A Note on the Security Framework of Two-key DbHtS MACs

Tingting Guo[1,2] and Peng Wang[1,2(✉)]

[1] SKLOIS, Institute of Information Engineering, CAS
w.rocking@gmail.com, guotingting@iie.ac.cn
[2] School of Cyber Security, University of Chinese Academy of Sciences

Abstract. Double-block Hash-then-Sum (DbHtS) MACs are a class of MACs achieve beyond-birthday-bound (BBB) security, including SUM-ECBC, PMAC_Plus, 3kf9 and LightMAC_Plus etc. Recently, Shen et al. (Crypto 2021) proposed a security framework for two-key DbHtS MACs in the multi-user setting, stating that when the underlying blockcipher is ideal and the universal hash function is regular and almost universal, the two-key DbHtS MACs achieve $2n/3$-bit security. Unfortunately, the regular and universal properties can not guarantee the BBB security of two-key DbHtS MACs. We propose three counter-examples which are proved to be $2n/3$-bit secure in the multi-user setting by the framework, but can be broken with probability 1 using only $\mathcal{O}(2^{n/2})$ queries even in the single-user setting. We also point out the miscalculation in their proof leading to such a flaw. However, we haven't found attacks against 2k-SUM-ECBC, 2k-PMAC_Plus and 2k-LightMAC_Plus proved $2n/3$-bit security in their paper.

Keywords: MAC · DbHtS · Beyond-birthday-bound security · Multi-user security.

1 Introduction

Message Authentication Code (MAC). MAC is a symmetric-key crypto primitive to ensure integrity of messages. Most of them follow the Hash-then-Encipher (HtE) framework:

$$\mathrm{HtE}[H,E](K_h,K,M) = E_K(H_{K_h}(M)).$$

When universal hash function H_{K_h} is a almost uiversal (AU) and E_K is a fixed-input-length PRF, such framework is a variable-input-length PRF [16] with birthday bound security (i.e., they break with $\mathcal{O}(2^{n/2})$ queries assuming the size of every block cipher is n bits). XCBC [4], PMAC [5, 14], HMAC [2], and NMAC [2] follow this framework. However, birthday bound security is always not enough for lightweight blockciphers (PRESENT [6], GIFT [1]), whose $n = 64$. Because in this case, the security is only 32 bits (i.e., secure within 2^{32} queries), which is practically vulnerable. So researchers make great efforts to improve the security strength of MAC.

Birthday-Birthday-Bound MACs. Plenty of MACs with beyond-birthday-bound security have been put forward. Such as SUM-ECBC [17], PMAC_Plus [18], 3kf9 [19], LightMAC_Plus [12], and so on. At FSE 2019, Datta et al. showed they all follow the Double-block Hash-then-Sum (DbHtS) framework [8], i.e., three-key DbHtS:

$$\text{DbHtS}[H, E]\,(K_h, K_1, K_2, M) = E_{K_1}(H^1_{K_{h,1}}(M)) \oplus E_{K_2}(H^2_{K_{h,2}}(M)),$$

where M is the massage, hash key $K_h = (K_{h,1}, K_{h,2})$, $H^1_{K_{h,1}}$ and $H^2_{K_{h,2}}$ are two universal hash functions and E_{K_1} and E_{K_2} are two blockciphers on n bits with two independent keys K_1, K_2 respectively. BBB MACs following three-key DbHtS have been proved with $2n/3$-bit security in their primary proofs [17–19, 12] and under the framework of three-key DbHtS proposed by Datta [8]. Later, Leurent et al. [11] showed the best attacks to them cost $\mathcal{O}(2^{3n/4})$ queries. Recently at EUROCRYPT 2020, Kim et al. [10] have proved the tight $3n/4$-bit security.

To facilitate key management, Datta et al. [8] also raised two-key DbHtS framework, that is to say, $K_1 = K_2$ in DbHtS framework. They showed two-key DbHtS MACs (2K-ECBC_Plus, 2K-PMAC_Plus, and 2K-LightMAC_Plus) under their framework are still $2n/3$-bit security.

Two-Key DbHtS in the Multi-User Setting. All the above MAC frameworks only considered a single user. We have put them in Table **??**. In practice, the adversary can attack multiple users. For instance, MACs are core elements of real-world security protocols such as TLS, SSH, and IPsec, which are used by lots of websites with plenty of daily active users. However, by a generic reduction, all above BBB results degrade to (or even worse than) the birthday bound in the multi-user setting [15].

So at Crypto 2021, Shen et al. [15] revisited the security of two-key DbHtS framework in the multi-user setting elaborately. Their framework (Theorem 1 in [15]) states when the underlying blockcipher is ideal and the two independent universal hash functions are both regular and almost universal, the two-key DbHtS MACs , including 2k-SUM-ECBC, achieve $2n/3$-bit security. They adjusted the proof of the framework for adapting to 2k-PMAC_Plus and 2k-LightMAC_Plus based on two dependent universal hash functions,stating they achieve $2n/3$-bit security, too.

Our Contributions. We show that *Theorem 1 in Shen et al.'s paper [15], giving the security of two-key DbHtS framework, has a critical flaw* by three counter-examples. According to their Theorem 1, these counter-examples are proved $2n/3$-bit security (ignoring the maximum message length and ideal-cipher queries) in the multi-user setting. However, they are all attacked successfully with only $\mathcal{O}(2^{n/2})$ queries even in the single-user setting. We also show clearly the miscalculation in their proof leading to such a flaw.

