A shortened version of this paper will appear under the same title in *ICICS 2005* (S. Qing, W. Mao, J. Lopez ed.), *LNCS*, Springer-Verlag.

On the Security of Encryption Modes of MD4, MD5 and HAVAL*

Jongsung Kim[1**], Alex Biryukov[1], Bart Preneel[1], and Sangjin Lee[2]

[1] Katholieke Universiteit Leuven, ESAT/SCD-COSIC, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium
{Kim.Jongsung, Alex.Biryukov, Bart.Preneel}@esat.kuleuven.be
[2] Center for Information Security Technologies(CIST), Korea University, Seoul, Korea
{sangjin}@cist.korea.ac.kr

Abstract. MD4 is a cryptographic hash function introduced in 1990 by Rivest. After MD4 was proposed, several hash functions such as MD5, HAVAL, RIPEMD, RIPEMD-160, SHA-1 and SHA-256 were designed based on the MD4 structure. In this paper, we cryptanalyze the compression functions of MD4, MD5 and 4-, 5-pass HAVAL in encryption modes. We exploit the recently proposed related-key rectangle and boomerang techniques to show non-randomness of MD4, MD5 and 4-, 5-pass HAVAL and to distinguish them from a randomly chosen cipher. The attacks are highly practical and have been confirmed by our experiments.

1 Introduction

Hash functions are an important type of cryptographic algorithms and are widely used in cryptography such as digital signature, data authentication, e-cash and many other applications. Hash functions are at work in the millions of transactions that take place on the internet every day. The purpose of the use of hash functions in many cryptographic protocols is to ensure their security as well as improve their efficiency. The most widely used hash functions are dedicated hash functions such as MD5 [14] and SHA-1 [24], which follow the design principle of MD4.

MD4 [13] is a cryptographic hash function introduced in 1990 by Rivest. It uses basic arithmetic operations and several Boolean functions which are suitable for fast software implementations on 32-bit processors. After MD4 was published, several hash functions based on the design philosophy of MD4 have been proposed: MD5 [14], HAVAL [22], RIPEMD [23], RIPEMD-160 [5], SHA-1 [24], SHA-256 [25], etc.

In 2004 and 2005 several important cryptanalytic articles [1, 2, 17–20] have been published that demonstrate collisions for the MD4-family of hash functions. Especially, a "precise" differential attack proposed by Wang et al. enables us to greatly improve previous known collision attacks of MD4, MD5, HAVAL, RIPEMD, SHA-0 and SHA-1 [17–20].

* This work was supported in part by the Concerted Research Action (GOA) Ambiorics 2005/11 of the Flemish Government and by the European Commission through the IST Programme under Contract IST2002507932 ECRYPT and in part by the MIC(Ministry of Information and Communication), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute of Information Technology Assessment).

** The first author was financed by a Ph.D. grant of the Katholieke Universiteit Leuven and supported by the Korea Research Foundation Grant funded by the Korean Government(MOEHRD) (KRF-2005-213-D00077).

Table 1. Distinguishing Attacks of Encryption Modes of MD4, MD5 and HAVAL

Primitive	Type of Attack	Number of Source Keys	Data Complexity	Number of Weak Keys	Paper
MD4	R	2	2^{69}RK-CP	·	This paper
	B†	**2**	$\mathbf{2^{18}}$**RK-CP/**$\mathbf{2^{18}}$**RK-ACC**	·	This paper
	B†	**2**	**2RK-CP/2RK-ACC**	$\mathbf{2^{320}}$	This paper
	R	4	2^{69}RK-CP	·	This paper
	B†	**4**	$\mathbf{2^{6}}$**RK-CP/**$\mathbf{2^{6}}$**RK-ACC**	·	This paper
	B†	**4**	**2RK-CP/2RK-ACC**	$\mathbf{2^{384}}$	This paper
MD5	D	1	2^{50}CP	·	[15]
	R	2	$2^{102.8}$RK-CP	·	This paper
	B	2	$2^{80.6}$RK-CP/$2^{78.6}$RK-ACC	·	This paper
	B†	**2**	**12RK-CP/12RK-ACC**	$\mathbf{2^{96}}$	This paper
	R	4	$2^{71.1}$RK-CP	·	This paper
	B†	**4**	$\mathbf{2^{13.6}}$**RK-CP/**$\mathbf{2^{11.6}}$**RK-ACC**	·	This paper
	B†	**4**	**6RK-CP/6RK-ACC**	$\mathbf{2^{352}}$	This paper
HAVAL (4 passes)	D	1	2^{127}CP	·	[21]
	R	2	$2^{148.5}$RK-CP	·	This paper
	B†	**2**	$\mathbf{2^{37.9}}$**RK-CP/**$\mathbf{2^{35.9}}$**RK-ACC**	·	This paper
	B†	**2**	$\mathbf{2^{12.3}}$**RK-CP/**$\mathbf{2^{12.3}}$**RK-ACC**	$\mathbf{2^{576}}$	This paper
	R	4	2^{133}RK-CP	·	This paper
	B†	**4**	$\mathbf{2^{11.6}}$**RK-CP/**$\mathbf{2^{9.6}}$**RK-ACC**	·	This paper
	B†	**4**	**32RK-CP/32RK-ACC**	$\mathbf{2^{896}}$	This paper
HAVAL (5 passes)	D	1	2^{170}CP	·	[21]
	R	2	$2^{188.6}$RK-CP	·	This paper
	B	2	$2^{127.9}$RK-CP/$2^{125.9}$RK-ACC	·	This paper
	R	4	$2^{158.5}$RK-CP	·	This paper
	B	4	2^{63}RK-CP/2^{61}RK-ACC	·	This paper

†: the attack can be implemented in a real time
D: Differential, B: Boomerang, R: Rectangle
RK: Related-Key, CP: Chosen Plaintexts, ACC: Adaptively Chosen Ciphertexts
Time complexity is the same as the amount of data complexity

There have been also several cryptanalytic articles that investigate non-randomness of the compression functions of MD5, HAVAL, SHA-1 and SHA-256 in encryption mode. The encryption modes of SHA-1 and SHA-256 have been proposed in the NESSIE project, which are called SHACAL-1 and SHACAL-2 [7], respectively. For the encryption modes of SHA-1 and SHA-256, the security has been checked against various block cipher cryptanalyses [3, 6, 8–12, 15, 16], while differential cryptanalysis has been applied to the encryption modes of MD5 and HAVAL [15, 21].

In this paper, we check the security of encryption modes of MD4, MD5 and HAVAL against the recently proposed related-key rectangle and boomerang attacks [4, 9, 10, 12], and we compare our results with the previous ones in terms of distinguishing attacks. Especially, we can distinguish the encryption modes of MD4, MD5 and 4-pass HAVAL from a randomly chosen cipher in practice by using a related-key boomerang attack. Furthermore, we can distinguish them more efficiently for a large class of weak keys (i.e., special subset of messages in hash mode). See Table 1 for a summary of our results and a comparison with the previous attacks.

The outline of this paper is as follows: in Sect. 2, we give a brief description of MD4, MD5 and HAVAL. In Sect. 3, we describe the methods of related-key rectangle and boomerang attacks. Section 4 presents our related-key rectangle and boomerang attacks on encryption modes of MD4, MD5 and HAVAL. We conclude in Sect. 5.

2 Description of MD4, MD5 and HAVAL

The MD4, MD5 and HAVAL hash functions are message digest algorithms which compress any arbitrary-bit length message into a hash value with a small and fixed bit-length. These hash functions are performed based on the well-known Davies-Meyer construction, which is described as follows. Before applying the hash function to a message M of arbitrary bit-length, it is divided into l-bit sub-messages $M_0, M_1, \cdots, M_{n-1}$, where l is specified. Then the t-bit hash value I_n for the message M is computed as follows:

$$I_0 = IV;\ I_{i+1} = \mathbf{com}(I_i, M_i) = E(I_i, M_i) + I_i \quad \text{for} \quad 0 \le i < n, \tag{1}$$

where IV is a t-bit fixed initial value, **com** is a compression function and E is an iterative step function. In MD4, MD5 and HAVAL, the function E is composed of 3, 4 or 5 passes and in each pass there are 16 or 32 steps that use only simple basic operations and Boolean functions on 32-bit words. The t-bit input I_i is loaded into $t/32$ 32-bit registers denoted $(A^0, B^0, \cdots)$ and the l-bit message block is divided into $l/32$ 32-bit words denoted $(X^0, X^1, \cdots, X^{l/32})$. The $t/32$ registers are updated through a number of steps. In each pass, every message word X^i is used exactly once in a specified order, and a fixed Boolean function f and 32-bit constants Ctr are used.

Table 2 shows the parameters of MD4, MD5 and HAVAL, and Fig. 1 shows the r-th step of MD4, MD5 and HAVAL. In Fig. 1, the rotation amount s^r is specified. See [13, 14, 22] for details.

