J. Math. Crypt. **3** (2009), 133–148 DOI 10.1515 / JMC.2009.006

Improved security analysis for OMAC as a pseudorandom function

Mridul Nandi

Communicated by Mike Burmester

Abstract. This paper shows that the advantage of any q-query adversary (which makes at most q queries) for distinguishing OMAC from a uniform random function is roughly $Lq^2/2^n$. Here L is the number of blocks of the longest query and n is the output size of the uniform random function. The so far best bound is roughly $\sigma^2/2^n = \mathrm{O}(L^2q^2/2^n)$ and hence our new bound is an improved bound. Our improved security analysis also works for OMAC1 and CMAC which has been recommended by NIST as a candidate of blockcipher based MAC.

Keywords. MAC, PMAC, distinguishing attack, pseudorandom function, random permutation.

AMS classification. 94A60.

1 Introduction

Pseudorandom functions or prf are an essential primitive in cryptography. A prf is a natural candidate of message authentication code or MAC[1]. It has been widely used in other constructions such as (strong) pseudorandom permutations or (s)prp, authenticated encryptions or AE, and even in public-key encryptions, for example, DHIES [1]. There are several candidates for prf. Cipher-Block-Chaining [2] or CBC is a method to obtain prf from an n-bit blockcipher $\pi : \{0,1\}^n \to \{0,1\}^n$. Given $(x_1, \ldots, x_\ell) \in (\{0,1\}^n)^\ell$, CBC^π or π^+ is defined as:

$$\pi^+(x_1, \ldots, x_\ell) = \pi(\pi(\ldots \pi(\pi(x_1) \oplus x_2) \ldots \oplus x_{\ell-1}) \oplus x_\ell).$$

There are different variants of CBC constructions, for example, XCBC [6], FCBC [6] TMAC [14], OMAC [11]. Among all these constructions, OMAC or One-key CBC MAC proposed by Iwata and Kurosawa is one of the best choices. Its use has been suggested in many applications such as EAX [4] (a secure authenticated encryption), TET [10] (a length-preserving tweakable strong pseudorandom permutation) etc. Given a padded[2] message $(x_1, \ldots, x_\ell) \in (\{0,1\}^n)^\ell$,

$$\mathrm{OMAC}^\pi(x_1, \ldots, x_\ell) = \pi^+(x_1, \ldots, x_{\ell-1}, c \cdot \pi(\mathbf{0}) \oplus x_\ell)$$

where $\mathbf{0} = 0^n$ and c is either c_0 or c_1 (see [8, 14, 11, 12] for the exact values of c_0 and c_1) and $\cdot$ denotes the Galois field multiplication over the set $\{0,1\}^n$. If the message-size is a multiple of n then $c = c_0$ otherwise, $c = c_1$. These two constants are chosen

[1] The Actual security notion for MAC is weaker than prf and is analogous with public-key digital signature.
[2] Padding is necessary only if the size of message is not multiple of n.

in such a way that the Galois field multiplication $c \cdot \pi(\mathbf{0})$ is efficiently computable. Only differences among OMAC [11], OMAC1 and CMAC [8] are choices of the constants c_0 and c_1. CMAC, which is equivalent to OMAC1 or OMAC, is considered as a recommended candidate among all sequential message authentication codes. This is mainly because of the key-size (a single key is sufficient). Moreover, it is also efficient (close to CBC-efficiency) when we have sequential invocations of block-ciphers and it can process messages of any size. A competitive prf is PMAC [7] proposed by Rogaway, which can be implemented in parallel. Jutla provides a class of DAG-based constructions [13] which also contains some interesting candidates of prf.

A recent development on improved prf-insecurity analysis

In recent years new developments on finding improved prf-insecurity analysis methodologies in some of the aforementioned constructions have been found. Intuitively, prf-insecurity is the maximum advantage of a distinguisher, where the advantage of a distinguisher $\mathcal{A}$ for a construction $\mathcal{D}$ is the success probability to distinguish $\mathcal{D}$ from the ideal random function or a uniform random function (whose values are independent and uniformly distributed). Note that in this paper, random means stochastic, whereas uniform random corresponds to the particular random which is uniformly distributed. A uniform random function or permutation corresponds to the ideal random function or permutation which are also familiar as random function or permutation in cryptography. We denote the advantage by $\mathbf{Adv}^{\text{prf}}_{\mathcal{D}}(\mathcal{A})$ and hence prf-insecurity for $\mathcal{D}$ as $\mathbf{Insec}^{\text{prf}}_{\mathcal{D}}(q,\sigma,L) := \max_{\mathcal{A}} \mathbf{Adv}^{\text{prf}}_{\mathcal{D}}(\mathcal{A})$ where the maximum is taken over all (q,σ,L)-distinguisher which make at most q queries having altogether at most σ blocks with L as the number of blocks for the longest query. A block means an element of the message space of the underlying blockcipher. For example, when AES-128 is used as underlying blockcipher, 128-bits are considered as a block. Here we use $\{0,1\}^n$ to denote the set of all possible blocks and the underlying block-cipher has message space $\{0,1\}^n$. If $\mathbf{Insec}^{\text{prf}}_{\mathcal{D}}(q,\sigma,L) < \epsilon$ then for any (q,σ,L)-distinguisher $\mathcal{A}$, the advantage for $\mathcal{A}$ distinguishing $\mathcal{D}$ from a uniform random function is at most ϵ. Thus a sharper upper bound of $\mathbf{Insec}^{\text{prf}}_{\mathcal{D}}(q,\sigma,L)$ guarantees stronger security.

(1) The first motivating result in this area can be found in [3] where an improved security analysis of CBC (for fixed length or for arbitrary length prefix-free messages) has been provided. The authors have shown that $\mathbf{Insec}^{\text{prf}}_{\text{CBC}}(q,\sigma,L) \leq 12Lq^2/2^n + 64q^2L^4/2^{2n}$. The second term becomes negligible or in the order of the first term if maximum number of blocks of a query, L, is small compared to 2^n. For example, if $L < 2^{n/3-1}$ then $\mathbf{Insec}^{\text{prf}}_{\text{CBC}}(q,\sigma,L) \leq 20Lq^2/2^n$.

(2) After this work, the improved analysis techniques were used on other constructions. In FSE-07 [16], improved bounds for XCBC, TMAC and PMAC have been provided. The prf-insecurity bounds for XCBC and TMAC are of form $\text{O}(Lq^2/2^n) + \text{O}(L^4q^2/2^{2n})$ and the bound for PMAC is $10Lq^2/2^n$. They have used the Maurer methodology [15] to obtain an improved bound.

(3) In [18], an improved bound of the form $O(q\sigma/2^n)$ for PMAC has been shown and it is mentioned that this bound is truly an improved bound because the original bounds $O(\sigma^2/2^n)$ can be better than the new bounds $O(Lq^2/2^n)$ (if the number of blocks of the longest query becomes significant compared with the number of queries). But this problem is not present when the bound is of the form $O(q\sigma/2^n)$. Moreover, we have $q\sigma/2^n = O(Lq^2/2^n)$ for any choices of q, σ and L.

(4) The original prf-insecurity bound for OMAC was $(5(L^2+1)q^2+1)/2^n$, presented in FSE-03 [11]. Later an improved bound $(4\sigma^2 + 1)/2^n$ for OMAC is presented in Indocrypt-03 [12].

