Revised: Block Cipher Based Hash Function Construction From PGV

Duo Lei, Guozhu Feng , Li Chao , and RuiLin Li

Department of Science, National University of Defense Technology,Changsha, China

Duoduolei@gmail.com

Abstract. Preneel, Govaerts, and Vandewalle[12] considered the 64 most basic ways to construct a hash function from a block cipher, and regarded 12 of these 64 schemes as secure. Black, Pogaway and Shrimpton[3] proved that, in black-box model, the 12 schemes that PGV singled out as secure really are secure and given tight upper and lower bounds on their collision resistance. And also they pointed out, by stepping outside of the Merkle-Damgard[5] approach to analysis, an additional 8 of the 64 schemes are just as collision resistant as the first group of schemes. In this paper we point out that the 12 compression functions that PGV singled out are free start collision resistant and others are not, the additional 8 compression functions are only fix start collision resistant as singled out by BRS, the hash functions based on those 20 schemes are fix start collision resistant, the upper bound of collision resistance and preimage resistant are given based on conditional probability of $P_{Y|X=x}(y)$, $P_{Y|K=k}(y)$ of compression function, not based on assumption of random oracle model, the bounds have more practical value than the bounds given by BRS. In view point of collision resistant, the best 4 schemes are not among the 12 schemes singled by PGV, and among the 8 schemes point out by BRS, and block cipher E itself is the best compression to build a collision resistant hash function. At the end of the paper, two recommend structure of block cipher based hash function are given, and a prove of their securities are also given.

Key Words: Hash Function, Block Cipher, M-D Construction

1 Introduction

Most of hash functions iterate a compression function by Merkle-Damgaard structure with constant IV[10]. A well known approach for building hash function is the compression function out of a block cipher which have been discussed sine Rabin[13] given the first model of that kind of structure. As pointed out by BRS the block cipher approach has been less widely used for variety of reasons, and the emergence of the AES[6] has somewhat modified this landscape, especially recently the MD5 and SHA1 were attacked[1][2][16][17].

The topics of building hash function based on block cipher had been systematically analyzed in paper [12][15][9][3] and [7]. The PGV paper considered turning a block cipher $E:\{0,1\}^n\times\{0,1\}^n\to\{0,1\}^n$ into a hash function $H:\{0,1\}^{n^{**}}\times\{0,1\}^n\to\{0,1\}^n$ using a compression function $F:\{0,1\}^n\times\{0,1\}^n\to\{0,1\}^n$ derived from E. For v is a fixed n-bit constant, PGV considerd all 64 compression functions F of the form $F(m_i,h_{i-1})=E_a(b)\oplus c$,where $a,b,c\in\{m_i,h_{i-1},m_i\oplus h_{i-1},v\}$,defined the iterated hash of F as $H(m_t\|\cdots\|m_1,IV)=h_t$, $h_0=IV$, $h_i=F(m_i,h_{i-1})$, $i\in[1,\cdots,t],|m_i|=n$. Of the 64 such schemes, the authors of [12] regarded 12 as secure.

The authors of [3] taken a more proof-centric look at the schemes from PGV, proved additional 8 schemes were collision resistant, divided the 20 schemes into two group where the group-1 was the 12

schemes picked by PGV and the group-2 was the new founded 8 schemes. For the new founded schemes, the hash function H immune to collision attack within the Merkle-Damgard paradigm, the compression functions were not immune to collision attack. The proves of collision resistant of group-2 used the assumptions of E was a random oracle model and H with fix start model. They also provided both upper and lower bounds for each scheme.

This paper provide the complexity of finding a collision or preimage based on the assumption of known conditional probability of block cipher E. We analyze the 64 schemes with M-D structure, the 12 compression functions that PGV singled out are immune to free start collision resistant and additional 8 compression functions singled out by BRS are immune to fix start collision resistant. All the 20 schemes based hash function are fix start collision resistant, and fix start preimage resistant (where the preimage of 8 schemes singled out by BRS based hash function can be found by meet in middle attack, but the complexity is $2^{n/2}$. On that condition we still call the hash is preimage resistant, for the complexity is same as birthday attack on collision).

