so H_1 is strongly collision-resistant, which illustrates that the security of H_1 is not less than the security of H_2.

In summary, the SFRI structure is equivalent to the MD structure in security. □

[***Theorem* 1**] If f is a secure one-way function, the MFRI structure is more secure than the MD structure.

Proof:

Suppose that f is a secure one-way compression function, namely it is infeasible in polynomial time to compute its preimage from an image of f.

If we can find a value set of $X_1, X_2, \ldots, X_n$ such that $(\lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_n) = \lhd Y_n$ and $(X_2 \oplus X_3 \oplus \ldots \oplus X_n) = X_n$ in polynomial time, then the MFRI structure will be reduced to the SFRI structure, and the security of MFRI will be equivalent to one of SFRI.

In what follows, we resort to reduction to absurdity.

Presume that a value set of $X_1, X_2, \ldots, X_n$ is found out in polynomial time, and satisfies $(\lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_n) = \lhd Y_n$ and $(X_2 \oplus X_3 \oplus \ldots \oplus X_n) = X_n$.

According to the definitions of the operators $\boxplus$ and $\oplus$, we see that $(\lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_{n-1}) = u2^k$ and $(X_2 \oplus X_3 \oplus \ldots \oplus X_{n-1}) = 0$, where u is a positive integer and k is the bit-length of Y_i.

Let $u = 1$ (if u equals other integers, it does not influence our discussion). In terms of the MFRI structure, the $X_1, X_2, \ldots, X_{n-1}$ satisfy the following simultaneous equations necessarily:

$$\begin{cases} Y_1 = f(IV, X_1) \\ Y_2 = f(Y_1, X_2) \\ \ldots\ldots \\ Y_{n-1} = f(Y_{n-2}, X_{n-1}) \\ (X_2 \oplus X_3 \oplus \ldots \oplus X_{n-1}) = 0 \\ (\lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_{n-1}) = 2^k \end{cases}$$

Transparently, this equation system contains only two equations substantially, and has $n - 1$ variables $X_1, X_2, \ldots, X_{n-1}$.

No matter how $X_1, X_2, \ldots, X_{n-1}$ are found or computed, it includes two steps:

- Firstly, determine values of any $n - 3$ variables among $X_1, X_2, \ldots,$ and X_{n-1};
- Secondly, compute the values of the other two variables according to the two constraints from the equation system.

Without loss of generality, assume that the values of $X_1, X_2, \ldots, X_{n-3}$ are determined.

Further, the values of X_{n-2} and X_{n-1} need to be sought.

According to $(X_2 \oplus X_3 \oplus \ldots \oplus X_{n-1}) = 0$, X_{n-2} can be expressed with the variable X_{n-1}, namely

$$X_{n-2} = X_2 \oplus X_3 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}.$$

According to $(\lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_{n-1}) = 2^k$, Substitution for $Y_1, Y_2, \ldots, Y_{n-2}, Y_{n-1}$ yields

$$\lhd f(IV, X_1) \boxplus \lhd f(Y_1, X_2) \boxplus \ldots \boxplus \lhd f(Y_{n-3}, X_{n-2}) \boxplus \lhd f(Y_{n-2}, X_{n-1}) = 2^k,$$

that is,

$$\lhd f(IV, X_1) \boxplus \lhd f(Y_1, X_2) \boxplus \ldots \boxplus \lhd f(Y_{n-3}, X_2 \oplus X_3 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}) \boxplus \lhd f(Y_{n-2}, X_{n-1}) = 2^k,$$

$$\lhd f(IV, X_1) \boxplus \lhd f(Y_1, X_2) \boxplus \ldots \boxplus \lhd f(Y_{n-3}, X_2 \oplus X_3 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}) \boxplus$$

$$\lhd f(f(Y_{n-3}, X_2 \oplus X_3 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}), X_{n-1}) = 2^k.$$

Clearly, seeking X_{n-1} from the above equation contains at least solving the problem on preimages of the one-way compression function f.

In terms of the preceding presumption, X_{n-1} can be evaluated in polynomial time. It indicates that a preimages of the function f can be found in polynomial time, namely the function f is not secure, which is in direct contradiction to the presupposition that f is a secure one-way compression function.