Improved prf-insecurity analysis of OMAC and its importance

The aforementioned results motivate us to search for an improved bound for OMAC. In the case of CBC, we need to consider only prefix-free messages. On the other hand, TMAC and XCBC have another independent key along with the blockcipher key. This independent key is used to mask the final output which eventually helps us to obtain the improved bound. Since OMAC does not satisfy any of the above properties, we need to be careful to obtain an improved bound for it. In this paper, we have provided the following prf-insecurity bound for OMAC as given in Theorem 4.6. Let ℓ_i denote the number of message blocks of the i^{th} query and $N = 2^n$.

$$\mathbf{Adv}^{\text{prf}}_{\text{OMAC}}(\mathcal{A}) \leq \frac{5q\sigma}{N} + \sum_{1\leq i<j\leq q} \frac{(\ell_i+\ell_j)^4}{N^2} \leq \frac{5q\sigma}{N} + \frac{8q(q-1)L^4}{N^2}.$$

By simplifying $\sum_{1\leq i<j\leq q}(\ell_i+\ell_j)^4$, we can rewrite the bound as (again Theorem 4.6)

$$\mathbf{Insec}^{\text{prf}}_{\text{OMAC}}(q,\sigma,L) \leq \frac{13q\sigma}{N} \text{ if } L < N^{1/3}.$$

The assumption $L < N^{1/3}$ is not restrictive and it holds in almost all applications. Now we briefly describe why we are getting an improved bound of the form $\frac{q\sigma}{N}$ instead of $\frac{\sigma^2}{N}$. When adversary is making q queries with total σ blocks then we have roughly σ intermediate inputs to the underlying blockcipher. Among which q inputs are final inputs. Final inputs are those intermediate inputs whose corresponding blockcipher output is the output of OMAC. The probability of collision between any final input and intermediate input is roughly $\frac{1}{2^n}$ on average[3]. The number of such pairs is roughly $q\sigma$. Thus, the probability of having collision between final inputs and intermediate inputs is $O(\frac{q\sigma}{2^n})$. Given that the final inputs are completely new (the complement of the above event) among all intermediate inputs, the probability distribution of the final output is very close to uniform. Hence it is difficult to distinguish output of OMAC from the output of a uniform random function. This is why we get an improved bound of the form $\frac{q\sigma}{2^n}$. If we consider that all intermediate inputs are distinct then it is likely that

[3]There are some pairs where the collision probability is more than $\frac{1}{2^n}$. We can still obtain the average as $\frac{1}{2^n}$ by estimating the numbers of such pairs.

we get the bound of the form $\frac{\sigma^2}{N}$. In case of the improved bound, collisions among all intermediate non-final inputs are allowed.

Note that this new bound is sharper than the till-date best known bound for OMAC provided $\sigma \geq 3.25 \times q$ and $L < 2^{n/3}$. Secondly, $q\sigma/2^n = \mathrm{O}(Lq^2/2^n)$ for any choices of q, σ and L, which is sharper than $L^2q^2/2^n$ in terms of order of function. So the new bound provides evidence that the number of queries has more significance than the query-length[4]. Our security analysis is valid for any non-zero, non-1^n distinct constants c_0 and c_1. Thus the same security analysis is true for OMAC1 and also CMAC.

The paper is organized as follows. We first provide the definition of prf and the measurement of prf-insecurity in Section 2. In the same section we state an important and useful theorem called the *strong interpolation theorem*. In Section 3, we provide an equivalent definition of OMAC based on intermediate inputs and outputs. We provide our new improved security analysis for OMAC in Section 4. In Section 5, we provide proofs of some of the statements which are not proved in Section 4. Finally, we conclude with possible future work.

2 Pseudorandom function and measurement of insecurity

Notation

We use the following notations in this paper. For any positive integer L, $A^{\leq L} = \bigcup_{i=0}^{L} A^i$, $A^* = \bigcup_{i=0}^{\infty} A^i$ and $A^+ = \bigcup_{i=1}^{\infty} A^i$. Note that $A^0 = \{\lambda\}$, where λ is the empty string. If $x \in \{0,1\}^i$ then we write $|x| = i$ and call the size of x as i. If the size is n (underlying block-cipher has domain on $\{0,1\}^n$) then we also call it as a block. Given any $M \in \{0,1\}^*$, the number of blocks of M is defined as $\|M\| = \lceil |M|/n \rceil$. Given a function $f : A \to B$ and q elements $x_1, \ldots, x_q \in A$, we denote the interpolation as $f(x_1, \ldots, x_q) := (f(x_1), \ldots, f(x_q)) \in B^q$. Given two positive integers a and b we denote $\mathbf{P}(a,b) = a(a-1)\ldots(a-b+1)$. By our convention $\mathbf{P}(a,0) = 1$. We denote $[0,t] := \{0,1,\ldots,t\}$. Let $x = (x_0, x_1, \ldots, x_t)$ be a vector (or tuple) of $t+1$ elements and $I = \{i_1, \ldots, i_s\} \subseteq [0,t]$ where $i_1 < \ldots < i_s$. We denote sub-vector $x_I = (x_{i_1}, \ldots, x_{i_s})$.

Definition 2.1 (Random function). Let $\mathrm{Func}(A,B)$ be the set of all functions from A to B and $\mathrm{Perm}(A)$ be the set of all permutations on A. A random function $\mathtt{X}$ from A to B is a random variable taking values on $\mathrm{Func}(A,B)$. It is called a random permutation on A if the random function has support on $\mathrm{Perm}(A) \subset \mathrm{Func}(A,A)$. Thus, $\mathtt{X}$ is a random permutation if $\Pr[\mathtt{X} \in \mathrm{Perm}(A)] = 1$.

A *uniform random function* or URF (the classical random function) [9] is the uniform random variable on $\mathrm{Func}(A,B)$ for some finite sets A and B. That is, $\Pr[\mathtt{G} = f] = \frac{1}{|B|^{|A|}}$ for all $f \in \mathrm{Func}(A,B)$. Similarly we define a *uniform random permutation* or URP $\mathtt{F}$ (the classical random permutation) on A as the uniform random variable on

[4] The bound is quadratic in q whereas linear in L.

Perm$(A) \subset$ Func(A, A). Given q distinct elements $x_1, \ldots, x_q \in A$ we can compute the joint distribution of $\mathtt{F}(x_1, \ldots, x_q)$ where $\mathtt{F}$ is either a uniform random function or a uniform random permutation on A. The following result is based on a straightforward counting of number of functions and permutations.

Proposition 2.2 (Interpolation probability for URF or URP). *Let $x_1, \ldots, x_q$ be q distinct elements. If $\mathtt{G}$ is a uniform random function then we have*

$$\Pr[\mathtt{G}(x_1, \ldots, x_q) = (y_1, \ldots, y_q)] = \frac{1}{|B|^q}.$$

If $\mathtt{F}$ is a uniform random permutation then the above probability is $\frac{1}{\mathbf{P}(|A|,q)}$ if $y_1, \ldots, y_q$ are distinct, otherwise the probability is zero.