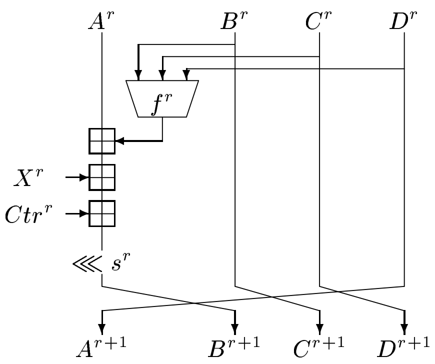
Encryption Modes of MD4, MD5 and HAVAL: Each of the steps described in Fig. 1 is an invertible function for each message word X^r. Hence, if we insert a secret key in the message part of M_i and a plaintext in the chaining value part of I_i, we get an invertible function from a compression function by removing the final addition with the previous chaining value. That is, $E(I_i, M_i)$ of Eq. (1) can be used in encryption mode $E(P, K)$, where P is a plaintext and K is a secret key. Therefore, according to Table 2, the encryption modes of MD4 and MD5 are 128-bit block ciphers with 512-bit keys and with 48 and 64 rounds, respectively, and the encryption mode of HAVAL is a 256-bit block cipher with 1024-bit keys and with 96, 128 or 160 round. In the encryption modes of MD4, MD5 and HAVAL, we use the terminology *rounds* instead of *steps* and we use the notation P and K for a plaintext and a key, respectively.

Table 2. Parameters of MD4, MD5 and HAVAL

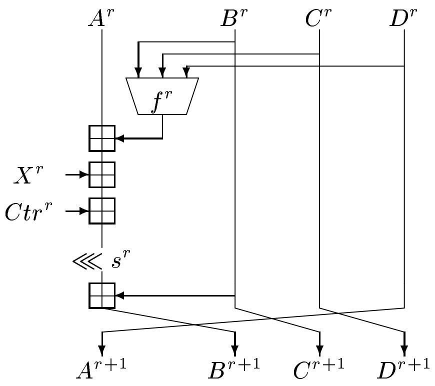
Hash Functions	Bit-Length of Message Block (l)	Bit-Length of Hash Value (t)	# of Passes	# of Steps in a Pass	Total # of Steps
MD4	512	128	3	16	48
MD5	512	128	4	16	64
HAVAL	1024	256	3,4 or 5	32	96, 128 or 160

3 Related-Key Rectangle and Boomerang Attacks

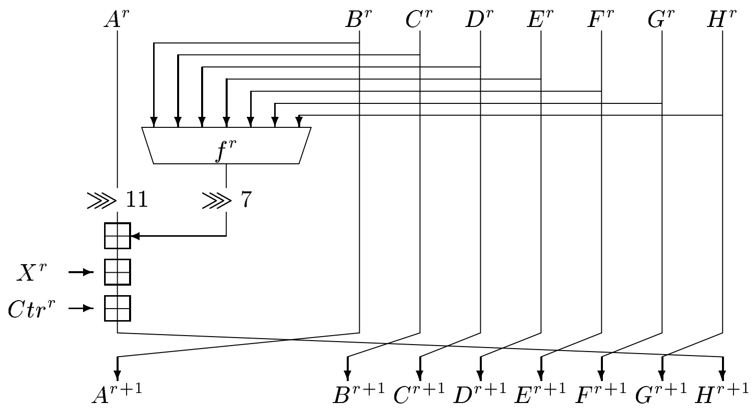
Related-key rectangle and boomerang attacks were presented in several papers [4, 9, 10, 12]. They exploit related-key rectangle and boomerang distinguishers based on



(One step of MD4)



(One step of MD5)



(One step of HAVAL)

Fig. 1. The r-th Step Functions of the Compression Functions of MD4, MD5 and HAVAL

2, 4 or 256 related keys. In this paper, we use related-key rectangle and boomerang distinguishers based on 2 or 4 related keys.

The following notations are used to facilitate the descriptions of related-key rectangle and boomerang distinguishers.

- $E : \{0,1\}^k \times \{0,1\}^n \rightarrow \{0,1\}^n$: a block cipher that uses $\{0,1\}^k$ and $\{0,1\}^n$ as key space and plaintext/ciphertext space, respectively.
- $E = E^1 \circ E^0$ (i.e., $E_K(P) = E_K^1 \circ E_K^0(P)$) : E is composed of E^0 and E^1 (E first performs E^0 and then E^1), where K is a master key and P is a plaintext.
- $p(\alpha, \beta, \Delta K)$: a probability of a related-key differential $\alpha \rightarrow \beta$ for E^0 under the related-key difference ΔK, i.e., $p(\alpha, \beta, \Delta K) = \Pr_{X,K}[E_K^0(X) \oplus E_{K \oplus \Delta K}^0(X \oplus \alpha) = \beta]$. Note that this is same as a probability of a related-key differential $\beta \rightarrow \alpha$ for $(E^0)^{-1}$ under the related-key difference ΔK.
- $q(\gamma, \delta, \Delta K)$: a probability of a related-key differential $\gamma \rightarrow \delta$ for E^1 under the related-key difference ΔK, i.e., $q(\gamma, \delta, \Delta K) = \Pr_{X,K}[E_K^1(X) \oplus E_{K \oplus \Delta K}^1(X \oplus \gamma) = \delta]$. Note that this is same as a probability of a related-key differential $\delta \rightarrow \gamma$ for $(E^1)^{-1}$ under the related-key difference ΔK.
- $p(D, \beta, \Delta K)$: a probability of a related-key truncated differential $\beta \rightarrow \alpha'$ for $(E^0)^{-1}$ under the related-key difference ΔK, where D is a nonempty set and $\alpha' \in D$, i.e., $p(D, \beta, \Delta K) = \Pr_{X,K}[(E_K^0)^{-1}(X) \oplus (E_{K \oplus \Delta K}^0)^{-1}(X \oplus \beta) \in D]$.
- $q(\gamma, D, \Delta K)$: a probability of a related-key truncated differential $\gamma \rightarrow \delta'$ for E^1 under the related-key difference ΔK, where D is a nonempty set and $\delta' \in D$, i.e., $q(\gamma, D, \Delta K) = \Pr_{X,K}[E_K^1(X) \oplus E_{K \oplus \Delta K}^1(X \oplus \gamma) \in D]$.

We first describe a related-key rectangle distinguisher based on two related keys. The related-key rectangle distinguisher works in the following process (refer to Fig. 2).

- Choose two plaintexts P_0 and P_1^* at random and compute two other plaintexts $P_0^* = P_0 \oplus \alpha$ and $P_1 = P_1^* \oplus \alpha$.
- With a chosen plaintext attack, obtain the corresponding ciphertexts $C_0 = E_K(P_0)$, $C_1 = E_K(P_1)$, $C_0^* = E_{K^*}(P_0^*)$ and $C_1^* = E_{K^*}(P_1^*)$, where $K \oplus K^* = \Delta K$.
- Check if $C_0 \oplus C_1^*$, $C_0^* \oplus C_1 \in D$.

What is the probability that the ciphertext quartet satisfies the last D test? The probability is computed as follows. Let X_0, X_1, X_0^* and X_1^* denote the encrypted values of P_0, P_1, P_0^* and P_1^* under E^0, respectively. Then the probabilities that $X_0 \oplus X_0^* = \beta$ and $X_1 \oplus X_1^* = \beta'$ are $p(\alpha, \beta, \Delta K)$ and $p(\alpha, \beta', \Delta K)$, respectively. In the above process we randomly choose two plaintexts P_0 and P_1^* and thus we expect $X_0 \oplus X_1^* = \gamma$ with probability 2^{-n}. Therefore, for any β, β' and γ, $X_0 \oplus X_0^* = \beta$, $X_1 \oplus X_1^* = \beta'$ and $X_0 \oplus X_1^* = \gamma$ (as in these cases $X_0^* \oplus X_1 = (X_0 \oplus \beta) \oplus (X_1^* \oplus \beta') = \beta \oplus \beta' \oplus \gamma$) hold with probability $p(\alpha, \beta, \Delta K) \cdot p(\alpha, \beta', \Delta K) \cdot 2^{-n}$. Since the probabilities of related-key truncated differentials $\gamma \rightarrow \delta(\in D)$ and $\beta \oplus \beta' \oplus \gamma \rightarrow \delta'(\in D)$ for E^1 under related-key difference ΔK are $q(\gamma, D, \Delta K)$ and $q(\gamma \oplus \beta \oplus \beta', D, \Delta K)$, the probability that the last D test in the above process is satisfied equals

$$Pr[REC\text{-}2] = \sum_{\beta, \beta', \gamma} p(\alpha, \beta, \Delta K) \cdot p(\alpha, \beta', \Delta K) \cdot 2^{-n} \cdot q(\gamma, D, \Delta K) \cdot q(\gamma \oplus \beta \oplus \beta', D, \Delta K)\,.$$

On the other hand, for a random cipher, the D test holds with probability $|D|^2 \cdot 2^{-2n}$ and thus if the above probability is larger than $|D|^2 \cdot 2^{-2n}$ for any triple $(\alpha, D, \Delta K)$, the related-key rectangle distinguisher based on two related keys can be used to distinguish E from a random cipher.

How many plaintext pairs are required to get at least two ciphertext quartets (this amount of quartets will be used in our attacks) that satisfy the D test? If the number of plaintext pairs (P_i, P_i^*) we collect is m, we can generate $m^2 \cdot 2^{-1}$ quartets and thus we have at least $m^2 \cdot 2^{-1} \cdot Pr[REC{-}2]$ ciphertext quartets which satisfy the D test. Therefore, in order to get at least 2 such quartets we need about $4 \cdot (Pr[REC{-}2])^{-1/2}$ chosen plaintext queries. It means that the number of required plaintexts to use this distinguisher is at least $2^{n/2}$. However, under an adaptive chosen plaintext and ciphertext attack we can make a related-key boomerang distinguisher which can remove the factor $2^{n/2}$ in the data requirement. The related-key boomerang distinguisher based on two related keys works as follows (refer to Fig. 2).

