2008－CSEC－42 (32)
2008 / 7 / 25

# MMO圧縮関数とMDP定義域拡大によるハッシュ関数

廣瀬 勝一† 桑門 秀典††

† 福井大学工学研究科電気・電子工学専攻 〒910–8507 福井市文京 3-9-1
†† 神戸大学大学院工学研究科 〒657–8501 兵庫県神戸市灘区六甲台町 1-1

**あらまし** 本稿では，ブロック暗号を用いた MMO 圧縮関数と MDP 定義域拡大より構成されるハッシュ関数 MDP-MMO の証明可能安全性を論じる．まず，理想暗号モデルを仮定して，MDP-MMO がランダムオラクルとの識別不能性 (indifferentiability) を満たすことを示す．次に，MDP で使用される置換に関する関連鍵攻撃の下でブロック暗号が擬似ランダム置換であれば，MDP-MMO を用いて構成される HMAC が擬似ランダム関数であることを示す．なお，HMAC に関する結果は，$(E_{IV}(K \oplus \texttt{opad}) \oplus K \oplus \texttt{opad}) \| (E_{IV}(K \oplus \texttt{ipad}) \oplus K \oplus \texttt{ipad})$ が擬似ランダムビット列生成器であるという仮定も要求する．ここで，$E$ はブロック暗号，$IV$ は MDP-MMO で定められた初期値，opad と ipad は HMAC で定められた系列である．この仮定は，$E$ のブロック暗号としての擬似ランダム性により保証されないものの，実現可能性の観点からは妥当であると考えられる．

**キーワード** ハッシュ関数, Matyas-Meyer-Oseas, 識別不能性, 理想暗号モデル, HMAC, 関連鍵攻撃

# Hash Function Using MMO Compression Function and MDP Domain Extension

Shoichi HIROSE† and Hidenori KUWAKADO††

† Graduate School of Engineering, University of Fukui, Fukui 910–8507, Japan
†† Graduate School of Engineering, Kobe University, Kobe 657–8501, Japan

**Abstract** This article discusses the provable security of a hash function using a block cipher. It assumes the construction using the Matyas-Meyer-Oseas (MMO) scheme for the compression function and the Merkle-Damgård with a permutation (MDP) for the domain extension transform. It is shown that this kind of hash function, MDP-MMO, is indifferentiable from the variable-input-length random oracle in the ideal cipher model. It is also shown that HMAC using MDP-MMO is a pseudorandom function if the underlying block cipher is a pseudorandom permutation under the related-key attack with respect to a permutation used in MDP. Actually, the latter result also assumes that the following function is a pseudorandom bit generator: $(E_{IV}(K \oplus \texttt{opad}) \oplus K \oplus \texttt{opad}) \| (E_{IV}(K \oplus \texttt{ipad}) \oplus K \oplus \texttt{ipad})$, where $E$ is the underlying block cipher, $IV$ is the fixed initial value of MDP-MMO, and opad and ipad are the binary strings used in HMAC. This assumption still seems reasonable for actual block ciphers, though it cannot be implied by the pseudorandomness of $E$ as a block cipher. The results of this article imply that the security of a hash function may be reduced to the security of the underlying block cipher to more extent with the MMO compression function than with the Davies-Meyer (DM) compression function, though the DM scheme is implicitly used by the widely used hash functions such as SHA-1 and MD5.

**Key words** Hash function, Matyas-Meyer-Oseas, Indifferentiability, Ideal cipher model, HMAC, Related-key attack

## 1. Introduction

### 1.1 Background

A hash function is one of the most important primitives in cryptography. It normally consists of a function with fixed input length. This component function is called a compression function. A domain-extension transform is also specified which describes how to apply the compression function to a given input of variable length.

The methods to construct a compression function are classified in two classes: dedicated methods and those using a block cipher. Compression functions of well-known hash functions such as SHA-1/256 are constructed with the dedicated methods. However, they are also regarded as Davies-Meyer functions using dedicated block ciphers known as SHACAL-1/2.

### 1.2 Contribution

The topic of this article is to reduce the security of a hash function to the security of the underlying block cipher. It assumes the construction using the Matyas-Meyer-Oseas (MMO) scheme [12] for the compression function and the Merkle-Damgård with a permutation (MDP) [8] for the domain extension transform. This kind of hash function is called MDP-MMO in this article. A message padding scheme with the MD-strengthening is also assumed for MDP-MMO.

This article mainly discusses two security properties of MDP-MMO: indifferentiability from the variable-input-length (VIL) random oracle and pseudorandomness of HMAC [2], [10] using MDP-MMO. Collision-resistance is also mentioned briefly. These results implies that the security of an iterated hash function may be reduced to the security of the underlying block cipher to more extent with the MMO compression function than with the Davies-Meyer (DM) compression function.

It is shown that MDP-MMO is indifferentiable from the VIL random oracle in the ideal cipher model. This work is motivated by the recent work of Gong et al. [7]. They claimed that hash functions indifferentiable from the VIL random oracle in the ideal cipher model can be constructed using the MMO compression function and the domain extension transforms in [6]. The contribution of the current article is to reconstruct the proof using the game playing technique. Also, notice that they did not consider MDP for domain extension.

Indifferentiability of an iterated hash function is often discussed on the assumption that the underlying compression function is a random oracle with fixed input length. Taking the structure of compression functions of widely used hash functions into consideration, it is not satisfactory. For example, DM and MMO compression functions are not indifferentiable from the fixed-input-length (FIL) random oracle [6], [11].

It is also shown that HMAC using MDP-MMO is a pseudorandom function (PRF) if the underlying block cipher is a pseudorandom permutation (PRP) under the related-key attack with respect to a permutation used in MDP. Actually, this result also requires that the following function is a pseudorandom bit generator (PRBG):

$$(E_{IV}(K \oplus \mathtt{opad}) \oplus K \oplus \mathtt{opad}) \| (E_{IV}(K \oplus \mathtt{ipad}) \oplus K \oplus \mathtt{ipad}) \ ,$$

where $E$ is the underlying block cipher, $IV$ is the fixed initial value of MDP-MMO, and opad and ipad are the binary strings used in HMAC. It does not seem difficult to design a block cipher with which the function shown above is PRBG, though it cannot be implied by the pseudorandomness of $E$ as a block cipher. It is because any adversary has no control over $IV$, ipad and opad.