Random function based on domain extension

A domain extension $\mathcal{D}$ is a mapping from Func(A, B) to Func$(\widetilde{A}, B)$ with $A \subset \widetilde{A}$. Now, any random function $\mathtt{F}$ on Func$(\widetilde{A}, B)$ induces a random function $\mathcal{D}(\mathtt{F}) := \mathcal{D}^{\mathtt{F}}$ on Func$(\widetilde{A}, B)$. In this paper we study the random function OMAC$^{\mathtt{F}}$ defined on Func$(\{0,1\}^{\leq nL}, \{0,1\}^n)$ where the *underlying random function* $\mathtt{F}$ is a uniform random permutation on Perm$(\{0,1\}^n)$ and L is the maximum number of blocks of a query.

A distinguisher $\mathcal{A}$ is nothing but an oracle algorithm which outputs 0 or 1. It can have an internal random coin $\mathtt{R}$. The oracle can be a function or a random function. Now we define advantage of a distinguisher at distinguishing two random functions and define prf-insecurity of a random function.

Definition 2.3 (Advantage and prf-Insecurity). The advantage of a distinguisher $\mathcal{A}_{\mathtt{R}}$ (a distinguisher $\mathcal{A}$ with random coins $\mathtt{R}$) at distinguishing two random functions $\mathtt{X}_1$ and $\mathtt{X}_2$ is defined as

$$\mathbf{Adv}_{\mathcal{A}_{\mathtt{R}}}(\mathtt{X}_1, \mathtt{X}_2) = \left|\Pr_{\mathtt{R},\mathtt{X}_1}[\mathcal{A}_{\mathtt{R}}^{\mathtt{X}_1} = 1] - \Pr_{\mathtt{R},\mathtt{X}_2}[\mathcal{A}_{\mathtt{R}}^{\mathtt{X}_2} = 1]\right|.$$

Let $\mathtt{G}$ be a uniform random function from $\{0,1\}^{\leq nL}$ to $\{0,1\}^n$. Then for a tuple of positive integers (q, σ, L) and a random function $\mathtt{X}$ we define,

$$\mathbf{Insec}_{\mathtt{X}}^{\mathrm{prf}}(q, \sigma, L) = \max_{\mathcal{A}} \mathbf{Adv}_{\mathcal{A}}(\mathtt{X}, \mathtt{G})$$

where the maximum is taken over all distinguishers making exactly q queries having altogether at most σ blocks with L as the number of blocks of a longest query. We call this type of distinguisher a (q, σ, L)-distinguisher.

Notational assumptions

In this paper we denote all q queries by $M_1, \ldots, M_q$ and there is no loss in assuming that all queries are distinct. We denote $\|M_i\| = \ell_i$ and hence $\sum_i \ell_i = \sigma$ and $\ell_i \leq L$ for all i. We use the notation $N = 2^n$.

A q-tuple $\mathbf{z} = (z_1, \ldots, z_q) \in C^q$ is called *block-wise distinct* if all z_i's are distinct where $z_i \in C$. Now we state a useful theorem which has been proven in [19, 17]. This is a general version of a theorem stated in [5].

Theorem 2.4 (Strong Interpolation Theorem). *Let* $\mathtt{G}$ *be a uniform random function with domain* $\{0,1\}^{\leq L}$ *and range* $\{0,1\}^n$ *and* $\mathtt{X}$ *be a random function with domain and range same as* $\mathtt{G}$*. Suppose for any block-wise distinct* $\mathbf{M} = (M_1, \ldots, M_q) \in (\{0,1\}^{\leq L})^q$*, block-wise distinct* $\mathbf{z} \in (\{0,1\}^n)^q$ *and for any* ε *(may depend on* $N = 2^n, q, \sigma$ *and* ℓ_i*'s) we have*

$$\Pr[\mathtt{X}(M_1) = z_1, \ldots, \mathtt{X}(M_q) = z_q] \geq \frac{(1-\varepsilon)}{N^q}.$$

Then we have $\mathbf{Adv}_{\mathcal{A}}(\mathtt{X}, \mathtt{G}) \leq \varepsilon + \frac{q(q-1)}{2N}$ *where* $\mathcal{A}$ *is a distinguisher making* q *queries with block length* $\ell_1, \ldots, \ell_q$*. Thus,* $\mathbf{Insec}^{\mathrm{prf}}_{\mathtt{X}}(q, \sigma, L) \leq \varepsilon + \frac{q(q-1)}{2N}$ *when* ε *does not depend on* ℓ_i*'s individually.*

Thus the computation of the interpolation probability $\Pr[\mathtt{X}(M_1) = z_1, \ldots, \mathtt{X}(M_q) = z_q]$ is important. Later we define the OMAC construction based on a uniform random permutation $\mathtt{F}$ and we compute the interpolation probability

$$\Pr[\mathrm{OMAC}^{\mathtt{F}}(M_1) = z_1, \ldots, \mathrm{OMAC}^{\mathtt{F}}(M_q) = z_q].$$

For a uniform random function $\mathtt{G}$, we have already stated the interpolation probability which is $\Pr[\mathtt{G}(\mathbf{M}) = \mathbf{z}] = \frac{1}{N^q}$ where $\mathbf{M} = (M_1, \ldots, M_q)$ and $\mathbf{z} = (z_1, \ldots, z_q)$.

3 One-key CBC MAC or OMAC

3.1 Definition of OMAC, OMAC1 and CMAC

In this paper, the Galois field $\mathbb{F}_{2^n}$ of order 2^n is defined on the set $\{0,1\}^n$. We denote $+$ and $\cdot$ for the field addition and multiplication and take $\mathbf{0}$ and $\mathbf{1}$ the additive and multiplicative identity respectively. Let $\pi \in \mathrm{Perm}(\mathbb{F}_{2^n})$ and $\pi^+ : \mathbb{F}_{2^n}^+ \to \mathbb{F}_{2^n}$ is defined as

$$\pi^+(m_1, \ldots, m_\ell) = \pi(\ldots(\pi(m_1) + m_2)\ldots + m_\ell)$$

where $m_1, \ldots, m_\ell \in \mathbb{F}_{2^n}$. The above function is also known as the CBC function. Now we define the OMAC function for arbitrary length, based on a n-bit permutation π. A pseudo-code is given in Figure 2 and an illustration for a three block message is provided in Figure 1. Given a message $M \in \{0,1\}^*$, we define $\mathbf{pad}(M) = \overline{M} \in (\{0,1\}^n)^+$ as

$$\overline{M} = \begin{cases} M^* & \text{if } n \nmid |M|, \\ M & \text{otherwise,} \end{cases}$$

where $M^* = M \parallel 10^i$ and $i = n \cdot \lceil \frac{|M|+1}{n} \rceil - |M| - 1$ (this is the smallest non-negative integer such that $|M10^i|$ is a multiple of n). We define the padding indicator constant as

$$\delta_M = \begin{cases} 1 & \text{if } n \nmid |M|, \\ 0 & \text{if } n \mid |M|. \end{cases}$$

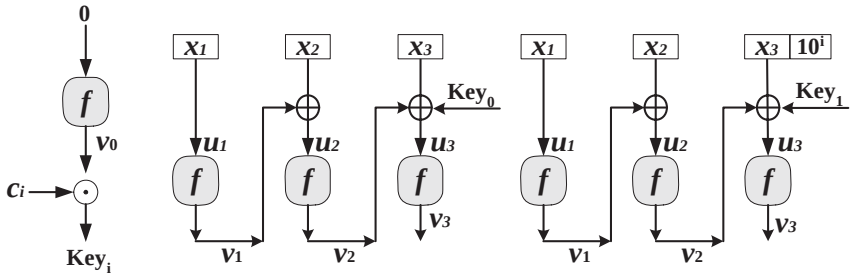

Figure 1. OMAC: $\text{Key}_i = c_i.f(\mathbf{0})$. Here the c_i's are distinct non-$\mathbf{0}$ and non-$\mathbf{1}$ constants such that $c_0+c_1 \neq 1$. The function f is the underlying blockcipher. The right most part is for incomplete message block. $\mathbf{0}$, u_1, u_2, u_3 are called intermediate inputs (inputs for f) and v_0, v_1, v_2, v_3 are called intermediate outputs (outputs for f). v_3 is the final output of OMAC.