The above discussion shows that we can not find a value set of $X_1, X_2, \ldots, X_n$ such that $(\lhd Y_1 \boxplus \lhd Y_2 \boxplus \ldots \boxplus \lhd Y_n) = \lhd Y_n$ and $(X_2 \oplus X_3 \oplus \ldots \oplus X_n) = X_n$ in polynomial time, namely the MFRI structure can not be reduced to the SFRI structure in polynomial time.

Therefore, the MFRI structure is more secure than the SFRI structure. By transitivity, the MFRI structure is also more secure than the MD structure. □

3.2 Resistance to the Multi-block Differential Attack

3.2.1 Brief Presentation of the Multi-block Differential Attack

Reference [4], [5] and [6] manifest the multi-block near differential attack on the hash functions MD4, MD5, SHA-0 and SHA-1. This attack consists of the following three steps:

(1) Find out a set of collision differential characteristics for M and M' which are expected to produce a collision.

(2) Derive a set of sufficient conditions which are described by the bits of the chaining variables, and ensure that the collision differential characteristics hold.

(3) Modify the random message M through the single-step / multi-step or single-message / multi-message method in order to make almost all the sufficient conditions be satisfied.

Assume that M is partitioned into n m-bit blocks $X_1, X_2, \ldots, X_n$, and the iterative outputs are $Y_1, Y_2, \ldots, Y_n = D$ in order.

Assume that M' is partitioned into n m-bit blocks $X'_1, X'_2, \ldots, X'_n$, and the iterative outputs are $Y'_1, Y'_2, \ldots, Y'_n = D'$, where $n \geq 2$.

According to reference [5] and [6], define the message differential as $\Delta X_i = X'_i - X_i$, and the iterative output differential as $\Delta Y_i = Y'_i - Y_i$.

Notice that a differential is computed by modular integer subtraction '–' in reference [5] and [6] while it is computed by exclusive or '⊕' in other references. Obviously, the combination of these two sorts of differentials can bring more information to attackers.

For the MD structure, assume that the collision differential characteristics are $\Delta X_1, \Delta X_2, \ldots, \Delta X_n$ and $0 = \Delta Y_0, \Delta Y_1, \Delta Y_2, \ldots, \Delta Y_n = 0$.

It is should be noted that because the same compression function f is used when two different messages are

hashed, the initial values of iteration are the same, namely $\Delta Y_0 = 0$. The $\Delta Y_n = 0$ indicates that the collision (M, M') is found out, and it is a goal which the attackers try to achieve.

ΔY_i is also the chaining variable difference. In terms of a concrete compression function, the attackers may set more detailed step-chaining variable differentials and round-chaining variable differentials [5][6][8].

3.2.2 MFRI Leading Block Modification to a Joint Event

In the MFRI structure, let the collision differential characteristics be $\Delta X_1, \Delta X_2, \ldots, \Delta X_n$, $0 = \Delta Y_0, \Delta Y_1, \Delta Y_2, \ldots, \Delta Y_n$, and the input chaining variable of the $(n + 1)$-th iteration be Y_n^{d}, then $Y_n^{\mathrm{d}} = \lrcorner\!\!\urcorner Y_1 \boxplus \lrcorner\!\!\urcorner Y_2 \boxplus \ldots \boxplus \lrcorner\!\!\urcorner Y_n$, where the superscript 'd' signifies timedelay. Therefore,

$$\begin{aligned}\Delta Y_n^{\mathrm{d}} &= (\lrcorner\!\!\urcorner Y'_1 \boxplus \lrcorner\!\!\urcorner Y'_2 \boxplus \ldots \boxplus \lrcorner\!\!\urcorner Y'_n) - (\lrcorner\!\!\urcorner Y_1 \boxplus \lrcorner\!\!\urcorner Y_2 \boxplus \ldots \boxplus \lrcorner\!\!\urcorner Y_n) \\ &= \Delta\lrcorner\!\!\urcorner Y_1 \boxplus \Delta\lrcorner\!\!\urcorner Y_2 \boxplus \ldots \boxplus \Delta\lrcorner\!\!\urcorner Y_n \\ &\neq \lrcorner\!\!\urcorner\Delta Y_1 \boxplus \lrcorner\!\!\urcorner\Delta Y_2 \boxplus \ldots \boxplus \lrcorner\!\!\urcorner\Delta Y_n.\end{aligned}$$

For example, when A = 11010100, A' = 00101011, and ΔA = 10101001, there are $\lrcorner\!\!\urcorner A$ = 00101011, $\lrcorner\!\!\urcorner A'$ = 11010100, and $\Delta\lrcorner\!\!\urcorner A$ = 01010111, and so $\Delta\lrcorner\!\!\urcorner A \neq \lrcorner\!\!\urcorner\Delta A$.

