Improved security analysis of OMAC

Mridul Nandi

CINVESTAV-IPN, Mexico City
mridul.nandi@gmail.com

Abstract. We present an improved security analysis of OMAC, the construction is widely used as a candidate of MAC or Pseudo Random Function (or PRF). In this direction, the first result was given in Crypto-05 where an improved security analysis of CBC (for fixed length or for arbitrary length prefix-free messages) had provided. Followed by this work, improved bounds for XCBC, TMAC and PMAC were found. The improved bounds are of the form $O(\frac{Lq^2}{2^n})$ where the original bounds are $O(\frac{\sigma^2}{2^n})$ which is roughly $O(\frac{L^2q^2}{2^n})$. Here, a distinguisher can make at most q queries having at most σ many blocks with L as the maximum block size. The original bound for OMAC was roughly $\frac{5L^2q^2}{2^n}$ shown in FSE-03 and the next improved bound was $\frac{4\sigma^2}{2^n}$ shown in Indocrypt-03. In this paper we have provided an improved bound (a similar form as provided for others) for OMAC and the bound we show is roughly $\frac{4q\sigma}{2^n} = O(\frac{Lq^2}{2^n})$.

1 Introduction

CBC or Cipher-Block-Chaining [2] is an way to obtain an pseudo random function of PRF given an underlying block cipher such as AES [6] which is usually modeled as Pseudo random permutation or PRP. There are different variants of CBC constructions [4, 7, 10]. Among all these constructions, OMAC [7] or One-Key MAC is the most widely used MAC or PRF. This is mainly because of the key-size (a single key is sufficient). It is also efficient when we have sequential invocations of block-ciphers. All the CBC constructions are sequential and hence OMAC is one of the best choice among the class. Besides the CBC family there are other constructions of PRF such as PMAC [5] which is parallelizable and DAG-based PRF [9, 15].

Recently there are some results on finding improved security analysis on some of the above constructions. The security analysis means for PRF-security analysis. Intuitively, the advantage of a distinguisher $\mathcal{A}$ for a construction $\mathcal{D}$ is the success probability to distinguish $\mathcal{D}$ with the ideal random function (which responses randomly and uniformly from the output space). We denote the advantage by $\mathbf{Adv}^{\mathrm{prf}}_{\mathcal{D}}(\mathcal{A})$. A PRF-construction

is secure if for any distinguisher $\mathcal{A}$, which is making at most q queries having at most σ many blocks with maximum block size L, the advantage $\mathbf{Adv}^{\mathrm{prf}}_{\mathcal{D}}(\mathcal{A})$ is small or negligible. We denote the maximum possible advantage by $\mathbf{Insec}^{\mathrm{prf}}_{\mathcal{D}}(q, \sigma, L)$ and call it by prf-insecurity. Thus, the main research in this direction is devoted to get a better bound for given a secure construction.

The first result was in Crypto-05 [1] where an improved security analysis of CBC (for fixed length or for arbitrary length prefix-free messages) had provided. They have shown that $\mathbf{Insec}^{\mathrm{prf}}_{\mathrm{CBC}}(q, \sigma, L) \leq \frac{12Lq^2}{2^n} + \frac{64L^4}{2^{2n}}$. The second term becomes negligible or in the order of the first term if maximum block size is small compare to 2^n. For example, if $L < 2^{n/3}$ then we have $\mathbf{Insec}^{\mathrm{prf}}_{\mathrm{CBC}}(q, \sigma, L) \leq \frac{20Lq^2}{2^n}$ [1]. After this work, the improved analysis for other constructions have got attentions by the researchers. In [11], improved bound for XCBC, TMAC and PAMC have been provided. Again, their bounds of the prf-insecurity are of form $\mathrm{O}(\frac{Lq^2}{2^n}) + \mathrm{O}(\frac{L^4q^2}{2^{2n}})$. In [12] an improved bound for PAMC was shown and the bound was $\mathrm{O}(\frac{q\sigma}{2^n})$. In that paper [12], it was mentioned that this form of bound is truly improved bound. The original bounds are of the form $\mathrm{O}(\frac{\sigma^2}{2^n})$ and these can be much better than the new bound $\mathrm{O}(\frac{Lq^2}{2^n}) + \mathrm{O}(\frac{L^4q^2}{2^{2n}})$ (if the maximum block size becomes significant). This problem is not present in case of the bound of the form $\mathrm{O}(\frac{q\sigma}{2^n})$.

