Security Reductions of Cryptographic Hash Functions

Shoichi Hirose

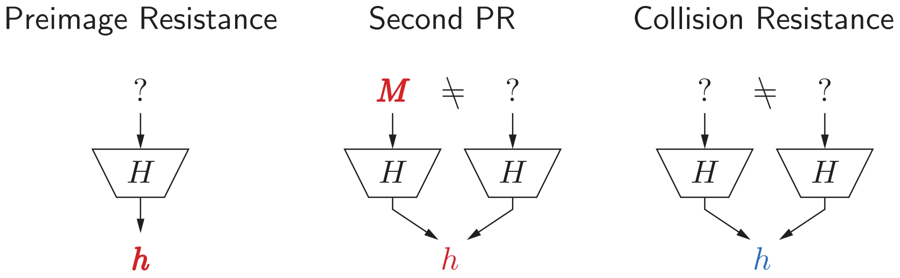
University of Fukui

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

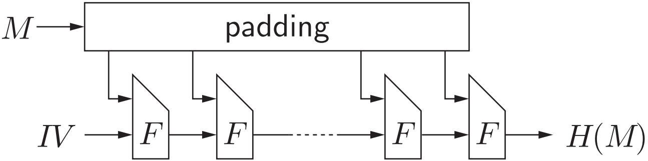


	PR	2ndPR	CR
Complexity	$O(2^n)$	$O(2^n)$	$O(2^{n/2})$

Iterated Hash Function (Merkle-Damgård)

- Compression function
  $F : \{0,1\}^n \times \{0,1\}^b \to \{0,1\}^n$
- Initial value $IV \in \{0,1\}^n$

Input $M \in \{0,1\}^*$



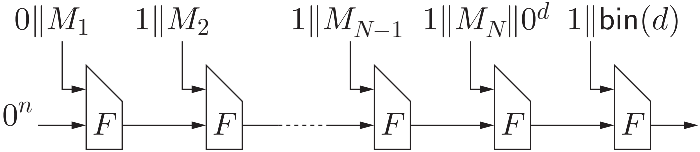
CR Preservation

$F : \{0,1\}^n \times \{0,1\}^b \rightarrow \{0,1\}^n$ 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

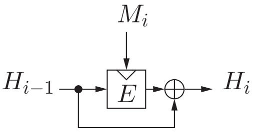
Customized (1990–)

- MDx family
  MD4, MD5; RIPEMD-160; SHA-1, SHA-224/256/384/512
- Whirlpool
- SHA-3 candidates

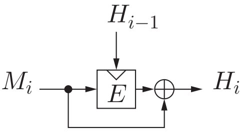
Using a block cipher

- Single block length (SBL): output-length = block-length
- Double block length (DBL): output-length $= 2 \times$ block-length

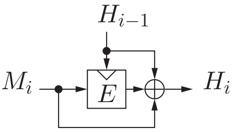
SHA-1/2	DM mode using a dedicated block cipher SHACAL-1/2
Whirlpool	MP mode using a dedicated block cipher W



Davies-Meyer



Matyas-Meyer-Oseas



Miyaguchi-Preneel

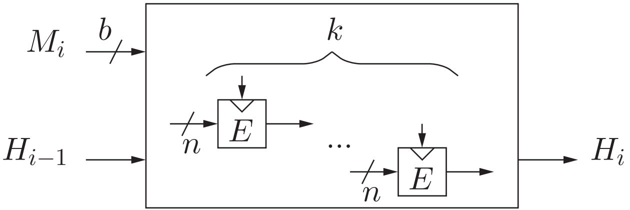
Outline

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

Rate

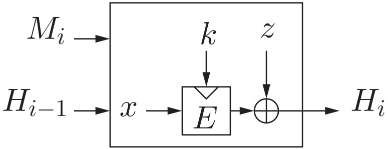
A measure of efficiency of a hash function using a block cipher E

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



PGV Model [Preneel, Govaerts, Vandewalle 93]