On considering of collision resistance, the best 4 schemes are not among the 12 schemes singled out by PGV and are among the 8 schemes singled out by BRS and the block cipher E is the best compression function, more precisely, we can tell the precise upper bound of preimage resistant and collision resistant of those 4 schemes, that of other 16 schemes are unpredictable and the block cipher E has much influence on the bounds. In this paper, we assume the padding is padding zero at end of the message, and the hash function is iterated by Merkle-Damgard paradigm.

2 Definition

2.1 Basic Definition

The notations of the probability it the paper is followed PhD paper of Christian Cachin[4].

Let $I_n \triangleq \{0,1\}^n$, the compression function $F: I_k \times I_n \to I_n$, $y = F(x_m, x_h)$, $x_h \in I_n$, $x_m \in I_k$, $y \in I_n$, in hash iteration, x_h is chaining value. Let M-D hash construction $H: I_{k\cdot *} \times I_n \to I_n$, $x \in I_n, m \in I_{k**}, z \in I_n$.

$$z = H(m,x) = H(m_* \| \cdots \| m_1, x) \triangleq F(m_*, F(\cdots, F(m_1, x))).$$

Let $G: I_l \times I_n \to I_n$, $z = G(m,x)$, $\Omega \triangleq \{(y,m,x) \mid m \in I_l, x \in I_n, z \in I_n\}$, $m \in I_l$, $x \in I_n$, $z \in I_n$, $\{(z,m,x)\}^G \triangleq \{(z,m,x) \mid (z,m,x) \in \Omega, z = G(m,x)\}$, $S_G \triangleq \max_{z_0, m_0} \#\{(z_0, m_0, x)\}^G$, $T_G \triangleq \max_{z_0, x_0} \#\{(z_0, m, x_0)\}^G$, $R_G \triangleq \max_{z_0} \#\{(z_0, m, x)\}^G$

A discrete random variable X is a mapping from the sample space Ω to an alphabet χ, X assigns a value $x \in X$ to each elementary event in the Ω and the probability distribution of X is the function:

$$P_X : \chi \to \Re : x \mapsto P[X = x] = \sum_{\omega \in \Omega : X(\omega) = x} P[\omega]$$

If the conditioning event involves another random variable Y defined on the same sample space, the conditional probability distribution of Y given that X takes on a value y is:

$$P_{Y|X=x}(y) = \frac{P_{XY}(x,y)}{P_X(x)}$$

whenever $P_X(x)$ is positive .

Theorem1: Let function $g: I_{k\cdot t} \times I_n \to I_n$ $y = g(m,x)$, the distributions of independent random variable M and X are $P_M(m)$ and $P_X(x)$, function $\chi_{g(m,x)}(y)$ is defined as:

$$\chi_{g(m,x)}(y) \triangleq \begin{cases} 1, & y = g(m,x) \\ 0, & y \neq g(m,x) \end{cases}$$

the random variable Y's distribution can be derived from X and M by:

$$P_Y(y) \triangleq \sum_{x \in I_n} \sum_{m \in I_{n \cdot t}} P_{MX}(m,x)\chi_{g(m,x)}(y) \triangleq \sum_{x \in I_n} \sum_{m \in I_{n \cdot t}} P_M(m)P_M(m)\chi_{g(m,x)}(y)$$

we call the probability of Y is derived probability of M and X.

Definition 1: (Random Oracles[3]). A fixed-size Random Oracle is a function $f : I_a \to I_b$ chosen uniformly at random from the set of all such functions.

Definition 2: The maximum advantage of collision attack and preimage attack are defined as follows:

1. Pseudo Preimage Attack:

$$\tilde{A}dv_H^{\Pr e}(A) \triangleq \max_{z_0} \Pr[z_0 \in I_n; \omega \leftarrow A^{F,H} : \omega \in \{(z_0, m, x)\}^H]$$

2. Fixed Start Preimage Attack:

$$\tilde{A}dv_H^{\mathrm{FixP}}(A) \triangleq \max_{z_0} \Pr[z_0, x_0 \in I_n; \omega \leftarrow A^{F,H} : \omega \in \{(z_0, m, x_0)\}^H]$$

3. Pseudo Collision Attack:

$$\tilde{A}dv_H^{Coll}(A) \triangleq \max_{z_0} \Pr[\omega, \omega' \leftarrow A^{F,H} : \quad \omega, \omega' \in \Lambda, \Lambda \subset \{\{(z_0, m, x)\}^F\}]$$

