Improved Security Analysis of PMAC

Avradip Mandal and Mridul Nandi
University of Waterloo, Canada

February 1, 2007

Abstract

In this paper we provide a simple and improved security analysis of **PMAC** [6], a Parallelizable MAC (Message Authentication Code) defined over arbitrary messages. A similar kind of result is shown by Bellare, Pietrzak and Rogaway [2] in Crypto-2005, where they have provided an improved bound for **CBC** MACs [3, 9, 11]. Our analysis idea is much more simpler to understand and is borrowed from [4, 13]. It shows that the advantage for any distinguishing attack for **PMAC** based on a random function is bounded by $\mathrm{O}(\frac{\sigma q}{2^n})$, where σ is the total number of blocks in all q queries made by the attacker. In the original paper [6], the bound is $\mathrm{O}(\frac{\sigma^2}{2^n})$.

Keywords : MAC, PMAC, Distinguishing attack, random function, pseudo random function.

1 Introduction

PMAC is a parallelizable Message Authentication Code unlike Cipher Block Chaining or CBC MACs [3] which are sequential based constructions. There are many literatures on CBC-MACs improving efficiency in performance as well as in key size. Some of them are XCBC [5], TMAC [11], OMAC [9]. Recently, Jutla [10] and Nandi [13] analyzed a wide class of tree based constructions, some of them can be implemented in parallel. All these constructions are based on pseudo random function or pseudo random permutation [12]. AES [7] is a believed to be a candidate of pseudo random permutation as well as pseudo random function. There are other constructions of MAC based on different universal hash families [8, 14, 15]. Now we provide definition of MAC and its security notions.

1.1 Message Authentication Codes (MAC) and its Security Notions

Definition of MAC

Message Authentication Code or MAC is a secret key version of digital signature. It is used as an authentication of a message. A MAC is a family of functions $\{F_k\}_{k \in \mathcal{K}}$ where $F_k : \mathcal{M} \to T$, $\mathcal{M}$ is the message space, T is the set of all tag space and $k \in \mathcal{K}$ is a secret key chosen uniformly from a key space. If $t = F_k(M)$ then t is called the *tag* of the message M. In this paper, we consider $T = \{0,1\}^n$ with a group addition $+$ and the identity element $\mathbf{0}$ and $\mathcal{M} = \{0,1\}^{\leq L} \triangleq \cup_{i \leq L} \{0,1\}^i$ for a sufficiently large integer L and a fixed integer n. A reasonable choice of parameters are $n = 128$ and $L = 2^{64}$.

Security Notions of MAC

There are two popular security notions for Message Authentication Code. Those are secure against *distinguishing attack* and secure against *forgery attack*. The distinguishing attack is a weaker attack than forgery. In other words, if a construction is secure against distinguishing attack then it is also secure against forgery attack with at least same security level. Thus, we mainly analyze the distinguishing attack security for PMAC.

1. Distinguishing Attack : Let Adversary $\mathcal{A}^O$ be an oracle algorithm where

- $O = F_k$, chosen uniformly from $\mathcal{F} = \{F_k : \mathcal{M} \to T;\ k \in \mathcal{K}\}$ (k is uniform on $\mathcal{K}$) or
- $O = F$, chosen uniformly from $\mathbf{Func}(\mathcal{M}, T) \stackrel{\Delta}{=} \{F;\quad F : \mathcal{M} \to T\}$ (or **Func** only).

The adversary can make at most q queries adaptively consisting of at most σ many blocks and runs in time at most t. Finally, it returns either 1 or 0. The *advantage for distinguishing attack* is computed as follows :

$$\mathbf{Adv}_{\mathcal{F},\ \mathbf{Func}}(\mathcal{A}) \stackrel{\Delta}{=} \big|\ \mathbf{Pr}[\mathcal{A}^{\mathcal{F}} = 1] - \mathbf{Pr}[\mathcal{A}^{\mathbf{Func}} = 1]\ \big|.$$

$$\mathbf{Adv}_{\mathcal{F},\mathbf{Func}}(q, \sigma, t) \stackrel{\Delta}{=} \mathbf{max}_{\mathcal{A}}\ \mathbf{Adv}_{\mathcal{F},\ \mathbf{Func}}(\mathcal{A} : q, \sigma, t)$$

where the maximum is taken over all distinguisher $\mathcal{A}$ with runtime at most t making at most q queries consisting of at most σ many blocks. For simplicity, we also denote $\mathbf{Adv}_{\mathcal{F}}(\mathcal{A})$ and $\mathbf{Adv}_{\mathcal{F}}(q, \sigma, t)$ in the places of $\mathbf{Adv}_{\mathcal{F},\ \mathbf{Func}}(\mathcal{A})$ and $\mathbf{Adv}_{\mathcal{F},\mathbf{Func}}(q, \sigma, t)$ respectively.