2 Preliminaries

Notation. For a finite set $\mathcal{X}$, let $X \xleftarrow{\$} \mathcal{X}$ denote sampling X from $\mathcal{X}$ uniformly and randomly. Let $|\mathcal{X}|$ be the size of the set $\mathcal{X}$. For a domain $\mathcal{X}$ and a range $\mathcal{Y}$, let $\mathsf{Func}(\mathcal{X}, \mathcal{Y})$ denote the set of all functions from $\mathcal{X}$ to $\mathcal{Y}$.

Multi-User Pseudorandom Function. Let $F : \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ be a function. The game $\mathbf{G}_F^{\mathrm{prf}}(\mathscr{A})$ about adversary $\mathscr{A}$ is defined as follows.

1. Initialize $K_1, K_2, \ldots \xleftarrow{\$} \mathcal{K}, f_1, f_2, \ldots \xleftarrow{\$} \mathsf{Func}(\mathcal{X}, \mathcal{Y})$, and $b \xleftarrow{\$} \{0, 1\}$;
2. $\mathscr{A}$ queries Eval function with (i, X) and get $\mathsf{Eval}(i, X)$, where $i \in \{1, 2, \ldots\}, X \in \mathcal{X}$, and
$$\mathsf{Eval}(i, X) = \begin{cases} F(K_i, X), & \text{if } b = 0, \\ f_i(X), & \text{if } b = 1; \end{cases}$$
3. $\mathscr{A}$ output $b' = b$.

Then the advantage of the adversary $\mathscr{A}$ against the multi-user Pseudorandom Function (PRF) security of F is

$$\mathrm{Adv}_F^{\mathrm{prf}}(\mathscr{A}) = 2 \Pr[\mathbf{G}_F^{\mathrm{prf}}(\mathscr{A})] - 1.$$

The H-Coefficient Technique. When considering interactions between an adversary $\mathscr{A}$ and an abstract system $\mathbf{S}$ which answers $\mathscr{A}$'s queries, let X_i denote the query from $\mathscr{A}$ to $\mathbf{S}$ and Y_i denote the response of X_i from $\mathbf{S}$ to $\mathscr{A}$. Then the resulting interaction can be recorded with a transcript $\tau = ((X_1, Y_1), \ldots, (X_q, Y_q))$. Let $p_{\mathbf{S}}(\tau)$ denote the probability that $\mathbf{S}$ produces τ. In fact, $p_{\mathbf{S}}(\tau)$ is the description of $\mathbf{S}$ and independent of the adversary $\mathscr{A}$. Then we describe the H-coefficient technique [7, 13]. Generically, it considers an adversary that aims at distinguishing a "real" system $\mathbf{S}_1$ from an"ideal" system $\mathbf{S}_0$. The interactions of the adversary with those two systems induce two transcript distributions D_1 and D_0 respectively. It is well known that the statistical distance $\mathsf{SD}(D_0, D_1)$ is an upper bound on the distinguishing advantage of $\mathscr{A}$.

Lemma 1. [7, 13] *Suppose that the set of attainable transcripts for the ideal system can be partitioned into good and bad ones. If there exists $\epsilon \geq 0$ such that $\frac{p_{\mathbf{S}_1}(\tau)}{p_{\mathbf{S}_0}(\tau)} \geq 1 - \epsilon$ for any good transcript τ, then*

$$\mathsf{SD}(D_0, D_1) \leq \epsilon + \Pr[D_0 \textit{ is bad}].$$

Regular and AU. Let $H : \mathcal{K}_h \times \mathcal{X} \to \mathcal{Y}$ be a hash function where $\mathcal{K}_h$ is the key space, $\mathcal{X}$ is the domain and $\mathcal{Y}$ is the range. Hash function H^i is said to be ϵ_1-regular if for any $X \in \mathcal{X}$, $Y \in \mathcal{Y}$,

$$\Pr[K_h \xleftarrow{\$} \mathcal{K}_h : H_{K_h}(X) = Y] \leq \epsilon_1.$$

And hash function H is said to be ϵ_2-AU if for any two distinct strings $X, X^{'} \in \mathcal{X}$,

$$\Pr[K_h \xleftarrow{\$} \mathcal{K}_h : H_{K_h}(X) = H_{K_h}(X^{'})] \leq \epsilon_2.$$

3 BBB-Security framework in [15]

Let $\mathcal{M}$ be the message space and $\mathcal{K}_h \times \mathcal{K}$ be the key space. Let blockcipher $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ and $\mathcal{K} = \{0,1\}^k$. Let hash function $H : \mathcal{K}_h \times \mathcal{M} \to \{0,1\}^n \times \{0,1\}^n$. The function H is consist of two n-bit hash functions H^1 and H^2, i,e., $H_{K_h}(M) = (H^1_{K_{h,1}}(M), H^2_{K_{h,2}}(M))$ where $K_h = (K_{h,1}, K_{h,2}) \in \mathcal{K}_{h,1} \times \mathcal{K}_{h,2}$ and $K_{h,1}, K_{h,2}$ are two independent keys. Then the two-key DbHtS framework in paper [15] (see Fig.1) is

$$\mathrm{DbHtS}[H,E]\,(K_h, K, M) = E_K\left(H^1_{K_{h,1}}(M)\right) \oplus E_K\left(H^2_{K_{h,2}}(M)\right).$$

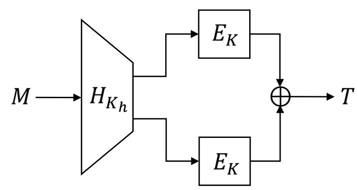

Fig. 1. The two-key DbHtS construction. Here H is a $2n$-bit hash function from $\mathcal{K}_h \times \mathcal{M}$ to $\{0,1\}^n \times \{0,1\}^n$, and E is a n-bit blockcipher from $\mathcal{K} \times \{0,1\}^n$ to $\{0,1\}^n$.