The above research are motivating to obtain an improved bound for OMAC. It is more likely to obtain this for OMAC as most of the others constructions from CBC family have got improved bounds. Only difficulty in the case of OMAC is the presence of a fixed input $\mathbf{0}$ and a single PRP is used trough out the constructions. In this context, we would like to make a note that improved analysis of TMAC and XCBC are mainly based on the presence of a second key which is being used just before the getting final output. In case of OMAC, we use different approach than the above. But some of ideas are very similar from the ideas provided in [1]. We mainly use the counting approach and the counting is based on finding solutions of some matrix-equations. In this paper, we have provided a prf-insecurity bound for OMAC as

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

where $N = 2^n$. Thus, we can write the bound as

1. $\frac{4q\sigma}{N} + \frac{8q(q-1)L^4}{N^2}$.
2. $\frac{10q\sigma}{N}$ if $L < N^{1/3}$.

Very recently a generalization of our approach has been provided [14]. In that paper, the improved security analysis of a wide class termed as affine domain extension is given. This class includes many other constructions such as DAG-based PRF.

Organization of the paper We first provide the definition of PRF and the measurement of PRF-insecurity in Section 2. In same section we state an important and useful theorem called as strong interpolation theorem. Then in Section 3 we provide the definition of OMAC with known security analysis of it. Then we provide our improved security analysis in Section 4. Finally we conclude with possible future work.

2 Pseudo random function and measurement of Insecurity

Random function. Random function is one of the common example of a random variable in cryptography. We first note that the random function defined in this paper is not same as what is defined classically. Like random variable, random function is a general object and uniform random function (which is classical random function) is actually a special random function which has uniform distribution on some set (similar to the uniform random variable). We denote $\mathrm{Func}(A, B)$ for the set of all functions from A to B and $\mathrm{Perm}(A)$ is the set of all permutations on A.

Definition 1. *A random function* $\mathtt{F}$ *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,* $\mathtt{F}$ *is a random permutation if* $\Pr[\mathtt{F} \in \mathrm{Perm}(A)] = 1$*.*

An **Uniform random function** or URF (the classical random function) is the uniform random variable on $\mathrm{Func}(A, B)$ for some finite sets A and B. That is, $\Pr[\mathtt{F} = f] = \frac{1}{|B|^{|A|}}$. Similarly we define **uniform random permutation** or URP (the classical random permutation) on A as the uniform random variable on $\mathrm{Perm}(A) \subset \mathrm{Func}(A, A)$. Given q distinct elements $x_1, \cdots, x_q \in A$ we can compute the joint distribution of $\mathtt{F}(x_1, \cdots, x_q) := (\mathtt{F}(x_1), \cdots, \mathtt{F}(x_q))$ where $\mathtt{F}$ is either uniform random function or uniform random permutation on A. We denote $\mathbf{P}(a, b) := a(a-1)\cdots(a-b+1)$ for two integers $0 < b \leq a$. We also define $\mathbf{P}(a, 0)$ as 1. The following result is based on simple counting of functions.