The definition of *block* is given later when we define PMAC. Intuitively, it is the number of n-bits in a padded message. A *random function* is a probability distribution on $\mathbf{Func}(\mathcal{M}, T)$. If the distribution is uniform then we say that it is an *uniform random function*. Note that, the uniform distribution on $\mathcal{K}$ induces a probability distribution on **Func**. Intuitively, if the advantage is high then the attacker $\mathcal{A}$ can distinguish the uniform random function and the random function $\mathcal{F}$ with high probability. If it is negligible, we sometimes say that the family $\mathcal{F}$ is a pseudo random function family.

2. MAC-forgery : In case of a MAC-forgery attack, an attacker makes successive queries M_i's for the oracle F_k (where k is secret and chosen uniformly from $\mathcal{K}$) and obtains responses $F_k(M_i)$'s. Let $(M_1, t_1 = F_k(M_1)), \cdots, (M_q, t_q = F_k(M_q))$ be all *query-responses*. If attacker can return a pair (M, t) such that $(M, t) \neq (M_i, t_i)$ for all i and t is a valid tag (i.e., $t = F_k(M)$) then we say that the attacker forges successfully. The probability for forging successfully a message-tag pair is the advantage for MAC-forgery attack.

If one can forge a message (say (M, t)) using this forgery attacker one can make a distinguishing attack (same as the forgery attacker except at the end it will submit the query M and checks whether the response is t or not). Thus a forgery attacker is much stronger, or equivalently secure against distinguishing attack is more stronger.

1.2 Known Results and Our Results

In [6], authors have shown that $\mathbf{Adv}_{\mathbf{PMAC}}(q, \sigma, t) \leq \frac{2(\sigma+1)^2}{N}$. In this paper we show that the advantage $\mathbf{Adv}_{\mathbf{PMAC}}(q, \sigma, t) \leq \frac{11\sigma(q-1)}{2N}$. When an attacker is making uniform message block queries the bound can be written as $\frac{11\ell q(q-1)}{2N}$ which is similar to the bound given in Crypto-05 [2] for CBC MACs. Note that, when an attacker has restriction on the total number of message blocks σ, then the upper bound of advantage is more if q is as large as possible. q can be at most σ, (that is, all message queries are single block length query) and in this case order of our bound is same as the order of original bound. The same thing we can say about the improved result in CBC-MACs in Crypto-05 [2]. The main results of this papers are the following theorems.

Theorem. Let $M^1, \cdots, M^q$ are distinct messages from $\mathcal{M}$ and $y^1, \cdots, y^q \in T$ (not necessarily distinct) then $\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q] \geq \frac{1-\epsilon}{N^q} = (1-\epsilon) \times \Pr[F(M^1) = y^1, \cdots, F(M^q) = y^q]$ where $\epsilon = \frac{11(q-1)\sigma}{2N}$ and F is an uniform random function on $\mathbf{Func}(\{0,1\}^{\leq L}, \{0,1\}^n)$.

Theorem. $\mathbf{Adv}_{\mathbf{PMAC}}(q, \sigma, t) \leq \frac{11(q-1)\sigma}{2N}$.

Organization of this paper

We have explained the MAC and it's security notions in this Section. We describe a slightly modified definition of PMAC in Section 2. Then we characterize a wide class of distinguishers in Section 3. Next, we give a detail security analysis of PMAC in Section 4. Finally we conclude.

2 Definition of PMAC

In this section we will describe PMAC. Later we will analyze the security of it. Before we define we would like to make the following important comments to the reader. The definition of PMAC we provide has a slight modification over the original definition. In the original definition, length of the message (possibly) with 10^s (for a suitably chosen s) is appended at the end of the message (this is called the *padding* and the message after padding is called padded message). In this paper, we consider a different (in fact, a simpler) padding which does not pad the length of the message. All other rules of padding and the definitions of PMAC are exactly same as the original one. There are some advantages in considering the modified definition.

1. First of all, it is more efficient as we may need one less invocation of underlying pseudo random function.

2. We do not have to keep the length of the messages. It reduces the internal memory requirement.

3. Finally, (and most importantly) it can be defined for any arbitrary messages. So, our definition of PMAC is defined over $\{0,1\}^*$. But for simplicity of our security analysis we will take $\{0,1\}^{\leq L}$ as a domain where L can be any large integer. Note that in the original definition L should be less than $n2^n$. Definitely, this choice of L is reasonably large enough in the current time. But it is always advantageous if we know that the same construction can be used to any arbitrary messages.