Model for SBL construction



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

$$x, k, z \in \{H_{i-1}, M_i, H_{i-1} \oplus M_i, const\}$$

- rate $= 1$
- $4^3 = 64$ modes

Security of PGV Modes

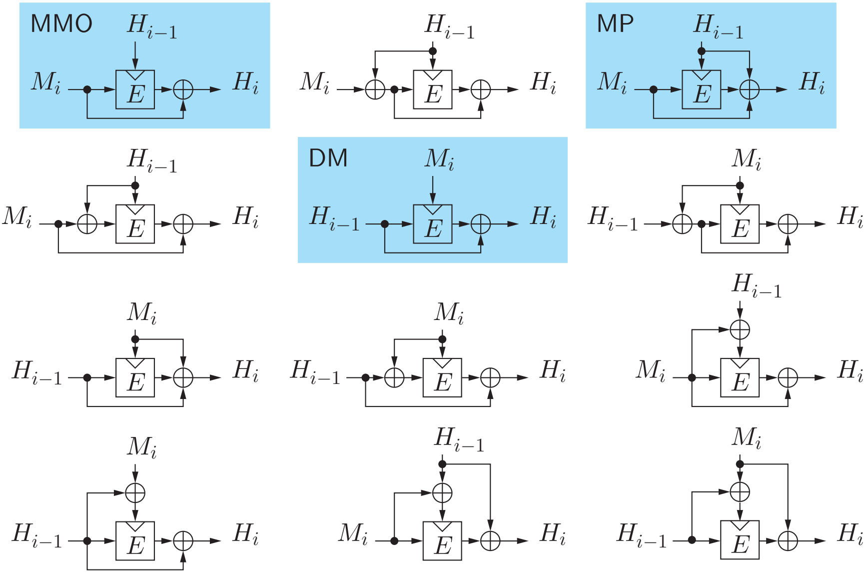
[Preneel, Govaerts and Vandewalle 93]

- Security analysis against several generic attacks
- 12 modes are collision-resistant (CR).

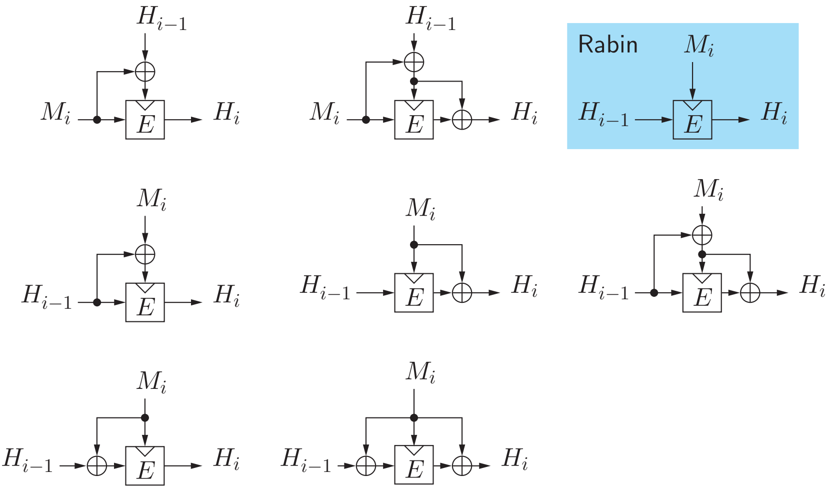
[Black, Rogaway and Shrimpton 02]

- 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



8 PGV Modes



Ideal Cipher Model

Let E be an (n, κ) block cipher:

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

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

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

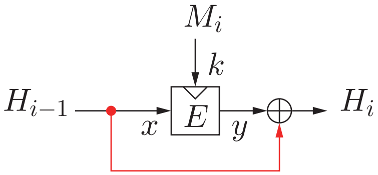
oracle	query	answer
E	(key, plaintext)	ciphertext
E^{-1}	(key, ciphertext)	plaintext

Provable security in the ideal cipher model

covers cryptanalysis not using intenal structure of E

Idea of the Proof

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



To compute $H_i = x \oplus y$, we ask

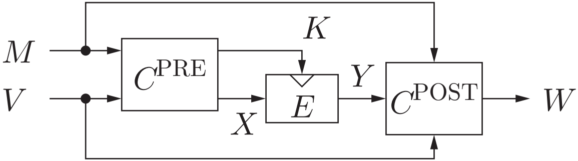
- (k, x) to E, and obtain random y, or
- (k, y) to E^{-1}, and obtain random x

In both cases, H_i is random.

Any collision attack is at most as effective as the birthday attack.

Stam Model (2009)

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



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

The compression function is CR and PR if

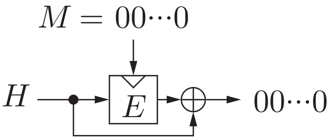
- C^{PRE} is bijective.
- For all M, V, $C^{\mathrm{POST}}(M, V, \cdot) : Y \mapsto W$ is bijective.
- For all K, Y, $C^{\mathrm{AUX}}(K, \cdot, Y) : X \mapsto W$ is bijective.

