

Sequential Hashing with Minimum Padding

Shoichi Hirose

University of Fukui

NIST LWC 2016 (2016/10/17-18, Gaithersburg)

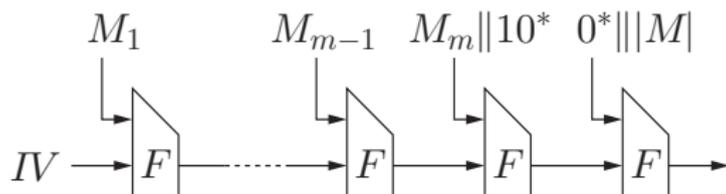
Background & Motivation

Hash function $H : \Sigma^* \rightarrow \Sigma^n$

Construction: FIL primitive + domain extension

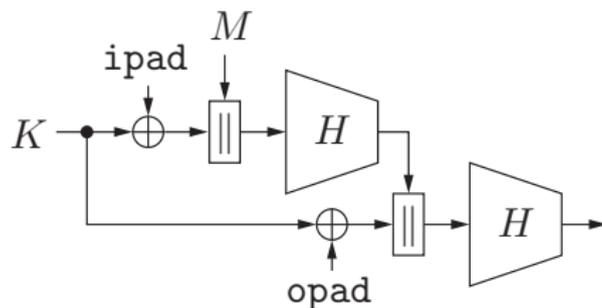
- Merkle-Damgård (Compression-function-based): SHA-2
- Sponge (Permutation-based): SHA-3

Strengthened MD



- Pros**
- Collision resistance is preserved.
- Cons**
- Length-extension property
 - The last message block may consist only of the padding sequence.

Cons degrade efficiency.



- Calls H twice to prevent length-extension attacks
- Not efficient for short messages

Overview of the Results

Domain extension scheme for sequential hashing

- with minimum padding (Padding sequence is as short as possible)
- free from length-extension

Security analysis of the domain extension scheme

- Collision resistance
- Indifferentiability from a random oracle (IRO)
- Pseudorandom function (PRF) of keyed-via-IV mode

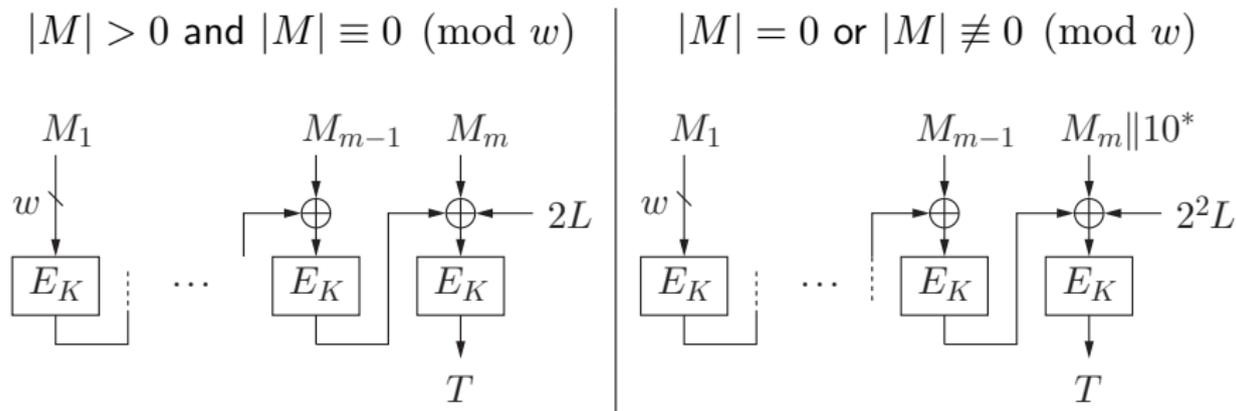
Application to sponge construction

- Indifferentiability from a random oracle

Minimum and Non Injective Padding

Minimum and non-injective padding is common for BC-based MAC

E.g.) CMAC

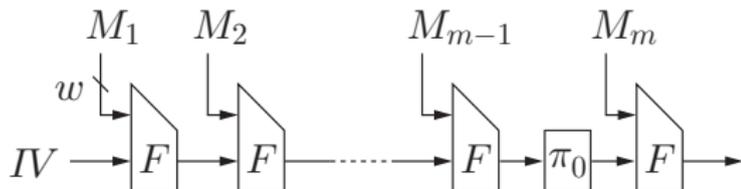


- $L = E_K(\mathbf{0})$
- $2L$ and 2^2L are used for
 - preventing the length-extension
 - separating the domain (Padding is not injective)