- Choose two plaintexts P_0 and P_0^* such that $P_0 \oplus P_0^* = \alpha$, and obtain the corresponding ciphertexts $C_0 = E_K(P_0)$ and $C_0^* = E_{K^*}(P_0^*)$, where $K \oplus K^* = \Delta K$.
- Compute other two ciphertexts $C_1 = C_0^* \oplus \delta$ and $C_1^* = C_0 \oplus \delta$, and obtain the corresponding plaintexts $P_1 = E_K^{-1}(C_1)$ and $P_1^* = E_{K^*}^{-1}(C_1^*)$.
- Check $P_1 \oplus P_1^* \in D$.

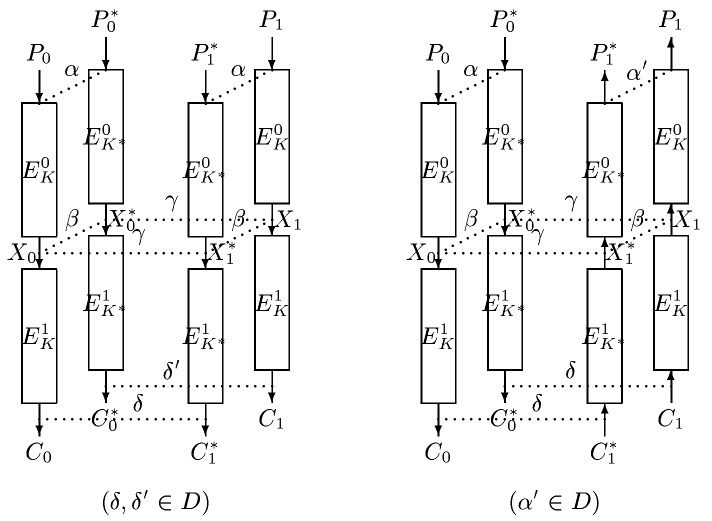


Fig. 2. Related-Key Rectangle Distinguisher (Left) and Related-Key Boomerang Distinguisher (Right) Based on Two Related Keys

Similarly, we can check the probability that the last α' test is satisfied. The probability that $X_0 \oplus X_0^* = \beta$ is $p(\alpha, \beta, \Delta K)$ (in the encryption direction) and the probabilities that $X_0^* \oplus X_1 = \gamma$ and $X_0 \oplus X_1^* = \gamma'$ are $q(\gamma, \delta, \Delta K)$ and $q(\gamma', \delta, \Delta K)$ (in the decryption direction), respectively. Therefore, for any β, γ and γ', $X_0 \oplus X_0^* = \beta$, $X_0^* \oplus X_1 = \gamma$ and $X_0 \oplus X_1^* = \gamma'$ (as in these cases $X_1 \oplus X_1^* = (X_0^* \oplus \gamma) \oplus (X_0 \oplus \gamma') = \gamma \oplus \gamma' \oplus \beta$) hold with probability $p(\alpha, \beta, \Delta K) \cdot q(\gamma, \delta, \Delta K) \cdot q(\gamma', \delta, \Delta K)$. Since the probability of related-key truncated differential $\gamma \oplus \gamma' \oplus \beta \rightarrow \alpha' (\in D)$ for $(E^0)^{-1}$ under related-key difference ΔK is $p(D, \gamma \oplus \gamma' \oplus \beta, \Delta K)$, the probability that satisfies the last D test in the above process is

$$Pr[BOO{-}2] = \sum_{\beta,\gamma,\gamma'} p(\alpha, \beta, \Delta K) \cdot q(\gamma, \delta, \Delta K) \cdot q(\gamma', \delta, \Delta K) \cdot p(D, \beta \oplus \gamma \oplus \gamma', \Delta K)\,.$$

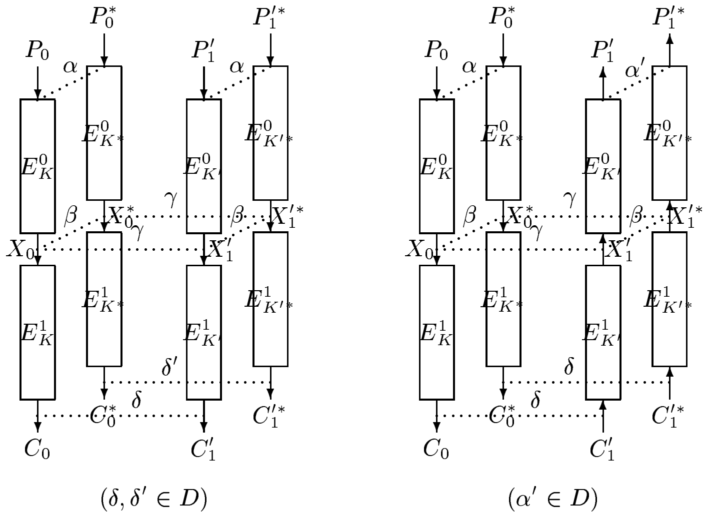


Fig. 3. Related-Key Rectangle Distinguisher (Left) and Related-Key Boomerang Distinguisher (Right) Based on Four Related Keys

Since for a random cipher, the D test holds with probability $|D| \cdot 2^{-n}$, $Pr[BOO{-}2] > |D| \cdot 2^{-n}$ must hold for the related-key boomerang distinguisher to work. Moreover, $2 \cdot (Pr[BOO{-}2])^{-1}$ chosen plaintext pairs and $2 \cdot (Pr[BOO{-}2])^{-1}$ adaptively chosen ciphertext pairs produce at least 2 quartets that satisfy the D test.

Related-key rectangle and boomerang distinguishers based on four related keys are the same as the previous distinguishers except for using four related keys K, K^*, K' and K'^* such that $K \oplus K^* = K' \oplus K'^* = \Delta K$ and $K \oplus K' = K^* \oplus K'^* = \Delta K'$ (see Fig. 3). Similarly, we can calculate the probabilities of related-key rectangle and boomerang distinguishers and the required data complexity. For a related-key rectangle distinguisher, the probability is

$$Pr[REC{-}4] = \sum_{\beta,\beta',\gamma} p(\alpha,\beta,\Delta K) \cdot p(\alpha,\beta',\Delta K) \cdot 2^{-n} \cdot q(\gamma, D, \Delta K') \cdot q(\gamma \oplus \beta \oplus \beta', D, \Delta K') \,.$$

If the number of plaintext pairs (P_i, P_i^*) and $(P_i', P_i'^*)$ we collect is m, respectively, we can generate m^2 quartets and thus we have at least $m^2 \cdot Pr[REC{-}4]$ ciphertext quartets which satisfy the D test. Therefore, in order to get at least 2 such quartets we need about $4 \cdot (Pr[REC{-}4])^{-1/2} \cdot 2^{1/2}$ chosen plaintext queries.

For a related-key boomerang distinguisher, the probability[1] is

$$Pr[BOO{-}4] = \sum_{\beta,\gamma,\gamma'} p\,(\alpha,\beta,\Delta K) \cdot q(\gamma,\delta,\Delta K') \cdot q(\gamma',\delta,\Delta K') \cdot p(D, \beta \oplus \gamma \oplus \gamma', \Delta K) \,.$$

So the data requirement to generate at least two good quartets is about $2 \cdot (Pr[BOO{-}4])^{-1}$ chosen plaintext pairs and $2 \cdot (Pr[BOO{-}4])^{-1}$ adaptively chosen ciphertext pairs.

Table 3 summarizes probabilities of related-key rectangle and boomerang distinguishers and the required data complexity, which is useful to calculate our data complexity of Table 1.

[1] If the set D has a single element α in $Pr[BOO{-}4]$ and the set D has a single element δ in $Pr[REC{-}4]$, it holds $Pr[REC{-}4] = 2^{-n} \cdot Pr[BOO{-}4]$. This relationship also holds between $Pr[BOO{-}2]$ and $Pr[REC{-}2]$. We use these relationships to estimate $Pr[REC{-}2]$ and $Pr[REC{-}4]$ in our attacks.

Table 3. Probabilities and Data Requirements for Related-Key Rectangle and Boomerang Distinguishers

Distinguisher	Probability	Data Complexity for Generating Two Good Quartets
2-Key Rectangle	$Pr[REC\text{–}2]$	$2^2 \cdot (Pr[REC\text{–}2])^{-1/2}$ RK-CP
4-Key Rectangle	$Pr[REC\text{–}4]$	$2^{2.5} \cdot (Pr[REC\text{–}4])^{-1/2}$ RK-CP
2-Key Boomerang	$Pr[BOO\text{–}2]$	$2^2 \cdot (Pr[BOO\text{–}2])^{-1}$ RK-CP/$2^2 \cdot (Pr[BOO\text{–}2])^{-1}$ RK-ACC
4-Key Boomerang	$Pr[BOO\text{–}4]$	$2^2 \cdot (Pr[BOO\text{–}4])^{-1}$ RK-CP/$2^2 \cdot (Pr[BOO\text{–}4])^{-1}$ RK-ACC

4 Related-Key Rectangle and Boomerang Attacks on Encryption Modes of MD4, MD5 and HAVAL

In this section, we present related-key rectangle and boomerang attacks on the encryption modes of MD4, MD5 and HAVAL. First, we present related-key rectangle and boomerang distinguishers of MD4 and show how to use them to distinguish MD4 from a random cipher. Second, we apply related-key rectangle and boomerang attacks to MD5 and HAVAL.