It can be said that the pseudorandomness of HMAC using MDP-MMO is almost reduced to the pseudorandomness of the underlying block cipher. Intuitively, it is because the chaining variables are fed into the block cipher via the key input and they are not disclosed to adversaries. On the other hand, if the Davies-Meyer compression function is used, then it is difficult to obtain the similar result. For this type of compression function, instead of the chaining variables, the message blocks are fed into the block cipher via the key input. They are selected and controlled fully by adversaries.

### 1.3 Related Work

Preneel et al. [15] presented a model of compression functions using a block cipher. It is called the PGV model and, for example, it includes DM and MMO compression functions.

Coron et al. [6] first discussed the indifferentiability of hash functions from the VIL random oracle. They presented four domain extension transforms: the Merkle-Damgård (MD) transform with prefix-free encoding, the MD transform dropping some output bits, and NMAC/HMAC-like transforms. Then, they showed that hash functions using them are indifferentiable from the VIL random oracle if the underlying compression functions are FIL random oracles. Moreover, they showed that hash functions using them and the DM compression function are indifferentiable from the VIL random oracle in the ideal cipher model.

Chang et al. [5] discussed the indifferentiability of hash functions from the VIL random oracle in the ideal cipher model. They assumed the compression functions using a block cipher in the PGV model [15] and the MD transform with prefix-free encoding for domain extension. They showed that the hash functions using 16 compression functions in the PGV model are indifferentiable from the VIL random oracle in the ideal cipher model. They also showed that the hash function using the MMO compression function is differentiable from the VIL random oracle.

On the other hand, as mentioned before, Gong, Lai and Chen claimed that it is possible to construct hash functions indifferentiable from the VIL random oracle in the ideal cipher model even with the MMO compression function [7].

Hirose, Park and Yun [8] proposed the MDP domain extension transform, and showed that a hash function using MDP and the DM compression function is indifferentiable from the VIL random oracle in the ideal cipher model. Ferguson had originally suggested an example of the MDP transform [9].

HMAC was first proposed by Bellare, Canetti and Krawczyk [2]. It was also shown in the same paper that HMAC is a PRF if the underlying compression function is a PRF with two keying strategies and the iterated hash function is weakly collision-resistant. Bellare proved that HMAC is a PRF under the sole assumption that the underlying compression function is a PRF with two keying strategies [1].

### 1.4 Organization

This article is organized as follows. Some notations and definitions are given in Section 2.. The definition of MDP-MMO is given in Section 3.. Section 4. is devoted to the indifferentiability of MDP-MMO from the VIL random oracle in the ideal cipher model. The security of HMAC using MDP-MMO as a PRF is discussed in Section 5..

## 2. Definitions

### 2.1 Notation

Let $\mathsf{Func}(Dom, Rng)$ be the set of all functions from $Dom$ to $Rng$, and $\mathsf{Perm}(Dom)$ be the set of all permutations on $Dom$.

Let $s \xleftarrow{\$} S$ represent that an element $s$ is selected from the set $S$ under the uniform distribution.

### 2.2 Pseudorandom Bit Generator

Let $g$ be a function such that $g : \{0,1\}^n \to \{0,1\}^l$, where $n < l$. Let $A$ be a probabilistic algorithm which outputs 0 or 1 for a given input in $\{0,1\}^l$. The prbg-advantage of $A$ against $g$ is defined as follows:

$$\mathrm{Adv}_g^{\mathrm{prbg}}(A) = \Big|\Pr[A(g(k)) = 1 \mid k \xleftarrow{\$} \{0,1\}^n] - \Pr[A(s) = 1 \mid s \xleftarrow{\$} \{0,1\}^l]\Big| ,$$

where the probabilities are taken over the coin tosses by $A$ and the uniform distributions on $\{0,1\}^n$ and $\{0,1\}^l$. $g$ is called a pseudorandom bit generator (PRBG) if $\mathrm{Adv}_g^{\mathrm{prbg}}(A)$ is negligible for any efficient $A$.

### 2.3 Pseudorandom Function

Let $f : Key \times Dom \to Rng$ be a keyed function or a function family. $f(k,\cdot)$ is often denoted by $f_k(\cdot)$. Let $A$ be a probabilistic algorithm which has oracle access to a function from $Dom$ to $Rng$. $A$ first asks elements in $Dom$ and obtains the corresponding elements in $Rng$ with respect to the function, and then outputs 0 or 1. The prf-advantage of $A$ against $f$ is defined as follows:

$$\mathrm{Adv}_f^{\mathrm{prf}}(A) = \Big|\Pr[A^{f_k} = 1 \mid k \xleftarrow{\$} Key] - \Pr[A^{\rho} = 1 \mid \rho \xleftarrow{\$} \mathsf{Func}(Dom, Rng)]\Big| ,$$

where the probabilities are taken over the coin tosses by $A$ and the uniform distributions on $Key$ and $\mathsf{Func}(Dom, Rng)$. $f$ is called a pseudorandom function (PRF) if $\mathrm{Adv}_f^{\mathrm{prf}}(A)$ is negligible for any efficient $A$.

Let $p : Key \times Dom \to Dom$ be a keyed permutation or a permutation family. The prp-advantage of $A$ against $p$ is defined similarly:

$$\mathrm{Adv}_p^{\mathrm{prp}}(A) = \Big|\Pr[A^{p_k} = 1 \mid k \xleftarrow{\$} Key] - \Pr[A^{\rho} = 1 \mid \rho \xleftarrow{\$} \mathsf{Perm}(Dom)]\Big| ,$$

where the probabilities are taken over the coin tosses by $A$ and the uniform distributions on $Key$ and $\mathsf{Perm}(Dom)$. $p$ is called a pseudorandom permutation (PRP) if $\mathrm{Adv}_p^{\mathrm{prp}}(A)$ is negligible for any efficient $A$.