Proposition 1. Interpolation probability for URF or URP
Let $x_1, \cdots, x_q$ be q distinct elements. If $\mathtt{F}$ is an uniform random function then we have

$$\Pr[\mathtt{F}(x_1, \cdots, x_q) = (y_1, \cdots, y_q)] = \frac{1}{|A|^q}.$$

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

Random function based on domain extension. A domain extension $\mathcal{D}$ is a mapping from $\mathrm{Func}(A, B)$ to $\mathrm{Func}(A, B)$ with $A \subset \widetilde{A}$. Now, any random function $\mathtt{F}$ on $\mathrm{Func}(\widetilde{A}, B)$ induces a random function $\mathcal{D}(\mathtt{F}) := \mathcal{D}^{\mathtt{F}}$ on $\mathrm{Func}(\widetilde{A}, B)$. In this paper we study $\mathrm{OMAC}^{\mathtt{F}}$ where the *underlying random function* $\mathtt{F}$ is an uniform random permutation. Thus, we have $A = B = \{0,1\}^n$ and $\widetilde{A} = \{0,1\}^{\leq nL}$ (sice we consider the distinguisher making whose block size is at most L) and for any $M \in \{0,1\}^*$ we define the number of blocks of M as $\lceil \frac{|M|}{n} \rceil := ||M||$.

Definition 2. Advantage and PRF-Insecurity
A distinguisher $\mathcal{A}$ is nothing but an oracle algorithm. It can have use random coin $\mathtt{R}$. Given a distinguisher $\mathcal{A}_{\mathtt{R}}$ (a distinguisher $\mathcal{A}$ with random coin $\mathtt{R}$), the advantage of $\mathcal{A}_{\mathtt{R}}$ between two random functions $\mathtt{F}$ and $\mathtt{G}$ is defined as

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

Let $\mathtt{G}$ be an uniform random function from $\{0,1\}^{\leq nL}$ to $\{0,1\}^n$. Then for (q, σ, L) we define,

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

where maximum is taken over all distinguishers making exactly q queries having altogether at most σ many blocks with the maximum block size at most L.

Strong interpolation Theorem

Definition 3. *A q-tuple message $\mathbf{M} = (M_1, \cdots, M_q) \in C^q$ is called* **block-wise distinct** *if all M_i's are distinct where $M_i \in C$.*

Now we state our useful theorem which has actually been proven in [15]. This is a general version of a theorem stated in [3]. Thus we skip the proof detail.

Theorem 1. Strong Interpolation Theorem
Suppose for any block-wise distinct $\mathbf{x} \in (\{0,1\}^{\leq L})^q$*, block-wise distinct* $\mathbf{y} \in (\{0,1\}^n)^q$ *and* ε *(depending on* N, q, σ *and* L*) we have*

$$\Pr[\mathtt{F}(\mathbf{x}) = \mathbf{y}] \geq \frac{(1-\varepsilon)}{N^q}.$$

Then we have $\mathbf{Insec}^{\mathrm{prf}}_{\mathtt{F}}(q, \sigma, L) \leq \varepsilon + \frac{q(q-1)}{2N}$.

Thus the computation of interpolation probability $\Pr[\mathtt{F}(\mathbf{x}) = \mathbf{y}]$ is important. Later we define OMAC construction and we compute the interpolation probability for it. For uniform random function $\mathtt{G}$, we have already computed the interpolation probability which is $\Pr[\mathtt{G}(\mathbf{x}) = \mathbf{y}] = \frac{1}{N^q}$.

3 One-Key MAC or OMAC

3.1 Definition of OMAC construction

In this paper, we identify $\mathbb{F}_{2^n}$ (the Galois field of size 2^n) and $\{0,1\}^n$. We denote $\mathbf{0}$ and $\mathbf{1}$ for the additive and multiplicative identity respectively. Let $\pi \in \mathrm{Perm}(\mathbb{F}_{2^n})$. Then we can define $\pi^+ : \mathbb{F}^+_{2^n} \to \mathbb{F}_{2^n}$ as

$$\pi^+(m_1, \cdots, m_\ell) = \pi(\cdots(\pi(x_1) + x_2)\cdots + m_\ell).$$

The above function is also known as CBC function. Now we define OMAC function for arbitrary length. So we need to define a padding rule. Given a message $M \in \{0,1\}^*$, we define $\mathbf{pad}(M) = \overline{M} \in (\{0,1\}^n)^+$ as

$$\left.\begin{array}{rl} \overline{M} = M^* & \text{if } n \nmid |M| \\ = M & \text{otherwise} \end{array}\right\}$$