Theorem 1 in [15]. Let E be modeled as an ideal blockcipher. Let H^1 and H^2 both satisfy ϵ_1-regular and ϵ_2-AU. Then Shen et al. [15] proved the security of two-key DbHtS in the multi-user setting as following, which is the core of their paper and they named it Theorem 1. For any adversary $\mathscr{A}$ that makes at most q evaluation queries and p ideal-cipher queries,

$$\begin{aligned}\mathrm{Adv}^{\mathrm{prf}}_{\mathrm{DbHtS}}(\mathscr{A}) \leq & \frac{2q}{2^k} + \frac{q(3q+p)(6q+2p)}{2^{2k}} + \frac{2qp\ell}{2^{n+k}} + \frac{2qp\epsilon_1}{2^k} + \frac{4qp}{2^{n+k}} \\ & + \frac{4q^2\epsilon_1}{2^k} + \frac{2q^2\ell\epsilon_1}{2^k} + 2q^3\left(\epsilon_1+\epsilon_2\right)^2 + \frac{8q^3\left(\epsilon_1+\epsilon_2\right)}{2^n} + \frac{6q^3}{2^{2n}}\end{aligned} \tag{1}$$

where ℓ is the maximal block length among these evaluation queries and assuming $p + q\ell \leq 2^{n-1}$.

An Overview of the Proof of Theorem 1 in [15]. They proved Theorem 1 based on H-coefficient technique. Let $\mathbf{S}_1$ be "real" system and $\mathbf{S}_0$ be "ideal" system. For $b \in \{0,1\}$, system $\mathbf{S}_b$ performs the following procedure.

1. Initialize $(K_h^1, K_1), \ldots, (K_h^u, K_u) \xleftarrow{\$} \mathcal{K}_h \times \mathcal{K}$ if $b = 1$; otherwise, initialize $f_1, \ldots, f_u \xleftarrow{\$} \mathsf{Func}(\mathcal{M}, \{0,1\}^n)$;
2. If an adversary $\mathscr{A}$ queries Eval function with (i, M), where $i \in \{1,2,\ldots\}$, $M \in \mathcal{M}$, return

$$\mathsf{Eval}(i, M) = \begin{cases} \mathrm{DbHtS}[H,E](K_h^i, K_i, M), & \text{if } b = 1, \\ f_i(M), & \text{if } b = 0; \end{cases}$$

3. If an adversary $\mathscr{A}$ queries Prim function with (J, X), where $J \in \mathcal{K}, X \in \{+,-\} \times \{0,1\}^n$, return

$$\mathsf{Prim}(J, X) = \begin{cases} E_J(x), & \text{if } X = \{+, x\}, \\ E_J^{-1}(y), & \text{if } X = \{-, y\}. \end{cases}$$

They called the query to Eval evaluation query and the query to Prim ideal-cipher query. For each query $T \leftarrow \mathsf{Eval}(i, M)$, they associated it with an entry $(eval, i, M, T)$. The query to Prim is similar to it. Transcript τ consisted of such entries. Then they defined bad transcripts, including fourteen cases. If a transcript is not bad then they said it's good. Let D_1 and D_0 be the random variables for the transcript distributions in the system $\mathbf{S}_1$ and $\mathbf{S}_0$ respectively. They firstly bounded the probability that D_0 is bad as follows. Let Bad_i be the event that the i-th case of bad transcripts happens. They calculated the probability $\Pr[\mathrm{Bad}_1], \ldots, \Pr[\mathrm{Bad}_{14}]$ in sequence. After summing up, they got

$$\begin{aligned}\Pr\left[D_0 \text{ is bad }\right] \leq& \sum_{i=1}^{14} \Pr\left[\mathrm{Bad}_i\right] \\ \leq& \frac{2q}{2^k} + \frac{q(3q+p)(6q+2p)}{2^{2k}} + \frac{2qp\ell}{2^{k+n}} + \frac{2qp\epsilon_1}{2^k} + \frac{4qp}{2^{n+k}} \\ &+ \frac{4q^2\epsilon_1}{2^k} + \frac{2q^2\ell\epsilon_1}{2^k} + 2q^3\left(\epsilon_1+\epsilon_2\right)^2 + \frac{8q^3\left(\epsilon_1+\epsilon_2\right)}{2^n}.\end{aligned}$$

Besides, they proved the transcript ratio $\frac{\mathrm{p}_{\mathbf{S}_1}(\tau)}{\mathrm{p}_{\mathbf{S}_0}(\tau)} \geq 1 - \frac{6q^3}{2^{2n}}$ for any good transcript τ. Thus they concluded Theorem 1 by Lemma 1.

4 Counter-Examples

We will show three counter-examples who are two-key DbHtS constructions and satisfy ϵ_1-regular and ϵ_2-AU are attacked in the single-user setting with fewer queries than the security claimed by Theorem 1 [15]. So they are counter-examples against the framework of Shen et al..

