Security Reductions of Cryptographic Hash Functions

Shoichi Hirose

University of Fukui

The first Asian Workshop on Symmetric Key Cryptography – ASK 2011
(2011/8/29-31, Nanyang Technological University)
Cryptographic Hash Function

\[ H : \{0, 1\}^* \rightarrow \{0, 1\}^n \]

Properties

<table>
<thead>
<tr>
<th>Preimage Resistance</th>
<th>Second PR</th>
<th>Collision Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complexity</th>
<th>PR</th>
<th>2ndPR</th>
<th>CR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(O(2^n))</td>
<td>(O(2^n))</td>
<td>(O(2^{n/2}))</td>
</tr>
</tbody>
</table>
Iterated Hash Function (Merkle-Damgård)

- Compression function
  \[ F : \{0, 1\}^n \times \{0, 1\}^b \rightarrow \{0, 1\}^n \]
- Initial value \( IV \in \{0, 1\}^n \)

Input \( M \in \{0, 1\}^* \)
\( F : \{0, 1\}^n \times \{0, 1\}^b \rightarrow \{0, 1\}^n \)  

Compression function

\( F \) is collision-resistant (CR) \( \Rightarrow \) \( H \) is CR

[Damgård 89]

1. If \( b \geq 2 \)

2. If \( b = 1 \), then prefix-free encoding is done for inputs.
Compression Function Construction

Customized (1990–)
- MD\textit{x} family
  - MD4, MD5; RIPEMD-160; SHA-1, SHA-224/256/384/512
- Whirlpool
- SHA-3 candidates

Using a block cipher
- Single block length (SBL): output-length = block-length
- Double block length (DBL): output-length = 2 \times \text{block-length}

SHA-1/2 DM mode using a dedicated block cipher SHACAL-1/2
Whirlpool MP mode using a dedicated block cipher W
Outline

- Hash function using block cipher
  - Single/Double-block-length constructions
- Multi-property preservation
- Security properties of hash-function family
- Cryptographic scheme using CR
A measure of efficiency of a hash function using a block cipher $E$

$$\text{rate} = \frac{b}{n \times k}$$

Shoichi Hirose (Univ Fukui)
Model for SBL construction

\[ E : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}^n \]

\[ x, k, z \in \{H_{i-1}, M_i, H_{i-1} \oplus M_i, \text{const}\} \]

- rate = 1
- \(4^3 = 64\) modes
Security of PGV Modes

[Preneel, Govaerts and Vandewalle 93]
- Security analysis against several generic attacks
- 12 modes are collision-resistant (CR).

[Black, Rogaway and Shrimpton 02]
- Provable security analysis in the ideal cipher model
- The same 12 modes are CR.
- Other 8 modes are CR with Merkle-Damgård domain extension.
12 PGV Modes

- **MMO**
  - $M_i \rightarrow E \rightarrow H_i \rightarrow H_i \rightarrow H_i$ with $H_{i-1}$

- **DM**
  - $M_i \rightarrow E \rightarrow H_i \rightarrow H_i \rightarrow H_i$ with $H_{i-1}$

- **MP**
  - $M_i \rightarrow E \rightarrow H_i \rightarrow H_i \rightarrow H_i$ with $H_{i-1}$

These diagrams illustrate different modes of operation for hash functions, showing how inputs ($M_i$) are processed through encryption ($E$) and finalization ($H_i$) to produce output.
8 PGV Modes

Ideal Cipher Model

Let $E$ be an $(n, \kappa)$ block cipher:

$$E : \{0, 1\}^\kappa \times \{0, 1\}^n \to \{0, 1\}^n.$$  

For each key $k$, $E(k, \cdot)$ is an **invertible random permutation**.

$E$ is evaluated by two kinds of **oracle queries**:

<table>
<thead>
<tr>
<th>oracle</th>
<th>query</th>
<th>answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$</td>
<td>(key, plaintext)</td>
<td>ciphertext</td>
</tr>
<tr>
<td>$E^{-1}$</td>
<td>(key, ciphertext)</td>
<td>plaintext</td>
</tr>
</tbody>
</table>

Provable security in the ideal cipher model

covers cryptanalysis not using internal structure of $E$
Idea of the Proof

The DM mode is CR in the ideal cipher model [Merkle 89]

\[
M_i \xrightarrow{k} \begin{array}{c}
H_{i-1} \\
x \oplus y
\end{array} \xrightarrow{E} H_i
\]

To compute \( H_i = x \oplus y \), we ask

- \((k, x)\) to \(E\), and obtain random \(y\), or
- \((k, y)\) to \(E^{-1}\), and obtain random \(x\)

In both cases, \(H_i\) is random.