4. Fixed Start Collision Attack:

$$\tilde{A}dv_H^{FixC}(A) \triangleq \max_{z_0} \Pr[x_0 \in I_n, \omega, \omega' \leftarrow A^{F,H} : \omega, \omega' \in \Lambda, \Lambda \subset \{\{(z_0, m, x_0)\}^H\}].$$

Definition3: [Black Box Model] function $g : I_{k \cdot t} \times I_n \to I_n \quad y = g(m,x)$ is a Black Box Model, if the probabilities of success of Game0 and Game1 are same, where $q \leq 2^{n/2}$:

Game0 (A, F, y_0, q)

For i=1,…,t do:

$A_{y_0} \to (m_i, x_i)$

A wins if exists i st $z_0 = G(m_i, x_i)$

Game1 (A, F, y_0, q)

$Q \leftarrow \varnothing$

For i=1,…,t do:

$A(y_0, Q) \to (m_i, x_i)$

$Q \to Q \cup (G(m_i, x_i), m_i, x_i)$

A wins if exists i st. $z_0 = G(m_i, x_i)$

If no special statement is given, $z = G(m, x_0)$ and $z = G(m_0, x)$ are not invertible.

3 The Security of Hash Function

Theorem 2 ([8]).For $y = F(x_m, x_h)$, and $y = H(m, x)$, then:

1. if $y = F(x_m, x_h)$ is black box model then:

$$\tilde{A}dv_H^{Pre}(q) = 2q \max\{\frac{S_H}{2^n}, \frac{T_H}{2^k}\};$$

$$\tilde{A}dv_H^{FixP}(q) = 2q \max\{\frac{T_H}{2^k}\};$$

$$\tilde{A}dv_H^{Coll}(q) = 1;$$

$$\tilde{A}dv_H^{FixC}(q) \le \max\{2q\frac{S_H}{2^n}, q(q-1)\frac{T_H - 1}{2^k}\}.$$

2. if, for any $x_{h_0} \in I_n$, $y = F(x_m, x_{h_0})$ is black box model then:

$$\tilde{A}dv_H^{FixP}(q) = 2q \max\{\frac{T_H}{2^k}\};$$

$$\tilde{A}dv_H^{FixC}(q) \le \max\{2q\frac{S_H}{2^n}, q(q-1)\frac{T_H - 1}{2^k}\}.$$

4 Collision Resistance of PGV Schemes

We assume block cipher $E: I_n \times I_n \to I_n$, $\overline{y} = E_k(x)$ is black box model, the security of 24 PGV schemes is summarized in tables 1 where we do not consider the constant value v. The functions are numbered in BRS[3], $F_i: I_n \times I_n \to I_n$, H_i are M-D construction with:

$$H_i: \{0,1\}^{n \cdot *} \times \{0,1\}^n \to \{0,1\}^n,$$

$$z = H_i(m, x) \triangleq F_i(m_*, F_i(\cdots, F_i(m_1, x)))$$

Where $m = m_* \| \cdots \| m_1$, $x \in I_n$.

The $F_1 \sim F_{12}$ are the group one schemes, which are immune to free start collision resistance and $F_{13} \sim F_{20}$ are the group two schemes which are not immune to free start collision resistance, immune to fix start collision resistance. In fact 24 schemes are driven from 12 compression function with different fix start and four of which are not immune to fix start collision resistance.

4.1 Probability of Compression Function

Lemma 1. If $\overline{y} = E_k(x)$ is a random oracle model, x, k is uniformly distributed in I_n then

$$T_E = 1 \text{ and } S_E = 1.$$

Theorem 3. Block cipher $\overline{y} = E_k(x)$ is a random oracle model, x and k are uniformly distributed in I_n, then for $y = F_i(x_m, x_h)$,

$$S_{F_i} = 1, \text{ and } \quad T_{F_i} = 1, \ 1 \le i \le 24.$$

Proof. Since $\overline{y} = E_k(x)$ is a random oracle, then x, k, $E_k(x)$ are independent from each other.