4.1 Counter-Example 1

Our first counter-example is a function with fixed input length. Let hash function

$$H_{K_h}(M) = (H^1_{K_1}(M), H^1_{K_2}(M)) = (M \oplus K_1, M \oplus K_2),$$

where M is the message from massage space $\{0,1\}^n$, $K_h = (K_1, K_2)$ and $K_1, K_2 \xleftarrow{\$} \{0,1\}^n$. Let blockcipher $E_K : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$. Then we define function $F : \{0,1\}^{2n} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ as

$$F[H, E](K_h, K, M) = E_K(H^1_{K_1}(M)) \oplus E_K(H^2_{K_2}(M)).$$

H^1 and H^2 are $\frac{1}{2^n}$-Regular and $\frac{1}{2^n}$-AU. It is easy to know that for any $M \in \{0,1\}^n, Y \in \{0,1\}^n$ and $i \in \{1,2\}$,

$$\Pr[K_i \xleftarrow{\$} \{0,1\}^n : M \oplus K_i = Y] \leq \frac{1}{2^n}.$$

And for any two distinct strings $M, M^{'} \in \{0,1\}^n$ and $i \in \{1,2\}$,

$$\Pr[K_i \xleftarrow{\$} \{0,1\}^n : M \oplus K_i = M^{'} \oplus K_i] = 0.$$

So hash functions H^1 and H^2 are both $\frac{1}{2^n}$-regular and $\frac{1}{2^n}$-AU.

$2n/3$-bit security according to [15]. According to Theorem 1 [15], function F is secure within $\mathcal{O}(2^{2n/3})$ evaluation queries assuming ideal-cipher queries is $\mathcal{O}(1)$ in the multi-user setting.

Attack with $\mathcal{O}(2^{n/2})$ query complexity. It is easy to know that for all keys in keyspace and messages in message space,

$$\begin{aligned} F[H,E](K_h, K, M \oplus K_1 \oplus K_2) &= E_K(M \oplus K_2) \oplus E_K(M \oplus K_1) \\ &= F[H,E](K_h, K, M). \end{aligned}$$

It means F has a period $s := K_1 \oplus K_2$. Based on this, there is an adversary $\mathscr{A}$ can distinguish F from random function f with only $\mathcal{O}(2^{n/2})$ evaluation queries as follows, which is contradictory to Theorem 1 [15].

1. $\mathscr{A}$ firstly makes $\mathcal{O}(2^{n/2})$ evaluation queries of distinct massages $M_1, M_2, \ldots$ chosen uniformly and randomly, and get $T_1, T_2, \ldots$;
2. $\mathscr{A}$ searches a message pair (M_i, M_j) for $M_i \neq M_j, M_i, M_j \in \{M_1, M_2, \ldots\}$ which makes (i) and (ii) hold.
   (i) $T_i = T_j$;
   (ii) After make another two evaluation queries with massages M' and $M' \oplus M_i \oplus M_j$ for $M' \notin \{M_i, M_j\}$, $\mathscr{A}$ gets two identical answers.

If the evaluation query is to F, one can expect on average that there exists one message pair (M_i, M_j) among $\mathcal{O}(2^{n/2})$ massages such that $M_i = M_j \oplus s$. Conditions (i) and (ii) in the second step of $\mathscr{A}$ filter out such pair. However, random function f has no period. If the evaluation query is to f, on average there exists one message pair (M_i, M_j) among $\mathcal{O}(2^{n/2})$ massages such that $T_i = T_j$. However, the probability of $f(M') = f(M' \oplus M_i \oplus M_j)$ for any $M' \notin \{M_i, M_j\}$ is only $1/2^n$. So $\mathscr{A}$ finds a pair (M_i, M_j) satisfying conditions (i) and (ii) with negligible probability. Thus $\mathscr{A}$ distinguish F from random function with probability $1 - 1/2^n$.

4.2 Counter-Example 2

Compared with the first counter-example with fixed input length, our second counter-example can handle variable-length input. We define the function of counter-example 2 the same as counter-example 1 except dealing with messages from $(\{0,1\}^n)^*$ and altering two hash functions H^1 and H^2 to

$$H^i_{K_i}(M) = M[1] \oplus M[2]K_i \oplus M[3]K_i^2 \oplus \ldots \oplus M[m]K_i^{m-1} \oplus |M|K_i^m, i = 1,2.$$

where $M = M[1] \parallel M[2] \parallel \ldots \parallel M[m]$ and every message block is n-bit. This example is a variant of PolyMAC [10].

H^1 and H^2 are $\frac{\ell}{2^n}$-Regular and $\frac{\ell}{2^n}$-AU. Assume the maximal block length of all evaluation queries is ℓ. Any equation of at most ℓ degree has at most ℓ roots. So it is easy to know that for any $M \in (\{0,1\}^n)^*, Y \in \{0,1\}^n$ and $i \in \{1,2\}$,

$$\Pr[K_i \xleftarrow{\$} \{0,1\}^n : H^i_{K_i}(M) = Y] \leq \frac{\ell}{2^n}.$$

And for any two distinct strings $M, M^{'} \in (\{0,1\}^n)^*$ and $i \in \{1,2\}$,

$$\Pr[K_i \xleftarrow{\$} \{0,1\}^n : H^i_{K_i}(M) = H^i_{K_i}(M^{'})] \leq \frac{\ell}{2^n}.$$

It means H^1 and H^2 are both $\frac{\ell}{2^n}$-regular and $\frac{\ell}{2^n}$-AU.