Now given $\pi \in \text{Perm}(\mathbb{F}_{2^n})$ we define the OMAC function as

$$\text{OMAC}^{\pi}(M) = \pi\big(\pi^{+}(m_1, \ldots, m_{\ell-1}) + m_\ell + c_\delta \cdot \pi(\mathbf{0})\big)$$

where $\overline{M} = (m_1, \ldots, m_\ell) \in \mathbb{F}_{2^n}^{\ell}, \delta = \delta_M \in \{0,1\}$ and c_0,c_1 are non-zero, non-$\mathbf{1}$ distinct constants such that $c_0 + c_1 \neq 1$ (which is indeed true for the original choices of these constants [8, 14, 11, 12]).

3.2 Equivalent definition of OMAC construction

While computing OMAC, the inputs of π are known as intermediate inputs. The last intermediate input or the final input is $\pi^{+}(m_1, \ldots, m_{\ell-1}) + m_\ell + c_\delta \cdot \pi(\mathbf{0})$. Similarly, the outputs of π are known as intermediate outputs and the final intermediate output is nothing but the output of OMAC. Now we write the definition of OMAC in terms of these intermediate inputs and outputs. Let $\pi \in \text{Perm}(\{0,1\}^n)$ and $\overline{M} = (m_1, \ldots, m_\ell)$ and denote $\delta = \delta_M$.

Definition 3.1. The values u_i, $0 \leq i \leq \ell$, (including $u_0 = \mathbf{0}$) are known as *intermediate input* and u_ℓ is known as the *final input*. Similarly the v_i's, $0 \leq i \leq \ell$, are known as *intermediate output* and v_ℓ is known as the *final output*.

One can observe that intermediate inputs are linear functions of message blocks and previous intermediate outputs. In fact, this type of linear relationship can be found in many constructions in all CBC-families and PMAC. We consider two column vectors $\mathbf{v}^{M,\pi} = (v_0, v_1, \ldots, v_\ell)^{\text{tr}}$ and $\mathbf{u}^{M,\pi} = (u_0, u_1, \ldots, u_\ell)^{\text{tr}}$ (these two vectors are

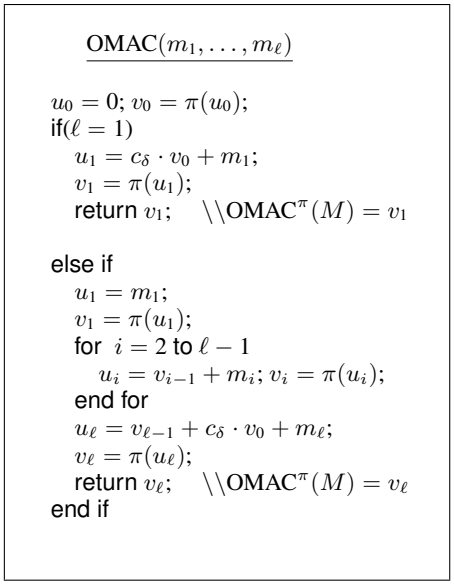

Figure 2. Definition of OMAC.

completely determined by π and M) and called the intermediate output vector and intermediate input vector respectively. Here "tr" means the transpose of a vector or a matrix. Now we represent the relationship between intermediate inputs and intermediate outputs by a matrix known as a *coefficient matrix* $\mathbf{A}^M_{(\ell+1)\times(\ell+2)}$. We have $\mathbf{A}^M \cdot \overline{\mathbf{v}}^{M,\pi} = \mathbf{u}^{M,\pi}$ where $\overline{\mathbf{v}}^{M,\pi} = \begin{pmatrix} \mathbf{1} \\ \mathbf{v}^{M,\pi} \end{pmatrix}$ and the coefficient matrix is

(1) if $\ell = 1$:

$$\mathbf{A}^M = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_1 & c_\delta & \mathbf{0} \end{pmatrix},$$

(2) if $\ell \geq 2$:

$$\mathbf{A}^M = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_1 & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_2 & \mathbf{0} & \mathbf{1} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ m_{\ell-1} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ m_\ell & c_\delta & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{1} & \mathbf{0} \end{pmatrix}.$$

Definition 3.2 (Equivalent definition of OMAC). Given a message M, compute the two unique vectors $\mathbf{u}^{M,\pi} = (u_0, u_1, \ldots, u_\ell)^{\text{tr}}$ and $\mathbf{v}^{M,\pi} = (v_0, v_1, \ldots, v_\ell)^{\text{tr}}$ such that

(1) $\mathbf{A}^M \cdot \overline{\mathbf{v}}^{M,\pi} = \mathbf{u}^{M,\pi}$ and

(2) $\pi(\mathbf{u}^{M,\pi}) = \mathbf{v}^{M,\pi}$.

Define $\text{OMAC}^\pi(M) := v_\ell$.

It is easy to verify that these two vectors are uniquely defined satisfying the above two properties because the coefficient matrix is a lower-triangular matrix. By $\mathbf{u}^M$ and $\mathbf{v}^M$, we mean the random variable for the vectors of intermediate inputs and intermediate outputs along with relation $\mathbf{A}^{\mathbf{M}} \cdot \overline{\mathbf{v}}^{\mathbf{M}} = \mathbf{u}^{\mathbf{M}}$. This relation is satisfied independently of π. Once we fix a permutation π, $\mathbf{u}^{M,\pi}$ and $\mathbf{v}^{M,\pi}$ are fixed vectors belonging to $\mathbb{F}_{2^n}^{\ell+1}$. In the next section, we extend this definition for q distinct messages.

4 Improved security analysis of OMAC

In this section, we provide the proof of our main theorem modulo some claims which are proved in the next section.

(1) We first compute the interpolation probability for OMAC based on a uniform random permutation $\mathtt{F}$ (a permutation is chosen uniformly from the set of all permutations on n-bits). In particular we show that (Proposition 4.5), given any q distinct messages $M_1, \ldots, M_q$ and q distinct n-bit outputs $z_1, \ldots, z_q$,

$$\Pr[\text{OMAC}^{\mathtt{F}}(M_1) = z_1, \ldots, \text{OMAC}^{\mathtt{F}}(M_q) = z_q] \geq (1 - \varepsilon) \times \frac{1}{N^q},$$

where $N = 2^n$ and $\varepsilon = 5q\sigma/N + 8q(q-1)L^4/N^2 - q(q-1)/2N$. This probability calculation is done by solving some matrix equations.