4.1 Cryptanalysis of MD4

In MD4 the message expansion algorithm is a linear function in each pass every message word is used exactly once in a specified order. It means that in the encryption mode of MD4 the key scheduling algorithm is the same linear function of the message expansion algorithm of MD4. We exploit the simple linear key scheduling algorithm in our distinguishers. The main idea behind our constructions of related-key rectangle and boomerang distinguishers based on two related keys is to give a difference in one key word whose interval between the first and third passes is as wide as possible. Let the round numbers involved in such a key word in the three passes be r_1, r_2 and r_3. Then we can make probability-one differentials for rounds $r_1 \sim r_2'$ and $r_2' \sim r_3$ by giving appropriate differences α and γ, respectively, where r_2' is a certain number between r_1 and r_2. Therefore, in order to find distinguishers with high probablilities we should find one key word for which the interval of $r_1 \sim r_3$ is as wide as possible.

In our observation giving a difference in the 3-rd key word provides the best probabilities to our distinguishers, which are described as follows. In MD4 there exist a related-key differential characteristic $(0, e_{31}, 0, 0) \rightarrow (0, 0, 0, 0)$ for rounds $0 \sim 27$ with probability 2^{-2} (denoted p) and a related-key differential characteristic $(e_{31}, 0, 0, 0) \rightarrow (e_2, e_{5,17,26,28}, e_{13,22}, e_{11})$ for rounds $28 \sim 47$ with probability 2^{-7} (denoted q) under key difference $\Delta K = (0, 0, 0, \Delta K^3 = e_{31}, 0, \cdots, 0)$, where e_i represents a 32-bit word that has $0's$ in all bit positions except for bit i and $e_{i_1, \cdots, i_k}$ represents $e_{i_1} \oplus \cdots \oplus e_{i_k}$ (in our notation the right most bit is referred to as the 0-th bit, i.e., the least significant bit). See Table 4 for more details. The notation used in Table 4 is essential in our distinguishing attacks. The *REC-2* and *BOO-2* rows represent probabilities which will be used in related-key rectangle and boomerang attacks, respectively and the BOO^W*-2* row represents a weak key class as well as a probability which will be used in a related-key boomerang attack under a weak key class. The notation $(r \rightarrow r')^{1 \text{ or } 2}$ means related-key differentials for rounds from r to r' (which have the fixed difference in round r or r' described in the table) used in our distinguishers. Here, the superscript 1 or 2 represents how many times related-key differentials are used in our distinguishers. Note that if $r > r'$ then the related-key differential works through decryption process.

Table 4. Related-Key Distinguishers of MD4 (Two Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔK^i	Prob.
0	0	e_{31}	0	0	0	1
1	0	0	e_{31}	0	0	2^{-1}
2	0	0	0	e_{31}	0	2^{-1}
3	e_{31}	0	0	0	$e_{31}(=\Delta K^3)$	1
4	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
27	0	0	0	0	0	1
	0	0	0	0		$p = 2^{-2}$
28	e_{31}	0	0	0	$e_{31}(=\Delta K^3)$	1
29	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
43	0	0	0	0	0	1
44	0	0	0	0	$e_{31}(=\Delta K^3)$	1
45	0	e_2	0	0	0	2^{-1}
46	0	e_{11}	e_2	0	0	2^{-2}
47	0	$e_{13,22}$	e_{11}	e_2	0	2^{-4}
	e_2	$e_{5,17,26,28}$	$e_{13,22}$	e_{11}		$q = 2^{-7}$
REC-2	$(0 \to 27)^2$, $(28 \to 45)^2$					Pr[REC-2] $\approx 2^{-134}$
BOO-2	$(0 \to 27)$, $(47 \to 28)^2$, $(27 \to 3)$					Pr[BOO-2] $\approx 2^{-16}$
BOO^W*-2*	Fixed $K^{0,1,2,7,11,15}$, $(3 \to 27)$, $(44 \to 28)^2$, $(27 \to 3)$					Pr[BOO-2] $= 1$

In order to estimate $Pr[BOO{-}2]$ we have carried out experiments on a number of related keys with 2^{23} chosen plaintext pairs and 2^{23} adaptively chosen ciphertext pairs each and we have observed 136, 115, 136, 125, 132, 130, 132, 131, 119, 144, $\cdots$ boomerangs returning for each related-key. This simulation result provides that the probability $Pr[BOO{-}2]$ is approximately 2^{-16} (which can be also calculated from the probabilities of related-key differential characteristics in Table 4). We can use the value of $Pr[BOO{-}2]$ or the probabilities of related-key differential characteristics in Table 4 to obtain the probability $Pr[REC{-}2]$.

We now present a distinguishing attack of the encryption mode of MD4 using a related-key rectangle distinguisher in Table 4. As stated in Table 4, in this attack we use $Pr[REC{-}2] \approx 2^{-134}$, which is derived from $p = 2^{-2}$ and $q' = 2^{-1}$ (the q' is the probability for rounds 28 $\sim$ 45 in Table 4). In order to use $p = 2^{-2}$ we should collect plaintext pairs (P_i, P_i^*) which satisfy not only the $(0, e_{31}, 0, 0)$ difference but also $c_{31} = d_{31} = 0$, where c_j and d_j represent the j-th bits of words C and D of P_i, respectively. Moreover, since we use $q' = 2^{-1}$ for rounds 28 $\sim$ 45 in our attack, our desired δ after round 47 can be any one of the differences which can be derived from the input difference of round 46, $(0, e_{11}, e_2, 0)$. It is easy to see that the number of all possible $\delta's$ is at most 2^{36}. We denote the set of all these possible $\delta's$ by $\mathcal{O}$. Next we describe our distinguishing attack on the encryption mode of MD4 using the related-key rectangle distinguisher.

1. Prepare 2^{68} plaintext pairs (P_i, P_i^*), $i = 0, 1, \cdots, 2^{68}-1$ with difference $(0, e_{31}, 0, 0)$ and $c_{31} = d_{31} = 0$.
2. With a chosen plaintext attack, obtain the 2^{68} corresponding ciphertext pairs (C_i, C_i^*), i.e., $C_i = E_K(P_i)$ and $C_i^* = E_{K^*}(P_i^*)$, where E is either MD4 or a randomly chosen cipher and $K \oplus K^* = (0, 0, 0, \Delta K^3 = e_{31}, 0, \cdots, 0)$.

3. If there exists at least one ciphertext quartet such that $C_i \oplus C_j^*, C_i^* \oplus C_j \in \mathcal{O}$ for $0 \le i \ne j \le 2^{68} - 1$, we identify E as MD4. Otherwise, we identify E as a randomly chosen cipher.

From the 2^{68} plaintext pairs we obtain 2^{135} quartets. Since our related-key rectangle distinguisher has a probability of $(2^{-2})^2 \cdot (2^{-1})^2 \cdot 2^{-128} = 2^{-134}$, if E is MD4, this attack will succeed with a probability of $1 - (1 - 2^{-134})^{2^{135}} \approx 0.86$. On the other hand, in case E is a randomly chosen cipher, the probability that each ciphertext quartet satisfies one of all possible δ's is less than $(\frac{2^{36}}{2^{128}})^2 = 2^{-184}$, so, in this case this attack will succeed with a probability of $(1 - 2^{-184})^{2^{135}} \approx 1$. Therefore, the success rate of this attack is about $\frac{1}{2} \cdot 0.86 + \frac{1}{2} \cdot 1 = 0.93$.

Based on the foregoing two related-key differentials we can also exploit a boomerang technique to distinguish MD4 from a randomly chosen cipher. In a boomerang technique we use $Pr[BOO\text{-}2] \approx 2^{-16}$. Since we use related-key differentials for rounds $27 \sim 3$ in the upper right in Fig. 2, our desired α before round 0 can be any one of the differences which can be derived from the input difference of round 3, $(e_{31}, 0, 0, 0)$, through the inverse direction. It is easy to see that the all possible $\alpha's$ are $(0, e_{31}, 0, 0), (e_{31}, e_{31}, 0, 0), (0, e_{31}, e_{31}, e_{31})$ and $(e_{31}, e_{31}, e_{31}, e_{31})$. We denote the set of all these possible $\alpha's$ by $\mathcal{I}$. Here is our distinguishing attack on the encryption mode of MD4 using the related-key boomerang distinguisher.