$2n/3$-bit security according to [15]. According to Theorem 1 [15], function F is secure within $\mathcal{O}(2^{2n/3})$ evaluation queries assuming ideal-cipher queries is $\mathcal{O}(1)$ and $\ell = \mathcal{O}(1)$ in the multi-user setting.

Attack with $\mathcal{O}(2^{n/2})$ query complexity. Fix any arbitrary string

$$M_{fix} := M[2] \| M[3] \| \ldots \| M[m] \in (\{0,1\}^n)^{m-1},$$

where $2 \leq m \leq \ell = O(1)$. Let

$$K_i^{'} := M[2]K_i \oplus M[3]K_i^2 \oplus \ldots M[m]K_i^{m-1} \oplus nmK_i^m, i = 1, 2.$$

Then it is easy to obtain for any keys in key space and $M[1] \in \{0,1\}^n$,

$$\begin{aligned} & F[H,E](K_h, K, (M[0] \oplus K_1^{'} \oplus K_2^{'}) \parallel M_{fix}) \\ =& E_K(M[0] \oplus K_2^{'}) \oplus E_K(M[0] \oplus K_1^{'}) \\ =& F[H,E](K_h, K, M[0] \parallel M_{fix}). \end{aligned}$$

It means F has a period $s := (K_1^{'} \oplus K_2^{'}) \parallel 0^{n(m-1)}$ for any $M \in \{0,1\}^n \times \{M_{fix}\}$. Based on this, there is an adversary $\mathscr{A}$ can distinguish F from random function f with only $\mathcal{O}(2^{n/2})$ evaluation queries as follows, which is contradictory to Theorem 1 [15].

1. $\mathscr{A}$ firstly makes $\mathcal{O}(2^{n/2})$ evaluation queries with distinct massages $M_1 \parallel M_{fix}$, $M_2 \parallel M_{fix}, \ldots$ where $M_1, M_2, \ldots \xleftarrow{\$} \{0,1\}^n$, and get $T_1, T_2, \ldots$;
2. $\mathscr{A}$ searches a pair (M_i, M_j) for $M_i \neq M_j, M_i, M_j \in \{M_1, M_2, \ldots\}$ which makes (i) and (ii) hold.
   (i) $T_i = T_j$;
   (ii) After make another two evaluation queries with massages $M' \parallel M_{fix}$ and $(M' \oplus M_i \oplus M_j) \parallel M_{fix}$ for $M' \notin \{M_i, M_j\}$, $\mathscr{A}$ gets two identical answers.

The same as counter-example 1, $\mathscr{A}$ distinguishes F from f with probability almost 1.

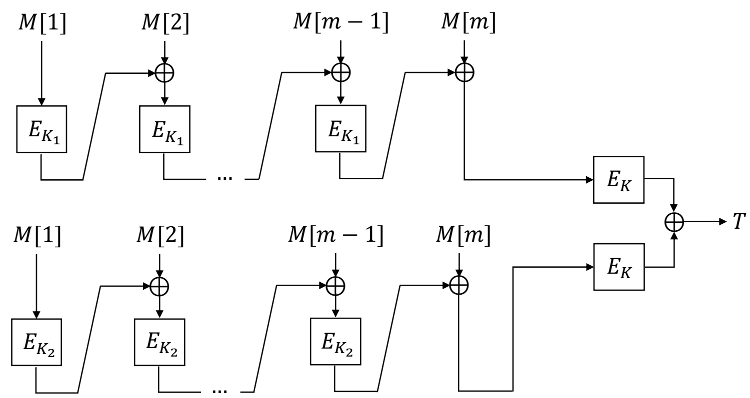

Fig. 2. The variant of 2k-SUM-ECBC. K_1, K_2, K_3 are three independent keys in $\{0,1\}^n$. E is a n-bit blockcipher from $\{0,1\}^n \times \{0,1\}^n$ to $\{0,1\}^n$.

4.3 Counter-Example 3

Unlike counter-examples 1 and 2, the third counter-example with hash functions based on block ciphers. It is a variant of 2k-SUM-ECBC [15]. Let $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a blockcipher with key $K \in \{0,1\}^n$. The two n-bit hash functions used in this function are two CBC MACs without the last cipherblocks, which we call as $\mathrm{CBC}^{'}$. They are keyed with two independent keys K_1 and K_2 respectively. And they deal with at least two message blocks respectively. For a message $M = M[1] \parallel M[2] \parallel \ldots \parallel M[m]$ where every message block is n-bit and $m \geq 2$, the $\mathrm{CBC}^{'}$ algorithm $\mathrm{CBC}^{'}[E](K, M)$ is defined as Y_m, where

$$\begin{aligned} Y_1 &= M[1], \\ Y_j &= E_K(Y_{j-1}) \oplus M[j], j = 2, \ldots, m. \end{aligned}$$

Let $K_h = (K_1, K_2)$. Then we define the function (see Fig.2) as

$$F[\mathrm{CBC}^{'}[E], E](K_h, K, M) = E_K(\mathrm{CBC}^{'}[E](K_1, M)) \oplus E_K(\mathrm{CBC}^{'}[E](K_2, M)).$$