Let $f : \{0,1\}^n \to \{0,1\}^n$ be a random function for some positive integer n. We write $N = 2^n$. Let $\mathcal{M} = \{0,1\}^{\leq L}$ for a sufficiently large integer L and $T = \{0,1\}^n$. Now we define a random function, known as **PMAC** function, $\mathrm{P}_f : \mathcal{M} \to T$ based on f. We first define a *padding rule* which makes message size a multiple of n if it is not so.

$$\left.\begin{array}{rcll} \mathbf{pad}(M) & = & M \parallel 10^s & \text{if } n \not| \, |M| \\ & = & M & \text{otherwise} \end{array}\right\} \quad (1)$$

where $s = n\lceil(|M|+1)/n\rceil - |M| - 1$. If $n \not| \, |M|$ then $|\mathbf{pad}(M)| = |M| + s + 1 = n\lceil(|M|+1)/n\rceil$, which is the smallest multiple of n strictly bigger than the size of $|M|$. Suppose for $M_1 \neq M_2$, $\mathbf{pad}(M_1) = \mathbf{pad}(M_2)$, then exactly one of these has size multiple of n (say $n \mid |M_2|$ and $n \not| \, |M_1|$) and $M_2 = \mathbf{pad}(M_1) = M_1 \parallel 10^s$.

Algorithm PMAC : $Y = \mathrm{P}_f(M)$

step-1 Write $\mathbf{pad}(M) = x_1 \parallel \cdots \parallel x_\ell \parallel z$, where $\ell \geq 0$ and $|x_1| = \cdots |x_\ell| = |z| = n$. \\ We say these x_i's and z as *blocks*. If $\ell = 0$, then $\mathbf{pad}(M)$ is nothing but z. Thus, $\ell + 1$ is the total number of message blocks for $\mathbf{pad}(M)$.

step-2 Compute $w = f(0)$. \\ Since f is a random function and kept secret the value of $f(0)$ has some distribution and can be used as a part of the key of the algorithm.

step-3 Compute $v_i = x_i + c_i \cdot w, 1 \leq i \leq \ell$. \\ c_i's are some fixed distinct nonzero constants as given in [6]. For our security analysis, we only need that $c_i \neq 0$ and they are distinct. $(\{0,1\}^n, +, \cdot)$ is any Galois field $GF(2^n)$. One can think $+$ as $\oplus$ as it is the simplest operation in both hardware and software.

step-4 Compute $w_i = f(v_i), 1 \leq i \leq \ell$.

step-5 Compute $v = z + \Delta + \sum_{1 \leq i \leq \ell} w_i$, where $\Delta = c \cdot w$ if $|M|$ is multiple of n, otherwise we set $\Delta = 0$. \\ Again, c is a nonzero fixed constant which is different from $c_1, c_2, \cdots$, and it is given in [6].

step-6 Finally, $Y \stackrel{\Delta}{=} \mathrm{P}_f(M) = f(v)$.

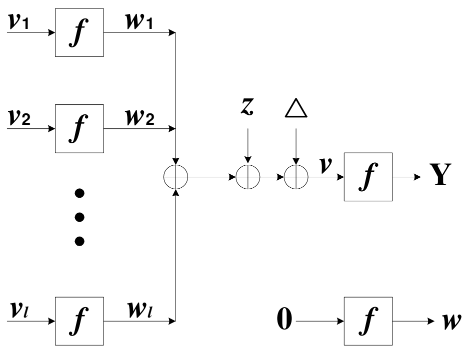

Figure 1: PMAC

3 Distinguishing Families of Functions or Random Functions

Suppose $\mathcal{A}$ distinguishes two random functions f and g which are probability distributions on $\mathbf{Func}(\mathcal{M}, T)$. The distinguisher $\mathcal{A}$ is an oracle algorithm and hence it can make several queries adaptively. The oracle can be either chosen from the distribution f or from the distribution g. Distinguisher is behaving as follows.

- First it chooses a random string r with some distribution (not necessarily uniform) on $\mathcal{R}$.
- Based on r it makes query $x_1 := x_1(r) \in \mathcal{M}$ and obtains $y_1 \in T$. Then it makes queries $x_2 = x_2(r, y_1) \in \mathcal{M}$ and obtains $y_2 \in T$ and so on.
- Based on all query-responses it outputs either 1 or 0.

Denote $\mathcal{A}_r$ as the distinguishing algorithm same as $\mathcal{A}$ after choosing the random string r. Thus, $\mathcal{A}_r$ is a deterministic algorithm.

$$\mathbf{Adv}_{f,g}(\mathcal{A}) = \Big| \sum_{r \in \mathcal{R}} (\Pr[\mathcal{A}_r^f = 1] - \Pr[\mathcal{A}_r^g = 1]) \times \Pr[r] \Big|$$

$$\leq \mathbf{max}_{r \in \mathcal{R}} \left| \Pr[\mathcal{A}_r^f = 1] - \Pr[\mathcal{A}_r^g = 1] \right| = \mathbf{Adv}_{f,g}(\mathcal{A}_{r^*}),$$

where the maximum takes place at $r = r^*$. So, now onwards we can assume that the distinguisher $\mathcal{A}$ is deterministic. We can also assume that all queries are distinct. This assumption is reasonable as if any attacker is making same query which has been asked before the response is determined with probability one for both oracles. Thus we can modify the attacker which skips the repetition query.