1. Prepare 2^{17} plaintext pairs (P_i, P_i^*), $i = 0, 1, \cdots, 2^{17}-1$ with difference $(0, e_{31}, 0, 0)$ and $c_{31} = d_{31} = 0$.
2. Obtain the 2^{17} corresponding ciphertext pairs (C_i, C_i^*), i.e., $C_i = E_K(P_i)$ and $C_i^* = E_{K^*}(P_i^*)$, where E is either MD4 or a randomly chosen cipher and $K \oplus K^* = (0, 0, 0, \Delta K^3 = e_{31}, 0, \cdots, 0)$.
3. Calculate $C_i' = C_i \oplus \delta$ and $C_i'^* = C_i^* \oplus \delta$, where $\delta = (e_2, e_{5,17,26,28}, e_{13,22}, e_{11})$, and obtain the 2^{17} corresponding plaintext pairs $(P_i', P_i'^*)$, i.e., $P_i' = E_K^{-1}(C_i')$ and $P_i'^* = E_{K^*}^{-1}(C_i'^*)$.
4. If there exists at least one plaintext pair such that $P_i' \oplus P_i'^* \in \mathcal{I}$ for $0 \le i \le 2^{17}-1$, we identify E as MD4. Otherwise, we identify E as a randomly chosen cipher.

Since our related-key boomerang distinguisher has a probability of $2^{-2} \cdot (2^{-7})^2 = 2^{-16}$, if E is MD4 this attack will succeed with a probability of $1 - (1 - 2^{-16})^{2^{17}} \approx 0.86$. In order to verify this estimation we have performed hundreds of simulations using 2^{18} chosen plaintext and adaptively chosen ciphertext pairs each (in each simulation we used randomly chosen related keys and plaintext/ciphetext pairs). In our simulations we could check that about 88 among 100 tests satisfy the above distinguishing attack on average. This result is quite similar to our estimation.

On the other hand, if E is a randomly chosen cipher, the probability that each plaintext pair satisfies one of the four α's is $\frac{4}{2^{128}} = 2^{-126}$, so, in this case this attack will succeed with a probability of $(1 - 2^{-126})^{2^{17}} \approx 1$. Therefore, the success rate of this attack is almost same as that of the related-key rectangle attack.

Moreover, we can increase the boomerang probability from 2^{-16} to 1 by using some weak key class. Assume that the first three and the last three round keys $K^0, K^1, K^2, K^7, K^{11}$ and K^{15} are fixed and known to the attacker. Then we can use $p' = 1$ for rounds $3 \sim 27$ and $q' = 1$ for rounds $44 \sim 28$ in our attack under the weak key class assumption. Following is our distinguishing attack on the encryption mode of MD4 with a weak key using the related-key boomerang distinguisher .

1. Choose one input pair (X, X^*) of round 3 with difference $(e_{31}, 0, 0, 0)$ and calculate the corresponding plaintext pair (P, P^*) by using the known keys $K^0, K^1, K^2, K^{*0}, K^{*1}, K^{*2}$ and MD4.
2. Obtain the corresponding ciphertext pair (C, C^*), i.e., $C = E_K(P)$ and $C^* = E_{K^*}(P^*)$, where E is either MD4 or a randomly chosen cipher and $K \oplus K^* = (0, 0, 0, \Delta K^3 = e_{31}, 0, \cdots, 0)$.
3. Calculate the corresponding input pair (Y, Y^*) of round 45 by using the known keys $K^7, K^{11}, K^{15}, K^{*7}, K^{*11}, K^{*15}$ and MD4, and calculate $Y' = Y \oplus (0, e_2, 0, 0)$ and $Y'^* = Y^* \oplus (0, e_2, 0, 0)$. Also calculate the corresponding ciphertext pair (C', C'^*) by using the known keys $K^7, K^{11}, K^{15}, K^{*7}, K^{*11}, K^{*15}$ and MD4.
4. Obtain the corresponding plaintext pair (P', P'^*), i.e., $P' = E_K^{-1}(C')$ and $P'^* = E_{K^*}^{-1}(C'^*)$.
5. Calculate the corresponding input pair (X', X'^*) of round 3 by using the known keys $K^0, K^1, K^2, K^{*0}, K^{*1}, K^{*2}$ and MD4.
6. If $X' \oplus X'^* = (e_{31}, 0, 0, 0)$, we identify E as MD4. Otherwise, we identify E as a randomly chosen cipher.

If E is MD4, this attack will succeed with probability one (we have checked with thousands of simulations that this attack always works in MD4), but if E is a randomly chosen cipher, this attack will succeed with probability $1 - 2^{-128}$. Therefore, the success rate of this attack is almost 1.

Similarly, we can construct related-key rectangle and boomerang distinguishers based on four related keys and distinguish MD4 from a randomly chosen cipher by using them. As a compensation of the use of four related keys, these attacks are more efficient than those with two related keys. See Table 5 in Appendix B for our distinguishers and Table 1 for our results. We have also performed a series of simulations to verify related-key boomerang attacks of the encryption mode of MD4 based on four related keys. Like the above distinguishing attacks based on two related keys, we could check that our simulation results follow our estimations.

4.2 Cryptanalysis of MD5 and HAVAL

Similarly, in the MD5 and HAVAL attacks, we first find consecutive two related-key differential characteristics with high probabilities which are independent of each other, and then we can estimate the probability $Pr[BOO\text{–}k]$ on the basis of those differential characteristics by a series of simulations, where k is 2 or 4. As for 5-pass HAVAL, we can carry out an experiment on a reduced-round variant (which is truncated for the first and the last several rounds) to get $Pr[BOO\text{–}k]$ for the reduced variant and then we can use the obtained value as well as probabilities for the truncated rounds of the consecutive two related-key differential characteristics (which were found in the first stage) to estimate $Pr[BOO\text{–}2]$ for the full 5-pass HAVAL. Once we get the probability $Pr[BOO-k]$, we can estimate the probability $Pr[REC-k]$ by using the relationship between them described in Section 3. See Appendix C and D for our distinguishers of MD5 and HAVAL.

The MD5 and HAVAL attacks are exactly the same as those of MD4 except for boomerang attack procedures. The boomerang attack works by finding not only a chosen plaintext pair but also an adaptively chosen ciphertext pair that satisfy a boomerang distinguisher. For MD5 and HAVAL, once we obtain a ciphertext pair by asking for the encryption of a chosen plaintext pair, we know whether or not the adaptively chosen ciphertexts can be a boomerang candidate. For example, consider the boomerang distinguisher of Table 6 in Appendix B. Assume that the ciphertext pair obtained by asking for the encryption of a chosen plaintext pair is (C, C^*) and (a_{31}, c_{31}, d_{31}) of C or $(a^*_{31}, c^*_{31}, d^*_{31})$ of C^* is in $\{(0,0,0),(0,1,0),(1,0,1),(1,1,1)\}$. Then the adaptively chosen ciphertext pair $(C \oplus \delta, C^* \oplus \delta)$ can not satisfy our boomerang distinguisher, where $\delta = (e_5, e_5, e_5, e_5)$. That is, in this case ΔA^{63} can't be of the form e_5 since the difference induced by the Boolean function of the last round is 0 for $(C \oplus \delta, C^* \oplus \delta)$. (Note that in the boomerang attacks of MD4 we can not use this procedure since the Boolean function used in the last round of MD4 is linear.) This is the reason why the required number of queries for decryption process is smaller than that for encryption process.

We have also performed a series of simulations to verify related-key boomerang attacks of the encryption modes of MD5 and HAVAL, which are indicated in Table 1 by the symbol †. According to our probabilities of related-key boomerang distinguishers in Appendix C and Appendix D and the data complexity in Table 1, related-key boomerang attacks work with a success rate of about 0.87, when E is MD5 or HAVAL. During our simulations, we have observed that the simulation results follow our estimation of success rate. As an example, we give in Appendix A a related-key quartet, a chosen plaintext pair and an adaptively chosen ciphertext pair of MD5 obtained by the boomerang distinguisher described in Table 7.

5 Conclusion

In this paper, we have applied the recently proposed related-key rectangle and boomerang attacks to the encryption modes of MD4, MD5 and HAVAL. The MD4, MD5 and HAVAL used in encryption modes are all vulnerable to those attacks, in particular, they can be broken by related-key boomerang attacks in a real time. The attacks have been experimentally tested and run milliseconds on a PC.

Our results show that one should be very careful when using existing hash functions in encryption mode.

References

1. E. Biham and R. Chen, *Near-Collisions of SHA-0*, Advances in Cryptology – Proceedings of CRYPTO 2004, LNCS 3152, pp. 290-305, Springer-Verlag, 2004.
2. E. Biham, R. Chen, A. Joux, P. Carribault, C. Lemuet and W. Jalby, *Collisions of SHA-0 and Reduced SHA-1*, Advances in Cryptology – Proceedings of EUROCRYPT 2005, LNCS 3494, pp. 22-35, Springer-Verlag, 2005.
3. E. Biham, O. Dunkelman and N. Keller, *Rectangle Attacks on 49-Round SHACAL-1*, Proceedings of Fast Software Encryption 2003, LNCS 2887, pp. 22-35, Springer-Verlag, 2003.
4. E. Biham, O. Dunkelman and N. Keller, *Related-Key Boomerang and Rectangle Attacks*, Advances in Cryptology – Proceedings of EUROCRYPT 2005, LNCS 3494, pp. 507-525, Springer-Verlag, 2005.
5. H. Dobbertin, A. Bosselaers and B. Preneel, *RIPEMD-160: A Strengthened Version of RIPEMD*, Proceedings of Fast Software Encryption 1996, LNCS 1039, pp. 71-82, Springer-Verlag, 1996.
6. H. Handschuh, L.R. Knudsen and M.J. Robshaw, *Analysis of SHA-1 in Encryption Mode*, Proceedings of CT-RSA 2001, LNCS 2020, pp. 70-83, Springer-Verlag, 2001.
7. H. Handschuh and D. Naccache, *SHACAL : A Family of Block Ciphers*, Submission to the NESSIE project, 2002.
8. J. Kim, D. Moon, W. Lee, S. Hong, S. Lee and S. Jung, *Amplified Boomerang Attack against Reduced-Round SHACAL*, Advances in Cryptology – ASIACRYPT 2002, LNCS 2501, pp. 243-253, Springer-Verlag, 2002.
9. J. Kim, G. Kim, S. Hong, S. Lee and D. Hong, *The Related-Key Rectangle Attack - Application to SHACAL-1*, Proceedings of Australian International Conference on Information Security and Privacy 2004, LNCS 3108, pp. 123-136, Springer-Verlag, 2004.