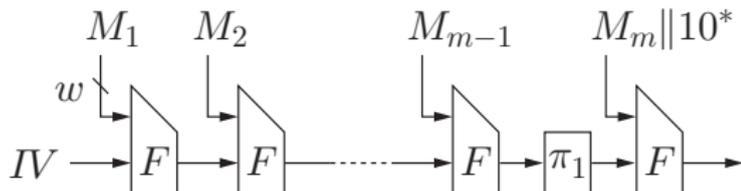
Proposed Domain Extension Scheme

For message $M = M_1 || M_2 || \dots || M_m$ such that

- 1 $|M| > 0$ and $|M| \equiv 0 \pmod{w}$



- 2 $|M| = 0$ or $|M| \not\equiv 0 \pmod{w}$



π_0 and π_1 are not cryptographic operations

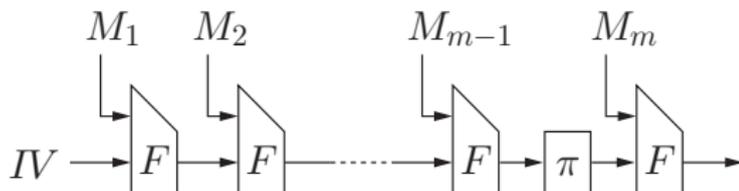
- Assumption: $\pi_0(v) \neq v \wedge \pi_1(v) \neq v \wedge \pi_0(v) \neq \pi_1(v)$ for any v
- E.g.) XOR with distinct non-zero constants

Collision-Resistance-preserving domain extension

- Merkle 1989
- Damgård 1989
- Nandi 2009: Variable-length encoding of the message length

Multi-property-preserving domain extension

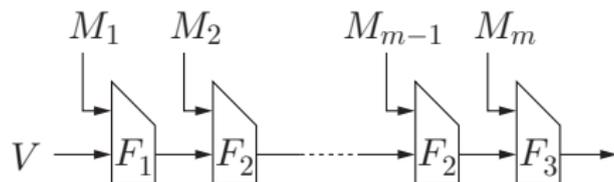
- EMD (Enveloped MD) [Bellare, Ristenpart 2006]
- MDP (MD with Permutation) [Hirose, Park, Yun 2007]



Cf.) Ferguson, Kelsey 2001 (Comment on Draft FIPS 180-2)

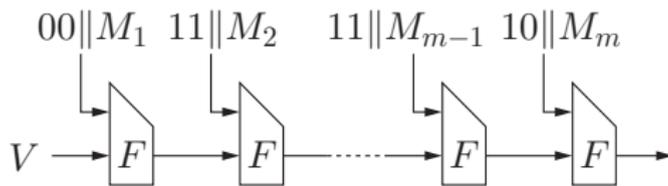
$$\pi(x) = x \oplus C$$

Suffix-Free Prefix-Free Hashing [BGKZ12]