This brings extra difficulties to the attackers who employ the differential analysis method.

From reference [5], [6] and [8], it is not difficult to understand that if there are not the multiple feedbacks, the modification to every block X_i is an independent event, and when the modification is made, it is feasible to consider ΔX_i and ΔY_i only relevant to the block X_i but not to other blocks. However, when the multiple feedbacks exist, due to

$$\Delta Y_n^{\mathrm{d}} = (\lrcorner\!\!\urcorner Y'_1 \boxplus \lrcorner\!\!\urcorner Y'_2 \boxplus \ldots \boxplus \lrcorner\!\!\urcorner Y'_n) - (\lrcorner\!\!\urcorner Y_1 \boxplus \lrcorner\!\!\urcorner Y_2 \boxplus \ldots \boxplus \lrcorner\!\!\urcorner Y_n),$$

the modification to every block X_i will influence the corresponding Y_i, and further influences ΔY_n^{d}. Thereby, the modification to every block X_i changes into an joint event from an individual independent event.

Assume that through message modification techniques the attackers can decrease the time complexity of a block near collision $O(2^{k_i})$. In terms of reference [5], [6] and [8], in the MD structure, the modification to every block is an independent event, and hence, the complexity of producing the n-block message collision is $O(2^{k_1} + 2^{k_2} + \ldots + 2^{k_n})$. However, in the MFRI structure, the modification to every block become a part of the joint event, and hence, the probability that two n-block messages produce a collision is $1 / (2^{k_1}\, 2^{k_2} \ldots 2^{k_n})$, namely, the complexity of producing the message collision increases to $O(2^{k_1 + k_2 + \ldots + k_n})$.

3.2.3 MFRI Leading Message Modification to an Endless Loop

To ensure that the differential characteristics being set holds, every block has a set of sufficient conditions derived from f and ΔY_i, where ΔY_i is a chaining variable differential. For the hash functions MD4 and MD5, the length of every chaining variable is 128 bits, is exactly one of four 32-bit words. Therefore, in fact, every chaining variable consists of the four word variables a, b, c, and d. Because every block-iteration consists of several round-iterations, and every round-iteration consists of several step-iterations, the variables a, b, c, d may be further divided into $a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, \ldots, a_s, b_s, c_s, d_s$ in every block-iteration process. For example, in MD4, s = 12, and in MD5, s = 16. The values of a_i, b_i, c_i, d_i in the sufficient conditions are expressed with 0, 1 or those prior to a_i, b_i, c_i, d_i. For SHA-1, its chaining variable is

composed of a, b, c, d, and e five word variables. The values of the sufficient conditions for a block collision are expressed with 0, 1 or those prior to a_i.

In the MFRI structure, because of feedback and $X_{n+1} = X_1$, the second modification to X_1 is needed. However, the second modification will surely influence the result of the first modification, that is, break the sufficient conditions satisfied and change the value of the chaining variable Y_1, and further, cause Y_2, Y_3, …, Y_n and the feedback Y_n^d to produce alteration. Thus, the attackers need to modify X_1 once more. In this way, the MFRI structure will lead the modification to X_1 and other blocks to an endless loop.

3.2.4 MFRI Leading Sufficient Conditions for Collision to Incompatibility

Due to $X_{n+1} = X_1$, for the MFRI structure, a key problem is whether the two sets of sufficient conditions described with the chaining variables respectively in the 1-st iteration and $(n + 1)$-th iteration are compatible or not. If they are compatible, the two sets of sufficient conditions can be deduced theoretically. If they are contrary to each other, the two sets of sufficient conditions is radically impossibly deduced, and thus it is not ensured that the differential characteristics will hold and that the collision for two messages can be found.