10. J. Kim, G. Kim, S. Lee, J. Lim and J. Song, *Related-Key Attacks on Reduced Rounds of SHACAL-2*, Proceedings of INDOCRYPT 2004, LNCS 3348, pp. 175-189, Springer-Verlag, 2004.
11. S. Hong, J. Kim, G. Kim, J. Sung, C. Lee and S. Lee, *Impossible Differential Attack on 30-Round SHACAL-2*, Proceedings of INDOCRYT 2003, LNCS 2904, pp. 97-106, Springer-Verlag, 2003.
12. S. Hong, J. Kim, S. Lee and B. Preneel, *Related-Key Rectangle Attacks on Reduced Versions of SHACAL-1 and AES-192*, Proceedings of Fast Software Encryption 2005, to appear.
13. R.L. Rivest, *The MD4 Message Digest Algorithm*, Advances in Cryptology – Proceedings of CRYPTO 1990, Springer-Verlag, 1991, 303-311.
14. R.L. Rivest, *The MD5 Message Digest Algorithm*, Request for Comments (RFC 1320), Internet Activities Board, Internet Privacy Task Force, 1992.
15. M.J.O. Saarinen, *Cryptanalysis of Block Ciphers Based on SHA-1 and MD5*, Proceedings of Fast Software Encryption 2003, LNCS 2887, pp. 36-44, Springer-Verlag, 2003.
16. Y. Shin, J. Kim, G. Kim, S. Hong and S. Lee, *Differential-Linear Type Attacks on Reduced Rounds of SHACAL-2*, Proceedings of Australian International Conference on Information Security and Privacy 2004, LNCS 3108, pp. 110-122, Springer-Verlag, 2004.
17. X. Wang and H. Yu, *How to Break MD5 and Other Hash Functions*, Advances in Cryptology – Proceedings of EUROCRYPT 2005, LNCS 3494, pp. 19-35, Springer-Verlag, 2005.
18. X. Wang, X. Lai, D. Feng, H. Chen and X. Yu, *Cryptanalysis of the Hash Functions MD4 and RIPEMD*, Advances in Cryptology – Proceedings of EUROCRYPT 2005, LNCS 3494, pp. 1-18, Springer-Verlag, 2005.
19. X. Wang, H. Yu and Y.L. Yin, *Efficient Collision Search Attacks on SHA-0*, Advances in Cryptology – Proceedings of CRYPTO 2005, LNCS 3621, pp. 1-16, Springer-Verlag, 2005.
20. X. Wang, Y.L. Yin and H. Yu, *Finding Collisions in the Full SHA-1*, Advances in Cryptology – Proceedings of CRYPTO 2005, LNCS 3621, pp. 17-36, Springer-Verlag, 2005.
21. H. Yoshida, A. Biryukov, C. De Cannière, J. Lano and B. Preneel, *Non-randomness of the Full 4 and 5-pass HAVAL*, Proceedings of SCN 2004, LNCS 3352, pp. 324-336, Springer-Verlag, 2005.
22. Y. Zheng, J. Pieprzyk and J. Seberry, *HAVAL-A One-way Hashing Algorithm with Variable Length of Output*, Advances in Cryptology – Proceedings of AUSCRYPT 1992, LNCS 718, pp. 83-104, Springer-Verlag, 1993.
23. RIPE, Integrity Primitives for Secure Information Systems, Final Report of RACE Integrity Primitives Evaluation(RIPE-RACE 1040), LNCS 1007, 1995.
24. U.S. Department of Commerce. *FIPS 180-1*: Secure Hash Standard, Federal Information Processing Standards Publication, N.I.S.T., April 1995.
25. U.S. Department of Commerce.*FIPS 180-2*: Secure Hash Standard ,Federal Information Processing Standards Publication, N.I.S.T., August 2002.

A An Example of Experimental Results: A Boomerang Quartet for MD5

– *Four Related Keys:*

$$\begin{aligned}
K\ &= 0x08b870e5\ 0x33b24180\ 0x7cec25d3\ 0x3c8b1b4d\ 0x50b44c2a\ 0x14b9206c\\
&\quad 0x4aa22bc5\ 0x51f907af\ 0x1e096337\ 0x2ee81e54\ 0x2c0734bb\ 0x74231c91\\
&\quad 0x55c31f6a\ 0x7cad2870\ 0x43f418b1\ 0x59917add\\
K^*\ &= 0x08b870e5\ 0x33b24180\ 0xfcec25d3\ 0x3c8b1b4d\ 0x50b44c2a\ 0x14b9206c\\
&\quad 0x4aa22bc5\ 0x51f907af\ 0x1e096337\ 0x2ee81e54\ 0x2c0734bb\ 0x74231c91\\
&\quad 0x55c31f6a\ 0x7cad2870\ 0x43f418b1\ 0x59917add\\
K'\ &= 0x08b870e5\ 0x33b24180\ 0x7cec25d3\ 0x3c8b1b4d\ 0x50b44c2a\ 0x14b9206c\\
&\quad 0x4aa22bc5\ 0x51f907af\ 0x1e096337\ 0x2ee81e54\ 0x2c0734bb\ 0xf4231c91\\
&\quad 0x55c31f6a\ 0x7cad2870\ 0x43f418b1\ 0x59917add\\
K'^*\ &= 0x08b870e5\ 0x33b24180\ 0xfcec25d3\ 0x3c8b1b4d\ 0x50b44c2a\ 0x14b9206c\\
&\quad 0x4aa22bc5\ 0x51f907af\ 0x1e096337\ 0x2ee81e54\ 0x2c0734bb\ 0xf4231c91\\
&\quad 0x55c31f6a\ 0x7cad2870\ 0x43f418b1\ 0x59917add
\end{aligned}$$

– *Chosen Plaintext Pair:*

$$\begin{aligned}
P\ &= 0x6a951691\ 0x44c50ce4\ 0x4f533b21\ 0x66c053b8\\
P^*\ &= 0x6a951691\ 0x44c50ce4\ 0xcf533b21\ 0x66c053b8
\end{aligned}$$

– *Corresponding Ciphertext Pair* $(C = MD5_K(P), C^* = MD5_{K^*}(P^*))$*:*
$C \ = 0xef54db89\ 0xdc642d4e\ 0x5b10bd8f\ 0xf8ab0cd7$
$C^* = 0xf989429c\ 0x8583799e\ 0xe3e1603f\ 0x81f0c43c$

– *Adaptively Chosen Ciphertext Pair* $(C' = C \oplus (0, e_9, e_9, e_9), C'^* = C^* \oplus (0, e_9, e_9, e_9))$*:*
$C' \ = 0xef54db89\ 0xdc642f4e\ 0x5b10bf8f\ 0xf8ab0ed7$
$C'^* = 0xf989429c\ 0x85837b9e\ 0xe3e1623f\ 0x81f0c63c$

– *Corresponding Plaintext Pair* $(P = MD5^{-1}_{K'}(C'), P'^* = MD5^{-1}_{K'^*}(C'^*))$*:*
$P' \ = 0x393c8bdc\ 0x2c6a7690\ 0x37d728f1\ 0xd778127f$
$P'^* = 0x393c8bdc\ 0x2c6a7690\ 0xb7d728f1\ 0xd778127f$

B Related-Key Distinguishers of MD4 and their Probabilities

Table 5. Related-Key Distinguishers of MD4 (Four Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔK^i	Prob.
0	0	e_{31}	0	0	0	1
1	0	0	e_{31}	0	0	2^{-1}
2	0	0	0	e_{31}	0	2^{-1}
3	e_{31}	0	0	0	$e_{31}(=\Delta K^3)$	1
4	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
27	0	0	0	0	0	1
28	0	0	0	0	$e_{31}(=\Delta K^3)$	1
	0	e_2	0	0		$p = 2^{-2}$
29	e_{31}	0	0	0	$e_{31}(=\Delta K^7)$	1
30	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
45	0	0	0	0	0	1
46	0	0	0	0	$e_{31}(=\Delta K^7)$	1
47	0	e_{10}	0	0	0	2^{-1}
	0	e_{25}	e_{10}	0	0	$q = 2^{-1}$
REC-4	$(0 \to 28)^2$, $(29 \to 46)^2$					Pr[REC-4] $\approx 2^{-132}$
BOO-4	$(0 \to 28)$, $(47 \to 29)^2$, $(28 \to 3)$					Pr[BOO-4] $\approx 2^{-4}$
BOO^W*-4*	Fixed $K^{0,1,2,15}$, $(3 \to 28)$, $(46 \to 29)^2$, $(28 \to 3)$					Pr[BOO-4] $= 1$

