How to Construct
Double-Block-Length Hash Functions

Shoichi Hirose
University of Fukui, Japan

24th Aug. 2006
Iterated Hash Function

• Compression function
  \[ F : \{0, 1\}^\ell \times \{0, 1\}^{\ell'} \rightarrow \{0, 1\}^\ell \]

• Initial value \( h_0 \in \{0, 1\}^\ell \)

Input \( m = (m_1, m_2, \ldots, m_l), m_i \in \{0, 1\}^{\ell'} \) for \( 1 \leq i \leq l \)

\[ H(m) = h_l \]
Motivation

How to construct a compression function using a smaller component?

E.g.) Double-block-length (DBL) hash function

- The component is a block cipher.
- output-length = 2 × block-length
- abreast/tandem Davies-Meyer, MDC-2, MDC-4, …

Cf.) Any single-block-length HF with AES is not secure.

- Output length is 128 bit.
- Complexity of birthday attack is $O(2^{64})$. 
Result

• Some plausible DBL HFs
  – Composed of a smaller compression function
    ∗ \( F(x) = (f(x), f(p(x))) \)
    \( p \) is a permutation satisfying some properties
    ∗ Optimally collision-resistant (CR) in the random oracle model
  – Composed of a block cipher with key-length > block-length
    ∗ AES with 192/256-bit key-length
    ∗ Optimally CR in the ideal cipher model

• A new security notion: Indistinguishability in the iteration

Def. (optimal collision resistance)
Any collision attack is at most as efficient as a birthday attack.
Related Work on Double-Block-Length Hash Function

- **Lucks 05**
  - $F(g, h, m) = (f(g, h, m), f(h, g, m))$
  - Optimally CR if $f$ is a random oracle

- **Nandi 05**
  - $F(x) = (f(x), f(p(x)))$, where $p$ is a permutation
  - Optimally CR schemes if $f$ is a random oracle
Other Related Work

Single block-length

- Preneel, Govaerts and Vandewalle 93
  PGV schemes and their informal security analysis
- Black, Rogaway and Shrimpton 02
  Provable security of PGV schemes in the ideal cipher model

Double block-length

- Satoh, Haga and Kurosawa 99
  Attacks against rate-1 HFs with a \((n, 2n)\) block cipher
- Hattori, Hirose and Yoshida 03
  No optimally CR rate-1 parallel-type CFs with a \((n, 2n)\) block cipher
DBL Hash Function Composed of a Smaller Compression Function

- $f$ is a random oracle
- $p$ is a permutation
  - Both $p$ and $p^{-1}$ are easy
  - $p \circ p$ is an identity permutation

\[
F(x) = (f(x), f(p(x)))
\]
\[
F(p(x)) = (f(p(x)), f(x))
\]

$f(x)$ and $f(p(x))$ is only used for $F(x)$ and $F(p(x))$.

We can assume that an adversary asks $x$ and $p(x)$ to $f$ simultaneously.
Collision Resistance

Th. 1 Let $F : \{0, 1\}^{2n+b} \rightarrow \{0, 1\}^{2n}$ and $F(x) = (f(x), f(p(x)))$.

Let $H$ be a hash function composed of $F$.

Suppose that

- $p(p(\cdot))$ is an identity permutation
- $p$ has no fixed points: $p(x) \neq x$ for $\forall x$

$$\text{Adv}^\text{coll}_H(q) \overset{\text{def}}{=} \text{success prob. of the optimal collision finder for } H$$

which asks $q$ pairs of queries to $f$.

Then, in the random oracle model, $\text{Adv}^\text{coll}_H(q) \leq \frac{q}{2^n} + \left(\frac{q}{2^n}\right)^2$.

Note) MD-strengthening is assumed in the analysis.
**Proof Sketch**

$F$ is CR $\Rightarrow$ $H$ is CR

Two kinds of collisions:

$$\Pr[F(x) = F(x') \mid x' = p(x)]$$

$$= \Pr[f(x) = f(x') \land f(p(x)) = f(p(x'))] = \left(\frac{1}{2^n}\right)^2$$

$$\Pr[F(x) = F(x') \mid x' = p(x)] = \Pr[f(x) = f(p(x))] = \frac{1}{2^n}$$

The collision finder asks $q$ pairs of queries to $f$: $x_j$ and $p(x_j)$ for $1 \leq j \leq q$.