Suppose that $\Lambda = \{a_{1,1}, a_{1,2}, \ldots, a_{1,32}\}$ is the set of 32 bits of the variable a_1, then the size of its power set is $|P(\Lambda)| = 2^{32}$. Let $P_{\cap NE}$ denote the probability that the intersection of any two nonempty subsets of Λ is empty, then

$$\begin{aligned} P_{\cap NE} &= (C_{32}^1 (2^{31} - 1) + C_{32}^2 (2^{30} - 1) + \ldots + C_{32}^{31}(2^1 - 1)) / (2C_{|P(\Lambda)|}^2) \\ &< (2^{23} - 1)(C_{32}^1 + C_{32}^2 + \ldots + C_{32}^{31}) / (2^{31} 2^{32} - 1) \\ &= (2^{23} - 1)(2^{32} - 2) / (2^{31} 2^{32} - 1) \approx 1/2^{12}. \end{aligned}$$

Thereby, the probability that the intersection of any two nonempty subsets of Λ is nonempty is greater than $(1 - 1/2^{12})$, which means that probability at least 1 bit of the condition variable a_1 produces overlap in the 1-st iteration and $(n + 1)$-th iteration is greater than $(1 - 1/2^{12})$.

We may as well suppose that $a_{1,1}$ overlaps, let $a_{1,1}^1$ denote the condition value of $a_{1,1}$ in the 1-th iteration, and let $a_{1,1}^{n+1}$ denote the condition value of $a_{1,1}$ in the $(n + 1)$-th iteration. If $a_{1,1}^1 = a_{1,1}^{n+1} = 0$ or 1, it indicates the two sets of sufficient conditions are compatible; otherwise the two sets of sufficient conditions are incompatible, that is, such two sets of sufficient conditions may impossibly exist simultaneously. Obviously, if the intersection contains only 1 bit, the probability of being incompatible is 1/2. If the intersection contains 2 bits, the probability of being incompatible is (1 – 1/4). Suppose that P_{a_1}, P_{b_1}, P_{c_1}, $P_{d_1} = 1 / 2$ or 1 represent respectively the probabilities that the intermediate chaining variables a_1, b_1, c_1, d_1 are condition-compatible in the 1-th iteration and $(n + 1)$-th iteration. Then, for a_1, b_1, c_1, and d_1, the probability that the conditions are incompatible is $(1 - 1 / (P_{a_1} P_{b_1} P_{c_1} P_{d_1}))$. For the other intermediate chaining variables a_2, b_2, c_2, d_2 …, there exist similar conclusions.

The above analysis manifests that in two different iterations of the same block, the probability that the condition bits produce overlap is close to 1, and the probability that the values of the overlapping bits are incompatible is greater than 1/2.

3.3 Ineffectiveness of the D-way Second Preimage Attack

Joux puts forward a attack method called D-way which is employed for seeking the second preimage of an output of a hash function based on the MD structure in reference [9]. For a given hash target value $Y = H(M) \in \{0, 1\}^k$, the attackers first find 2^d collisions on d-block messages $M_1, M_2, \ldots, M_{2^d}$ making $H_d = H(M_1) = H(M_2) = \ldots = H(M_{2^d})$. Then, find the block X_{d+1} such that $f(H_d, X_{d+1}) = Y$. In this way, the attackers succeed in seeking the second preimage with the message M. In terms of reference [9], the time complexity of this attack is $O(d\,2^{k/2} + 2^k)$.

For a hash function based on the MFRI structure, because there are $X_{d+1} = X_1$, $X_{d+2} = X_2 \oplus X_3 \oplus \ldots \oplus X_d$, and $X'_{d+1} = X'_1$, $X'_{d+2} = X'_2 \oplus X'_3 \oplus \ldots \oplus X'_d$, even though $\Delta Y_1 = \Delta Y_2 = \ldots = \Delta Y_d = 0$, it can not be ensured that $\Delta Y_{d+1} = 0$ and $\Delta Y_{d+2} = 0$. That is, it is intractable to find out two d-block messages M_1 and M_2 for collision by the birthday attack [9]. Therefore, the D-way method is ineffective on hash functions based on the MFRI structure.