We give the prove of the most famous model $F_1 : y = E_{x_h}(x_m) \oplus x_m$.

$$
\begin{aligned}
& P_{Y|X_h=x_h}(y) \\
&= \sum_{x_m} P_{X_m}(x_m)\, P_{Y|X_m=x_m, X_h=x_h}(y = E_{x_h}(x_m) \oplus x_m) \\
&= \sum_{x_m} P_{X_m}(x_m)\, P_{Y|X_m=x_m, X_h=x_h}(y = x_m \oplus u, u = E_{x_h}(x_m)) \\
&= \sum_{u} \sum_{x_m} \frac{1}{2^n} \chi_{(x_m,u)}(y = x_m \oplus u) \chi_{F_1(x_m,x_h)}(u) \\
&\le \sum_{u} \chi_{(x_{m_0},u)}(y = x_{m_0} \oplus u) P_{U|X_h=x_h}(u) = \frac{1}{2^n}
\end{aligned}
$$

And also

$$
\begin{aligned}
& P_{Y|X_m=x_m}(y) \\
&= \sum_{x_h} P_{X_h}(x_h)\, P_{Y|X_m=x_m, X_h=x_h}(y \oplus x_m = E_{x_h}(x_m)) \\
&= \sum_{x_h} P_{X_h}(x_h) P_{Y|X_m=x_m, X_h=x_h}(y' = E_{x_h}(x_m)) \\
&= \frac{1}{2^n}
\end{aligned}
$$

Other models can be proved in similar way.□

Theorem 4. $\forall k_0$, $\overline{y} = E(x, k_0)$ is a black box model, then $F_i, 1 \le i \le 12$ are black box model.

Proof. We give the prove of F_1, others can be proved in similar way.

$F_1 : y = E_{x_h}(x_m) \oplus x_m \Leftrightarrow E_{x_h}^{-1}(y \oplus x_m) = x_m$, for given y and x_m, if we can get x_h, that implies for given plaintext x_m and cipher text $y \oplus x_m$, we get the key x_h with $y \oplus x_m = E_{x_h}(x_m)$, since E is a block cipher, the best way to find x_h is exhaustive search; for given y and x_h, if we can get x_m by direct computation means for given y, x_h, we can get the cipher text $x_m \oplus y$ and plaintext x_m directly, but the computation we can do is E and E^{-1}, the condition is possible only when E is a linear function, for $y = E_{x_h}(x_m) \oplus x_m$, E is not a linear function, so the only way to find the x_m is exhaustive search, collision resistant can be proved in similar way.

Theorem 5. $\forall x_{h_0}, y = F_i(x_m, x_{h_0})$ is black box model.

Proof.

$-F_{13} : \forall y$, $x = E_k^{-1}(y)$, let $x_m \triangleq x$, $x_h \triangleq x \oplus k,$, then $E_{x_h \oplus x_m}(x_m) \oplus x_h \oplus x_m = y$, but $\forall y$ and x_{h_0}, the way to find x_m satisfy $E_{x_{h_0} \oplus x_m}(x_m) = y$ is exhaustive search, so $y = F_{13}(x_m, x_{h_0})$ is balck box model;

$-F_{14}$: $\forall y$, $x = E_k^{-1}(y \oplus k)$, let $x_h \triangleq x \oplus k$, $x_m \triangleq x$, then $E_{x_h \oplus x_m}(x_m) \oplus x_h \oplus x_m = y$, but $\forall y$ and x_{h_0}, we can't find x_m satisfy $E_{x_{h_0} \oplus x_m}(x_m) \oplus x_{h_0} \oplus x_m = y$ directly.

$-F_{15} : \forall y$, $x = E_k^{-1}(y \oplus k)$, let $x_h \triangleq x$, $x_m \triangleq k$, then $E_{x_m}(x_h) = y$, but $\forall y$ and x_{h_0}, we

can't find x_m satisfy $E_{x_m}(x_{h_0}) = y$;

$-F_{16}$: $\forall y, x = E_k^{-1}(y)$, let $x_m \triangleq k \oplus x$, $x_h \triangleq x$, then $E_{x_m \oplus x_h}(x_h) = y$, but $\forall y$ and x_{h_0}, we can't find x_m satisfy $E_{x_m \oplus x_{h_0}}(x_h) = y$;

$-F_{17}$: $\forall y$, $x = E_k^{-1}(y \oplus k)$, let $x_h \triangleq x$, $x_m \triangleq k \oplus x$, then $E_{x_m}(x_h) \oplus x_m = y$, but $\forall y$ and x_{h_0}, we can't find x_m satisfy $E_{x_m}(x_{h_0}) \oplus x_m = y$;