$\mathrm{CBC}^{'}[E]$ is $\left(\frac{2\sqrt{\ell}}{2^n} + \frac{16\ell^4}{2^{2n}}\right)$-Regular and $\left(\frac{2\sqrt{\ell}}{2^n} + \frac{16\ell^4}{2^{2n}}\right)$-AU. For any two different message $M, M^{'} \in (\{0,1\}^n)^*$ with at most ℓ blocks and the adversary making no ideal-cipher query, Ballare et al. [3] and Jha and Nandi [9] show that for $i \in \{1, 2\}$,

$$\begin{aligned} &\Pr[K_i \xleftarrow{\$} \{0,1\}^n : E_K(\mathrm{CBC}^{'}[E](K_i, M)) = E_K(\mathrm{CBC}^{'}[E](K_i, M^{'}))] \\ \leq &\frac{2\sqrt{\ell}}{2^n} + \frac{16\ell^4}{2^{2n}}. \end{aligned}$$

Blockcipher E_K is a permutation. So

$$\Pr[K_i \xleftarrow{\$} \{0,1\}^n : \mathrm{CBC}^{'}[E](K_i, M) = \mathrm{CBC}^{'}[E](K_i, M^{'})] \leq \frac{2\sqrt{\ell}}{2^n} + \frac{16\ell^4}{2^{2n}}.$$

It means $\mathrm{CBC}^{'}$ is $\left(\frac{2\sqrt{\ell}}{2^n}+\frac{16\ell^4}{2^{2n}}\right)$-AU. Let $M = X[1] \parallel (X[2] \oplus Y) \parallel Z \in (\{0,1\}^n)^* \times \{0,1\}^n \times \{0,1\}^n$ and $M^{'} = 0^n \parallel Z \in \{0,1\}^n \times \{0,1\}^n$. Then

$$\begin{aligned}
&\Pr[K_i \xleftarrow{\$} \{0,1\}^n : \mathrm{CBC}^{'}[E](K_i, X[1] \parallel X[2]) = Y] \\
=&\Pr[K_i \xleftarrow{\$} \{0,1\}^n : \mathrm{CBC}^{'}[E](K_i, M) = \mathrm{CBC}^{'}[E](K_i, M^{'})] \\
\leq&\frac{2\sqrt{\ell}}{2^n}+\frac{16\ell^4}{2^{2n}}.
\end{aligned}$$

So $\mathrm{CBC}^{'}$ is $\left(\frac{2\sqrt{\ell}}{2^n}+\frac{16\ell^4}{2^{2n}}\right)$-regular.

$2n/3$-bit security according to [15]. According to Theorem 1 [15], function F is secure within $\mathcal{O}(2^{2n/3})$ evaluation queries assuming no ideal-cipher queries and $\ell = \mathcal{O}(1)$ in the multi-user setting.

Attack with $\mathcal{O}(2^{n/2})$ query complexity. Fix any arbitrary string $M_{fix} \in (\{0,1\}^n)^{m-1}$ where $2 \leq m \leq \ell = O(1)$. Let

$$s = \mathrm{CBC}^{'}[E](K_1, M_{fix} \parallel 0^n) \oplus \mathrm{CBC}^{'}[E](K_2, M_{fix} \parallel 0^n)).$$

Then it is easy to obtain for any keys in key space and $M[m] \in \{0,1\}^n$,

$$\begin{aligned}
&F[\mathrm{CBC}^{'}[E], E](K_h, K, M_{fix} \parallel (M[m] \oplus s)) \\
=&E_K(\mathrm{CBC}^{'}[E](K_2, M_{fix} \parallel 0^n) \oplus M[m]) \oplus \\
&E_K(\mathrm{CBC}^{'}[E](K_1, M_{fix} \parallel 0^n) \oplus M[m]) \\
=&E_K(\mathrm{CBC}^{'}[E](K_2, M_{fix} \parallel M[m])) \oplus E_K(\mathrm{CBC}^{'}[E](K_1, M_{fix} \parallel M[m])) \\
=&F[\mathrm{CBC}^{'}[E], E](K_h, K, M_{fix} \parallel M[m]).
\end{aligned}$$

It means F has a period $s := 0^{n(m-1)} \parallel s$ for any $M \in \{M_{fix}\} \times \{0,1\}^n$. So there is an adversary $\mathscr{A}$ distinguishes F from random function with only $\mathcal{O}(2^{n/2})$ evaluation queries when considering single user similar as counter-example 2.

5 The Flaw of the Proof of Theorem 1 in [15]

In section 3, we have shown the procedure of how Shen et al. [15] proved Theorem 1 based on H-coefficient technique. However, we find they make a critical flaw when they were calculating $\Pr[\mathrm{Bad}_9]$ in their proof, which leads to existing our counter-examples. We now show it.

Assume there are u users and the adversary make q_i evaluation queries to the i-th user in all. Let $(eval, i, M_a^i, T_a^i)$ be the entry obtained when the adversary makes the a-th query to user i. During the computation of entry $(eval, i, M_a^i, T_a^i)$, let Σ_a^i and Λ_a^i be the internal outputs of hash function H in "real" system $\mathbf{S}_1$, namely $\Sigma_a^i = H^1_{K_{h,1}}\left(M_a^i\right)$ and $\Lambda_a^i = H^2_{K_{h,2}}\left(M_a^i\right)$ respectively. The ninth bad event is

"There is an entry $(eval, i, M_a^i, T_a^i)$ *such that either* $\Sigma_a^i = \Sigma_b^i$ *or* $\Sigma_a^i = \Lambda_b^i$, *and either* $\Lambda_a^i = \Lambda_b^i$ *or* $\Lambda_a^i = \Sigma_b^i$ *for some entry* $(eval, i, M_a^i, T_a^i)$*."*