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 also define

$$\left.\begin{array}{rl} \delta_M = 1 & \text{if } n \nmid |M| \\ = 0 & \text{if } n \mid |M| \end{array}\right\}$$

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

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

where $\overline{M} = (m_1, \cdots, m_\ell) \in \mathbb{F}^\ell_{2^n}, \delta = \delta_M \in \{0,1\}$ 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 constant [7]).

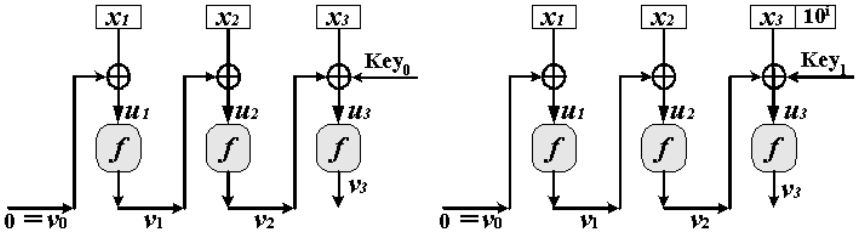

Fig. 1. OMAC: $\text{Key}_i = c_i.f(\mathbf{0})$. Here c_i's are distinct non-$\mathbf{0}$ and non-$\mathbf{1}$ constant such that $c_0 + c_1 \neq 1$. The function f is the underlying block cipher and v_3 is the final output of OMAC.

3.2 Known Security analysis of OMAC

In [7], the OMAC is proposed and there it had been shown that

$$\mathbf{Insec}^{\text{prf}}_{\text{OMAC}}(q, \sigma, L) \leq \frac{(5L^2+1)q^2+1}{2^n}.$$

Later, in [8], the bound was improved to $\frac{4\sigma^2+1}{2^n}$.

4 Improved security analysis of OMAC

We can define the OMAC function in the following equivalent way for $\ell \geq 2$.

1. $u_0 = 0$, $v_0 = \pi(u_0)$.
2. $u_1 = m_1$ and $v_1 = \pi(u_1)$.
3. $u_i = v_{i-1} + m_i$, $v_i = \pi(u_i)$ for $2 \leq i \leq \ell - 1$.
4. $u_\ell = v_{\ell-1} + c_\delta \cdot v_0 + m_\ell$ and $v_\ell = \pi(u_\ell)$.
5. $\text{OMAC}^\pi(M) = v_\ell = \text{OMAC}^\pi(M)$.

For $\ell = 1$, we have

1. $u_0 = 0$, $v_0 = \pi(u_0)$.
2. $u_1 = c_\delta \cdot v_0 + m_1$ and $v_1 = \pi(u_1)$.
3. $\text{OMAC}^\pi(M) = v_1$.

Definition 4. *The values u_i's (including $u_0 = \mathbf{0}$) are known as* **intermediate input**, $0 \leq i \leq \ell$ *and u_ℓ is known as the* **final input**. *Similarly, v_i's are known as* **intermediate output** *and v_ℓ is known as the* **final output**, $0 \leq i \leq \ell$.