4 Performance Analysis of the Hash Ring-iterative Structure

For the same message M, MFRI is two f mapping operations, n reverse code operations, n modular addition operations and $(n - 1)$ exclusive OR operations more than the MD structure. Reverse code, modular addition and exclusive OR are fundamental operations, and they can not expend too much time. Hence, the MFRI structure has comparatively fast operation speed. The two extra variables in memory space need to be increased respectively for the feedforward and feedback values. The initial values of these two variables may be set to zero. Then, the feedforward variable admit $X_2, X_3, \ldots,$ and X_n one by one by exclusive OR, and the feedback variable admit $\leftrightarrows Y_1, \leftrightarrows Y_2, \ldots$ and $\leftrightarrows Y_n$ one by one by modular addition.

At present, the tree structure and zipper structure for hash functions are also believed to be more secure than the MD structure [10][11]. However, the MFRI structure is more applicable than the tree structure since the compression mapping f in any existing hash function may be transplanted into the MFRI structure with no change, and is more efficient than the zipper structure since the number of time of operation on the mapping f in the MFRI structure is roughly half as many as in the zipper structure.

5 Conclusions

In this paper, we have proposed the ring-iterative structures with feedback, and proved that the MFRI structure is more secure than the MD structure.

At the time every iteration output is fed back, first to do a reversal transform is important, which makes it impossible that ΔY_n^{d} is derived directly from $\Delta Y_1, \Delta Y_2, \ldots,$ and ΔY_n.

It is known from section 3.1 that there is some comparability between the MFRI structure and the MD structure. The last two extra blocks in MFRI may be regarded as an extension of the MD padding. Therefore, for the same compression f, if the hash output of the MD structure is uniform, independent and random, the hash output of the MFRI structure is also uniform, independent and random. If the MD structure can cause the avalanche effect of the hash output, the MFRI structure can also cause the avalanche effect of the hash output.

It is should be noted that for the input message X, only if it has at least two blocks does the MFRI structure

take effect.

For the existing hash functions ─ MD5 and SHA-1 for example, if their compression functions are extracted and transplanted into the MFRI structure, the preceding analysis shows that the existing attack methods will be ineffective on the newly forming hash functions.

Acknowledgment

The authors would like to thank Zhaozhi Zhang, and Xinxin Niu for their important suggestions, corrections, and encouragements.

References

[1] A. J. Menezes, P. Van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, London, 1997, ch. 9.

[2] Ivan Damgard, "A design principle for hash functions," *Advances in Cryptology: CRYPTO 89*, Springer-Verlag, 1989, pp. 416–427.

[3] Ralph Merkle, "One way hash functions and DES," *Advances in Cryptology: CRYPTO 89*, Springer-Verlag, 1989, pp. 428–446.

[4] Eli Biham, Rafi Chen, and Antoine Joux etc, "Collisions of SHA-0 and Reduced SHA-1," *Advances in Cryptology-EUROCRYPT 2005*, Springer-Verlag, 2005, pp. 36–57.

[5] Xiaoyun Wang, Yiqun Lisa Yin, and Hongbo Yu, "Finding collisions in the full SHA-1," *Advances in Cryptology—CRYPTO '05*, Springer-Verlag, 2005, pp. 17–36.

[6] Xiaoyun Wang and Hongbo Yu, "How to Break MD5 and Other Hash Functions," *Advances in Cryptology – EUROCRYPT '05*, Springer-Verlag, 2005, pp. 19–35.

[7] Praveen Gauravaram, William Millan, Ed Dawson, and Kapali Viswanathan, "Constructing Secure Hash Functions by Enhancing Merkle-Damgård Construction," *Australasian Conference on Information Security and Privacy '06*, LNCS, v4058, Springer-Verlag, 2006, pp. 407-420.

[8] Xiaoyun Wang, Dengguo Feng, Xuejia Lai, and Hongbo Yu, *Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEND*, rump seession of Crypto '04, e-print, 2004.

[9] Antoine Joux, "Multicollisions in Iterated Hash Functions. Application to Cascaded Constructions," *Advances in Cryptology-CRYPTO '04*, Springer-Verlag, 2004, pp. 306–316.

[10] Oded Goldreich, *Foundations of Cryptography Volume II, Basic Applications*, Cambridge University Press, Cambridge, 2004, pp. 521–523.

[11] Moses Liskov, *Constructing Secure Hash Functions from Weak Compression Functions: The Case for Non-Streamable Hash Functions*, available: www.cs.wm.edu/~mliskov/hash.pdf, 2006.