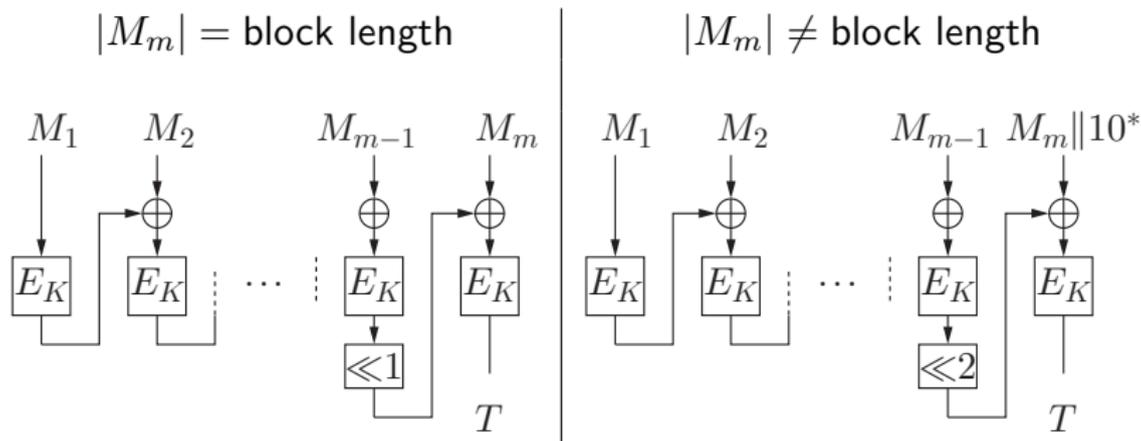


- IV is variable; without MD strengthening
- Needs three CFs
 - F_1 provides prefix-freeness; F_3 provides suffix-freeness
- Satisfies IRO
- Assumes injective padding

Cf.)



- Padding-length = $O(|M|)$



- XOR with constants does not work
- Requires at least two message blocks

Lemma

Any collision pair for $H_{IV}^{F, \{\pi_0, \pi_1\}}$ implies

- a collision pair,
- a $\{\pi_0, \pi_1\}$ -pseudo-collision pair, or
- a preimage of IV , $\pi_0^{-1}(\pi_1(IV))$, or $\pi_1^{-1}(\pi_0(IV))$

for F

Proof: Backward induction

$\{\pi_0, \pi_1\}$ -pseudo-collision pair for F :

$$(V, X) \text{ and } (V', X') \text{ s.t. } \pi_0(F(V, X)) = \pi_1(F(V', X'))$$

Theorem

The collision resistance of $H_{IV}^{F, \{\pi_0, \pi_1\}}$ is reduced to

- the collision resistance
- the $\{\pi_0, \pi_1\}$ -pseudo-collision resistance, and
- the everywhere preimage resistance

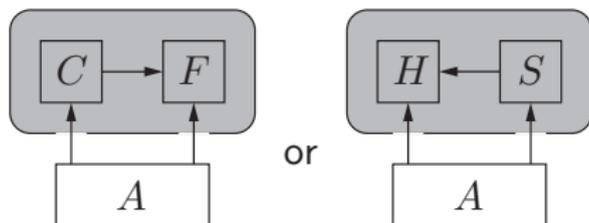
of F .

Everywhere preimage resistance of h :

$$\text{Adv}_h^{\text{epre}}(A) = \max_{Y \in \mathcal{Y}} \{\Pr[M \leftarrow A(h) : h(M) = Y]\}$$

Definition of Indifferentiability from a Random Oracle

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



- C is hashing mode of F
- F is FIL ideal primitive
 - Random oracle
 - Ideal block cipher
- H is VIL RO
- Simulator S tries to mimic F with access to oracle H

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

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

Indifferentiability from a Random Oracle (IRO)

Theorem

Suppose that CF $F : \Sigma^n \times \Sigma^w \rightarrow \Sigma^n$ is chosen uniformly at random. Then, for HF $H_{IV}^{F, \{\pi_0, \pi_1\}}$, there exists a simulator S of F s.t., for any adversary A making

- at most q queries to its FIL oracle
- queries to its VIL oracle which cost at most σ message blocks in total,

$$\text{Adv}_{H_{IV}^{F, \{\pi_0, \pi_1\}}, S}^{\text{indiff}}(A) \leq \frac{5(\sigma + q)^2}{2^n} + \frac{3\sigma q}{2^n - 6q + 1},$$

and S makes at most q queries.