$-F_{18}$: $\forall y, x = E_k^{-1}(y \oplus k)$, let $x_h \triangleq x$, $x_m \triangleq k \oplus x$, then $E_{x_m \oplus x_h}(x_h) \oplus x_m \oplus x_h = y$, but $\forall y$ and x_{h_0}, we can't find x_m satisfy $E_{x_m \oplus x_{h_0}}(x_{h_0}) \oplus x_m \oplus x_{h_0} = y$;

$-F_{19}$: $\forall y$ computes $x = E_k^{-1}(y)$, let $x_h \triangleq x \oplus k$, $x_m \triangleq k$, then $E_{x_m}(x_m \oplus x_h) = y$, but $\forall y$ and x_{h_0}, we can't find x_m satisfy $E_{x_m}(x_m \oplus x_{h_0}) = y$;

$-F_{20}$: $\forall y$ compute $x = E_k^{-1}(y \oplus k)$, let $x_h \triangleq x$, $x_m \triangleq k$, then $E_{x_m}(x_m \oplus x_h) \oplus x_m = y$, but $\forall y$ and x_{h_0}, we can't find x_m satisfy $E_{x_m}(x_m \oplus x_{h_0}) \oplus x_m = y$; □

Theorem 6. F_i, $21 \le i \le 24$ are not fix start preimage resistant and fix start collision resistant.

Proof.

$-F_{21}$: $\forall y, x_{h_0}$, let $x_m \triangleq E_{x_{h_0}}^{-1}(y)$ then $E_{x_{h_0}}(x_m) = y$;

$-F_{22}$: $\forall y, x_{h_0}$, let $x_m \triangleq E_{x_{h_0}}^{-1}(y \oplus x_{h_0})$ then $y = E_{x_{h_0}}(x_m) \oplus x_{h_0}$;

$-F_{23}$: $\forall y, x_{h_0}$, let $x_m \triangleq E_{x_{h_0}}^{-1}(y) \oplus x_{h_0}$ then $y = E_{x_{h_0}}(x_m \oplus x_{h_0})$;

$-F_{23}$: $\forall y, x_{h_0}$, let $x_m \triangleq E_{x_{h_0}}^{-1}(y \oplus x_{h_0}) \oplus x_{h_0}$ then $E_{x_{h_0}}(x_m \oplus x_{h_0}) \oplus x_{h_0} = y$.

4.2 Collision Resistance of Hash Function

Theorem 7. If block cipher E is a random oracle model, if the compression function is $y = F_i(x_m, x_h)$, $1 \le i \le 20$, then the Hash function $H_i(m, IV)$, constant value $IV \in I_n$ is preimage resistance and collision resistant, the bound of collision resistance and preimage resistance are shown as bellow.

– For H_i, $1 \le i \le 12$,

$$\tilde{A}dv_{H_i}^{FixP}(q) \le \frac{2q}{2^n}, \qquad \tilde{A}dv_{H_i}^{FixC}(q) \le \frac{q(q-1)}{2^n}.$$

– For H_i, $13 \le i \le 20$,

$$\tilde{A}dv_{H_i}^{FixP}(q) \le \frac{q(q-1)}{2^n}, \qquad \tilde{A}dv_{H_i}^{FixC}(q) \le \frac{q(q-1)}{2^n}$$

– For H_i, $21 \le i \le 24$,

$$\tilde{A}dv_{H_i}^{FixP}(q) = 1, \qquad \tilde{A}dv_{H_i}^{FixC}(q) = 1.$$

Proof. From Lemma 1, Theorem 2, Theorem 3 Theorem 4 and Theorem5 and Theorem6, we can get the conclusions directly.