They defined this event bad for the reason that the appearance of such entry $(eval, i, M_a^i, T_a^i)$ is easy used to distinguish systems $\mathbf{S}_1$ and $\mathbf{S}_0$. We call the event of either $\Sigma_a^i = \Sigma_b^i$ or $\Sigma_a^i = \Lambda_b^i$ as event 1, and the event of either $\Lambda_a^i = \Lambda_b^i$ or $\Lambda_a^i = \Sigma_b^i$ as event 2. Then we can regard the simultaneous events 1 and 2 as one of the following 4 events:

- Event 3: $\Sigma_a^i = \Sigma_b^i$ and $\Lambda_a^i = \Lambda_b^i$;
- Event 4: $\Sigma_a^i = \Sigma_b^i$ and $\Lambda_a^i = \Sigma_b^i$;
- Event 5: $\Sigma_a^i = \Lambda_b^i$ and $\Lambda_a^i = \Lambda_b^i$;
- Event 6: $\Sigma_a^i = \Lambda_b^i$ and $\Lambda_a^i = \Sigma_b^i$.

In "real" system $\mathbf{S}_1$, event 4 or 5 leads to $T_a^i = 0^n$; event 3 or 6 leads to $T_a^i = T_b^i$. However in "ideal" system $\mathbf{S}_0$ these happen with negligible probability by the randomness of random function f_i. Thus it is easy distinguish these two systems.

When calculating $\Pr[\mathrm{Bad}_9]$, Shen et al. [15] regarded that the event 1 is independent from event 2 when $K_{h,1}^i, K_{h,2}^i$ are independent from each other. So by H^1, H^2 are both ϵ_1-regular and ϵ_2-AU, they thought the probability of event 1 (resp. event 2) is at most $\epsilon_1 + \epsilon_2$. Note that for each user, there are at most q_i^2 pairs of (a, b). So they summed among u users and got

$$\Pr[\mathrm{Bad}_9] \le \Sigma_{i=1}^u q_i^2(\epsilon_1 + \epsilon_2)^2 \le q^2(\epsilon_1 + \epsilon_2)^2.$$

In fact, even if $K_{h,1}^i, K_{h,2}^i$ are independent of each other, the event 1 and event 2 may not be independent, which has been shown in counter-examples 1-3. We regard the ninth event as the union set of events 3,4,5 and 6. Event 3 holds with probability at most ϵ_2^2 by the assumption that H^1 and H^2 are ϵ_2-AU. Event 4 holds with probability at most $\epsilon_1\epsilon_2$ by the assumption that H^1 is ϵ_2-AU and H^2 is ϵ_1-regular. Event 5 holds with probability at most $\epsilon_1\epsilon_2$ by the assumption that H^1 is ϵ_1-regular and H^2 is ϵ_2-AU. For event 6,

$$\begin{aligned}
&\Pr[K_{h,1}^i \xleftarrow{\$} \mathcal{K}_{h,1}, K_{h,2}^i \xleftarrow{\$} \mathcal{K}_{h,2} : \Sigma_a^i = \Lambda_b^i, \Lambda_a^i = \Sigma_b^i] \\
=&\Pr[K_{h,1}^i \xleftarrow{\$} \mathcal{K}_{h,1}, K_{h,2}^i \xleftarrow{\$} \mathcal{K}_{h,2} : \Sigma_a^i = \Lambda_b^i | \Lambda_a^i = \Sigma_b^i] \\
&\cdot \Pr[K_{h,1}^i \xleftarrow{\$} \mathcal{K}_{h,2}, K_{h,2}^i \xleftarrow{\$} \mathcal{K}_{h,1} : \Lambda_a^i = \Sigma_b^i] \\
\le&\epsilon_3\epsilon_1
\end{aligned}$$

by the assumption that H^2 is ϵ_1-regular and let

$$\epsilon_3 = \Pr[K_{h,1}^i \xleftarrow{\$} \mathcal{K}_{h,1}, K_{h,2}^i \xleftarrow{\$} \mathcal{K}_{h,2} : \Sigma_a^i = \Lambda_b^i | \Lambda_a^i = \Sigma_b^i].$$

So we sum among u users and got

$$\Pr[\mathrm{Bad}_9] \le \Sigma_{i=1}^u q_i^2(\epsilon_2^2 + 2\epsilon_1\epsilon_2 + \epsilon_3\epsilon_1) \le q^2(\epsilon_2^2 + 2\epsilon_1\epsilon_2 + \epsilon_3\epsilon_1).$$

For counter-examples 1-3, it is easy to get $\epsilon_3 = 1$. So for these cases, $\Pr[\mathrm{Bad}_9] \le q^2(\epsilon_2^2 + 2\epsilon_1\epsilon_2 + \epsilon_1)$. If we substitute our $\Pr[\mathrm{Bad}_9]$ for that in paper [15], we get the security of proofs of counter-examples 1-3 should be within $\mathcal{O}(2^{n/2})$ evaluation queries assuming ideal-cipher queries are $\mathcal{O}(1)$ and the maximal block length of all evaluation queries is $\mathcal{O}(1)$, which is consistent with attacks.