### 2.4 Pseudorandom Function under Related-Key Attack

Pseudorandom functions under related-key attacks are first formalized by Bellare and Kohno [3]. In this article, we only consider a related-key attack with respect to a permutation $\pi$ as in [8]. We will refer to this type of related-key attack as the $\pi$-related-key attack and formalize in the following way. Let $A$ be a probabilistic algorithm which has oracle access to a pair of functions from $Dom$ to $Rng$. The prf-rka-advantage of $A$ against $f$ under the $\pi$-related-key attack is given by

$$\mathrm{Adv}_{\pi,f}^{\mathrm{prf\text{-}rka}}(A) = \Big|\Pr[A^{f_k, f_{\pi(k)}} = 1 \mid k \xleftarrow{\$} Key] - \Pr[A^{\rho,\rho'} = 1 \mid \rho, \rho' \xleftarrow{\$} \mathsf{Func}(Dom, Rng)]\Big| ,$$

where the probabilities are taken over the coin tosses by $A$ and the uniform distributions on $Key$ and $\mathsf{Func}(Dom, Rng)$. $f$ is called a $\pi$-RKA-secure PRF if $\mathrm{Adv}_{\pi,f}^{\mathrm{prf\text{-}rka}}(A)$ is negligible for any efficient $A$.

For a permutation, the prp-rka-advantage of an adversary and the $\pi$-RKA-secure PRP can also be defined similarly.

### 2.5 Computationally Almost Universal Function Family

Computationally almost universal function families are formalized by Bellare in [1]. Let $f : Key \times Dom \to Rng$ be a function family. Let $A$ be a probabilistic algorithm which takes no inputs and produces a pair of elements in $Dom$. The au-advantage of $A$ against $f$ is defined as follows:

$$\mathrm{Adv}_f^{\mathrm{au}}(A) = \Pr[f_k(M_1) = f_k(M_2) \wedge M_1 \neq M_2 \mid (M_1, M_2) \leftarrow A \wedge k \xleftarrow{\$} Key] ,$$

where the probabilities are taken over the coin tosses by $A$ and the uniform distribution on $Key$. $f$ is called a computationally almost universal function family if $\mathrm{Adv}_f^{\mathrm{au}}(A)$ is negligible for any efficient $A$.

### 2.6 Indifferentiability

The notion of indifferentiability is introduced by Maurer et al. [13] as a generalized notion of indistinguishability. Then, it is tailored to security analysis of hash functions by Coron et al. [6].

Let $C$ be an algorithm with oracle access to an ideal primitive $\mathcal{F}$. In the setting of this article, $C$ is an algorithm to construct a hash function using $\mathcal{F}$ with fixed input length. Let $\mathcal{H}$ be the VIL random oracle and $S$ be a simulator which has oracle access to $\mathcal{H}$. $S^{\mathcal{H}}$ tries to behave like $\mathcal{F}$ in order to convince an adversary that $\mathcal{H}$ is $C^{\mathcal{F}}$. Let $A$ be an adversary with access to two oracles. The indiff-advantage of $A$ against $C$ with respect to $S$ is given by

$$\mathrm{Adv}_{C,S}^{\mathrm{indiff}}(A) = \left|\Pr[A^{C^{\mathcal{F}},\mathcal{F}} = 1] - \Pr[A^{\mathcal{H},S^{\mathcal{H}}} = 1]\right| ,$$

where the probabilities are taken over the coin tosses by $A$, $C$ and $S$ and the distributions of ideal primitives. $C^{\mathcal{F}}$ is said to be indifferentiable from $\mathcal{H}$ if there exists a simulator $S^{\mathcal{H}}$ such that $\mathrm{Adv}_{C,S}^{\mathrm{indiff}}(A)$ is negligible for any efficient $A$.

### 2.7 Ideal Cipher Model

A block cipher with block length $n$ and key length $\kappa$ is called an $(n,\kappa)$ block cipher. Let $E : \{0,1\}^{\kappa} \times \{0,1\}^{n} \to \{0,1\}^{n}$ be an $(n,\kappa)$ block cipher. Then, $E(K,\cdot) = E_K(\cdot)$ is a permutation for every $K \in \{0,1\}^{\kappa}$. An $(n,\kappa)$ block cipher $E$ is called an ideal cipher if $E_K$ is a truly random permutation for every $K$.

The lazy evaluation of an ideal cipher is described as follows. The encryption oracle $E$ receives a pair of a key and a plaintext as a query, and returns a randomly selected ciphertext. On the other hand, the decryption oracle $D$ receives a pair of a key and a ciphertext as a query, and returns a randomly selected plaintext. The oracles $E$ and $D$ share a table of triplets of keys, plaintexts and ciphertexts, which are produced by the queries and the corresponding replies. Referring to the table, they select a reply to a new query under the restriction that $E_K$ is a permutation for every $K$.

## 3. MDP with MMO Compression Function

We denote concatenation of sequences by $\|$. For sequences $M_1, M_2, \ldots, M_k$, we often denote $M_1\|M_2\|\cdots\|M_k$ simply by $M_1M_2\cdots M_k$.

Let $\mathcal{B} = \{0,1\}^n$. Let $\mathcal{B}^+ = \cup_{i=1}^{\infty}\mathcal{B}^i$ and $\mathcal{B}^{\leq k} = \cup_{i=1}^{k}\mathcal{B}^i$.

Let $E : \mathcal{B}\times\mathcal{B} \to \mathcal{B}$ be an $(n,n)$ block cipher. The Matyas-Meyer-Oseas (MMO) compression function [14] $F : \mathcal{B}\times\mathcal{B} \to \mathcal{B}$ with $E$ is defined as follows: $F(s,x) = E_s(x) \oplus x$, where $s$ is a chaining variable and $x$ is a message block.