(2) Note that for a uniform random function the above interpolation probability is $\frac{1}{N^q}$. Now a distinguisher has only seen messages M_i's and its outputs z_is which occur with almost equal probability in both cases (as stated above). Hence, it is hard to distinguish OMAC from a uniform function. A more formal statement is known as strong interpolation theorem (see Theorem 2.4) and by this theorem the insecurity of OMAC is bounded by ε stated above.

We have already defined a coefficient matrix $\mathbf{A}^M$ for a given message M. Now we extend our definition for a tuple of q messages $\mathbf{M} = (M_1, \ldots, M_q)$. First we think all intermediaries for q messages as random variables. Note that the intermediate inputs are linearly related to the intermediate outputs. The linear relation is captured by the coefficient matrix $\mathbf{A}^{\mathbf{M}}$ as in a single message. Let us consider an example for two messages.

Example 4.1. Let the padded messages $\overline{M_1} = (m_1^1, m_2^1, m_3^1)$ and $\overline{M}_2 = (m_1^2, m_2^2)$ then define the coefficient matrix for the pair $\mathbf{M} = (M_1, M_2)$ is

$$\mathbf{A}^{\mathbf{M}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_1^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_2^1 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_3^1 & c_\delta & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_1^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_2^2 & c_{\delta'} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \end{pmatrix}.$$

Let $(u_0^1 = \mathbf{0}, u_1^1, u_2^1, u_3^1)$ and $(v_0, v_1^1, v_2^1, v_3^1)$ denote the intermediate inputs and outputs respectively for the message M_1. Similarly, we have $(u_0^1 = \mathbf{0}, u_1^2, u_2^2)$ and (v_0, v_1^2, v_2^2) for the message M_2. Note that the first intermediate output variable is same for both messages (in fact, it corresponds to the output of $\mathbf{0}$). Define the combined intermediate input and output vectors as

$$\mathbf{u}^{\mathbf{M}} = (\mathbf{0}, u_1^1, u_2^1, u_3^1, u_1^2, u_2^2)^{\mathsf{tr}}, \quad \mathbf{v}^{\mathbf{M}} = (v_0, v_1^1, v_2^1, v_3^1, v_1^2, v_2^2)^{\mathsf{tr}}.$$

Let $F = \{t_1 := \ell_1, t_2 := \ell_1 + \ell_2, \ldots, t_q := \sum_{i=1}^q \ell_i = t - 1\}$, known as the *set of the final input indices*.

Definition 4.2 (Coefficient matrix for a tuple of message $\mathbf{M} = (M_1, \ldots, M_q)$). Let $M_1, M_2, \ldots, M_q$ be q distinct messages with $\|M_i\| = \ell_i$. We denote m_j^i, u_j^i and v_j^i, $1 \le j \le \ell_i$, $1 \le i \le q$ for the j^{th} message block, intermediate input and output respectively for the i^{th} message. We define the combined intermediate input and intermediate output as

$$\mathbf{u}^{\mathbf{M}} = (\mathbf{0}, u_1^1, \ldots, u_{\ell_1}^1, \ldots, u_1^q, \ldots, u_{\ell_q}^q)_{t\times 1}^{\mathsf{tr}},$$
$$\mathbf{v}^{\mathbf{M}} = (v_0, v_1^1, \ldots, v_{\ell_1}^1, \ldots, v_1^q, \ldots, v_{\ell_q}^q)_{t\times 1}^{\mathsf{tr}}$$

where $t = \ell_1 + \ldots + \ell_q + 1 := \sigma + 1$. Now, define the combined coefficient matrix $\mathbf{A}^{\mathbf{M}}_{t\times(t+1)} = ((a_{\alpha,\beta}))_{0\le\alpha\le\sigma}^{0\le\beta\le t}$ which represents the linear relationship between these two formal variables $\mathbf{u}^{\mathbf{M}}$ and $\mathbf{v}^{\mathbf{M}}$ as $\mathbf{A}^{\mathbf{M}} \cdot \mathbf{v}^{\mathbf{M}} = \mathbf{u}^{\mathbf{M}}$:

$$a_{i,j} = \begin{cases} m_{j'}^{i'} & \text{if } j = 0, i = t_{i'-1} + j', 0 < j' \le \ell_{i'}, \\ c_{\delta_{M_{i'}}} & \text{if } j = 1, i = t_{i'}, \\ 1 & \text{if } i = j \notin F, \\ 0 & \text{otherwise .} \end{cases}$$

One can easily observe that for each t_j, $\mathbf{A}_{\cdot t_j} = \mathbf{0}^t$ where $\mathbf{A}^{\mathbf{M}} = (\alpha^{\mathbf{M}}\ \mathbf{A}^{\mathbf{M}}_{\cdot 1}\ \ldots\ \mathbf{A}^{\mathbf{M}}_{\cdot t})$ and $\alpha^{\mathbf{M}} = (\mathbf{0}, m_1^1, \ldots, m_{\ell_1}^1, \ldots, m_1^q, \ldots, m_{\ell_q}^q)^{\mathsf{tr}}$. In other words, the final outputs have no effect on the intermediate inputs. The first row is zero and it corresponds to the input $\mathbf{0}$ and the first column is $\alpha^{\mathbf{M}}$ corresponding to all message blocks including the constant zero block. Now for any permutation π we have

(1) $\mathbf{A}^{\mathbf{M}} \cdot \overline{\mathbf{v}}^{\mathbf{M},\pi} = \mathbf{u}^{\mathbf{M},\pi}$ and

(2) $\mathbf{v}^{\mathbf{M},\pi} = \pi(\mathbf{u}^{\mathbf{M},\pi})$,

(3) $(\mathrm{OMAC}^{\pi}(M_1), \ldots, \mathrm{OMAC}^{\pi}(M_q)) = \mathbf{v}_F^{\mathbf{M},\pi}$ (the sub-vector indexed by the set F).

We define two bad sets. From now onwards, we fix q distinct messages $M_1, \ldots, M_q$ and a q-tuple $(z_1, \ldots, z_q) \in \mathbb{F}_{2^n}^q$ so that z_i's are distinct and $\|M_i\| = \ell_i$ and $\sum_i \ell_i = \sigma = t - 1$, $L = \max_i \ell_i$. Moreover, we denote by u_j^i's and v_j^i's for the intermediate inputs and outputs respectively, $1 \le i \le q$, $0 \le j \le \ell_i$ where $u_0^i = \mathbf{0}$ and $v_0^i = v_0$. All these variables take values from $\mathbb{F}_{2^n}$ once we fix a permutation π.

Definition 4.3 (Bad events). We say that the event $\mathtt{Bad}_1$ is true (or π satisfies $\mathtt{Bad}_1$) if for some $(i, \ell_i) \neq (i', j')$, $u_{\ell_i}^i = u_{j'}^{i'}$. In other words, there is a collision between final input and intermediate input. We say that the event $\mathtt{Bad}_2$ is true (or π satisfies $\mathtt{Bad}_2$) if for some (i, j), $1 \le j < \ell_i$, $1 \le i \le q$, $v_j^i \in \{z_1, \ldots, z_q\}$. In other words, some of the intermediate outputs (except final outputs) are from the set $\{z_1, \ldots, z_q\}$.