Why Discuss CR in the Ideal Cipher Model?

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

$$E_k(x) = \begin{cases} x & \text{if } k = 00\cdots0 \text{ or } 11\cdots1 \\ R_k(x) & \text{otherwise} \quad (R_k \text{ is a random permutation}) \end{cases}$$

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



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**.

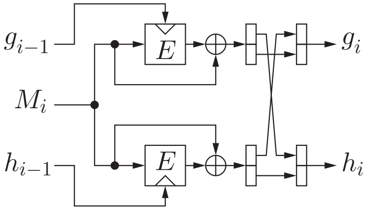
- Output length is 128 bit.
- Complexity of birthday attack $\approx 2^{64}$.

Goal: DBL hash function using a block cipher with block-size n

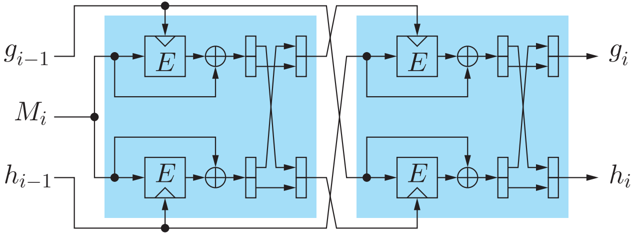
- Complexity of collision attack $\approx 2^n$

[Brachtl, Coppersmith, et.al. 88]
Using an (n, n) block cipher

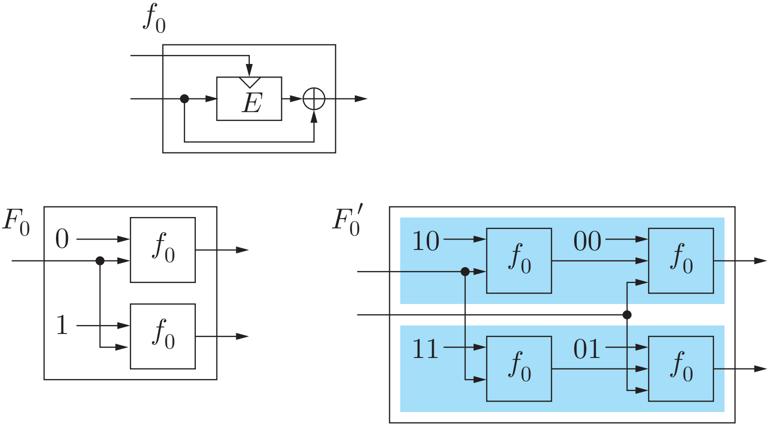
MDC-2
rate $= 1/2$



MDC-4
rate $= 1/4$



Using DES or an (n, n) block cipher

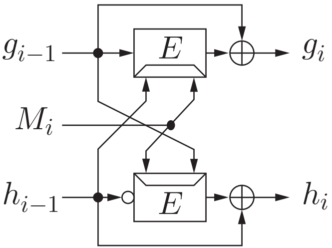


Constants are fed into the key of E.

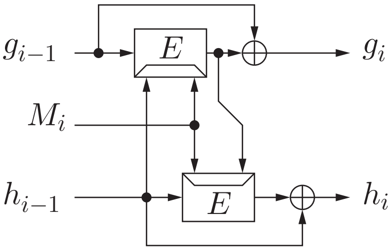
rate < 0.276

DBL Compression Functions: Abreast-/Tandem-DM

[Lai, Massey 92]
Using an $(n, 2n)$ block cipher (n-bit plaintext, $2n$-bit key)

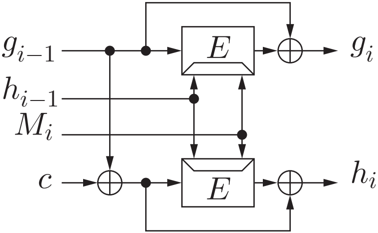


abreast Davies-Meyer
rate $= 1/2$



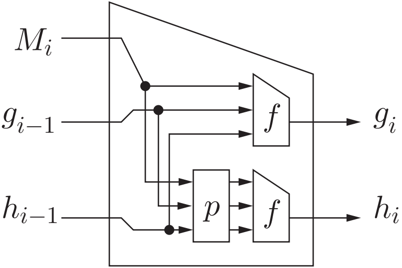
tandem Davies-Meyer
$1/2$

DBL Compression Functions: Hirose 06



- c is a non-zero constant
- 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 ($p = p^{-1}$)