We denote $\mathbf{v}^{M,\pi} = (v_0, v_1, \cdots, v_\ell)$ and $\mathbf{u}^{M,\pi} = (u_0, u_1, \cdots, u_\ell)$ for the intermediate output vector and intermediate input vector respectively. Now we represent the above relation between intermediate inputs and intermediate outputs by a matrix known as **coefficient matrix** $\mathbf{A}^M_{(\ell+1)\times(\ell+2)}$ as $\mathbf{A}^M \cdot \overline{\mathbf{v}}^{M,\pi} = \mathbf{u}^{M,\pi}$ 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 $\ell \geq 2$:

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

We can combine these linear relationship for two distinct messages M, M' also. Since the first row (corresponding to the intermediate input $\mathbf{0}$) is always zero, we ignore the first row for the second message. For example, if $\overline{M} = (m_1, m_2, m_3)$ and $\overline{M'} = (m'_1, m'_2)$ then the coefficient matrix for the pair $\mathbf{M} = (M, M')$ is

$$\mathbf{A}^{\mathbf{M}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_2 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m_3 & c_\delta & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m'_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ m'_2 & c_{\delta'} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \end{pmatrix}.$$

Similarly, we define $\mathbf{u} = \mathbf{u}^{\mathbf{M},\pi} = (u_0, u_1, \cdots, u_\ell, u'_1, \cdots, u'_{\ell'}) = (u_0, u_1, \cdots, u_{\ell+\ell'})$ and $\mathbf{v} = \mathbf{v}^{\mathbf{M},\pi} = (v_0, v_1, \cdots, v_\ell, v'_1, \cdots, v'_{\ell'}) = (v_0, v_1, \cdots, v_{\ell+\ell'})$ and we have the following relationship

$$\mathbf{A}^{\mathbf{M}}_{t\times(t+1)} \cdot \mathbf{v}^{\mathbf{M},\pi} = \mathbf{u}^{\mathbf{M},\pi} \quad \text{and} \quad \pi(\mathbf{u}) = \mathbf{v} \tag{1}$$

where $t = \ell + \ell' + 1$. In general, for the tuple of q distinct messages $\mathbf{M} = (M_1, \cdots, M_q)$ we have coefficient matrix $\mathbf{A}^{\mathbf{M}}_{t\times(t+1)}$ and the intermediate input and output vectors as $\mathbf{u}^{\mathbf{M},\pi}_{t\times 1}$ and $\mathbf{v}^{\mathbf{M},\pi}_{t\times 1}$ where $t = \ell_1 + \cdots \ell_q + 1$ and $\overline{M_i} \in \mathbb{F}_{2^n}^{\ell_i}$, $1 \leq i \leq q$. We also have the relationship as in Equation 1.

Interpolation of OMAC. Let $J = \{i_1, \cdots, i_s\} \subset [0, t]$ be a subset of indices such that $i_1 < \cdots < i_s$ and $\mathbf{x} = (x_0, \cdots, x_t)$ be a $(t+1)$-tuple. Now we define a sub-tuple $\mathbf{x}_J = (x_{i_1}, \cdots, x_{i_s})$. Let $F = \{\ell_1, \ell_1 + \ell_2, \cdots, \sum_{i=1}^{q} \ell_i = t-1\}$ be a subset of indices known as the set of the final input indices. Now it is easy to check that

$$(\mathrm{OMAC}(M_1), \cdots, \mathrm{OMAC}(M_q)) = \mathbf{v}_F^{\mathbf{M},\pi}$$

where $\mathbf{A}^{\mathbf{M}}_{t\times(t+1)} \cdot \mathbf{v}^{\mathbf{M},\pi} = \mathbf{u}^{\mathbf{M},\pi}$ and $\pi(\mathbf{u}) = \mathbf{v}$.

One can easily observe that for each $t_j = \sum_{i=1}^{j} \ell_j$, $\mathbf{A}_{\cdot t_j} = \mathbf{0}^t$ where $\mathbf{A}^{\mathbf{M}} = (\alpha^{\mathbf{M}}\ \mathbf{A}^{M}_{\cdot 0}\ \cdots \mathbf{A}_{\cdot t-1})$. That is the final outputs have no effect on the intermediate inputs. We rewrite the Equation 1 as

$$\mathbf{u} = \mathbf{A}' \cdot \mathbf{v}_I \quad \text{and} \quad \pi(\mathbf{u}_I) = \mathbf{v}_I \tag{2}$$

where $\mathbf{A}'$ is the matrix obtained after removing the columns $\mathbf{A}_{\cdot t_j}$, $1 \le j \le q$ from the coefficient matrix $\mathbf{A}$ and $I = [0, t-1] \setminus F$.