Lemma 2. Let E is compression function, then $\overline{y} = E_k(x \oplus k)$, $\overline{y} = E_k(x) \oplus k$, $\overline{y} = E_k(x \oplus k) \oplus k$ then

$$P_{\overline{Y}|K=k}(\overline{y}) = \frac{1}{2^n}.$$

Theorem 8. Let block cipher $E: I_n \times I_n \to I_n$ has no weakness then:

– For $i \in \{15,17,19,20\}$:

$$\tilde{A}dv_{H_i}^{FixP}(q) \le q(q-1)\frac{T_{F_i}}{2^n}, \qquad \tilde{A}dv_{H_i}^{FixC}(q) \le q(q-1)\frac{T_{F_i}}{2^n}$$

– For $i \in \{13,14,16,18\}$:

$$\tilde{A}dv_{H_i}^{FixP}(q) \le \max\{q(q-1)\frac{T_{F_i}}{2^n}, 2q\frac{S_{H_i}}{2^n}\}, \qquad \tilde{A}dv_{H_i}^{FixC}(q) \le \max\{q(q-1)\frac{T_{F_i}}{2^n}, 2q\frac{S_{H_i}}{2^n}\}.$$

– For $i \in \{1,2,3,4,5,6,7,8\}$:

$$\tilde{A}dv_{H_i}^{FixP}(q) \le \max\{2q\frac{T_{F_i}}{2^n}, 2q\frac{S_{H_i}}{2^n}\}, \qquad \tilde{A}dv_{H_i}^{FixC}(q) \le \max\{q(q-1)\frac{T_{F_i}}{2^n}, 2q\frac{S_{H_i}}{2^n}\}.$$

5 Design of Secure Hash Function

From above discussion, we can make a conclusion that, if we can design a iterated hash function with $S_H = 2^{-n}$ and $T_H = 2^{-n}$, then the hash function is a idea model.

Definition4: An iterated hash function structure is defined as $H^{B1}: I^{n\cdot *} \times I^n \to I^n$

$$H^{B1}(m,x) \triangleq E_{H(m_*\|\cdots\|m_1,x)}(h_* \oplus \cdots \oplus h_1 \oplus h_0),$$

Where $x \in \{0,1\}^n$, $m = m_* \| \cdots \| m_1$, $m \in \{0,1\}^{n\cdot *}$, $h_i = E_{m_i}(h_{i-1}), h_0 = x$, $z = H(m,x)$,

$$\overline{z} = H^{B1}(m,x).$$

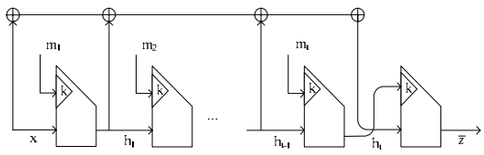

Fig.2 The design of H^{B1} Hash Function

The structure of H^{B1} is similar to the structure of 3c[19], but the attentions on the structure are different.

Definition5. An iterated hash function structure is defined as $H^{B2}: I^{n\cdot *} \times I^n \to I^n$

$$z, z' \in I_n, \quad z = E_{m_t}(\cdots E_{m_1}(x)), \quad z' = H(m,x), y = F(x_m, x_h) = E_{x_h}(x_m),$$
$$\overline{z} = H^{B2}(m,x) = E_z(z'), m = m_t \| \cdots \| m_1 \in \{0,1\}^{n\cdot t}.$$

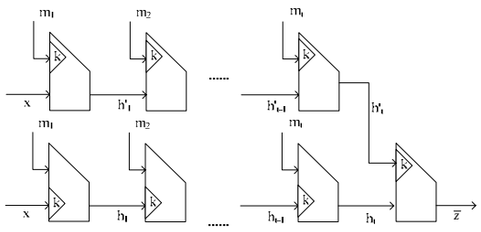

Fig. 3. The design of H^{B2} Hash Function

Theorem 9. For $\overline{z} = H^{B1}(m,x)$, $\overline{z} = H^{B2}(m,x)$, if x, k and y are independent from each other, and $f(m,x) = h_t \oplus \cdots \oplus h_1 \oplus h_0$ is independent from x and m, and $y = F(x_m, x_h)$ then:

$$S_{H^{B1}} = 1,\ T_{H^{B1}} = T_E\ , S_{H^{B2}} = 1,\quad \text{and}\ T_{H^{B2}} = T_E\,.$$