Any tuple $((M^1, y^1), \cdots, (M^q, y^q))$ is said to be a *transcript* of the attacker $\mathcal{A}$ if $M^1 = x_1(\cdot), M^2 = x_2(M^1, y^1), \cdots, M^q = x_q(M^1, y^1, \cdots, M^{q-1}, y^{q-1})$. Now we state a theorem which would be used to obtain an upper bound of the advantage. Different versions of the theorem have been proven in [4, 13].

Theorem 1. *Suppose* $\Pr[f(M^1) = y^1, \cdots f(M^q) = y^q] \geq (1 - \epsilon) \times \Pr[g(M^1) = y^1, \cdots g(M^q) = y^q]$ *for each distinct* $M^1, \cdots, M^q \in \mathcal{M}$ *and any* $y^1, \cdots, y^q \in T$. *Then for any attacker* $\mathcal{A}$ *making at most* q *queries has advantage* $\mathbf{Adv}_{f,g}(\mathcal{A}) \leq \epsilon$.

Proof. Let S_1 be the set of all tuples $((M^1, y^1), \cdots, (M^q, y^q))$ such that it is a transcript and $\mathcal{A}$ outputs 1. Note that the set S_1 does not depend on f and g. Only the probability distribution the transcript appears when $\mathcal{A}$ interacts with the oracle f or g depends on f or g respectively. Thus,

$$\mathbf{Adv}_{f,g}(\mathcal{A}) = | \sum_{((M^1,y^1),\cdots,(M^q,y^q)) \in S_1} \Pr[f(M^1) = y^1, \cdots, f(M^q) = y^q]$$

$$- \sum_{((M^1,y^1),\cdots,(M^q,y^q)) \in S_1} \Pr[g(M^1) = y^1, \cdots, g(M^q) = y^q] \; |$$

$$\leq \epsilon \times \sum_{((M^1,y^1),\cdots,(M^q,y^q)) \in S_1} \Pr[g(M^1) = y^1, \cdots, g(M^q) = y^q] \leq \epsilon.$$

The inequality holds due to the given condition. $\square$

4 Improved Security Analysis of PMAC

We are interested in computing the probability

$$\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q], \quad y^i \in \{0,1\}^n, M^i \text{ are distinct} .$$

The probability is computed under the probability distribution of f, an uniform random function, and it is known as **interpolation probability**. Denote $\mathbb{M} = \{M^1, \cdots, M^q\}$ and $\ell_j = \| \mathbf{pad}(M^j) \|$ (the number of message blocks), $1 \leq j \leq q$. For each $1 \leq j \leq q$, we denote all variables in the computation of $\mathrm{P}_f(M^j)$ with a superscript j, that is, $x_i^j, z^j, v_i^j, w_i^j, \Delta^j, v^j, Y^j$, $1 \leq i \leq \ell_j$. Among them, x_i^j and z^j (sometimes Δ^j when $|M^j|$ is not multiple of n) are not random variables and fixed. All other variables are random variables with a distribution induced from the distribution of uniform random function. Sometime we also write them as $w[f], v_i^j[f], w_i^j[f], v^j[f], \Delta^j[f], Y^j[f]$ to show the dependency with f. We call

- $0, v_i^j$ as *intermediate inputs* and v^j as a *final input*,
- w, w_i^j as *intermediate outputs* and Y^j as a *final output*.

Note that, intermediate and final inputs are really inputs of f while computing $\mathrm{P}_f(M^j)$ and intermediate and final outputs are outputs of f. We will show that for some small ϵ, the interpolation probability $\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q] \geq \frac{(1-\epsilon)}{N^q}$.

Definition 2. *An m-tuple $(a_1, a_2, \ldots, a_m)$ is new in an r-tuple $(b_1, b_2, \ldots, b_r)$ if for all $1 \leq i \leq m$ and $1 \leq j \leq r$ we have $a_i \neq b_j$ and a_i's are distinct. Note that m can be equal to one and in this case, we say that a_1 is new in $(b_1, b_2, \ldots, b_r)$.*

Let us denote the event D that all final inputs are distinct and different from all other intermediate inputs. More precisely, $(v^1, \cdots, v^q)$ is new in $(0, v_1^1, \cdots, v_{\ell_1}^1, v_1^2, \cdots, v_{\ell_q}^q)$. Now we prove that the interpolation probability conditioned on D is $1/N^q$. Intuitively, it is clear that the value of $(f(v^1), \cdots, f(v^q))$ follows uniform distribution condition on that v^j's are not occurred as an intermediate inputs which is assured by the event D. Now we have a more precise proof of the above statement.