Any collision attack is at most as effective as the birthday attack.
$E : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}^n$

$$C^{\text{AUX}}(K, X, Y) = C^{\text{POST}}(C^{-\text{PRE}}(K, X), Y)$$

The compression function is CR and PR if

- $C^{\text{PRE}}$ is bijective.
- For all $M, V$, $C^{\text{POST}}(M, V, \cdot) : Y \rightarrow W$ is bijective.
- For all $K, Y$, $C^{\text{AUX}}(K, \cdot, Y) : X \rightarrow W$ is bijective.
Why Discuss CR in the Ideal Cipher Model?

An almost ideal cipher may not produce a CR compression function.

\[
E_k(x) = \begin{cases} 
    x & \text{if } k = 00 \cdots 0 \text{ or } 11 \cdots 1 \\
    R_k(x) & \text{otherwise (} R_k \text{ is a random permutation)}
\end{cases}
\]

There is a trivial collision of DM compression function using \( E \):

\[
\begin{align*}
E_M &= \begin{array}{c} 00 \cdots 0 \\ 00 \cdots 0 \end{array} \\
H &\rightarrow E \\
00 \cdots 0 \\
\end{align*}
\]

\[
\begin{align*}
E_M &= \begin{array}{c} 11 \cdots 1 \\ 00 \cdots 0 \end{array} \\
H &\rightarrow E \\
00 \cdots 0 \\
\end{align*}
\]

Similar examples can be constructed for 12 CR modes in PGV model.

[Simon 98]

A CR HF cannot be constructed with a black-box OW permutation.
DBL Hash Function: Motivation

Any SBL hash function using AES is not secure.
- Output length is 128 bit.
- Complexity of birthday attack $\approx 2^{64}$.

Goal: DBL hash function using a block cipher with block-size $n$
- Complexity of collision attack $\approx 2^n$
[Brachtl, Coppersmith, et.al. 88]  
Using an \((n, n)\) block cipher

MDC-2  
rate = \(1/2\)

\[
\begin{align*}
g_{i-1} &\quad \rightarrow \quad E \quad \rightarrow \quad g_i \\
M_i &\quad \rightarrow \quad E \quad \rightarrow \quad g_i \\
h_{i-1} &\quad \rightarrow \quad E \quad \rightarrow \quad h_i
\end{align*}
\]

MDC-4  
rate = \(1/4\)

\[
\begin{align*}
g_{i-1} &\quad \rightarrow \quad E \quad \rightarrow \quad g_i \\
M_i &\quad \rightarrow \quad E \quad \rightarrow \quad g_i \\
h_{i-1} &\quad \rightarrow \quad E \quad \rightarrow \quad h_i
\end{align*}
\]
Using DES or an \((n, n)\) block cipher

Constants are fed into the key of \(E\).

rate < 0.276
[Lai, Massey 92]
Using an \((n, 2n)\) block cipher \((n\text{-bit plaintext, } 2n\text{-bit key})\)

abreast Davies-Meyer
rate = \(1/2\)

tandem Davies-Meyer
1/2
DBL Compression Functions: Hirose 06

\[ g_{i-1} \rightarrow E \rightarrow g_i \]
\[ h_{i-1} \leftrightarrow \]
\[ M_i \]
\[ c \rightarrow E \rightarrow h_i \]

- \( c \) is a non-zero constant
- \( \text{rate} = \begin{cases} 1/2 & \text{with 256-bit key} \\ 1/4 & \text{with 192-bit key} \end{cases} \)
- **only one key scheduling**

Note) Based on [Nandi 05]. \( p \) is involution \((p = p^{-1})\)
Output length: $2n$

<table>
<thead>
<tr>
<th>Attack</th>
<th>MDC-2</th>
<th>ab-DM</th>
<th>ta-DM</th>
<th>Hir</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collision</td>
<td>$\Omega(2^{0.6n})$ (1)</td>
<td>$\Theta(2^n)$ (2)</td>
<td>$\Omega(2^n/n)$ (3)</td>
<td>$\Theta(2^n)$ (4)</td>
</tr>
<tr>
<td>Preimage</td>
<td>$O(2^n)$ (5)</td>
<td>$\Theta(2^{2n})$ (6)</td>
<td>$\Theta(2^{2n})$ (6,7)</td>
<td>$\Theta(2^{2n})$ (6)</td>
</tr>
</tbody>
</table>