The MDP transform [8] of $F$ with a permutation $\pi$ is denoted by $F_{\pi}^{\circ} : \mathcal{B}\times\mathcal{B}^+ \to \mathcal{B}$ and defined as follows: For $s \in \mathcal{B}$ and $M_1M_2\cdots M_k$ ($M_i \in \mathcal{B}$),

(1) $s_0 = s$,
(2) $s_i = F(s_{i-1}, M_i)$ for $1 \leqq i \leqq k-1$,
(3) $s_k = F(\pi(s_{k-1}), M_k)$,
(4) $F_{\pi}^{\circ}(s, M_1M_2\cdots M_k) \stackrel{\mathrm{def}}{=} s_k$.

The following padding function pad : $\{0,1\}^* \to \cup_{i=2}^{\infty}\mathcal{B}^i$ is also prepared:

$$\mathsf{pad}(M) = M\|10^{\ell}\|bin(|M|) ,$$

where

- $\ell$ is the minimum non-negative integer such that $|M| + \ell \equiv 0 \pmod n$,
- $bin(|M|)$ is the $(n-1)$-bit binary representation of $|M|$.

Now, MDP-MMO is a scheme to construct a hash function using a block cipher $E : \mathcal{B}\times\mathcal{B} \to \mathcal{B}$, a permutation $\pi : \mathcal{B} \to \mathcal{B}$ and an initial value $IV \in \mathcal{B}$ defined as follows:

$$\mathsf{MDP\text{-}MMO}[E,\pi,IV](M) \stackrel{\mathrm{def}}{=} F_{\pi}^{\circ}(IV, \mathsf{pad}(M)) .$$

A diagram of MDP-MMO is shown in Figure 1.

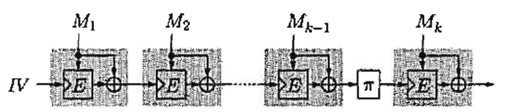


Fig. 1 $\mathsf{MDP\text{-}MMO}[E,\pi,IV](M)$. $\mathsf{pad}(M) = M_1M_2\cdots M_k$.

## 4. Security of MDP-MMO

### 4.1 Collision Resistance

It is easy to see that $\mathsf{MDP\text{-}MMO}[E,\pi,IV]$ is collision-resistant (CR) if its compression function is CR, that is, it is difficult to compute a pair of distinct $(S,X)$ and $(S',X')$ such that $E_S(X) \oplus X = E_{S'}(X') \oplus X'$ for the underlying block cipher $E$. The pseudorandomness of a block cipher cannot imply the property. It is easy to find a counterexample. However, it seems still reasonable to assume that a well-designed block cipher such as AES has this property.

The CR of MDP-MMO can also be proved in the ideal cipher model using the technique by Black et al. in [4].

### 4.2 Indifferentiability from Random Oracle

In this section, we show that $\mathsf{MDP\text{-}MMO}[E,\pi,IV]$ is indifferentiable from the VIL random oracle in the ideal cipher model. The following theorem states the indifferentiability of MDP-MMO in the ideal cipher model. The proof is omitted due to page limitation.

**Theorem 1** *Let $E$ be an $(n,n)$ block cipher. Let $\pi$ be a permutation and $P_{\pi}$ be the set of its fixed points. Let $A$ be an adversary that asks at most $q_{\mathrm{V}}$ queries to the VIL oracle, $q_{\mathrm{F}_{\mathcal{E}}}$ queries to the FIL encryption oracle and $q_{\mathrm{F}_{\mathcal{D}}}$*

*queries to the FIL decryption oracle. Let $l$ be the maximum number of message blocks in a VIL query. Suppose that $lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}} \leqq 2^{n-1}$. Then, in the ideal cipher model,*

$$\begin{aligned}
&\mathrm{Adv}^{\mathrm{indiff}}_{\mathrm{MDP\text{-}MMO},S_E,S_D}(A) \\
&\leqq \frac{7\,(lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}})^2 + 5\,(lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}})\,q_{\mathrm{F}_\mathcal{D}} - q_{\mathrm{F}_\mathcal{D}}^2}{2^{n+1}} \\
&\quad + \frac{2lq_{\mathrm{V}}(q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}})}{2^{n-1} - 3(lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}}) - |P_\pi|} \\
&\quad + \frac{(4\,|P_\pi| + 5)\,(lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}}) + 21\,q_{\mathrm{F}_\mathcal{D}}}{2^{n+1}} \ .
\end{aligned}$$

*where the simulators $S_E$ and $S_D$ are given in Figure 2. $S_E$ is a simulator for the encryption oracle, and $S_D$ for the decryption oracle. $S_E$ makes at most $q_{\mathrm{F}_\mathcal{E}}$ queries and runs in time $O(q_{\mathrm{F}_\mathcal{E}}(q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}}))$. $S_D$ makes at most $q_{\mathrm{F}_\mathcal{D}}$ queries and runs in time $O(q_{\mathrm{F}_\mathcal{D}}(q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}}))$.*

For Theorem 1, suppose that $\pi$ has no fixed points. Also suppose that $lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}} \leqq 2^{n-3}$, $lq_{\mathrm{V}} \geqq 1$, $q_{\mathrm{F}_\mathcal{E}} \geqq 1$, and $q_{\mathrm{F}_\mathcal{D}} \geqq 1$. Then, a looser but simpler bound is obtained:

$$\mathrm{Adv}^{\mathrm{indiff}}_{\mathrm{MDP\text{-}MMO},S_E,S_D}(A) \leqq \frac{(lq_{\mathrm{V}} + q_{\mathrm{F}_\mathcal{E}} + q_{\mathrm{F}_\mathcal{D}})^2}{2^{n-3}} \ .$$

$S_E$ and $S_D$ simulate the ideal cipher using lazy evaluation. In Figure 2, $\mathcal{P}(s)$ and $\mathcal{C}(s)$ represent the set of plaintexts and that of ciphertexts, respectively, which are available for the reply to the current query with the key $s$. Both of them are initially $\{0,1\}^n$, and their elements are deleted one by one as the simulation proceeds.