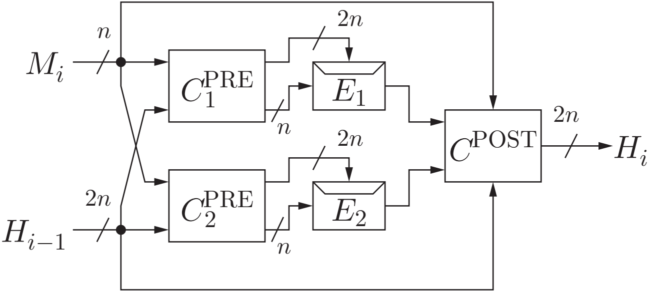
Security (Number of Oracle Queries)

Output length: $2n$

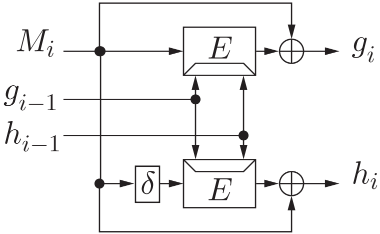
Attack	MDC-2	ab-DM	ta-DM	Hir
Collision	$\Omega(2^{0.6\,n})^{(1)}$	$\Theta(2^n)^{(2)}$	$\Omega(2^n/n)^{(3)}$	$\Theta(2^n)^{(4)}$
Preimage	$O(2^n)^{(5)}$	$\Theta(2^{2n})^{(6)}$	$\Theta(2^{2n})^{(6,7)}$	$\Theta(2^{2n})^{(6)}$

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 $= 0^{2n}$.

Özen-Stam Model (2010)



Jonsson & Robshaw 05

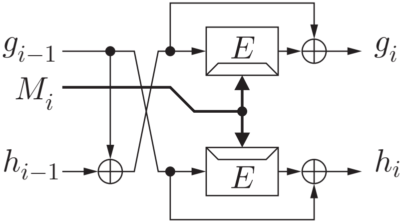


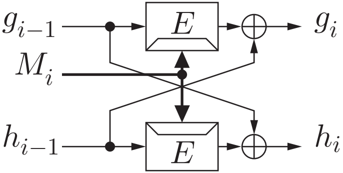
r	$00\|r'$	$01\|r'$	$10\|r'$	$11\|r'$
$\delta(r)$	$01\|r'$	$10\|r'$	$11\|r'$	$00\|r'$

$$\delta(r) = \delta((a)_2 \| r') = (a + 1 \bmod 4)_2 \| r'$$

Constructions As Efficient As MDC-2

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





- rate $= \dfrac{\kappa}{2n}$ with an (n, κ) block cipher
- As secure as MDC-2?

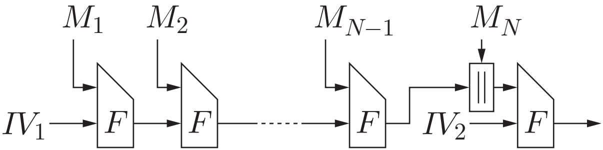
Multi-Property Preservation

Introduced by [Bellare, Ristenpart 06]

Security reduction to compression function

Security properties: CR, PRO (IRO), PRF

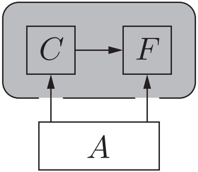
EMD (Enveloped Merkle-Damgård)

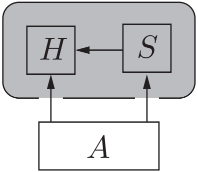


For PRF, IV_1 and IV_2 are replaced by independent secret keys.

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 adver A can tell apart

$$(C^F, F) \quad \text{and} \quad (H, S^H)$$

Multi-Property Preservation

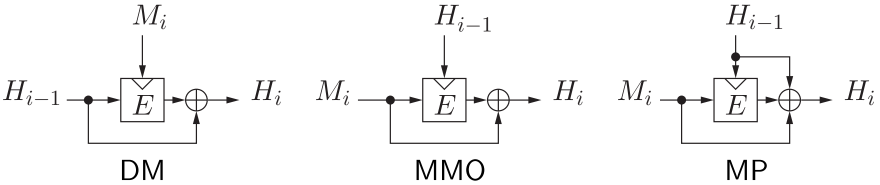
For block-cipher-based construction

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.