1. [Steinberger 06]
2. [Fleischmann, Gorski, Lucks 09], [Lee, Kwon 09]
3. [Lee, Stam, Steinberger 10]
4. [Hirose 06]
5. Requires $O(2^n)$ memory [Knudsen, Mendel, Rechberger, Thomsen 09]
6. [Lee, Stam, Steinberger 11]
7. $O(2^n)$ if digest $= 0^{2n}$.
\[
M_i \rightarrow E \rightarrow g_i \\
g_{i-1} \rightarrow E \rightarrow h_{i-1} \\
h_{i-1} \rightarrow \delta \rightarrow E \rightarrow h_i
\]

\[
\begin{array}{c|cccc}
  r & 00\|r' & 01\|r' & 10\|r' & 11\|r' \\
\delta(r) & 01\|r' & 10\|r' & 11\|r' & 00\|r'
\end{array}
\]

\[
\delta(r) = \delta((a)_{2}\|r') = (a + 1 \mod 4)_{2}\|r'
\]
Constructions As Efficient As MDC-2

[Satoh, Haga, Kurosawa 99], [Hattori, Hirose, Yoshida 03]

\[
\begin{align*}
g_{i-1} &\quad E & g_i \\
M_i &\quad E & g_i \\
h_{i-1} &\quad E & h_i
\end{align*}
\]

- rate = $\frac{\kappa}{2n}$ with an $(n, \kappa)$ block cipher
- As secure as MDC-2?
Multi-Property Preservation

Introduced by [Bellare, Ristenpart 06]

Security reduction to compression function

Security properties: CR, PRO (IRO), PRF

EMD (Enveloped Merkle-Damgård)

For PRF, $IV_1$ and $IV_2$ are replaced by independent secret keys.
Indifferentiability from RO (IRO)

[Maurer, Renner, Holenstein 04], [Coron, Dodis, Malinaud, Puniya 05]

- $H$ is VIL RO
- $F$ is FIL ideal primitive
  - Ideal block cipher
  - Random oracle

- $C$ is hash function construction using $F$
- Simulator $S$ tries to mimic $F$ with access to oracle $H$

**Definition**

$C^F$ is **indiff. from VIL RO (IRO)** if no efficient adversary $A$ can tell apart

$$(C^F, F') \text{ and } (H, S^H)$$
Multi-Property Preservation

For block-cipher-based construction

Security reduction to underlying block cipher

E.g.) DM, MMO, and MP are not IRO in the ideal cipher model.
E.g.) DM is not good for PRF since a message block is fed to the key.

Block ciphers are not designed for such usage!

\[ M_i \xrightarrow{H_{i-1}} \xrightarrow{E} H_i \]

DM

\[ M_i \xrightarrow{H_{i-1}} \xrightarrow{E} H_i \]

MMO

\[ M_i \xrightarrow{H_{i-1}} \xrightarrow{E} H_i \]

MP
Multi-Property Preservation

MMO seems best among PGV [Hirose, Kuwakado 08, 09]
Using MDP domain extension [Hirose, Park, Yun 07]

\[
\begin{array}{cccc}
M_1 & M_2 & \cdots & M_{N-1} & M_N \\
& F & \rightarrow & F & \rightarrow & \ldots & \rightarrow & F & \rightarrow & F
\end{array}
\]

If \( F \) is MMO, then

1. Hash function is CR and IRO in the ideal cipher model.
2. KIV mode is PRF if \( E \) is PRP under related-key attacks \( \text{wrt} \ \pi \).

Cf.) MMO is adopted by Skein (a SHA-3 finalist).
An interesting example:

\[ E_k(x) = E_{k \oplus d}(x \oplus d) \oplus d \]

This mode is not a PRF if \( E \) satisfies

for some const \( d \neq 0^n \) (DES has this property for \( d = 1^n \)).
Permutation-Based Schemes: Impossibility

[Black, Cochran, Shrimpton 05]

\[
\begin{array}{ccc}
M_1 & M_2 & M_i \\
F & F & F \\
IV & F & F \\
\end{array}
\]

\[
\begin{array}{c}
C_{in} \rightarrow E_K \rightarrow C_{out} \\
\end{array}
\]

\(K\) is fixed

Collision can be found with \(O(n + \log n)\) queries.
Collision can be found with $2^{(1-(m-r/2)/k)n}$ queries in the ideal permutation model.