C Related-Key Distinguishers of MD5 and their Probabilities

Table 6. Related-Key Distinguishers of MD5 (Two Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔK^i	Prob.
0	e_{17}	0	e_2	$e_{7,12}$	0	2^{-2}
1	$e_{7,12}$	e_{24}	0	e_2	0	2^{-6}
2	e_2	e_{19}	e_{24}	0	0	2^{-4}
3	0	0	e_{19}	e_{24}	0	2^{-2}
4	e_{24}	0	0	e_{19}	0	2^{-2}
5	e_{19}	e_{31}	0	0	0	2^{-2}
6	0	0	e_{31}	0	0	2^{-1}
7	0	0	0	e_{31}	0	2^{-1}
8	e_{31}	0	0	0	$e_{31}(=\Delta K^8)$	1
9	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
26	0	0	0	0	0	1
27	0	0	0	0	$e_{31}(=\Delta K^8)$	2^{-1}
28	0	e_{19}	0	0	0	2^{-2}
29	0	e_{19}	e_{19}	0	0	2^{-3}
	0	$e_{19,28}$	e_{19}	e_{19}	0	$p = 2^{-26}$
30	$e_{17,31}$	e_{31}	e_{31}	0	0	2^{-1}
31	0	0	e_{31}	e_{31}	0	2^{-1}
32	e_{31}	0	0	e_{31}	0	1
33	e_{31}	0	0	0	$e_{31}(=\Delta K^8)$	1
34	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
55	0	0	0	0	0	1
56	0	0	0	0	$e_{31}(=\Delta K^8)$	2^{-1}
57	0	e_5	0	0	0	2^{-2}
58	0	e_5	e_5	0	0	2^{-2}
59	0	e_5	e_5	e_5	0	2^{-2}
60	e_5	e_5	e_5	e_5	0	2^{-3}
61	e_5	e_5	e_5	e_5	0	2^{-3}
62	e_5	e_5	e_5	e_5	0	2^{-3}
63	e_5	e_5	e_5	e_5	0	2^{-3}
	e_5	e_5	e_5	e_5		$q = 2^{-21}$
REC-2	$(0 \to 29)^2$, $(30 \to 62)^2$					$\Pr[\text{REC-2}] \approx 2^{-201.6}$
BOO-2	$(0 \to 29)$, $(63 \to 30)^2$, $(29 \to 1)$					$\Pr[\text{BOO-2}] \approx 2^{-78.6}$
BOO^W-2	Fixed $K^{0,1,2,3,4,5,6,7,8,9,11,13,15}$ $(8 \to 29)$, $(55 \to 30)^2$, $(29 \to 8)$					$\Pr[\text{BOO-2}] \approx 2^{-1.6}$

Table 7. Related-Key Distinguishers of MD5 (Four Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔK^i	Prob.
0	0	0	e_{31}	0	0	1
1	0	0	0	e_{31}	0	2^{-1}
2	e_{31}	0	0	0	$e_{31}(=\Delta K^2)$	1
3	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
28	0	0	0	0	0	1
29	0	0	0	0	$e_{31}(=\Delta K^2)$	2^{-1}
30	0	e_8	0	0	0	2^{-2}
	0	e_8	e_8	0		$p = 2^{-4}$
31	$e_{11,31}$	e_{31}	e_{31}	0	0	2^{-1}
32	0	0	e_{31}	e_{31}	0	1
33	e_{31}	0	0	e_{31}	0	1
34	e_{31}	0	0	0	$e_{31}(=\Delta K^{11})$	1
35	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
60	0	0	0	0	0	1
61	0	0	0	0	$e_{31}(=\Delta K^{11})$	2^{-1}
62	0	e_9	0	0	0	2^{-2}
63	0	e_9	e_9	0	0	2^{-2}
	0	e_9	e_9	e_9		$q = 2^{-6}$
REC-4	$(0 \rightarrow 30)^2$, $(31 \rightarrow 62)^2$					$\Pr[\text{REC-4}] \approx 2^{-137.1}$
BOO-4	$(0 \rightarrow 30)$, $(63 \rightarrow 31)^2$, $(30 \rightarrow 2)$					$\Pr[\text{BOO-4}] \approx 2^{-11.6}$
BOO^W*-4*	Fixed $K^{0,1,2,9,11}$, $(2 \rightarrow 30)$, $(60 \rightarrow 31)^2$, $(30 \rightarrow 2)$					$\Pr[\text{BOO-4}] \approx 2^{-0.6}$

D Related-Key Distinguishers of HAVAL and their Probabilities

Table 8. Related-Key Distinguishers of 4-pass HAVAL (Two Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔE^i	ΔF^i	ΔG^i	ΔH^i	ΔK^i	Prob.
0	0	e_{21}	0	0	0	0	0	0	0	1
1	e_{21}	0	0	0	0	0	0	0	0	2^{-1}
2	0	0	0	0	0	0	0	e_{10}	0	2^{-1}
3	0	0	0	0	0	0	e_{10}	0	0	2^{-1}
4	0	0	0	0	0	e_{10}	0	0	0	2^{-1}
5	0	0	0	0	e_{10}	0	0	0	0	2^{-1}
6	0	0	0	e_{10}	0	0	0	0	0	2^{-1}
7	0	0	e_{10}	0	0	0	0	0	0	2^{-1}
8	0	e_{10}	0	0	0	0	0	0	0	2^{-1}
9	e_{10}	0	0	0	0	0	0	0	$e_{31}(=\Delta K^9)$	1
10	0	0	0	0	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
50	0	0	0	0	0	0	0	0	0	1
51	0	0	0	0	0	0	0	0	$e_{31}(=\Delta K^9)$	1
52	0	0	0	0	0	0	0	e_{31}	0	2^{-1}
53	0	0	0	0	0	0	e_{31}	0	0	2^{-1}
54	0	0	0	0	0	e_{31}	0	0	0	2^{-1}
55	0	0	0	0	e_{31}	0	0	0	0	2^{-1}
56	0	0	0	e_{31}	0	0	0	0	0	2^{-1}
57	0	0	e_{31}	0	0	0	0	0	0	2^{-1}
58	0	e_{31}	0	0	0	0	0	0	0	2^{-1}
59	e_{31}	0	0	0	0	0	0	0	0	2^{-1}
60	0	0	0	0	0	0	0	e_{20}	0	2^{-1}
61	0	0	0	0	0	0	e_{20}	0	0	2^{-1}
62	0	0	0	0	0	e_{20}	0	0	0	2^{-1}
63	0	0	0	0	e_{20}	0	0	0	0	2^{-1}
	0	0	0	e_{20}	0	0	0	0		$p = 2^{-20}$
64	0	e_{10}	0	0	0	0	0	0	0	2^{-1}
65	e_{10}	0	0	0	0	0	0	0	$e_{31}(=\Delta K^9)$	1
66	0	0	0	0	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
118	0	0	0	0	0	0	0	0	0	1
119	0	0	0	0	0	0	0	0	$e_{31}(=\Delta K^9)$	1
120	0	0	0	0	0	0	0	e_{31}	0	2^{-1}
121	0	0	0	0	0	0	e_{31}	0	0	2^{-1}
122	0	0	0	0	0	e_{31}	0	0	0	2^{-1}
123	0	0	0	0	e_{31}	0	0	0	0	2^{-1}
124	0	0	0	e_{31}	0	0	0	0	0	2^{-1}
125	0	0	e_{31}	0	0	0	0	0	0	2^{-1}
126	0	e_{31}	0	0	0	0	0	0	0	2^{-1}
127	e_{31}	0	0	0	0	0	0	0	0	2^{-1}
	0	0	0	0	0	0	0	e_{20}	0	$q = 2^{-9}$
REC-2	$(0 \to 63)^2$, $(64 \to 124)^2$									Pr[REC-2] $\approx 2^{-292.9}$
BOO-2	$(0 \to 63)$, $(127 \to 64)^2$, $(63 \to 7)$									Pr[BOO-2] $\approx 2^{-35.9}$
BOO^W-2	Fixed $K^{0,1,2,3,4,5,6,7,8,15,16,22,23,25}$, $(9 \to 63)$, $(119 \to 64)^2$, $(63 \to 9)$									Pr[BOO-2] $\approx 2^{-10.3}$

Table 9. Related-Key Distinguishers of 5-pass HAVAL (Two Related Keys) – Extension of the distinguishers for 4-pass HAVAL