Let $(s_i, x_i, y_i)$ be the triplet determined by the $i$-th query of the adversary and the corresponding answer, where $E_{s_i}(x_i) = y_i$. Thus, for the MMO compression function, $s_i$ is a chaining variable, and $x_i$ is a message block. The triplets naturally defines a graph which initially consists of a single node labeled by the initial value $IV$ and grows as the simulation proceeds. $(s_i, x_i, y_i)$ adds two nodes labeled by $s_i$ and $z_i = x_i \oplus y_i$, and an edge labeled by $x_i$ from $s_i$ to $z_i$. The additions avoid duplication of nodes with the same labels.

The simulators use two sets $\mathcal{V}$ and $\mathcal{T}$. $\mathcal{V}$ keeps all the labels of the nodes with outgoing edge(s) in the graph. $\mathcal{T}$ keeps all the labels of the nodes reachable from the node labeled by $IV$ following the paths. The procedure getnode($s$) returns the sequence of labels of the edges on the path from the node labeled by $IV$ to the node labeled by $s$.

The simulators select a reply not simply from $\mathcal{C}(s)$ or $\mathcal{P}(s)$ but from $\mathcal{C}(s) \setminus S_{\mathrm{bad}}$ or $\mathcal{P}(s) \setminus S_{\mathrm{bad}}$. It prevents most of the events which make the simulators fail. For example, since $\{y \,|\, x \oplus y \in \mathcal{T}\} \subset S_{\mathrm{bad}}$, every node in $\mathcal{T}$ has a unique path from the node labeled by $IV$. Thus, $\tilde{M}$ is uniquely identified at the lines 204 and 304.

The most critical work of the simulators is to reply to the decryption query related to the final invocation of the compression function in MDP-MMO$[E, \pi, IV](M)$ for some $M$. Let $(s, x)$ be the query to $S_D$. In order to reply to $(s, x)$ properly, the simulator $S_D$ has to ask $M$ to the VIL random oracle $H$ and return $H(M) \oplus x$. Owing to the padding scheme pad, there exist only two possibilities for $M$, $M^{(0)}$ and $M^{(1)}$, which correspond to the message blocks $\tilde{M}$ fed to the compression functions before the permutation $\pi$. Thus, $S_D$ can accomplish the work.

## 5. Security of HMAC Using MDP-MMO

In this section, we discuss the pseudorandomness of HMAC using the MDP-MMO hash function (HMAC-MDP-MMO). This function is defined as follows:

$$\begin{aligned}
&\mathrm{HMAC}[E, \pi, IV](K, M) = \\
&\qquad H((K \oplus \mathtt{opad}) \| H((K \oplus \mathtt{ipad}) \| M)) \ ,
\end{aligned}$$

where $H$ is MDP-MMO$[E, \pi, IV]$ and $K$ is a secret key. A diagram of HMAC-MDP-MMO is given in Figure 3. Let us call $H((K \oplus \mathtt{ipad})\|\cdot)$ inner hashing and $H((K \oplus \mathtt{opad})\|\cdot)$ outer hashing.

The proof is similar to the one given by Bellare in [1].

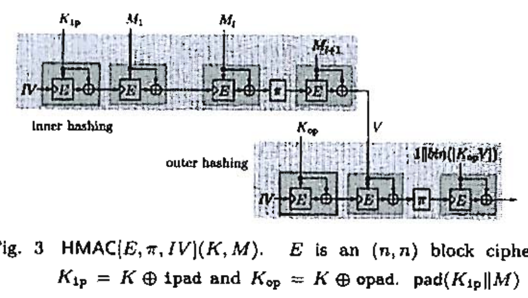


Fig. 3 HMAC$[E, \pi, IV](K, M)$. $E$ is an $(n, n)$ block cipher. $K_{\mathrm{ip}} = K \oplus \mathtt{ipad}$ and $K_{\mathrm{op}} = K \oplus \mathtt{opad}$. $\mathrm{pad}(K_{\mathrm{ip}} \| M) = K_{\mathrm{ip}} M_1 \cdots M_{l+1}$.

First, the compression function construction is considered. The following lemma says that the MMO compression function is a PRF when keyed via the chaining variable if the underlying block cipher is a PRP under the chosen plaintext attack up to the birthday bound. The proof is easy and omitted.

**Lemma 1** *Let $E$ be an $(n, n)$ block cipher and $F$ be a function such that $F_K(x) = E_K(x) \oplus x$.*

- *Let $A_F$ be a prf-adversary against $F$ which runs in time at most $t$ and asks at most $q$ queries. Then, there exists a prp-adversary $A_E$ against $E$ such that*

$$\mathrm{Adv}^{\mathrm{prf}}_F(A_F) \leqq \mathrm{Adv}^{\mathrm{prp}}_E(A_E) + \frac{q(q-1)}{2^{n+1}} \ ,$$

*where $A_E$ runs in time at most $t + O(q)$ and asks at most $q$ queries.*

**Initialize:**

$\mathcal{V} \leftarrow \emptyset$
$\mathcal{T} \leftarrow \{IV\}$
$\mathcal{P}(s) \leftarrow \{0,1\}^n$
$\mathcal{C}(s) \leftarrow \{0,1\}^n$

**Interface $\mathcal{E}(s,x)$:**

**if** $s \in \mathcal{T}$ **then**
  $E_s(x) \xleftarrow{\$} \mathcal{C}(s) \setminus S_{\mathrm{bad}}$
  $\mathcal{T} \leftarrow \mathcal{T} \cup \{E_s(x) \oplus x\}$
**else if** $\pi^{-1}(s) \in \mathcal{T}$ **then**
  $\tilde{M} \leftarrow \mathrm{getnode}(\pi^{-1}(s))$
  **if** $x \in \{lb(M^{(0)}), lb(M^{(1)})\}$ **then**
    **if** $x = lb(M^{(0)})$ **then**
      $E_s(x) \leftarrow \mathsf{H}(M^{(0)}) \oplus lb(M^{(0)})$
    **else**
      $E_s(x) \leftarrow \mathsf{H}(M^{(1)}) \oplus lb(M^{(1)})$
    **if** $E_s(x) \notin \mathcal{C}(s)$ **then**
      **return fail**
  **else**
    $E_s(x) \xleftarrow{\$} \mathcal{C}(s)$
**else**
  $E_s(x) \xleftarrow{\$} \mathcal{C}(s)$
$\mathcal{V} \leftarrow \mathcal{V} \cup \{s\}$
$\mathcal{P}(s) \leftarrow \mathcal{P}(s) \setminus \{x\}$
$\mathcal{C}(s) \leftarrow \mathcal{C}(s) \setminus \{E_s(x)\}$
**return** $E_s(x)$