Definition 5. *Given $\pi \in \mathrm{Perm}(\mathbb{F}_{2^n})$ 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 induced equivalence relation if there is a permutation* π *such that* $\mathfrak{R}^\pi = \mathfrak{R}$*.*

Note that, any equivalence relation may not be an induced equivalence relation. A tuple $(i_1, \cdots, i_s)$ 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 6. *Suppose $(i_1, \cdots, i_s)$ is the tuple of representatives of $\mathfrak{R}$ on $[0, t-1]$. Now we define a new $t \times (s+1)$ matrix $\mathbf{B} := \mathbf{A}^{\mathfrak{R}} = (\alpha^{\mathbf{M}}\ \mathbf{B}_{\cdot 1} \cdots \mathbf{B}_{\cdot s})$ where $\mathbf{B}_{\cdot j} = \sum_{i \in [i_j]} \mathbf{A}_{\cdot i}$. If $\mathbf{v}$ satisfies $\mathfrak{R}$, we consider a new s-vector $(w_1, \cdots, w_s) = \mathbf{w} = \mathbf{v}^{\mathfrak{R}}$ 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 7. *We define* **rank** *of a permutation π (also rank of the induced relation $\mathfrak{R}^\pi$) as the rank of the following 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[1] 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 $(N-s)!$ many permutations π (check it!) such that $\mathbf{v}^{\mathbf{M},\pi} = \mathbf{v}$. Thus we have the following result.

Proposition 2. *Given a relation $\mathfrak{R}$ of rank r and of size s, there are at most $N! \times \frac{1}{P(N-s+r,r)}$ many permutations π such that $\mathfrak{R}^\pi = \mathfrak{R}$.*

Proposition 3. *The number of relations of rank r is at most $\binom{t}{2}^r$.*

Proof of Proposition 3 is given in [14]. In [1], it has been studied in terms of graphs for CBC constructions. A very similar analysis will work here. Now from the above propositions we can prove the following corollary. A similar corollary is also given for CBC in [1]. But here we have a modified bound which is obtained by applying inequality carefully.

Corollary 1. *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 $N! \times \frac{(\ell+\ell')^4}{N^2}$.*

An element i is called single in $\mathfrak{R}$ if $[i] = \{i\}$. A set is called single if every element is single. Now it is easy to see that for any distinct $M \neq M'$ and the induced relation $\mathfrak{R}_0$ of rank zero (there are exactly one such) satisfies the following property : both ℓ and $\ell+\ell'$ are single elements in $\mathfrak{R}_0$. In fact, one can write down the relation $\mathfrak{R}_0$.

Proposition 4. *Let $\overline{M} = (m_1, \cdots, m_\ell)$ and $\overline{M'} == (m'_1, \cdots, m'_{\ell'})$. 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 $(m_1, \cdots, m_{\ell-1})$ and $(m'_1, \cdots, m'_{\ell'-1})$ have exactly $p \geq 1$ common prefix blocks then $(1, \ell+1), \cdots (p, \ell+p) \in \mathfrak{R}_0$.*

Now we study the number of valid relations of rank one such that $F = \{\ell, \ell+\ell'\}$ is not single. We do it by considering 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 $\mathbf{B} = \mathbf{A}^{\mathfrak{R}}$. But, the second element in $\mathbf{B}_{\ell+\ell'} - \mathbf{B}_{i'}$ is not zero (either

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

$c_{\delta'} - c_\delta$ or $c_{\delta'} - 1$ or $c_{\delta'}$) where as that of $\mathbf{B}_i - \mathbf{B}_j$ is zero. Thus, the rank should be more than one. Hence only possible valid relation of rank one such that F is not single is that one with the basis (i, j) where either i or $j \in F$. Thus, the number of such relations is at most $2(\ell + \ell')$.