<table>
<thead>
<tr>
<th>m</th>
<th>r</th>
<th>k</th>
<th># of queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>$2^{n/4}$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>$2^{n/2}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m</th>
<th>r</th>
<th>k</th>
<th># of queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>4</td>
<td>$2^{n/2}$</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>5</td>
<td>$2^{3n/5}$</td>
</tr>
</tbody>
</table>
[Gauravaram, Knudsen, Matusiewicz, Mendel, Rechberger, Schläffer, Thomsen 09]

IRO in the ideal permutation model

Security Reductions of Hash Functions

Number of queries = $\Theta(2^{\ell/2})$
Permutation-Based Schemes: Sponge

[Bertoni, Daemen, Peeters, van Assche 07]

\[ M_1 \quad M_2 \quad M_N \]

\[ r \]

\[ c \]

\[ \pi \]

\[ z_1 \]

\[ z_2 \]

absorbing squeezing

[Bertoni, Daemen, Peeters, van Assche 08]
IRO in the ideal permutation model

Number of queries = \( \Theta(2^{c/2}) \)
Permutation-Based Schemes: JH

[Wu 09]

\[ M_1 \xrightarrow{\pi} M_2 \xrightarrow{\pi} M_N \xrightarrow{\pi} IV \]

chop

[Bhattacharyya, Mandal, Nandi 10]
IRO in the ideal permutation model

\[
\text{Number of queries} = \Omega(2^{\ell/3}) \quad (\Omega(2^{\ell/2}) \ [\text{CRYPTO 11 rump}])
\]
Security Properties of Hash-Function Family

[Rogaway, Shrimpton 04]

Hash-function Family $H : \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{Y}$

<table>
<thead>
<tr>
<th>Property</th>
<th>Key</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>random</td>
<td>random</td>
</tr>
<tr>
<td>ePre</td>
<td>random</td>
<td>fixed</td>
</tr>
<tr>
<td>aPre</td>
<td>fixed</td>
<td>random</td>
</tr>
<tr>
<td>Sec</td>
<td>random</td>
<td>random</td>
</tr>
<tr>
<td>eSec</td>
<td>random</td>
<td>fixed</td>
</tr>
<tr>
<td>aSec</td>
<td>fixed</td>
<td>random</td>
</tr>
<tr>
<td>Coll</td>
<td>random</td>
<td>—</td>
</tr>
</tbody>
</table>

“a” means always.
“e” means everywhere.
Second Preimage Resistance

\[
\text{Adv}^\text{Sec}_H(A) = \Pr \left[ \begin{array}{c}
K \leftarrow \mathcal{K}; M \leftarrow \{0, 1\}^m : M \neq M' \land \\
M' \leftarrow A(K, M)
\end{array} \right] H_K(M) = H_K(M')
\]

\[
\text{Adv}^\text{eSec}_H(A) = \max_{M \in \{0, 1\}^m} \left\{ \Pr \left[ \begin{array}{c}
K \leftarrow \mathcal{K} \\
M' \leftarrow A(K, M)
\end{array} \right] H_K(M) = H_K(M') : M \neq M' \land \right. 
\]

\[
\text{Adv}^\text{aSec}_H(A) = \max_{K \in \mathcal{K}} \left\{ \Pr \left[ \begin{array}{c}
M \leftarrow \{0, 1\}^m \\
M' \leftarrow A(K, Y)
\end{array} \right] H_K(M) = H_K(M') : M \neq M' \land \right. 
\]

eSec is also called universal one-wayness (UOW) [Naor, Yung 89].
Another two-stage definition [Naor, Yung 89]

1. An adversary first selects input $M$.
2. $K$ is selected uniformly at random.

It is difficult to compute $M'$ such that $H_K(M) = H_K(M') \land M \neq M'$.

Signature scheme using UOW hash-function family [Naor, Yung 89]

A UOW hash-function family is constructed from

- any one-way permutation [Naor, Yung 89].
- any one-way function [Rompel 90], [Katz, Koo 05].
Merkle-Damgård does not work [Bellare, Rogaway 97].