**Interface $\mathcal{D}(s,x)$:**

**if** $s \in \mathcal{T}$ **then**
  $D_s(x) \xleftarrow{\$} \mathcal{P}(s) \setminus S_{\mathrm{bad}}$
  $\mathcal{T} \leftarrow \mathcal{T} \cup \{D_s(x) \oplus x\}$
**else if** $\pi^{-1}(s) \in \mathcal{T}$ **then**
  $\tilde{M} \leftarrow \mathrm{getnode}(\pi^{-1}(s))$
  **if** $x = \mathsf{H}(M^{(0)}) \oplus lb(M^{(0)})$ **then**
    $D_s(x) \leftarrow lb(M^{(0)})$
  **else if** $x = \mathsf{H}(M^{(1)}) \oplus lb(M^{(1)})$ **then**
    $D_s(x) \leftarrow lb(M^{(1)})$
  **else**
    $D_s(x) \xleftarrow{\$} \mathcal{P}(s) \setminus \{lb(M^{(0)}), lb(M^{(1)})\}$
**else**
  $D_s(x) \xleftarrow{\$} \mathcal{P}(s)$
$\mathcal{V} \leftarrow \mathcal{V} \cup \{s\}$
$\mathcal{P}(s) \leftarrow \mathcal{P}(s) \setminus \{D_s(x)\}$
$\mathcal{C}(s) \leftarrow \mathcal{C}(s) \setminus \{x\}$
**return** $D_s(x)$

Fig. 2 Pseudocode for the simulators $S_E$ and $S_D$. $S_{\mathrm{bad}} = \{y \mid y \in \{0,1\}^n \wedge x \oplus y \in \mathcal{V} \cup \mathcal{T} \cup \pi^{-1}(\mathcal{V} \cup \mathcal{T}) \cup \pi(\mathcal{T}) \cup P_\pi\}$. $\mathrm{pad}(M^{(0)}) = \tilde{M} \| lb(M^{(0)})$, and $\mathrm{pad}(M^{(1)}) = \tilde{M} \| lb(M^{(1)})$. $\tilde{M} = M^{(0)} \| 10^{\ell}$ $(0 \le \ell \le n-2)$ and $lb(M^{(0)}) = 0 \| bin(|M^{(0)}|)$. $\tilde{M} = M^{(1)}$ and $lb(M^{(1)}) = 1 \| bin(|M^{(1)}|)$.

- *Let $\pi$ be a permutation. Let $A_{\pi,F}$ be a prf-rka-adversary against $F$ with respect to $\pi$ which runs in time at most $t$ and asks at most $q$ queries. Then, there exists a prp-rka-adversary $A_{\pi,E}$ against $E$ with respect to $\pi$ such that*

$$\mathrm{Adv}^{\mathrm{prf\text{-}rka}}_{\pi,F}(A_{\pi,F}) \le \mathrm{Adv}^{\mathrm{prp\text{-}rka}}_{\pi,E}(A_{\pi,E}) + \frac{q(q-1)}{2^{n+1}} ,$$

*where $A_{\pi,E}$ runs in time at most $t+O(q)$ and asks at most $q$ queries.*

The following lemma is on the inner hashing. It says that, if the compression function $F$ is a $\pi$-RKA secure PRF, then the MDP composition of $F$ and $\pi$ is computationally almost universal. The proof is omitted due to page limitation.

**Lemma 2** *Let $F : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^\kappa$ be a function family, and let $A_{F^\circ_\pi}$ be an au-adversary against $F^\circ_\pi$. Suppose that $A_{F^\circ_\pi}$ outputs two messages with at most $\ell_1$ and $\ell_2$ blocks, respectively. Then, there exists a prf-rka-adversary $A_{\pi,F}$ against $F$ with respect to $\pi$ such that*

$$\mathrm{Adv}^{\mathrm{au}}_{F^\circ_\pi}(A_{F^\circ_\pi}) \le (\ell_1+\ell_2-1)\,\mathrm{Adv}^{\mathrm{prf\text{-}rka}}_{\pi,F}(A_{\pi,F}) + \frac{1}{2^\kappa} ,$$

*where $A_{\pi,F}$ runs in time at most $O((\ell_1+\ell_2)T_F)$ and makes at most 2 queries. $T_F$ represents the time required to compute $F$.*

Lemma 2 requires a $\pi$-RKA secure compression function. However, the assumption does not seem severe since adversaries are allowed to make only at most 2 queries to the oracles.

The following lemma is on the outer hashing. It says that, if the compression function is a PRF, then the outer-hashing function is also a PRF. The proof is omitted.

**Lemma 3** *Let $F : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^\kappa$ be a function family. Let $\hat{F}^2 : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^\kappa$ be a function family defined by*

$$\hat{F}^2(K,X) = F(\pi(F(K,X)), 1\|bin(\kappa+n)) ,$$

*where $K \in \{0,1\}^\kappa$ and $X \in \{0,1\}^n$. Let $A_{\hat{F}^2}$ be a prf-adversary against $\hat{F}^2$ that runs in time at most $t$ and makes at most $q$ queries. Then, there exist prf-adversaries $A_F$ and $A'_F$ against $F$ such that*

$$\mathrm{Adv}^{\mathrm{prf}}_{\hat{F}^2}(A_{\hat{F}^2}) \le \mathrm{Adv}^{\mathrm{prf}}_F(A_F) + q\,\mathrm{Adv}^{\mathrm{prf}}_F(A'_F) ,$$

*where $A_F$ runs in time at most $t+O(qT_F)$ and makes at most $q$ queries, and $A'_F$ runs in time $t+O(qT_F)$ and makes at most 1 query. $T_F$ represents the time required to compute $F$.*

The following lemma is Lemma 3.2 in [1]. It says that $h(K_o, G(K_1,\cdot))$ is a PRF if $h(K_o,\cdot)$ is a PRF and $G(K_1,\cdot)$ is computationally almost universal, where $K_o$ and $K_1$ are secret keys chosen uniformly and independently of each other.