Lemma 3. $\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q \mid D] = \frac{1}{N^q}$.

Proof. Let $\mathcal{F}_D$ denotes the set of all functions from $\mathcal{F}$ which satisfies the event D.

$$\mathcal{F}_D = \{f_0 \in \mathcal{F} : (v^1[f_0], \cdots, v^q[f_0]) \text{ is new in } (0, v_1^1[f_0], \cdots, v_{\ell_q}^q[f_0])\}.$$

Let $\mathcal{F}_{D_1} = \{f_0 \in \mathcal{F} : (v^1[f_0], \cdots, v^q[f_0]) \text{ is new in } (0, v_1^1[f_0], \cdots, v_{\ell_q}^q[f_0]) \wedge Y^j[f_0] = y^j, 1 \leq j \leq q\}$. Thus, $\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q \mid D] = |\mathcal{F}_{D_1}|/|\mathcal{F}_D|$. Now consider the mapping α from $\mathcal{F}_D$ to $\mathcal{F}_{D_1}$ as follows,

$$\left.\begin{array}{rcll} \alpha(f_0)(x) & = & f_0(x) & \text{if } x \neq v^j[f_0] \text{ for all } j \\ & = & y^j & \text{if } x = v^j[f_0] \text{ for some } j \end{array}\right\} \tag{2}$$

Now α is an N^q onto one mapping. That is, for every $f_1 \in \mathcal{F}_{D_1}$, there exists exactly N^q many f_0's such that $\alpha(f_0) = f_1$. Given f_1, f_0's are exactly same as f_1 except that it can take any N^q possible values on $v^j[f_1]$'s. This is well defined since the values of $f_1((v^j[f_1])$'s do not have any effect on the whole computations of $\mathrm{P}_{f_1}(M^j)$'s except the final output. Thus, $|\mathcal{F}_D| = N^q|\mathcal{F}_{D_1}|$ and hence, $\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q \mid D] = \frac{1}{N^q}$. $\square$

Now we would give a lower bound of $\Pr[D]$, equivalently, an upper bound of $\Pr[\overline{D}]$. Let D^{j_1,j_2} be the event that (v^{j_1}, v^{j_2}) is new in $(0, v_1^{j_1}, \cdots, v_{\ell_{j_1}}^{j_1}, v_1^{j_2}, \cdots, v_{\ell_{j_2}}^{j_2})$, $j_1 \neq j_2$. Now it is easy to check that $\overline{D} = \cup_{1\le j_1<j_2\le q} \overline{D^{j_1,j_2}}$. Thus if $\Pr[\overline{D^{j_1,j_2}}] \le \delta$ for some δ and all choices of $j_1 < j_2$, then $\Pr[D] \ge (1 - \binom{q}{2}\delta)$. Without loss of generality, we compute $\Pr[D^{1,2}]$ for the message M^1 and M^2. We have several cases depending on the messages M^1 and M^2.

Lower bound of $\Pr[D^{1,2}]$

Case-1 : $\ell_1 = \ell_2 = \ell$ **(say) and** $x_1^1 = x_1^2, \cdots, x_\ell^1 = x_\ell^2, z^1 \neq z^2$**.**

Let us denote $v_1 = v_1^1 = v_1^2, \cdots, v_\ell = v_\ell^1 = v_\ell^2$ and $w_1 = w_1^1 = w_1^2, \cdots, w_\ell = w_\ell^1 = w_\ell^2$. We choose the $(\ell+1)$-tuple $(w, w_1, \cdots, w_\ell)$ such that (v_1, v^1, v^2) is *new* in $(0, v_2, \cdots, v_\ell)$.

- Let A be the event such that v_1 is new in $(0, v_2, \cdots, v_\ell)$ and $\Delta^1 + z^1 \neq \Delta^2 + z^2$. Hence, for $2 \le i \le \ell$, $w \neq -\frac{x_1^1}{c_1}, -\frac{x_1^1 - x_i^1}{c_1 - c_i}, \frac{z^2 - z^1}{c}$ (assume that $|M^1|$ is a multiple of n and $|M^2|$ is not, if both are multiple or not multiple of n then always $\Delta^1 + z^1 \neq \Delta^2 + z^2$). So $\Pr[A] = \frac{N-\ell-1}{N} = 1 - \frac{\ell+1}{N}$.
- Let B be the event such that (v^1, v^2) is new in $(0, v_1, \cdots, v_\ell)$. So,
  - $w_1 + z^1 + (w_2 + \cdots + w_\ell) + \Delta^1 \neq v_i, 0$,
  - $w_1 + z^2 + (w_2 + \cdots + w_\ell^2) + \Delta^2 \neq v_i, 0$ and
  - $w_1 + z^1 + (w_2 + \cdots + w_\ell) + \Delta^1 \neq w_1 + z^2 + (w_2 + \cdots + w_\ell^2) + \Delta^2$. This is always true given that A is true.