Theorem 10. For $\overline{z} = H^{B1}(m,x)$, $\overline{z} = H^{B2}(m,x)$, if x, k and y are independent from each other, and $f(m,x) = h_t \oplus \cdots \oplus h_1 \oplus h_0$ is independent from x and m, and $y = F(x_m, x_h)$ then:

$$\tilde{A}dv^{FixP}_{H^{B1}}(q) \le 2q\frac{T_E}{2^n},\qquad ,\ \tilde{A}dv^{FixC}_{H^{B1}}(q) \le q(q-1)\frac{T_E}{2^n}$$

$$\tilde{A}dv^{FixP}_{H^{B2}}(q) \le 2q\frac{T_E}{2^n},\qquad \tilde{A}dv^{FixC}_{H^{B2}}(q) \le q(q-1)\frac{T_E}{2^n}$$

6 Conclusion

The theorem 7 needs block cipher is random oracle, or else (x and $E_k(x)$) or ($E_k(x)$ and k) is not independent, respectively. If block cipher E is not assumed as a random oracle, then in $13 \le i \le 20$, the $F_{15}, F_{17}\ F_{19}\ F_{20}$ with properities of $P_{Y|K=k}(y) = \frac{1}{2^n}$,which means in group 2, the best compression function is $F_{15}, F_{17}\ F_{19}\ F_{20}$. And also the F_{15} is the block cipher E itself. In group 1, the possibility of $Q_i = 1$ is very low, then the preimage resistant and collision resistant of these schemes are unpredictable, different block cipher may result in a totally different results, for in design of block cipher does not consider these properties and it is difficult to consider.

Table 1. Summary of results. Columen1 is the number of hash function which are given by BRS[3]. Column2 is the compression functions to build hash function, and column 3 is the compression function F_i, column 4 is the count of y for constant x_h and column 5 is the count of y for constant x_m:

i	$\bar{y} =$	$y =$	$2^n P_{Y\|X_h}$	$2^n P_{Y\|X_m}$
15	$E_k(x)$	$E_{x_m}(x_h)$	P_{15}	1
21	$E_k(x)$	$E_{x_h}(x_m)$	1	Q_{21}
19	$E_k(x \oplus k)$	$E_{x_m}(x_h \oplus x_m)$	P_{19}	1
23	$E_k(x \oplus k)$	$E_{x_h}(x_h \oplus x_m)$	1	Q_{23}
5	$E_k(x) \oplus x$	$E_{x_m}(x_h) \oplus x_h$	P_5	Q_5
1	$E_k(x) \oplus x$	$E_{x_h}(x_m) \oplus x_m$	P_1	Q_1
17	$E_k(x) \oplus k$	$E_{x_m}(x_h) \oplus x_m$	P_{17}	1
22	$E_k(x) \oplus k$	$E_{x_h}(x_m) \oplus x_h$	1	Q_{22}
7	$E_k(x) \oplus x \oplus k$	$E_{x_m}(x_h) \oplus x_h \oplus x_m$	P_7	Q_7
3	$E_k(x) \oplus x \oplus k$	$E_{x_h}(x_m) \oplus x_h \oplus x_m$	P_3	Q_3
8	$E_k(x \oplus k) \oplus x$	$E_{x_m}(x_h \oplus x_m) \oplus x_h$	P_8	Q_8
4	$E_k(x \oplus k) \oplus x$	$E_{x_h}(x_h \oplus x_m) \oplus x_m$	P_4	Q_4
20	$E_k(x \oplus k) \oplus k$	$E_{x_m}(x_h \oplus x_m) \oplus x_m$	P_{20}	1
24	$E_k(x \oplus k) \oplus k$	$E_{x_h}(x_h \oplus x_m) \oplus x_h$	1	Q_{24}
6	$E_k(k \oplus x) \oplus x \oplus k$	$E_{x_m}(x_h \oplus x_m) \oplus x_h \oplus x_m$	P_6	Q_6
2	$E_k(k \oplus x) \oplus x \oplus k$	$E_{x_h}(x_h \oplus x_m) \oplus x_h \oplus x_m$	P_2	Q_2
16	$E_{k \oplus x}(x)$	$E_{x_h \oplus x_m}(x_h) \oplus v$	P_{16}	Q_{16}
13	$E_{k \oplus x}(x)$	$E_{x_h \oplus x_m}(x_m) \oplus v$	P_{13}	Q_{13}
10	$E_{k \oplus x}(x) \oplus x$	$E_{x_h \oplus x_m}(x_h) \oplus x_h$	P_{10}	Q_{10}
9	$E_{k \oplus x}(x) \oplus x$	$E_{x_h \oplus x_m}(x_m) \oplus x_m$	P_9	Q_9
12	$E_{k \oplus x}(x) \oplus k$	$E_{x_h \oplus x_m}(x_h) \oplus x_m$	P_{12}	Q_{12}
11	$E_{k \oplus x}(x) \oplus k$	$E_{x_h \oplus x_m}(x_m) \oplus x_h$	P_{11}	Q_{11}
18	$E_{k \oplus x}(x) \oplus x \oplus k$	$E_{x_h \oplus x_m}(x_h) \oplus x_h \oplus x_m$	P_{18}	Q_{18}
14	$E_{k \oplus x}(x) \oplus x \oplus k$	$E_{x_h \oplus x_m}(x_m) \oplus x_h \oplus x_m$	P_{14}	Q_{14}