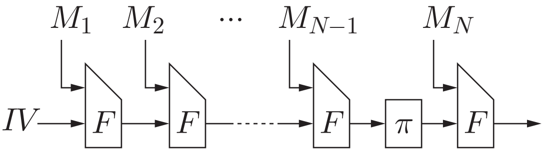
Block ciphers are not designed for such usage!



Multi-Property Preservation

MMO seems best among PGV [Hirose, Kuwakado 08, 09]

Using MDP domain extension [Hirose, Park, Yun 07]



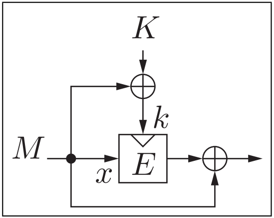
If F is MMO, then

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 wrt π.

Cf.) MMO is adopted by Skein (a SHA-3 finalist).

Multi-Property Preservation

An interesting example:



is one of the 12 secure PGV modes.

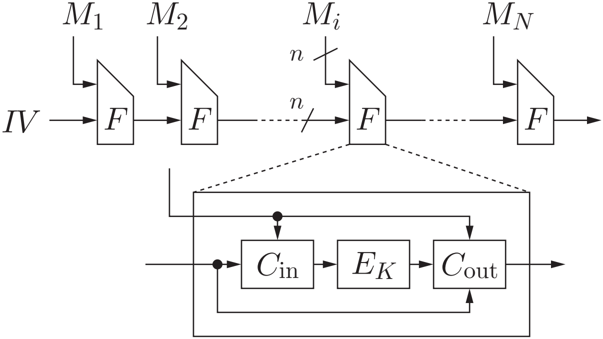
This mode is not a PRF if E satisfies

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

for some const $d \neq 0^n$ (DES has this property for $d = 1^n$).

Permutation-Based Schemes: Impossibility

[Black, Cochran, Shrimpton 05]

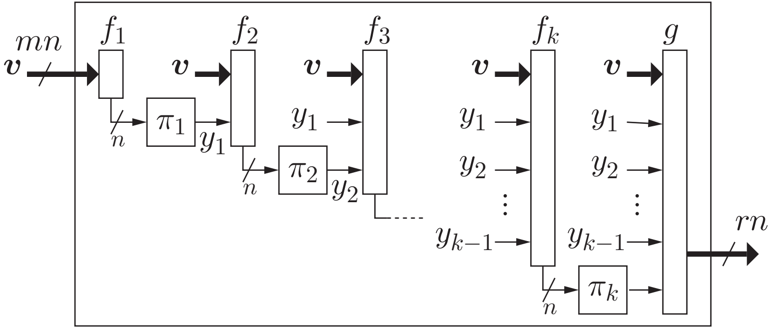


K is fixed

Collision can be found with $O(n + \log n)$ queries.

Permutation-Based Schemes: Security/Efficiency Tradeoff

[Rogaway, Steinberger 08]



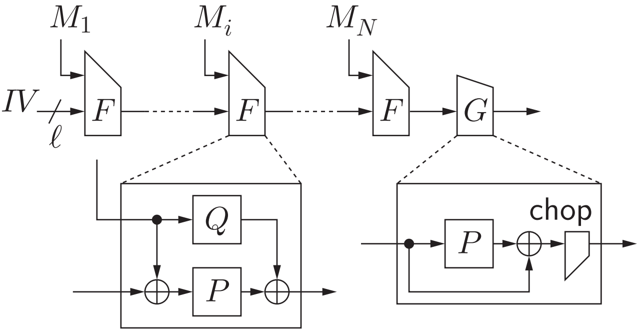
Collision can be found with $2^{(1-(m-r/2)/k)n}$ queries in the ideal permutation model.

m	r	k	# of queries
2	1	2	$2^{n/4}$
2	1	3	$2^{n/2}$

m	r	k	# of queries
3	2	4	$2^{n/2}$
3	2	5	$2^{3n/5}$

Permutation-Based Schemes: Grøstl

[Gauravaram, Knudsen, Matusiewicz, Mendel, Rechberger, Schläffer, Thomsen 09]