  Thus, we get $\mathbf{Pr}[B \mid A] \ge \frac{N-2\ell-2}{N} = (1 - \frac{2\ell+2}{N})$. Note that w_1 is the output of v_1 which is new in $(0, v_2, \cdots, v_\ell)$.
- Now, $A \cap B \subseteq D^{1,2}$ and hence $\Pr[D^{1,2}] \ge (1 - \frac{\ell+1}{N})(1 - \frac{2\ell+2}{N}) \ge 1 - \frac{3\ell+3}{N}$.

Case-2 : $\ell_1 = \ell_2 = \ell$ **(say) and** $x_1^1 = x_1^2, \cdots, x_\ell^1 = x_\ell^2, z^1 = z^2$**.**

This case can happen only if $\mathbf{pad}(M^1) = M^1 = M^2 \parallel 10^s = \mathbf{pad}(M^2)$ (there is one more similar case where $|M^2|$ is a multiple of n and $|M^1|$ is not). We denote $v_1 = v_1^1 = v_1^2, \cdots, v_\ell = v_\ell^1 = v_\ell^2$ and $w_1 = w_1^1 = w_1^2, \cdots, w_\ell = w_\ell^1 = w_\ell^2$. We choose the $(\ell+1)$-tuple $(w, w_1, \cdots, w_\ell)$ such that (v_1, v^1, v^2) is *new* in $(0, v_2, \cdots, v_\ell)$.

- Let A be the event such that v_1 is new in $(0, v_2, \cdots, v_\ell)$ and $\Delta^1 + z^1 \neq \Delta^2 + z^2$. Hence, for $2 \le i \le \ell$, $w \neq -\frac{x_1^1}{c_1}, -\frac{x_1^1 - x_i^1}{c_1 - c_i}, \frac{z^2 - z^1}{c}$. So $\Pr[A] = \frac{N-\ell-1}{N} = 1 - \frac{\ell+1}{N}$.
- Let B be the event such that (v^1, v^2) is new in $(0, v_1, \cdots, v_\ell)$. So,
  - $w_1 + z^1 + (w_2 + \cdots + w_\ell) + \Delta^1 \neq v_i, 0$,
  - $w_1 + z^2 + (w_2 + \cdots + w_\ell^2) + \Delta^2 \neq v_i, 0$ and

– $w_1 + z^1 + (w_2 + \cdots + w_\ell) + \Delta^1 \neq w_1 + z^2 + (w_2 + \cdots + w_\ell^2) + \Delta^2$. This is always true given that A is true.

Thus, we get $\Pr[B \mid A] \geq \frac{N-2\ell-2}{N} = (1 - \frac{2\ell+2}{N})$. Note that w_1 is the output of v_1 which is new in $(0, v_2, \cdots, v_\ell)$.

- Now, $A \cap B \subseteq D^{1,2}$ and hence $\Pr[D^{1,2}] \geq (1 - \frac{\ell+1}{N})(1 - \frac{2\ell+2}{N}) \geq 1 - \frac{3\ell+3}{N}$.

Case-3: $x_1^1 x_2^1 \ldots x_\ell^1 \neq x_1^2 x_2^2 \ldots x_\ell^2$**.**

Without loss of generality we can assume $x_1^1 \neq x_1^2$. Choose $(w, w_1^1, w_1^2, \ldots, w_\ell^1, w_\ell^2)$-tuple (some of them may be equal), such that (v_1^1, v_1^2, v^1, v^2) is new in $(0, v_2^1, v_2^2, \ldots, v_\ell^1, v_\ell^2)$.