Example

\[
h : \{0, 1\}^n \times \{0, 1\}^{m+n+c} \rightarrow \{0, 1\}^{n+c}
\]

\[
h_k(x, y, z) = \begin{cases} 
  k || f_k(x, y, z) & \text{if } y \neq k \\
  1^n || 1^c & \text{if } y = k
\end{cases}
\]

where \( f : \{0, 1\}^n \times \{0, 1\}^{m+n+c} \rightarrow \{0, 1\}^c \).

\( f \) is UOW \( \Rightarrow \) \( h \) is UOW
\[ h_k(x, y, z) = \begin{cases} 
    k\|f_k(x, y, z) & \text{if } y \neq k \\
    1^n\|1^c & \text{if } y = k 
\end{cases} \]

For any \( M \in \{0, 1\}^m \)
\[
\mu(i) = \text{largest integer } \mu \text{ such that } 2^\mu | i
\]

\(k\) and \(K_0, K_1, \ldots, K_{\lceil \log N \rceil}\) are selected uniformly at random.

**Theorem**

\(h\) is UOW \(\Rightarrow\) the family above is UOW
Shoup’s scheme is optimal among the following type [Mironov 01]

\[
M_1 \xrightarrow{h_k} K_{\gamma(1)} \quad M_2 \xrightarrow{h_k} K_{\gamma(2)} \quad M_i \xrightarrow{h_k} K_{\gamma(i)} \quad M_N \xrightarrow{h_k} K_{\gamma(N)}
\]

**Theorem**

For any \( \gamma \),

The family above is UOW \( \Rightarrow |\gamma(\{1, 2, \ldots, N\})| > \log N \)
OW permutation $p : \{0, 1\}^\ell \rightarrow \{0, 1\}^\ell$

$f : \mathcal{K} \times \{0, 1\}^\ell \rightarrow \{0, 1\}^{\ell - 1}$

$H : \mathcal{K} \times \{0, 1\}^\ell \rightarrow \{0, 1\}^{\ell - 1}$ such that $H_k = f_k \circ p$

**Theorem**

$f$ is a universal hash-function family $\Rightarrow H$ is UOW
Cascade of UOW Hash-Function Family

\[ h^{(i)} : \mathcal{K}_i \times \{0, 1\}^{\ell_i} \rightarrow \{0, 1\}^{\ell_{i+1}} \quad (1 \leq i \leq m) \]

\[ H : (\mathcal{K}_1 \times \cdots \times \mathcal{K}_m) \times \{0, 1\}^{\ell_1} \rightarrow \{0, 1\}^{\ell_{m+1}} \text{ such that} \]

\[ H_{(k_1, \ldots, k_m)} = h^{(m)}_{k_m} \circ h^{(m-1)}_{k_{m-1}} \circ \cdots \circ h^{(1)}_{k_1} \]

**Theorem**

\( h^{(1)}, \ldots, h^{(m)} \) are UOW \( \Rightarrow \) \( H \) is UOW
Cryptographic Schemes Using CR

CR hash function $H$

S selects input $x$ uniformly at random, and sends $y = H(x)$ to R.

- R has no knowledge on $x$ other than $x \in H^{-1}(y)$ even if R is computationally unbounded.
- Computationally bounded S does not have $x'(\neq x)$ s.t. $y = H(x')$.

Examples using the property above:

- Fail-stop signature [Damgård, Pedersen, Pfitzmann 93]
- Non-interactive string commitment statistically secure against computationally unbounded receiver [Halevi, Micali 96]
Non-interactive string commitment [Halevi, Micali 96]

\[ H : \{0, 1\}^* \rightarrow \{0, 1\}^\ell \text{ CR HF} \]
\[ F = \{ f \mid f : \{0, 1\}^{O(n+\ell)} \rightarrow \{0, 1\}^n \} \text{ Universal HF family} \]

**Commit**  For a committed string \( x \in \{0, 1\}^n \),

1. S selects uniformly at random \( f \in F \) and \( w \in \{0, 1\}^{O(n+\ell)} \) satisfying \( x = f(w) \).
2. computes \( y = H(w) \).
3. sends \( f, y \) to R.

**Open**  S sends \( w \) to R.
Non-interactive string commitment [Halevi, Micali 96]

Committed string $x$

Commit $f$ and $y$

Open $w$

Statistically secure against computationally unbounded $R$
Conclusion

- Hash function using block cipher
  - Single/Double-block-length constructions
- Multi-property preservation
- Security properties of hash-function family
- Cryptographic scheme using CR