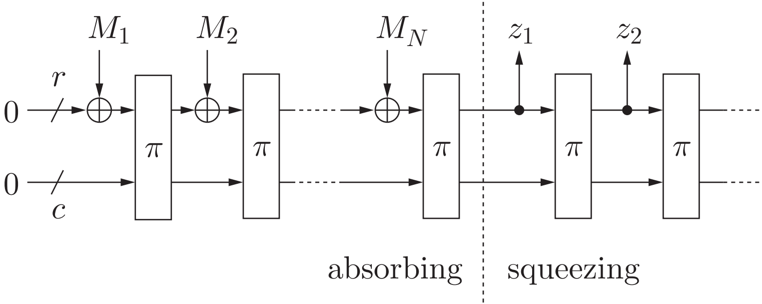


[Andreeva, Mennink, Preneel 10]
IRO in the ideal permutation model

Number of queries $= \Theta(2^{\ell/2})$

Permutation-Based Schemes: Sponge

[Bertoni, Daemen, Peeters, van Assche 07]

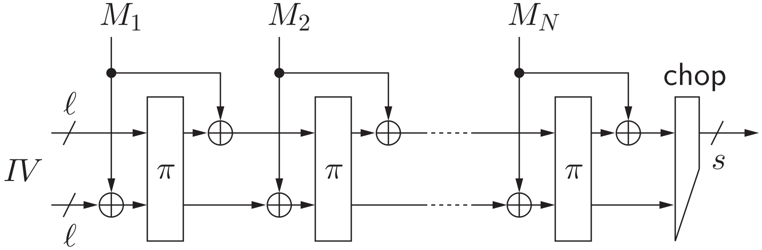


[Bertoni, Daemen, Peeters, van Assche 08]
IRO in the ideal permutation model

Number of queries $= \Theta(2^{c/2})$

Permutation-Based Schemes: JH

[Wu 09]



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

Number of queries $= \Omega(2^{\ell/3})$ $\quad$ ($\Omega(2^{\ell/2})$ [CRYPTO 11 rump])

Security Properties of Hash-Function Family

[Rogaway, Shrimpton 04]

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

Property	Key	Challenge
Pre	random	random
ePre	random	fixed
aPre	fixed	random
Sec	random	random
eSec	random	fixed
aSec	fixed	random
Coll	random	—

"a" means always.
"e" means everywhere.

Second Preimage Resistance

$$\mathrm{Adv}_H^{\mathrm{Sec}}(A) = \Pr\left[\begin{array}{l} K \xleftarrow{\$} \mathcal{K}; M \xleftarrow{\$} \{0,1\}^m \\ M' \xleftarrow{\$} A(K,M) \end{array} : \begin{array}{l} M \neq M' \wedge \\ H_K(M) = H_K(M') \end{array}\right]$$

$$\mathrm{Adv}_H^{\mathrm{eSec}}(A) = \max_{M \in \{0,1\}^m} \left\{ \Pr\left[\begin{array}{l} K \xleftarrow{\$} \mathcal{K} \\ M' \xleftarrow{\$} A(K,M) \end{array} : \begin{array}{l} M \neq M' \wedge \\ H_K(M) = H_K(M') \end{array}\right]\right\}$$

$$\mathrm{Adv}_H^{\mathrm{aSec}}(A) = \max_{K \in \mathcal{K}} \left\{ \Pr\left[\begin{array}{l} M \xleftarrow{\$} \{0,1\}^m \\ M' \xleftarrow{\$} A(K,Y) \end{array} : \begin{array}{l} M \neq M' \wedge \\ H_K(M) = H_K(M') \end{array}\right]\right\}$$

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

Universal One-Wayness (UOW)

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 $H_K(M) = H_K(M') \wedge M \neq M'$.

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].

Domain Extension for UOW Hash-Function Family

Merkle-Damgård does not work [Bellare, Rogaway 97].

Example

$h : \{0,1\}^n \times \{0,1\}^{m+n+c} \to \{0,1\}^{n+c}$

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

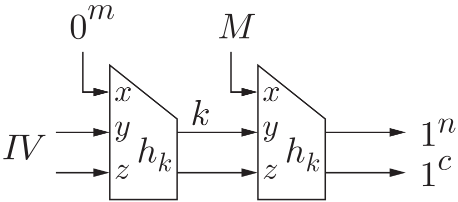
where $f : \{0,1\}^n \times \{0,1\}^{m+n+c} \to \{0,1\}^c$.