- Let A denote the event that (v_1^1, v_1^2) is new in $(0, v_2^1, v_2^2, \ldots, v_\ell^1, v_\ell^2)$.
Hence $w \neq \frac{x_1^1 - x_i^1}{c_i - c_1}, \frac{x_1^2 - x_i^2}{c_i - c_1}, \frac{x_1^1 - x_j^2}{c_j - c_1}, \frac{x_1^2 - x_j^1}{c_j - c_1}, -\frac{x_1^1}{c_1}, -\frac{x_1^2}{c_1}$ for $2 \leq i, j \leq \ell$. So $\Pr[A] \geq (1 - \frac{4\ell-2}{N})$

- Let B_1 denote the event that v^1 is new in $(0, v_1^1, v_1^2, \ldots, v_\ell^1, v_\ell^2)$. Hence $w_1^1 \neq -(z^1 + w_2^1 + \cdots w_\ell^1), -(z^1 + w_2^1 + \cdots w_\ell^1) + v_i^1, -(z^1 + w_2^1 + \cdots w_\ell^1) + v_i^2$ for $1 \leq i \leq \ell$. So $\Pr[B_1 \mid A] \geq (1 - \frac{2\ell+1}{N})$

- Let B_2 denote the event that v^2 is new in $(0, v_1^1, v_1^2, \ldots, v_\ell^1, v_\ell^2, v^1)$. Hence $w_1^2 \neq -(z^2 + w_2^2 + \cdots w_\ell^2), -(z^2 + w_2^2 + \cdots w_\ell^2) + v_i^1, -(z^2 + w_2^2 + \cdots w_\ell^2) + v_i^2, -(z^2 + w_2^2 + \cdots w_\ell^2) + w_1^1 + (z^1 + w_2^1 + \cdots + w_\ell^1)$ for $1 \leq i \leq \ell$. So $\Pr[B_2 \mid B_1 \cap A] \geq (1 - \frac{2\ell+2}{N})$.

- Now, $A \cap B_1 \cap B_2 \subseteq D^{1,2}$ and hence $\Pr[D^{1,2}] \geq (1 - \frac{4\ell-2}{N})(1 - \frac{2\ell+1}{N})(1 - \frac{2\ell+2}{N}) \geq 1 - \frac{8\ell+1}{N}$.

Case-4 : $\ell_1 \neq \ell_2$**.**

Assume $\ell_2 > \ell_1$. Choose $(w, w_1^1, \cdots, w_{\ell_1}^1, w_1^2, \cdots, w_{\ell_2}^2)$-tuple (some of them may be equal), such that $(v_1^1, v_{\ell_2}^2, v^1, v^2)$ is new in $(0, v_2^1, \cdots, v_{\ell_1}^1, v_1^2, \cdots, v_{\ell_2}^2)$ (if $x_1^1 \neq x_1^2$) or $(0, v_2^1, \cdots, v_{\ell_1}^1, v_2^2, \cdots, v_{\ell_2}^2)$ (if $x_1^1 = x_1^2$, in this case $v_1^1 = v_1^2$). We assume that $x_1^1 \neq x_1^2$. The other case is very similar to this.

- Let **A** denote the event that $(v_1^1, v_{\ell_2}^2)$ is new in $(0, v_2^1, \cdots, v_{\ell_1}^1, v_1^2, v_2^2, \cdots, v_{\ell_2-1}^2)$. Hence $w \neq \frac{x_1^1 - x_i^1}{c_i - c_1}, \frac{x_1^1 - x_j^2}{c_j - c_1}, -\frac{x_1^1}{c_1}$ for $2 \leq i \leq \ell_1, 1 \leq j \leq \ell_2$ and $w \neq \frac{x_{\ell_2}^2 - x_i^1}{c_i - c_{\ell_2}}, \frac{x_{\ell_2}^2 - x_j^2}{c_j - c_{\ell_2}}, -\frac{x_{\ell_2}^2}{c_{\ell_2}}$ for $1 \leq i \leq \ell_1, 1 \leq j \leq \ell_2 - 1$. So $\Pr[A] \geq (1 - \frac{2(\ell_1+\ell_2)}{N})$.

- Let B_1 denote the event that $(v_1^1, v_{\ell_2}^2, v^1)$ is new in $(0, v_2^1, \cdots, v_{\ell_1}^1, v_1^2, v_2^2, \cdots, v_{\ell_2-1}^2)$. Hence we have $w_1^1 \neq -(z^1 + w_2^1 + \cdots + w_{\ell_1}^1), -(z^1 + w_2^1 + \cdots + w_{\ell_1}^1) + v_i^1, -(z^1 + w_2^1 + \cdots + w_{\ell_1}^1) + v_j^2$ for $1 \leq i \leq \ell_1, 1 \leq j \leq \ell_2$. So $\Pr[B_1 | A] \geq (1 - \frac{\ell_1+\ell_2+1}{N})$.

- Let B_2 denote the event that $(v_1^1, v_{\ell_2}^2, v^1, v^2)$ is new in $(0, v_2^1, \cdots, v_{\ell_1}^1, v_1^2, v_2^2, \cdots, v_{\ell_2-1}^2)$. Hence we have $w_{\ell_2}^2 \neq -(z^2 + w_1^2 + \cdots + w_{\ell_2-1}^2), -(z^2 + w_1^2 + \cdots + w_{\ell_2-1}^2) + v_i^1, -(z^2 + w_1^2 + \cdots + w_{\ell_2-1}^2) + v_j^2, -(z^2 + w_1^2 + \cdots + w_{\ell_2-1}^1) + w_1^1 + (z^1 + w_2^1 + \cdots + w_\ell^1)$ for $1 \leq i \leq \ell_1, 1 \leq j \leq \ell_2$. So $\Pr[B_2 | A \cap B_1] \geq (1 - \frac{\ell_1+\ell_2+2}{N})$.