Case-A : $\delta_M = \delta_{M'}$

Suppose we have $(\ell + \ell', i') \in \mathfrak{R}$ where $i \notin F$. Then by similar reason, the basis should contain the pair whose one element is from F. So there are at most $2(\ell + \ell')$ many 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 [1], we know that there are at most $d(|\ell - \ell'|)$ many relations of rank one containing the pair $(\ell + 1, \ell' + 1)$. Here, the function $d(m)$ means the number of factors of m. Thus, the total number of relations of rank one such that F is not single is at most $3(\ell + \ell')$.

Proposition 5. *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 $\overline{M} = (m_1, \cdots, m_\ell)$, $\overline{M'} = (m'_1, \cdots, m'_{\ell'})$, $\delta = \delta_M$ 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

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

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

$$\mathrm{New}[M_1, \cdots, M_q] = \cap_{1 \leq i < j \leq q} \mathrm{New}[M_i, M_j].$$

We denote $\mathtt{Bad}_1 = \overline{\mathrm{New}[M_1, \cdots, M_q]}$ the complement of the event New. From the above discussion and by using Corollary 1 we have the following results.

Proposition 6. *If $\mathtt{F}$ is an 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,*

$$\Pr[\overline{\mathrm{New}[M, M']}] \leq \frac{3(\ell + \ell')}{N} + \frac{(\ell + \ell')^4}{N^2}.$$

Corollary 2. *For q distinct messages $M_1, \cdots M_q$ with $\overline{M_i} \in \mathbb{F}_{2^n}^{\ell_i}$ we have,*

$$\Pr[\mathtt{Bad}_1] = \Pr[\overline{\mathrm{New}[M_1, \cdots, M_q]}] \leq \frac{3(q-1)\sigma}{N} + \frac{8q(q-1)L^4}{N^2}$$

where $L = \mathbf{max}_{1 \leq i \leq q}\ \ell_i$. Moreover, if $L \leq N^{1/3}$ we have $\Pr[\mathtt{Bad}_1] \leq \frac{9q\sigma}{N}$.

Let $M_1, \cdots, M_q$ be q distinct messages such that $\overline{M_i} \in \mathbb{F}_{2^n}^{\ell_i}$ and $\sum_{j=1}^{q} \ell_j = \sigma$. Let $z_1, \cdots, z_q$ be q distinct elements from $\mathbb{F}_{2^n}$. We define an event $\mathtt{Bad}_2$ as $v_j^{M_k} = z_i$ for some $1 \leq i \leq q$ and $1 \leq j < \ell_k$, $1 \leq k \leq q$. Thus, the set of all intermediate outputs are not disjoint from the set $\{z_1, \cdots, z_q\}$. Finally we define $\mathtt{Bad} = \mathtt{Bad}_1 \cup \mathtt{Bad}_2$.

Proposition 7.

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

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

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

Proposition 8.

$$\Pr[\mathrm{OMAC}^{\mathtt{F}}(M_1) = z_1, \cdots, \mathrm{OMAC}^{\mathtt{F}}(M_q) = z_q | \overline{\mathtt{Bad}}] \geq \frac{1}{\mathbf{P}(N,q)}.$$

Proof. It is easy to see that for a fixed input vector $\mathbf{w}$ such that $\Pr[\Pi$ 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 $\mathbf{v}_I^{\mathtt{F}} = \mathbf{w}] \geq \frac{1}{\mathbf{P}(N,q)}$. □

Corollary 3. *Given any q distinct messages $M_1, \cdots, M_q$ and q distinct elements $z_1, \cdots, z_q \in \mathbb{F}_{2^n}$, we have,*

$$\Pr[\mathrm{OMAC}^{\mathtt{F}}(M_1) = z_1, \cdots, \mathrm{OMAC}^{\mathtt{F}}(M_q) = z_q] \geq \frac{1-\varepsilon}{\mathbf{P}(N,q)}$$

where $\varepsilon = \frac{4q\sigma}{N} + \frac{8q(q-1)L^4}{N^2} - \frac{q(q-1)}{2N}$.