Secure if $\sigma + q = o(2^{n/2})$

IRO in the Ideal Cipher Model

The CF $F : \Sigma^n \times \Sigma^w \rightarrow \Sigma^n$ is the Davies-Meyer mode of a BC E

- E is chosen uniformly at random

Theorem

For the hash function $H_{IV}^{F, \{\pi_0, \pi_1\}}$, there exists a simulator S of E s.t., for any adversary A making

- at most q_e queries to its FIL encryption oracle
- at most q_d queries to its FIL decryption oracle
- queries to its VIL oracle which cost at most σ message blocks in total,

$$\text{Adv}_{H_{IV}^{F, \{\pi_0, \pi_1\}}, S}^{\text{indiff}}(A) \leq \frac{12(\sigma + q_e + q_d)^2}{2^n} + \frac{3\sigma(q_e + q_d)}{2^n - 6(q_e + q_d) - 5},$$

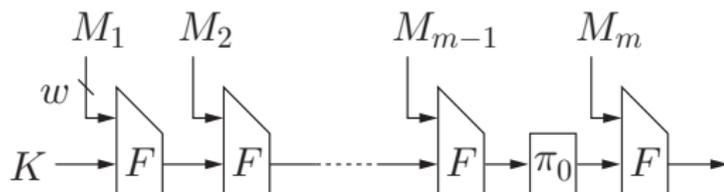
and S makes at most q_e queries.

Secure if $\sigma + q_e + q_d = o(2^{n/2})$

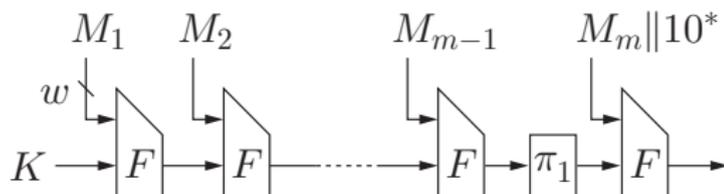
Keyed via IV mode of $H_{IV}^{F, \{\pi_0, \pi_1\}}$

For message M such that

- 1 $|M| > 0$ and $|M| \equiv 0 \pmod{w}$,



- 2 $|M| = 0$ or $|M| \not\equiv 0 \pmod{w}$,



Theorem

Let A be any adversary against KIV mode of $H_{IV}^{F, \{\pi_0, \pi_1\}}$:

- A runs in time at most t and makes at most q queries
- The length of each query is at most ℓw

Then, there exists an adversary B against F such that

$$\text{Adv}_{H_{IV}^{F, \{\pi_0, \pi_1\}}}^{\text{prf}}(A) \leq \ell q \text{Adv}_{\{id, \pi_1, \pi_2\}, F}^{\text{prf-rka}}(B) .$$

B runs in time at most $t + O(\ell q T_F)$ and makes at most q queries.

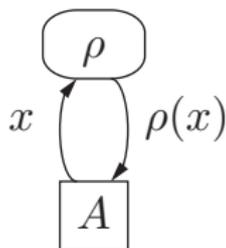
$H^{F, \{\pi_0, \pi_1\}}$ is PRF $\iff F$ is PRF against $\{id, \pi_1, \pi_2\}$ -restricted RKAs

Definition of PRF

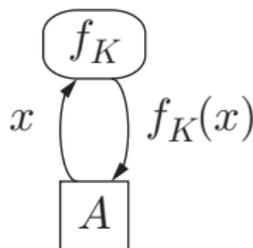
A keyed function $f : \mathcal{K} \times \mathcal{D} \rightarrow \mathcal{R}$ is PRF

if f is indistinguishable from uniform random function $\rho \Psi \mathcal{D} \rightarrow \mathcal{R}$

- Adversary makes queries to f_K or $\rho \Psi$
- Secret key $K \in \mathcal{K}$ is chosen uniformly at random



ideal world



real world

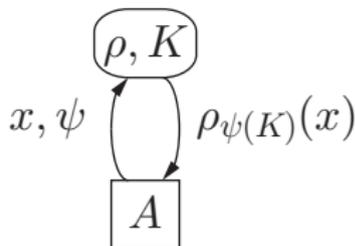
$$\text{Adv}_f^{\text{prf}}(A) = \left| \Pr[A^{f_K} = 1] - \Pr[A^{\rho \Psi} = 1] \right|$$