References

1. E.Biham and R.Chen. Near-Collisions of SHA-0 and SHA-1. In Selected Areas in Cryptography-SAC 2004.

2. E.Biham and R.Chen. Near-Collisions of SHA-0,In Advances in Cryptology CRYPTO'2004, LNCS 3152, pp290-305, 2004.

3. J.Black, P.Rogaway, and T.Shrimpton, "Black-box analysis of the block-cipher-based hash function constructions from PGV". In Advances in Cryptology -CRYPTO'02, volume 2442 of Lecture Notes in Computer Science. Springer-Verlag,2002.pp.320-335.

4. C.Chchin. Entropy Measures and Uncoditional Security in Cryptography, PHD thesis.

5. I.Damg^{o}ard. A design principle for hash functions. In G. Brassard, editor, Advances in CRYPTO' 89, LNCS 435, 1990.

6. J.Daemen and V.Rijmen: The Design of Rijndael: AES The Advanced Encryption Standard. Springer, 2002.

7. X.Lai and J.L.Massey: Hash functions based on block ciphers. In Advances in Cryptology Eurocrypt'92, Lecture Notes in Computer Science, Vol. 658. Springer-Verlag, Berlin Hei-delberg New York (1993) 55-70.

8. D.Lei, The Security Proof of Iterated Hash Structure. http:\\eprint.iacr.org\2006\147.pdf

9. C. H. Meyer and S. M. Matyas. Cryptography: a New Dimension in Data Security.Wiley & Sons, 1982.

10. B.Preneel, V. Rijmen, A.Bosselaers: Recent Developments in the Design of Conventional Cryptographic Algorithms. In State of the Art and Evolution of Computer Security and Industrial Cryptography. Lecture Notes in Computer Science, Vol 1528. Springer-Verlag, Berlin Heidelberg New York(1998) 106-131.

11. B.Preneel: The State of Cryptographic Hash Functions. In Lectures on Data Security, Lecture Notes in Computer Science, Vol. 1561. Springer-Verlag, Berlin Hei-delberg New York (1999) 158-182.

12. B. Preneel, R. Govaerts, and J.Vandewalle, " Hash functions based on block ciphers,", In Advances in Cryptology -CRYPTO'93, Lecture Notes in Computer Science,pages 368-378. Springer-Verlag, 1994.

13. M. O. Rabin. Digitalized Signatures. In R. A. Demillo, D. P. Dopkin, A. K. Jones, and R. J. Lipton, editors, Foundations of Secure Computation, pages 155-166, New York, 1978. Academic Press.

14. C.E. Shannon. "Communication theory of secrecy systems,", Bell System Technical Journal,

28:656C715, 1949.

15. B. Van Rompay,Analysis and design of cryptographic hash functions, MAC algorithms and block cipher, K. U. Leuven, Juni 2004.

16. X.Wang, H.Yu, How to Break MD5 and Other Hash Functions, EUROCRYPT'2005, Springer-Verlag, LNCS 3494, pp19-35, 2005.

17. X. Wang, X. Lai, D.Feng and H.Yu., Cryptanalysis of the Hash Functions MD4

and RIPEMD, EUROCRYPT 2005, Springer-Verlag,LNCS 3494, pp1-18, 2005.

18. P.Gauravaram, W.Millan, J. Gonzalez Neito and E. Dawson: 3C-A Provably Secure Pseudorandom Function and Message Authentication Code. A New mode of operation for Cryptographic Hash Function. The preliminary draft version of this work is available at eprint-2005/390 .