$$\text{Adv}_{H}^{\text{coll}}(q) \leq \frac{q}{2^n} + \left(\frac{q}{2^n}\right)^2$$
Collision Resistance: A Better Bound

**Th. 2** Let $H$ be a hash function composed of $F : \{0, 1\}^{2n+b} \rightarrow \{0, 1\}^{2n}$. Suppose that

- $p(p(\cdot))$ is an identity permutation
- $p(g, h, m) = (p_{cv}(g, h), p_m(m))$
  - $p_{cv}$ has no fixed points
  - $p_{cv}(g, h) = (h, g)$ for $\forall (g, h)$

Then, in the random oracle model,

$$\text{Adv}_{\text{coll}}^H(q) \leq 3 \left( \frac{q}{2^n} \right)^2$$
Proof Sketch

Two kinds of collisions:

\[
\Pr[F(x) = F(x') | x' = p(x)] = \left( \frac{1}{2^n} \right)^2
\]

\[
\Pr[F(x) = F(x') | x' = p(x)] = \frac{1}{2^n}
\]

However,

\[
F(x) = F(x') \land x' = p(x) \Rightarrow F(w') = p_{cv}(F(w)) \land w' = p(w)
\]

\[
\Pr[F(w') = p_{cv}(F(w)) | w' = p(w)] = \left( \frac{1}{2^n} \right)^2
\]

\[
\text{Adv}^\text{coll}_H(q) \leq 3 \left( \frac{q}{2^n} \right)^2 = \left( \frac{q}{2^n} \right)^2 + 2 \left( \frac{q}{2^n} \right)^2
\]
**Th. 1 vs. Th. 2**

The difference between the upper bounds is significant.

E.g.) $n = 128, \ q = 2^{80}$

**Th. 1** \( \text{Adv}_H^{\text{coll}}(q) \leq \frac{q}{2^n} + \left( \frac{q}{2^n} \right)^2 \approx 2^{-48} \)

**Th. 2** \( \text{Adv}_H^{\text{coll}}(q) \leq 3 \left( \frac{q}{2^n} \right)^2 \approx 2^{-94} \)

E.g.) A permutation \( p \) satisfying the properties in Th. 2

\[
p(g, h, m) = (g \oplus c_1, h \oplus c_2, m), \text{ where } c_1 = c_2
\]
DBL Hash Function Composed of a Block Cipher

\[ F = g_{i-1} \rightarrow e \rightarrow g_i \]
\[ h_{i-1} \rightarrow e \rightarrow h_i \]
\[ m_i \rightarrow e \rightarrow h_i \]
\[ c \rightarrow e \rightarrow h_i \]

\[ c \text{ is a non-zero constant.} \]

Cf.)

such that

\[ f = \]
\[ p(g, h, m) = (g \oplus c, h, m) \]
DBL Hash Function Composed of a Block Cipher

\[ F = g_{i-1} \rightarrow e \rightarrow g_i \]

- can be constructed using AES with 192/256-bit key
- requires only one key scheduling

\[ F \text{ is simpler than abreast Davies-Meyer and tandem Davies-Meyer} \]
Collision Resistance

Th. 3 Let $H$ be a HF composed of $F : \{0, 1\}^{2n+b} \rightarrow \{0, 1\}^{2n}$ such that

$$F = \begin{array}{c}
\begin{array}{c}
g_{i-1} \\
h_{i-1} \\
m_i \\
c
\end{array}
\end{array} \quad \begin{array}{c}
\begin{array}{c}
e \\
\text{e}
\end{array}
\end{array} \quad \begin{array}{c}
\begin{array}{c}
g_i \\
h_i
\end{array}
\end{array}$$

$$\text{Adv}^{\text{coll}}_H(q) \overset{\text{def}}{=} \text{success prob. of the optimal collision finder for } H$$
which asks $q$ pairs of queries to $(e, e^{-1})$.

Then, in the ideal cipher model, for $1 \leq q \leq 2^{n-2}$,

$$\text{Adv}^{\text{coll}}_H(q) \leq 3 \left( \frac{q}{2^{n-1}} \right)^2$$
A Few More Examples of Compression Functions

for AES with 256-bit key

for AES with 192-bit key
Conclusion

• Some plausible DBL HFs
  – composed of
    a smaller compression function or a block cipher

\[ p \circ p \text{ is an identity permutation} \]
  – optimally collision-resistant

• A new security notion: Indistinguishability in the iteration