We shall show that if bad event sets are not true the output distributions of OMAC is close to uniform. We compute the probability of the bad event sets for π uniformly chosen from $\mathrm{Perm}(\{0,1\}^n)$ i.e., the uniform random permutation. Let $\mathtt{F}$ be the uniform random permutation on $\{0,1\}^n$. We first compute the probability of bad events. The proof is given in Section 5.

Proposition 4.4.

$$\Pr[\mathtt{Bad}_1] \le \frac{4(q-1)\sigma}{N} + \sum_{1 \le i < j \le q} \frac{(\ell_i + \ell_j)^4}{N^2} := \varepsilon_1.$$

$$\Pr[\mathtt{Bad}_2] \le \frac{(\sigma - q + 1)(q+1)}{N} := \varepsilon_2.$$

Proposition 4.5.

$$\Pr[\mathrm{OMAC}^{\mathtt{F}}(M_1) = z_1, \ldots, \mathrm{OMAC}^{\mathtt{F}}(M_q) = z_q | \ \overline{\mathtt{Bad}_1 \cup \mathtt{Bad}_2}\ ] \ge \frac{1}{N^q}.$$

$$\Pr[\mathrm{OMAC}^{\mathtt{F}}(M_1) = z_1, \ldots, \mathrm{OMAC}^{\mathtt{F}}(M_q) = z_q] \ge \frac{1 - \varepsilon_1 - \varepsilon_2}{N^q}.$$

Proof. It is easy to see that for a fixed input vector $\mathbf{w}$ such that $\Pr[\mathtt{F}$ is good and $\mathbf{v}_I^{\mathtt{F}} = \mathbf{z}] > 0$ we have $\Pr[\mathbf{v}_F^{\mathtt{F}} = \mathbf{z} \,|\overline{\mathtt{Bad}}$ and $\mathbf{v}_I^{\mathtt{F}} = \mathbf{w}] \ge \frac{1}{\mathbf{P}(N,q)}$. □

Now we simplify the sum in the definition of ε_1. One can write $\sum_{1 \le i < j \le q} (\ell_i + \ell_j)^4 = (q-1)\sum_i \ell_i^4 + 4\sum_i \ell_i^3(\sigma - \ell_i) + 3\sum_i \ell_i^2(\sum_j \ell_j^2 - \ell_i^2)$. Suppose $L \le N^{1/3}$. Since $\ell_i^4 \le L^3 \ell_i$, $(q-1)\sum_i \ell_i^4 \le Nq\sigma$. Similarly, $\ell_i^2 \le \ell_i \cdot L$, we have $(\sum_i \ell_i^2)^2 \le L^2\sigma^2 \le Nq\sigma$ ($\sigma \le Lq$). Thus,

$$\frac{\sum_{1 \le i < j \le q} (\ell_i + \ell_j)^4}{N^2} \le \frac{8q\sigma}{N}.$$

By using the above inequality and the strong interpolation theorem one can obtain our following main theorem. The second inequality of the first part is straightforward by substituting $\ell_i + \ell_j \le 2L$.

Theorem 4.6 (Improved security bound for OMAC). *For any distinguisher $\mathcal{A}$ making at most q queries having at most σ blocks such that the number of blocks of a longest query is at most L, the prf-advantage of $\mathcal{A}$ for* OMAC *is*

$$\mathbf{Adv}^{\text{prf}}_{\text{OMAC}}(\mathcal{A}) \le \frac{5q\sigma}{N} + \sum_{1\le i<j\le q} \frac{(\ell_i+\ell_j)^4}{N^2} \le \frac{5q\sigma}{N} + \frac{8q(q-1)L^4}{N^2}.$$

Moreover if $L \le N^{1/3}$, then we have

$$\mathbf{Insec}^{\text{prf}}_{\text{OMAC}}(q,\sigma,L) \le \frac{13q\sigma}{N}.$$

5 Proof of Proposition 4.4

The proof of Proposition 4.4 needs a few more definitions and notations. We mainly want to count the number of permutations π such that $\mathbf{A}^{\mathbf{M}} \cdot \overline{\mathbf{v}}^{\mathbf{M},\pi} = \mathbf{u}^{\mathbf{M},\pi}$, $\pi(\mathbf{u}^{\mathbf{M},\pi}) = \mathbf{v}^{\mathbf{M},\pi}$ and π satisfies $\mathtt{Bad}_1$ or $\mathtt{Bad}_2$. We denote $\mathbf{v}^{\mathbf{M},\pi} := \mathbf{v} = (v_0 = \pi(\mathbf{0}), v_1, \ldots, v_\sigma)$ and similarly, $\mathbf{u}^{\mathbf{M},\pi} := \mathbf{u} = (u_0 = \mathbf{0}, u_1, \ldots, u_\sigma)$. Now we define an equivalence relation on intermediate input which characterizes all intermediate collisions on input (equivalently output since π is permutation). In [3], authors considered directed graphs for improved security analysis of CBC which is another equivalent representation of an equivalence relation. In general, it would not be easy to handle with a directed graph.

Definition 5.1. Given $\pi \in \text{Perm}(\mathbb{F}_{2^n})$ we can define an *induced equivalence relation* $\mathfrak{R} = \mathfrak{R}^\pi$ on $[0, t-1]$ as $(i,j) \in \mathfrak{R}$ if and only if $u_i = u_j$ (equivalently $v_i = v_j$). We also say that $\mathbf{u}$ (equivalently $\mathbf{v}$) satisfies $\mathfrak{R}$. An equivalence relation $\mathfrak{R}$ is also called an induced equivalence relation if there is a permutation π such that $\mathfrak{R}^\pi = \mathfrak{R}$.

Note that an equivalence relation may not be an induced equivalence relation. A tuple $(i_1, \ldots, i_s)$ is called the tuple of representatives of $\mathfrak{R}$ on $[0, t-1]$ if $0 = i_1 < i_s \le t-1$ and $\mathfrak{R}$ has s distinct equivalence classes $[i_j]$'s such that i_j is minimum in the class $[i_j]$. Given that the induced relation is $\mathfrak{R}$, we can modify the equation $\mathbf{A} \cdot \overline{\mathbf{v}} = \mathbf{u}$ into $\mathbf{A}^{\mathfrak{R}} \cdot \overline{\mathbf{v}_{\mathfrak{R}}} = \mathbf{u}$ where the matrix $\mathbf{A}^{\mathfrak{R}}$ and the vector $\mathbf{v}_{\mathfrak{R}}$ are defined as follows.

Definition 5.2. Suppose $(i_1, \ldots, i_s)$ is the tuple of representatives of $\mathfrak{R}$ on $[0, t-1]$. Now we define a new $t \times (s+1)$ matrix $\mathbf{B} := \mathbf{A}^{\mathfrak{R}} = (\alpha^{\mathbf{M}}\ \mathbf{B}_{.1} \ldots \mathbf{B}_{.s})$ where $\mathbf{B}_{.j} = \sum_{i\in[i_j]} \mathbf{A}_{.i}$. If $\mathbf{v}$ satisfies $\mathfrak{R}$, we consider a new s-vector $(w_1, \ldots, w_s) = \mathbf{w} = \mathbf{v}^{\mathfrak{R}}$ such that $w_j = v_{i_j}$.