f is UOW $\Rightarrow$ h is UOW

Domain Extension for UOW Hash-Function Family

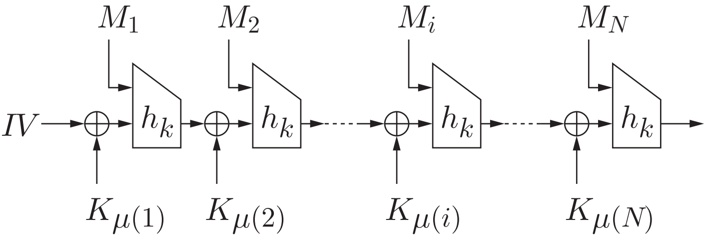
$$h_k(x, y, z) = \begin{cases} k \| f_k(x, y, z) & \text{if } y \neq k \\ 1^n \| 1^c & \text{if } y = k \end{cases}$$

For any $M \in \{0, 1\}^m$



Domain Extension for UOW Hash-Function Family

[Shoup 00]



$\mu(i) =$ largest integer μ such that $2^\mu | i$

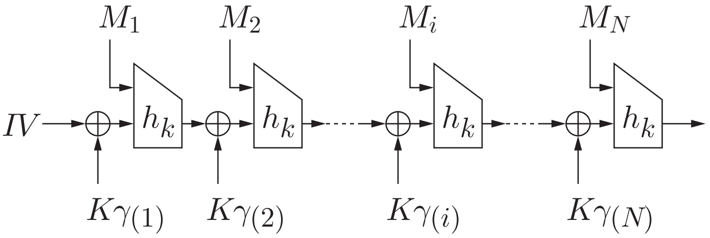
k and $K_0, K_1, \ldots, K_{\lfloor \log N \rfloor}$ are selected uniformly at random.

Theorem

h *is UOW* $\Rightarrow$ *the family above is UOW*

Domain Extension for UOW Hash-Function Family

Shoup's scheme is optimal among the following type [Mironov 01]



Theorem

For any γ,

The family above is UOW $\Rightarrow |\gamma(\{1, 2, \ldots, N\})| > \log N$

UOW Hash-Function Family from OW Permutation

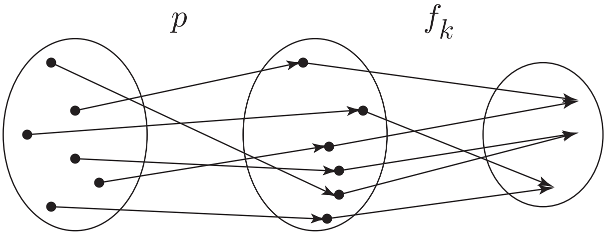
OW permutation $p : \{0,1\}^\ell \to \{0,1\}^\ell$

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

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

Theorem

f is a universal hash-function family $\Rightarrow$ H is UOW



Cascade of UOW Hash-Function Family

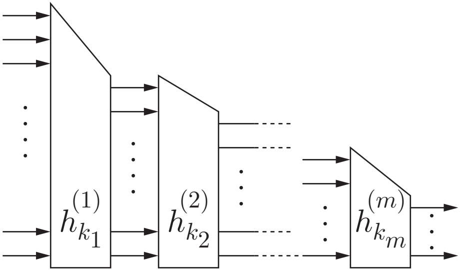
$h^{(i)} : \mathcal{K}_i \times \{0,1\}^{\ell_i} \to \{0,1\}^{\ell_{i+1}} \qquad (1 \le i \le m)$

$H : (\mathcal{K}_1 \times \cdots \times \mathcal{K}_m) \times \{0,1\}^{\ell_1} \to \{0,1\}^{\ell_{m+1}}$ 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

$h^{(1)}$, ..., $h^{(m)}$ *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]

$H : \{0,1\}^* \to \{0,1\}^\ell$ CR HF
$F = \{f \mid f : \{0,1\}^{O(n+\ell)} \to \{0,1\}^n\}$ 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.

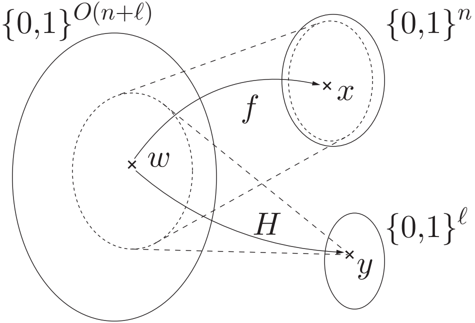
Open S sends w to R.

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