PRF against related key attacks

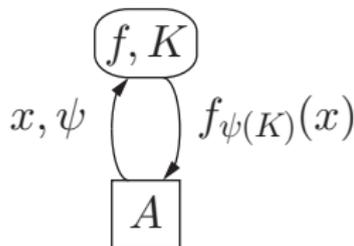
$f : \mathcal{K} \times \mathcal{D} \rightarrow \mathcal{R}$ is PRF against Ψ -restricted RKAs if

f is indistinguishable from uniform random keyed function $\rho \Psi \mathcal{K} \times \mathcal{D} \rightarrow \mathcal{R}$

- Ψ is a set of related-key deriving functions
- Secret key $K \in \mathcal{K}$ is chosen uniformly at random
- Adversary makes queries to $f_{\psi(K)}$ or $\rho \Psi_{\psi(K)}$ for any $\psi \in \Psi$



ideal world



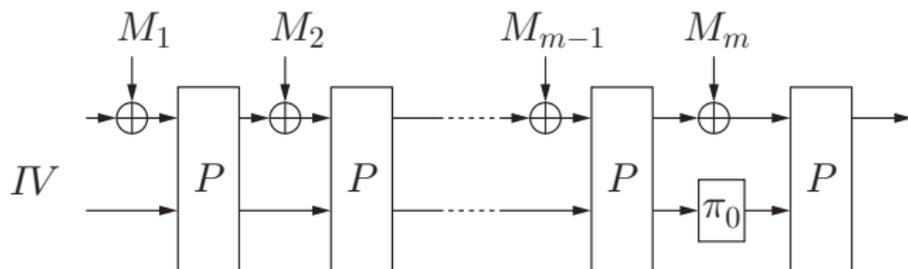
real world

$$\text{Adv}_{,f}^{\text{prf-rka}}(A) = \left| \Pr[A^{(f_{\psi(K)})} \in \mathcal{E}] - \Pr[A^{(\rho \Psi_{\psi(K)})} \in \mathcal{E}] \right|$$

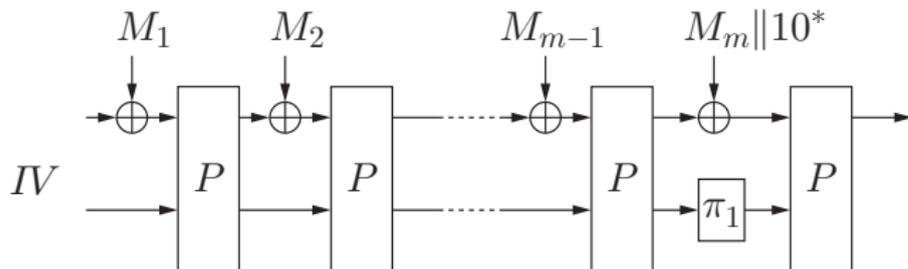
Application to Sponge Construction

For message M such that

- 1 $|M| > 0$ and $|M| \equiv 0 \pmod{w}$,



- 2 $|M| = 0$ or $|M| \not\equiv 0 \pmod{w}$,



IRO in the Ideal Permutation Model

The permutation $P : \Sigma^b \rightarrow \Sigma^b$ is chosen uniformly at random

- $b = r + c$ and c is capacity of sponge construction

Theorem

For the hash function $G_{IV}^{P, \{\pi_0, \pi_1\}}$, there exists a simulator S of P s.t., for any adversary A making

- at most q_f queries to its FIL forward oracle
- at most q_b queries to its FIL backward oracle
- queries to its VIL oracle which cost at most σ message blocks in total,

$$\text{Adv}_{G_{IV}^{P, \{\pi_0, \pi_1\}}, S}^{\text{indiff}}(A) \leq \frac{12(\sigma + q_f + q_b)^2}{2^c} + \frac{3\sigma(q_f + q_b)}{2^c - 6(q_f + q_b) - 5},$$

and S makes at most q_f queries.

Secure if $\sigma + q_f + q_b = o(2^{c/2})$

Domain extension scheme for sequential hashing

- with minimum padding
- free from length-extension

Security analysis of the domain extension scheme

- Collision resistance
- Indifferentiability from a random oracle
 - in the random oracle model
 - in the ideal cipher model with Davies-Meyer CF
- Pseudorandom function by keyed-via-IV

Application to sponge construction

- IRO in the ideal permutation model

Our proposal may be useful for lightweight hashing.