- Now, $A \cap B_1 \cap B_2 \subseteq D^{1,2}$ and hence $\Pr[D^{1,2}] \geq (1 - \frac{2(\ell_1+\ell_2)}{N})(1 - \frac{\ell_1+\ell_2+1}{N})(1 - \frac{\ell_1+\ell_2+2}{N}) \geq 1 - \frac{4(\ell_1+\ell_2)+3}{N}$.

Theorem 4. *Let $M^1, \cdots, M^q$ are distinct messages from $\mathcal{M}$ and $y^1, \cdots, y^q \in T$ (not necessarily distinct) then* $\Pr[\mathrm{P}_f(M^1) = y^1, \cdots, \mathrm{P}_f(M^q) = y^q] \geq \frac{1-\epsilon}{N^q} = (1-\epsilon) \times \Pr[F(M^1) = y^1, \cdots, F(M^q) = y^q]$ *where* $\epsilon = \frac{11(q-1)\sigma}{2N}$ *and F is an uniform random function on* $\mathbf{Func}(\{0,1\}^{\leq L}, \{0,1\}^n)$.

Proof. From the above four cases we can say that for any two messages M^{j_1} and M^{j_2}, $\Pr[\overline{D^{j_1,j_2}}] \leq \frac{4(\ell_{j_1}+\ell_{j_2})+3}{N}$. Thus, $\Pr[\overline{D}] \leq \sum_{1 \leq j_1 < j_2 \leq q} \frac{4(\ell_{j_1}+\ell_{j_2})+3}{N} = \frac{4(q-1)\sum_j \ell_j}{N} + \frac{3q(q-1)}{2N} \leq \frac{11(q-1)\sigma}{2N}$.

Corollary 5. $\mathbf{Adv}_{\mathbf{PMAC}}(q, \sigma, t) \leq \frac{11(q-1)\sigma}{2N}$.

5 Conclusion

This paper provides a simpler and improved upper bound $\mathrm{O}(q\sigma/N)$ for the distinguishing advantage of PMAC. We have used the proof idea taken from [4, 13]. This idea has unifying nature in proving indistinguishability. The security analysis is made on a slight modification of PMAC (without length padding). The security analysis holds for the original PMAC definition also. So, one can use PMAC for arbitrary length messages also. As a future research work, we hope our security analysis can be extended to have an improved bound on a general class given in [10, 13].

References

[1] M. Bellare, A. Boldyreva, L. Knudsen and C. Namprempre. On-Line Ciphers and the Hash-CBC constructions. Advances in Cryptology - CRYPTO 2001. Lecture Notes in Computer Science, Volume **2139**, pp 292-309.

[2] 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.

[3] 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.

[4] Daniel J. Bernstein. A short proof of the unpredictability of cipher block chaining (2005). URL: http://cr.yp.to/papers.html#easycbc.

[5] 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.

[6] 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.

[7] J. Daemen and V. Rijmen. Resistance Against Implementation Attacks. A Comparative Study of the AES Proposals. In Proceedings of the Second AES Candidate Conference (AES2), Rome, Italy, March 1999. Available at http://csrc.nist.gov/encryption/aes/aes_ home.htm.

[8] H. Krawczyk. LFSR-based hashing and authenticating. Advances in Cryptology, CRYPTO 1994, Lecture Notes in Computer Science, Volume **839**, pp 129-139, Springer-Verlag 1994.

[9] 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.

[10] 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.

[11] 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.

[12] M. Luby and C. Rackoff. How to construct pseudo-random permutations from pseudo-random functions. Advances in Cryptology, CRYPTO' 85, Lecture Notes in Computer Science, Volume **218**, pp 447, Springer-Verlag 1985.

[13] M. Nandi. A Simple and Unified Method of Proving Indistinguishability. Indocrypt 2006, Lecture Notes in Computer Science, Volume **4329**, pp 317-334.

[14] P. Rogaway. Bucket Hashing and Its Application to Fast Message Authentication. Advances in Cryptology, CRYPTO 1995, Lecture Notes in Computer Science, Volume **963**, pp 29-42, Spronger-Verlag, 1995.

[15] D. R. Stinson. On the connections between universal hashing, combinatorial designs and error-correcting codes. Congressus Numerantium **114**, 1996, pp 7-27.