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔE^i	ΔF^i	ΔG^i	ΔH^i	ΔK^i	Prob.
128	0	0	0	0	0	0	0	e_{20}	0	2^{-1}
129	0	0	0	0	0	0	e_{20}	0	0	2^{-1}
130	0	0	0	0	0	e_{20}	0	0	0	2^{-1}
131	0	0	0	0	e_{20}	0	0	0	0	2^{-1}
132	0	0	0	e_{20}	0	0	0	0	0	2^{-1}
133	0	0	e_{20}	0	0	0	0	0	0	2^{-1}
134	0	e_{20}	0	0	0	0	0	0	0	2^{-1}
135	e_{20}	0	0	0	0	0	0	0	0	2^{-1}
136	0	0	0	0	0	0	0	e_9	0	2^{-1}
137	0	0	0	0	0	0	e_9	0	0	2^{-1}
138	0	0	0	0	0	e_9	0	0	0	2^{-1}
139	0	0	0	0	e_9	0	0	0	0	2^{-1}
140	0	0	0	e_9	0	0	0	0	0	2^{-1}
141	0	0	e_9	0	0	0	0	0	0	2^{-1}
142	0	e_9	0	0	0	0	0	0	0	2^{-1}
143	e_9	0	0	0	0	0	0	0	0	2^{-1}
144	0	0	0	0	0	0	0	e_{30}	0	2^{-1}
145	0	0	0	0	0	0	e_{30}	0	e_{31}=($= \Delta K^9$)	2^{-1}
146	0	0	0	0	0	e_{30}	0	e_{31}	0	2^{-2}
147	0	0	0	0	e_{30}	0	e_{31}	0	0	2^{-2}
148	0	0	0	e_{30}	0	e_{31}	0	0	0	2^{-2}
149	0	0	e_{30}	0	e_{31}	0	0	0	0	2^{-2}
150	0	e_{30}	0	e_{31}	0	0	0	0	0	2^{-2}
151	e_{30}	0	e_{31}	0	0	0	0	0	0	2^{-2}
152	0	e_{31}	0	0	0	0	0	e_{19}	0	2^{-2}
153	e_{31}	0	0	0	0	0	e_{19}	0	0	2^{-2}
154	0	0	0	0	0	e_{19}	0	e_{20}	0	2^{-2}
155	0	0	0	0	e_{19}	0	e_{20}	0	0	2^{-2}
156	0	0	0	e_{19}	0	e_{20}	0	0	0	2^{-2}
157	0	0	e_{19}	0	e_{20}	0	0	0	0	2^{-2}
158	0	e_{19}	0	e_{20}	0	0	0	0	0	2^{-2}
159	e_{19}	0	e_{20}	0	0	0	0	0	0	2^{-2}
	0	e_{20}	0	0	0	0	0	e_8		$q = 2^{-54}$
REC-2	$(0 \rightarrow 63)^2$, $(64 \rightarrow 155)^2$									Pr[REC-2] $\approx 2^{-373.1}$
BOO-2	$(0 \rightarrow 63)$, $(159 \rightarrow 64)^2$, $(63 \rightarrow 4)$									Pr[BOO-2] $\approx 2^{-125.9}$

Table 10. Related-Key Distinguishers of 4-pass HAVAL (Four Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔE^i	ΔF^i	ΔG^i	ΔH^i	ΔK^i	Prob.
0	0	0	e_{10}	0	0	0	0	0	0	1
1	0	e_{10}	0	0	0	0	0	0	0	2^{-1}
2	e_{10}	0	0	0	0	0	0	0	$e_{31}(=\Delta K^2)$	1
3	0	0	0	0	0	0	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
59	0	0	0	0	0	0	0	0	0	1
60	0	0	0	0	0	0	0	0	$e_{31}(=\Delta K^2)$	1
61	0	0	0	0	0	0	0	e_{31}	0	2^{-1}
62	0	0	0	0	0	0	e_{31}	0	0	2^{-1}
63	0	0	0	0	0	e_{31}	0	0	0	2^{-1}
	0	0	0	0	e_{31}	0	0	0		$p = 2^{-4}$
64	0	0	0	0	0	e_{10}	0	0	0	2^{-1}
65	0	0	0	0	e_{10}	0	0	0	0	2^{-1}
66	0	0	0	e_{10}	0	0	0	0	0	2^{-1}
67	0	0	e_{10}	0	0	0	0	0	0	2^{-1}
68	0	e_{10}	0	0	0	0	0	0	0	2^{-1}
69	e_{10}	0	0	0	0	0	0	0	$e_{31}(=\Delta K^{17})$	1
70	0	0	0	0	0	0	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
123	0	0	0	0	0	0	0	0	0	1
124	0	0	0	0	0	0	0	0	$e_{31}(=\Delta K^{17})$	1
125	0	0	0	0	0	0	0	e_{31}	0	2^{-1}
126	0	0	0	0	0	0	e_{31}	0	0	2^{-1}
127	0	0	0	0	0	e_{31}	0	0	0	2^{-1}
	0	0	0	0	e_{31}	0	0	0	0	$q = 2^{-8}$
REC-4	$(0 \to 63)^2$, $(64 \to 124)^2$									$\Pr[\text{REC-4}] \approx 2^{-261.0}$
BOO-4	$(0 \to 63)$, $(127 \to 64)^2$, $(63 \to 2)$									$\Pr[\text{BOO-4}] \approx 2^{-9.6}$
BOO^W*-4*	Fixed $K^{0,1,15,24}$, $(2 \to 63)$, $(124 \to 64)^2$, $(63 \to 2)$									$\Pr[\text{BOO-4}] \approx 2^{-3}$

Table 11. Related-Key Distinguishers of 5-pass HAVAL (Four Related Keys)

Round (i)	ΔA^i	ΔB^i	ΔC^i	ΔD^i	ΔE^i	ΔF^i	ΔG^i	ΔH^i	ΔK^i	Prob.
0	0	0	e_{10}	0	0	0	0	0	0	1
1	0	e_{10}	0	0	0	0	0	0	0	2^{-1}
2	e_{10}	0	0	0	0	0	0	0	$e_{31}(=\Delta K^2)$	1
3	0	0	0	0	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
59	0	0	0	0	0	0	0	0	0	1
60	0	0	0	0	0	0	0	0	$e_{31}(=\Delta K^2)$	1
61	0	0	0	0	0	0	0	e_{31}	0	2^{-1}
62	0	0	0	0	0	0	e_{31}	0	0	2^{-1}
63	0	0	0	0	0	e_{31}	0	0	0	2^{-1}
64	0	0	0	0	e_{31}	0	0	0		2^{-1}
65	0	0	0	e_{31}	0	0	0	0	0	2^{-1}
66	0	0	e_{31}	0	0	0	0	0	0	2^{-1}
67	0	e_{31}	0	0	0	0	0	0	0	2^{-1}
68	e_{31}	0	0	0	0	0	0	0	0	2^{-1}
69	0	0	0	0	0	0	0	e_{20}	0	2^{-1}
70	0	0	0	0	0	0	e_{20}	0	0	2^{-1}
71	0	0	0	0	0	e_{20}	0	0	0	2^{-1}
72	0	0	0	0	e_{20}	0	0	0	0	2^{-1}
73	0	0	0	e_{20}	0	0	0	0	0	2^{-1}
74	0	0	e_{20}	0	0	0	0	0	0	2^{-1}
75	0	e_{20}	0	0	0	0	0	0	0	2^{-1}
76	e_{20}	0	0	0	0	0	0	0	0	2^{-1}
77	0	0	0	0	0	0	0	e_9	0	2^{-1}
78	0	0	0	0	0	0	e_9	0	0	2^{-1}
79	0	0	0	0	0	e_9	0	0	0	2^{-1}
80	0	0	0	0	e_9	0	0	0	0	2^{-1}
81	0	0	0	e_9	0	0	0	0	0	2^{-1}
82	0	0	e_9	0	0	0	0	0	0	2^{-1}
83	0	e_9	0	0	0	0	0	0	0	2^{-1}
84	e_9	0	0	0	0	0	0	0	0	2^{-1}
85	0	0	0	0	0	0	0	e_{30}	0	2^{-1}
86	0	0	0	0	0	0	e_{30}	0	0	2^{-1}
87	0	0	0	0	0	e_{30}	0	0	0	2^{-1}
88	0	0	0	0	e_{30}	0	0	0		2^{-1}
89	0	0	0	e_{30}	0	0	0	0	0	2^{-1}
90	0	0	e_{30}	0	0	0	0	0	0	2^{-1}
91	0	e_{30}	0	0	0	0	0	0	0	2^{-1}
92	e_{30}	0	0	0	0	0	0	0	0	2^{-1}
93	0	0	0	0	0	0	0	e_{19}	0	2^{-1}
94	0	0	0	0	0	0	e_{19}	0	0	2^{-1}
	0	0	0	0	0	e_{19}	0	0	0	$p = 2^{-33}$
95	0	0	e_{10}	0	0	0	0	0	0	2^{-1}
96	0	e_{10}	0	0	0	0	0	0	0	2^{-1}
97	e_{10}	0	0	0	0	0	0	0	$e_{31}(=\Delta K^4)$	1
98	0	0	0	0	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
155	0	0	0	0	0	0	0	0	0	1
156	0	0	0	0	0	0	0	0	$e_{31}(=\Delta K^4)$	1
157	0	0	0	0	0	0	0	e_{31}	0	2^{-1}
158	0	0	0	0	0	0	e_{31}	0	0	2^{-1}
159	0	0	0	0	0	e_{31}	0	0	0	2^{-1}
	0	0	0	0	e_{31}	0	0	0	0	$q = 2^{-5}$
REC-4	$(0 \to 94)^2$, $(95 \to 156)^2$									Pr[REC-4] $\approx 2^{-312}$
BOO-4	$(0 \to 94)$, $(159 \to 95)^2$, $(94 \to 2)$									Pr[BOO-4] $\approx 2^{-61}$