We also say that $\mathbf{B}$ (or $\mathbf{A}^{\mathfrak{R}}$) is obtained by merging $\mathbf{A}$ w.r.t. $\mathfrak{R}$. In this new terminology, $\mathbf{B} \cdot \overline{\mathbf{w}} = \mathbf{u}$ where $\mathbf{w}$ is block-wise distinct.

Definition 5.3. We define the *rank* of a permutation π (also the rank of the induced relation $\mathfrak{R}^\pi$) as the rank of the set of vectors $\mathcal{V} = \{\mathbf{B}_i - \mathbf{B}_j : (i,j) \in \mathfrak{R}\}$.

Since $\mathbf{u}$ satisfies the relation $\mathfrak{R}$, the vector $\mathbf{w}$ must be a solution for $\mathcal{V}$. The number of block-wise distinct solutions[5] is at most $\mathbf{P}(N, s-r)$ where $r := \text{rank}(\mathcal{V}) := \text{rank}(\mathfrak{R})$. Given any such solutions $\mathbf{w}$ (that uniquely determine $\mathbf{v}$ also) there are at most $(N-s)!$ permutations π (s outputs of π are already determined) such that $\mathbf{v}^{\mathbf{M},\pi} = \mathbf{v}$. Thus, given a relation $\mathfrak{R}$ of rank r and of size s, there are at most $(N-s)! \times \mathbf{P}(N, s-r) \leq N! \times \frac{1}{P(N-s+r,r)}$ permutations π such that $\mathfrak{R}^\pi = \mathfrak{R}$.

Lemma 5.4. *Given a relation $\mathfrak{R}$ of rank r and of size s, there are at most $N! \times \frac{1}{P(N-s+r,r)}$ permutations π such that $\mathfrak{R}^\pi = \mathfrak{R}$.*

Lemma 5.5. *The number of relations of rank r is at most $\binom{t}{2}^r$.*

In [3], a similar lemma has been proved for CBC constructions. A very similar analysis will work here and hence we omit the proof.

Corollary 5.6. *Let $q = 2$, $\mathbf{M} = (M_1, M_2)$ and $\|M\| = \ell, \|M'\| = \ell'$ such that $(\ell + \ell')^2 \leq N$. Then, the number of permutations of rank at least two is at most $N! \times \frac{(\ell+\ell')^4}{N^2}$.*

From Lemma 5.4 and 5.5, one can show the corollary. A similar result was also stated in the case of CBC [3]. An element i is called single in $\mathfrak{R}$ if $[i] = \{i\}$. A set is called single if every element is a single element. Now it is easy to see that for any distinct $M \neq M'$ and the induced relation $\mathfrak{R}_0$ of rank zero (there is exactly one such) the following property holds: both ℓ and $\ell + \ell'$ are single elements in $\mathfrak{R}_0$. In fact, one can write down the relation $\mathfrak{R}_0$.

Lemma 5.7. *Let $\overline{M} = (m_1, \ldots, m_\ell)$ and $\overline{M'} = (m'_1, \ldots, m'_{\ell'})$. If $m_1 = \mathbf{0}$ then $(0, 1) \in \mathfrak{R}_0$ and similarly, if $m'_1 = \mathbf{0}$ then $(0, \ell + 1) \in \mathfrak{R}_0$. If $(m_1, \ldots, m_{\ell-1})$ and $(m'_1, \ldots, m'_{\ell'-1})$ have exactly $p \geq 1$ common prefix blocks then $(1, \ell+1), \ldots (p, \ell+p) \in \mathfrak{R}_0$.*

The relation $\mathfrak{R}_0$ corresponds to the trivial collision which hold for any permutation. This is due to the choice of message blocks. For example, if we know that two messages share a common prefix then the intermediate input and output up to the common part are identical independent of the underlying permutation π. Now we study the number of valid relations of rank one such that $F = \{\ell, \ell + \ell'\}$ is not single. We consider two cases.

Case-A: $\boldsymbol{\delta_M \neq \delta_{M'}}$. Suppose F is not single in a valid relation $\mathfrak{R}$ of rank one and say $(\ell + \ell', i') \in \mathfrak{R}$. Let $\mathbf{B}_i - \mathbf{B}_j$ be an independent vector for $\mathcal{V}$ such that $i, j \notin F$ and $\mathbf{B} = \mathbf{A}^{\mathfrak{R}}$. Then the second element in $\mathbf{B}_{\ell+\ell'} - \mathbf{B}_{i'}$ is not zero (either $c_{\delta'} - c_\delta$ or $c_{\delta'} - 1$ or $c_{\delta'}$) whereas that of $\mathbf{B}_i - \mathbf{B}_j$ is zero. Thus, the rank should be more than one. Here the only possible valid relation of rank one such that F is not single, is the one with the basis (i, j) where either i or $j \in F$. Thus, the number of such relations is at most $2(\ell + \ell')$.

[5] This is a straightforward generalization of a well-known linear algebra fact which says that the number of all solutions is exactly N^{s-r} if there is one such solution.

Case-B: $\delta_M = \delta_{M'}$. Suppose we have $(\ell+\ell', i') \in \mathfrak{R}$ where $i \notin F$. Then by similar reasoning, the basis should contain the pair whose one element is from F. So there are at most $2(\ell + \ell')$ such relations.

Now we consider the case when $(\ell + \ell', \ell) \in \mathfrak{R}$. This implies that $\mathrm{CBC}(\overline{M}) = \mathrm{CBC}(\overline{M}')$. Since $\delta_M = \delta_{M'}$, $\overline{M} \neq \overline{M}'$. Now as in Lemma 13 of [3], we know that there are at most $d(|\ell - \ell'|)$ relations of rank one containing the pair $(\ell + 1, \ell' + 1)$. Here $d(m)$ is the number of factors of m. Thus, the total number of relations of rank one such that F is not single is at most $3(\ell + \ell')$.

Lemma 5.8. *For $q = 2$, the number of induced relations of rank one such that $\{\ell, \ell+\ell'\}$ is not single is at most $3(\ell + \ell')$.*

Let $M \neq M'$ and let $\overline{M} = (m_1, \ldots, m_\ell)$, $\overline{M'} = (m'_1, \ldots, m'_{\ell'})$, $\delta = \delta_M$ and $\delta' = \delta_{M'}$. We denote the intermediate inputs and outputs by u_i, v_i, u'_i and v'_i. Let $\mathrm{New} := \mathrm{New}[M, M']$ be the event that

$$u_\ell \neq u'_{\ell'} \quad \text{and} \quad \{u_\ell, u'_{\ell'}\} \cap \{u_1, \ldots, u_{\ell-1}, u'_1, \ldots, u'_{\ell'-1}, \mathbf{0}\} = \emptyset.$$

In this case, we also say that the final inputs are new. One can similarly define the event New for q distinct messages $M_1, \ldots, M_q$. An easy exercise shows that

$$\mathrm{New}[M_1, \ldots, M_q] = \bigcap_{1 \le i < j \le q} \mathrm{New}[M_i, M_j].$$

Now it is easy to see that $\mathtt{Bad}_1 = \overline{\mathrm{New}[M_1, \ldots, M_q]}$ is the complement of the event New. From the above discussion and by using Corollary 5.6 we have the following results.