**Lemma 4 (Lemma 3.2 in [1])** *Let $h : \{0,1\}^\mu \times \{0,1\}^n \to \{0,1\}^\mu$ and $G : \{0,1\}^\kappa \times \mathcal{D} \to \{0,1\}^n$ be function families. Let $hG : \{0,1\}^{\mu+\kappa} \times \mathcal{D} \to \{0,1\}^\mu$ be defined by $hG(K_o\|K_1, M) = h(K_o, G(K_1,M))$ for $K_o \in \{0,1\}^\mu$, $K_1 \in \{0,1\}^\kappa$ and $M \in \mathcal{D}$. Let $A_{hG}$ be a prf-adversary against $hG$ that runs in time at most $t$ and makes at most $q\,(\ge 2)$ queries each of whose lengths is at most $d$. Then, there exist a prf-adversary $A_h$ against $h$ and an au-adversary $A_G$ against $G$ such that*

$$\mathrm{Adv}^{\mathrm{prf}}_{hG}(A_{hG}) \le \mathrm{Adv}^{\mathrm{prf}}_h(A_h) + \frac{q(q-1)}{2}\mathrm{Adv}^{\mathrm{au}}_G(A_G) ,$$

*where $A_h$ runs in time at most $t$ and makes at most $q$ queries, and $A_G$ runs in time $O(T_G(d))$ and the two messages it outputs have length at most $d$. $T_G(d)$ is the time to compute $G$ on a $d$-bit input.*

The following theorem is on the pseudorandomness of the NMAC-like function made from $\mathsf{HMAC}[E,\pi,IV](K,\cdot)$ by replacing the first calls of the compression function in inner and outer hashing with two secret keys chosen uniformly and independently of each other. The theorem states that the security of the function as a PRF is reduced to the security of the underlying block cipher as a PRP under the related-key attack with respect to $\pi$. It directly follows from Lemmas 1 through 4.

**Theorem 2** *Let $\mathcal{B} = \{0,1\}^n$ and $E$ be an $(n,n)$ block cipher. Let $F : \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ be a function such that $F_K(x) = E_K(x) \oplus x$. Let $\hat{F}^2 F^\circ_\pi : \mathcal{B}^2 \times \mathcal{B}^+ \to \mathcal{B}$ be defined by $\hat{F}^2 F^\circ_\pi(K_o\|K_1, M) = \hat{F}^2(K_o, F^\circ_\pi(K_1, M))$ for $K_o, K_1 \in \mathcal{B}$ and $M \in \mathcal{B}^+$. Let $A_{\hat{F}^2 F^\circ_\pi}$ be a prf-adversary against $\hat{F}^2 F^\circ_\pi$ that runs in time at most $t$ and makes at most $q\,(\ge 2)$ queries each of which has at most $\ell$ blocks. Then, there exist prp-adversaries $A_E$ and $A'_E$ against $E$ and a prp-rka-adversary $A_{\pi,E}$ against $E$ with respect to $\pi$ such that*

$$\mathrm{Adv}^{\mathrm{prf}}_{\hat{F}^2 F^\circ_\pi}(A_{\hat{F}^2 F^\circ_\pi}) \le \mathrm{Adv}^{\mathrm{prp}}_E(A_E) + q\,\mathrm{Adv}^{\mathrm{prp}}_E(A'_E) + \ell\, q^2 \mathrm{Adv}^{\mathrm{prp\text{-}rka}}_{\pi,E}(A_{\pi,E}) + \frac{(2\ell+3)q^2}{2^{n+1}} ,$$

*where $A_E$ runs in time at most $t+O(qT_E)$ and makes at most $q$ queries, $A'_E$ runs in time at most $t+O(qT_E)$ and makes at most 1 query, and $A_{\pi,E}$ runs in time $O(\ell T_E)$ and makes at most 2 queries. $T_E$ represents the time required to compute $E$.*

The following lemma says that, even if the secret key of a PRF is replaced by the output of a PRBG, the resulting function remains a PRF. The proof is easy and omitted.

**Lemma 5** *Let $g : \{0,1\}^{\kappa} \to \{0,1\}^{\kappa'}$ be a function and $G : \{0,1\}^{\kappa'} \times \mathcal{D} \to \{0,1\}^n$ be a function family. Let $Gg : \{0,1\}^{\kappa} \times \mathcal{D} \to \{0,1\}^n$ be a function family defined by $Gg(K,M) = G(g(K),M)$ for $K \in \{0,1\}^{\kappa}$ and $M \in \mathcal{D}$. Let $A_{Gg}$ be a prf-adversary against $Gg$ that runs in time at most $t$ and makes at most $q$ queries of length at most $d$. Then, there exist a prbg-adversary $A_g$ against $g$ and a prf-adversary $A_G$ against $G$ such that*

$$\mathrm{Adv}^{\mathrm{prf}}_{Gg}(A_{Gg}) \leqq \mathrm{Adv}^{\mathrm{prbg}}_{g}(A_g) + \mathrm{Adv}^{\mathrm{prf}}_{G}(A_G) \ ,$$

*where $A_g$ runs in time at most $t + O(qT_G(d))$, and $A_G$ runs in time $t$ and makes at most $q$ queries of length at most $d$.*

Now, we can obtain the result on the pseudorandomness of HMAC-MDP-MMO simply by combining Theorem 2 and Lemma 5.