Now by using Strong Interpolation theorem we can get our main result of the paper.

Theorem 2. (**Improved security bound for OMAC**)
For any distinguisher $\mathcal{A}$ making at most q queries having at most σ many blocks such that the maximum block size is at most L then the PRF-advantage of $\mathcal{A}$,

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

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

$$\mathbf{Adv}^{\mathrm{prf}}_{\mathrm{OMAC}}(\mathcal{A}) \leq \frac{10q\sigma}{N}.$$

5 Conclusion and future work

In this paper we have provided an improved PRF-insecurity bound which is roughly $\frac{4q\sigma}{2^n}$. 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 for generalized constructions. We also hope that this idea is useful to obtain improved and more appealing security analysis for other indistinguishability security notions such as online cipher [13], PRP in modes of operation etcetera.

References

1. M. Bellare, K. Pietrzak and P. Rogaway. Improved Security Analysis for CBC MACs. Advances in Cryptology - CRYPTO 2005. Lecture Notes in Computer Science, Volume **3621**, pp 527-545.

2. M. Bellare, J. Killan and P. Rogaway. The security of the cipher block chanining Message Authentication Code. Advances in Cryptology - CRYPTO 1994. Lecture Notes in Computer Science, Volume **839**, pp 341-358.

3. Daniel J. Bernstein. A short proof of the unpredictability of cipher block chaining (2005). URL: http://cr.yp.to/papers.html#easycbc. ID 24120a1f8b92722b5e1 5fbb6a86521a0.

4. J. Black and P. Rogaway. CBC MACs for arbitrary length messages. Advances in Cryptology - CRYPTO 2000. Lecture Notes in Computer Science, Volume **1880**, pp 197-215.

5. J. Black and P. Rogaway. A Block-Cipher Mode of Operations for Parallelizable Message Authentication. Advances in Cryptology - Eurocrypt 2002. Lecture Notes in Computer Science, Volume **2332**, pp 384-397.

6. Joan Daemen and Vincent Rijmen The Design of Rijndael: AES The Advanced Encryption Standard. Springer 2002. http://csrc.nist.gov/CryptoToolkit/aes/rijndael/Rijndael-ammended.pdf

7. T. Iwata and K. Kurosawa. OMAC : One-Key CBC MAC. Fast Software Encryption, 10th International Workshop, FSE 2003. Lecture Notes in Computer Science, Volume **2887**, pp 129-153.

8. T. Iwata and K. Kurosawa. Stronger Security Bounds for OMAC, TMAC, and XCBC. Progress in Cryptology - INDOCRYPT 2003. Lecture Notes in Computer Science, Volume **2904**, pp 402-415.

9. C. S. Jutla. PRF Domain Extension using DAG. Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006. Lecture Notes in Computer Science, Volume **3876** pp 561-580.

10. K. Kurosawa and T. 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, Volume **2612**, pp 33-49.

11. K. Minematsu and T. Matsushima Improved Security Bounds for PMAC, TMAC, and XCBC. Fast Software Encryption 2007.

12. M. Nandi and A. Mandal Improved Security Analysis of PMAC. Available in http://eprint.iacr.org/2007/031

13. M. Nandi A Simple Security Analysis of Hash-CBC and a New Efficient One-Key Online Cipher. Available in http://eprint.iacr.org/2007/031

14. M. Nandi An improved security analysis of Affine Domain Extension. preprint version.

15. M. Nandi A Simple and Unified Method of Proving Indistinguishability. Progress in Cryptology - INDOCRYPT 2006. Lecture Notes in Computer Science, Volume **4329**, pp 317-334.