Lemma 5.9. *If $\mathtt{F}$ is a uniform random permutation then for any two distinct messages $M \neq M'$ such that $\overline{M} \in \mathbb{F}_{2^n}^{\ell}$ and $\overline{M'} \in \mathbb{F}_{2^n}^{\ell'}$ we have,*

$$\Pr[\overline{\mathrm{New}[M, M']}] \le \frac{4(\ell + \ell')}{N} + \frac{(\ell + \ell')^4}{N^2}.$$

The first part of Proposition 4.4 is a corollary of the above lemma by summing over all possible pairs of messages. The second part is proved in the following lemma.

Lemma 5.10.

$$\Pr[\mathtt{Bad}_2] \le \frac{(\sigma - q + 1)(q + 1)}{N}$$

where $\sigma = \sum_{j=1}^{q} \ell_j = t - 1$.

Proof. We define an event $E_j : v_{i_j}^{\mathtt{F}} \notin \mathbf{z}$, $1 \le j \le \sigma - q$ where $I = \{i_1, i_1, \ldots, i_{\sigma+1-q}\}$ and $i_0 < \ldots < i_{\sigma+1-q}$. $E_{\le j} = \bigcup_{s=1}^{i} E_s$. Now, it is easy to see that $\Pr[E_{i+1} \mid E_{\le i}] \ge \frac{N-q-i}{N-i}$ and hence $\Pr[E_{\le t-q}] \ge \prod_{i=0}^{\sigma-q} \frac{N-q-i}{N-i} \ge 1 - \frac{(\sigma-q+1)(q+1)}{N}$. Thus, $\Pr[\mathtt{Bad}_2] \le \frac{(\sigma-q+1)(q+1)}{N}$. □

6 Conclusion and future work

In this paper we have provided an improved prf-insecurity bound which is $\frac{13q\sigma}{2^n}$. This improved bound suggests that OMAC is a strong design for prf or MAC. The idea of the proof can be used for the improved security analysis of other constructions of MAC including DAG-based constructions. We also hope that this idea is useful to obtain an improved and more appealing security analysis for other indistinguishability security notions such as online cipher, PRP or SPRP, authenticated encryption modes of operation etc. It would be interesting to see a distinguishing attack for MAC which achieves this security bound $\Omega(Lq^2/2^n)$ where L is not constant or one can try to further reduce the bound to $O(q^2/2^n)$ + some small terms.

Acknowledgments. We would like to thank the anonymous reviewers for their valuable comments on earlier drafts of this paper.

References

[1] Michel Abdalla, Mihir Bellare, and Phillip Rogaway, *The Oracle Diffie-Hellman Assumptions and an Analysis of DHIES*. CT-RSA (David Naccache, ed.), Lecture Notes in Computer Science 2020, pp. 143–158. Springer, 2001.

[2] Mihir Bellare, Joe Kilian, and Phillip Rogaway, *The Security of Cipher Block Chaining*. CRYPTO (Yvo Desmedt, ed.), Lecture Notes in Computer Science 839, pp. 341–358. Springer, 1994.

[3] Mihir Bellare, Krzysztof Pietrzak, and Phillip Rogaway, *Improved Security Analysis for CBC MACs*. Advances in Cryptology – Crypto 2005, Lecture Notes in Computer Science 3621, pp. 527–545. Springer, Berlin, 2005.

[4] Mihir Bellare, Phillip Rogaway, and David Wagner, *The EAX Mode of Operation*. FSE (Bimal K. Roy and Willi Meier, eds.), Lecture Notes in Computer Science 3017, pp. 389–407. Springer, 2004.

[5] Daniel J. Bernstein, *A short proof of the unpredictability of cipher block chaining (2005)*, Available at `http://cr.yp.to/papers.html#easycbc`.

[6] John Black and Phillip Rogaway, *CBC MACs for arbitrary length messages*. Advances in Cryptology – Crypto 2000, Lecture Notes in Computer Science 1880, pp. 197–215. Springer, Berlin, 2000.

[7] ———, *A Block-Cipher Mode of Operation for Parallelizable Message Authentication*. Advances in Cryptology – Eurocrypt 2002, Lecture Notes in Computer Science 2332, pp. 384–397. Springer, Berlin, 2002.

[8] Morris Dworkin., *Recommendation for Block Cipher Modes of Operation: The CMAC Mode for Authentication.*, `http://csrc.nist.gov/publications/nistpubs/index.html#sp800-38B`.

[9] Oded Goldreich, Shafi Goldwasser, and Silvio Micali, *How to construct random functions*, J. ACM 33 (1986), pp. 792–807.

[10] Shai Halevi, *Invertible Universal Hashing and the TET Encryption Mode*. CRYPTO (Alfred Menezes, ed.), Lecture Notes in Computer Science 4622, pp. 412–429. Springer, 2007.

[11] Tetsu Iwata and Kaoru Kurosawa, *OMAC: One-Key CBC MAC*. Fast Software Encryption, 10th International Workshop – FSE 2003, Lecture Notes in Computer Science 2887, pp. 129–153. Springer, Berlin, 2003.

[12] ———, *Stronger Security Bounds for OMAC, TMAC, and XCBC*. INDOCRYPT (Thomas Johansson and Subhamoy Maitra, eds.), Lecture Notes in Computer Science 2904, pp. 402–415. Springer, 2003.

[13] Charanjit S. Jutla, *PRF Domain Extension using DAG*. Theory of Cryptography: Third Theory of Cryptography Conference – TCC 2006, Lecture Notes in Computer Science 3876, pp. 561–580. Springer, Berlin, 2006.

[14] Kaoru Kurosawa and Tetsu Iwata, *TMAC: Two-Key CBC MAC*. Topics in Cryptology – CT-RSA 2003: The Cryptographers' Track at the RSA Conference 2003, Lecture Notes in Computer Science 2612, pp. 33–49. Springer, Berlin, 2003.

[15] Ueli M. Maurer, *Indistinguishability of Random Systems*. Advances in Cryptology – Eurocrypt 2002, Lecture Notes in Computer Science 2332, pp. 110–132. Springer, Berlin, 2002.

[16] Kazuhiko Minematsu and Toshiyasu Matsushima, *New Bounds for PMAC, TMAC, and XCBC*. FSE (Alex Biryukov, ed.), Lecture Notes in Computer Science 4593, pp. 434–451. Springer, 2007.

[17] Mridul Nandi, *A Simple and Unified Method of Proving Indistinguishability*. Progress in Cryptology – Indocrypt 2006, Lecture Notes in Computer Science 4329, pp. 317–334. Springer, Berlin, 2006.

[18] Mridul Nandi and Avradip Mandal, *Improved security analysis of PMAC*, Journal of Mathematical Cryptology 2 (2008), pp. 149–162.

[19] Serge Vaudenay, *Decorrelation: A Theory for Block Cipher Security*. Journal of Cryptology 16(4), Lecture Notes in Computer Science, pp. 249–286. Springer-Verlag, New York, 2003.

Received 28 February, 2008; revised 26 May, 2009

Author information

Mridul Nandi, Indian Statistical Institute, Kolkata, India.
Email: mridul.nandi@gmail.com