6 Conclusion

In this paper, we find a critical flaw of the security framework of two-key DbHtS in the multi-user setting raised by Shen et al. [15] by three counter-examples. We also present the reason of existing such a flaw. This is due to the fact that the authors overlooked the dependence of $H_{K_{h_1}}(M_1) = H_{K_{h_2}}(M_2)$ and $H_{K_{h_2}}(M_1) = H_{K_{h_1}}(M_2)$ when K_{h_1}, K_{h_2} are independent and M_1, M_2 are two different messages in the proof of Theorem 1 [15]. In their paper, they also stated 2k-SUM-ECBC, 2k-PMAC_Plus, and 2k-LightMAC_Plus all achieve $2n/3$-bit security. For 2k-SUM-ECBC based on two independent CBC MACs, the probability ϵ_3 is about $\frac{1}{2^n}$. So if we substitute our $\Pr[\mathrm{Bad}_9]$ for that in paper [15], 2k-SUM-ECBC still achieves $2n/3$ security. The two universal hash functions of 2k-PMAC_Plus or 2k-LightMAC_Plus are dependent, they adjusted the concrete proof of these two MACs from the framework. So we haven't found attacks against these three MACs.

References

1. Banik, S., Pandey, S.K., Peyrin, T., Sasaki, Y., Sim, S.M., Todo, Y.: GIFT: A small present - towards reaching the limit of lightweight encryption. In: CHES 2017. vol. 10529, pp. 321–345. Springer (2017), https://doi.org/10.1007/978-3-319-66787-4_16
2. Bellare, M.: New proofs for NMAC and HMAC: security without collision-resistance. In: CRYPTO 2006, Proceedings. LNCS, vol. 4117, pp. 602–619. Springer (2006), https://doi.org/10.1007/11818175_36
3. Bellare, M., Pietrzak, K., Rogaway, P.: Improved security analyses for CBC macs. In: CRYPTO 2005. vol. 3621, pp. 527–545. Springer (2005), https://doi.org/10.1007/11535218_32
4. Black, J., Rogaway, P.: CBC macs for arbitrary-length messages: The three-key constructions. In: CRYPTO 2000, Proceedings. LNCS, vol. 1880, pp. 197–215. Springer (2000), https://doi.org/10.1007/3-540-44598-6_12
5. Black, J., Rogaway, P.: A block-cipher mode of operation for parallelizable message authentication. In: EUROCRYPT 2002, Proceedings. LNCS, vol. 2332, pp. 384–397. Springer (2002), https://doi.org/10.1007/3-540-46035-7_25
6. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: CHES 2007. vol. 4727, pp. 450–466. Springer (2007), https://doi.org/10.1007/978-3-540-74735-2_31
7. Chen, S., Steinberger, J.P.: Tight security bounds for key-alternating ciphers. In: EUROCRYPT 2014. vol. 8441, pp. 327–350. Springer (2014), https://doi.org/10.1007/978-3-642-55220-5_19
8. Datta, N., Dutta, A., Nandi, M., Paul, G.: Double-block hash-then-sum: A paradigm for constructing BBB secure PRF. IACR Trans. Symmetric Cryptol. **2018**(3), 36–92 (2018), https://doi.org/10.13154/tosc.v2018.i3.36-92

9. Jha, A., Nandi, M.: Revisiting structure graph and its applications to CBC-MAC and EMAC. IACR Cryptol. ePrint Arch. p. 161 (2016), http://eprint.iacr.org/2016/161
10. Kim, S., Lee, B., Lee, J.: Tight security bounds for Double-Block Hash-then-Sum MACs. In: EUROCRYPT 2020. vol. 12105, pp. 435–465. Springer (2020), https://doi.org/10.1007/978-3-030-45721-1_16
11. Leurent, G., Nandi, M., Sibleyras, F.: Generic attacks against beyond-birthday-bound macs. In: CRYPTO 2018. vol. 10991, pp. 306–336. Springer (2018), https://doi.org/10.1007/978-3-319-96884-1_11
12. Naito, Y.: Blockcipher-based macs: Beyond the birthday bound without message length. In: ASIACRYPT 2017. vol. 10626, pp. 446–470. Springer (2017), https://doi.org/10.1007/978-3-319-70700-6_16
13. Patarin, J.: The "coefficients h" technique. In: SAC 2008. vol. 5381, pp. 328–345. Springer (2008), https://doi.org/10.1007/978-3-642-04159-4_21
14. Rogaway, P.: Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC. In: ASIACRYPT 2004, Proceedings. LNCS, vol. 3329, pp. 16–31. Springer (2004), https://doi.org/10.1007/978-3-540-30539-2_2
15. Shen, Y., Wang, L., Gu, D., Weng, J.: Revisiting the security of dbhts macs: Beyond-birthday-bound in the multi-user setting. In: CRYPTO 2021. vol. 12827, pp. 309–336. Springer (2021), https://doi.org/10.1007/978-3-030-84252-9_11
16. Shoup, V.: Sequences of games: a tool for taming complexity in security proofs. IACR Cryptol. ePrint Arch. p. 332 (2004), http://eprint.iacr.org/2004/332
17. Yasuda, K.: The sum of CBC macs is a secure PRF. In: CT-RSA 2010. vol. 5985, pp. 366–381. Springer (2010), https://doi.org/10.1007/978-3-642-11925-5_25
18. Yasuda, K.: A new variant of PMAC: beyond the birthday bound. In: CRYPTO 2011. vol. 6841, pp. 596–609. Springer (2011), https://doi.org/10.1007/978-3-642-22792-9_34
19. Zhang, L., Wu, W., Sui, H., Wang, P.: 3kf9: Enhancing 3gpp-mac beyond the birthday bound. In: ASIACRYPT 2012. vol. 7658, pp. 296–312. Springer (2012), https://doi.org/10.1007/978-3-642-34961-4_19