**Corollary 1** *Let $E$ be an $(n,n)$ block cipher. Let $g_E : \{0,1\}^n \to \{0,1\}^{2n}$ be a function such that $g_E(K) = (E_{IV}(K_{\mathrm{op}}) \oplus K_{\mathrm{op}}) \| (E_{IV}(K_{\mathrm{ip}}) \oplus K_{\mathrm{ip}})$, where $K_{\mathrm{op}} = K \oplus \mathtt{opad}$ and $K_{\mathrm{ip}} = K \oplus \mathtt{ipad}$. Let $A$ be a prf-adversary against $\mathsf{HMAC}[E,\pi,IV]$ that runs in time at most $t$ and makes at most $q\,(\geqq 2)$ queries each of which has at most $\ell$ blocks. Then, there exist prp-adversaries $A_E$ and $A'_E$ against $E$, a prp-rka-adversary $A_{\pi,E}$ against $E$ with respect to $\pi$ and a prbg-adversary $A_{g_E}$ such that*

$$\mathrm{Adv}^{\mathrm{prf}}_{\mathsf{HMAC}[E,\pi,IV]}(A) \leqq \mathrm{Adv}^{\mathrm{prbg}}_{g_E}(A_{g_E}) + \mathrm{Adv}^{\mathrm{prp}}_{E}(A_E) + q\,\mathrm{Adv}^{\mathrm{prp}}_{E}(A'_E) + \ell\, q^2 \mathrm{Adv}^{\mathrm{prp\text{-}rka}}_{\pi,E}(A_{\pi,E}) + \frac{(2\ell+3)q^2}{2^{n+1}} \ ,$$

*where $A_{g_E}$ runs in time at most $t + O(q\,\ell T_E)$, $A_E$ runs in time at most $t + O(qT_E)$ and makes at most $q$ queries, $A'_E$ runs in time at most $t + O(qT_E)$ and makes at most 1 query, and $A_{\pi,E}$ runs in time $O(\ell T_E)$ and makes at most 2 queries.*

Actually, we have not completely reduced the security of HMAC-MDP-MMO as a PRF to the security of the underlying block cipher as a PRP under the related-key attack with respect to $\pi$. It is easy to see that the function $g_E$ in Corollary 1 may not be a PRBG in general even if $E$ is a PRP. However, it does not seem so difficult to design a block cipher $E$ such that $g_E$ is a PRBG. This is because $IV$ is a fixed initial value chosen by the designer of the hash function and the block cipher. Furthermore, `ipad` and `opad` are fixed sequences given by HMAC. Any adversary has no control over them.

## Acknowledgements

The authors would like to thank Dr. Yoshida and Dr. Ideguchi at Hitachi, Ltd. and Prof. Ohta and Dr. Wang at The University of Electro-Communications for their valuable discussions and comments on this research. This research was supported by the National Institute of Information and Communications Technology, Japan.

## References

[1] M. Bellare. New proofs for NMAC and HMAC: Security without collision-resistance. In *CRYPTO 2006 Proceedings, Lecture Notes in Computer Science 4117*, pages 602–619, 2006. The full version is "Cryptology ePrint Archive: Report 2006/043" at `http://eprint.iacr.org/`.

[2] M. Bellare, R. Canetti, and H. Krawczyk. Keying hash functions for message authentication. In *CRYPTO '96 Proceedings, Lecture Notes in Computer Science 1109*, pages 1–15, 1996.

[3] M. Bellare and T. Kohno. A theoretical treatment of related-key attacks: RKA-PRPs, RKA-PRFs, and applications. In *EUROCRYPT 2003 Proceedings, Lecture Notes in Computer Science 2656*, pages 491–506, 2003.

[4] J. Black, P. Rogaway, and T. Shrimpton. Black-box analysis of the block-cipher-based hash-function constructions from PGV. In *CRYPTO 2002 Proceedings, Lecture Notes in Computer Science 2442*, pages 320–335, 2002.

[5] D. Chang, S. Lee, M. Nandi, and M. Yung. Indifferentiable security analysis of popular hash function with prefix-free padding. In *ASIACRYPT 2006 Proceedings, Lecture Notes in Computer Science 4284*, pages 283–298, 2006.

[6] J.-S. Coron, Y. Dodis, C. Malinaud, and P. Puniya. Merkle-Damgård revisited: How to construct a hash function. In *CRYPTO 2005 Proceedings, Lecture Notes in Computer Science 3621*, pages 430–448, 2005.

[7] Z. Gong, X. Lai, and K. Chen. A synthetic indifferentiability analysis of some block-cipher-based hash functions. Cryptology ePrint Archive, Report 2007/465, 2007. `http://eprint.iacr.org/`.

[8] S. Hirose, J. H. Park, and A. Yun. A simple variant of the Merkle-Damgård scheme with a permutation. In *ASIACRYPT 2007 Proceedings, Lecture Notes in Computer Science 4833*, pages 113–129, 2007.

[9] J. Kelsey. A comment on draft FIPS 180-2. Public Comments on the Draft Federal Information Processing Standard (FIPS) Draft FIPS 180-2, Secure Hash Standard (SHS), 2001. `http://csrc.nist.gov/CryptoToolkit/shs/dfips-180-2-comments1.pdf`.

[10] H. Krawczyk, M. Bellare, and R. Canetti. HMAC: Keyed-hashing for message authentication. Network Working Group RFC 2104, 1997.

[11] H. Kuwakado and M. Morii. Compression functions suitable for the multi-property-preserving transform. Cryptology ePrint Archive, Report 2007/302, 2007. `http://eprint.iacr.org/`.

[12] S. M. Matyas, C. H. Meyer, and J. Oseas. Generating strong one-way functions with cryptographic algorithm. *IBM Technical Disclosure Bulletin*, 27:5658–5659, 1985.

[13] U. M. Maurer, R. Renner, and C. Holenstein. Indifferentiability, impossibility results on reductions, and applications to the random oracle methodology. In *Proceedings of the First Theory of Cryptography Conference (TCC '04), Lecture Notes in Computer Science 2951*, pages 21–39, 2004.

[14] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1996.

[15] B. Preneel, R. Govaerts, and J. Vandewalle. Hash functions based on block ciphers: A synthetic approach. In *CRYPTO '93 Proceedings, Lecture Notes in Computer Science 773*, pages 368–378, 1994.