# Constructing Rate-1 MACs from Related-Key Unpredictable Block Ciphers: PGV Model Revisited

Liting Zhang, Wenling Wu, Peng Wang, Lei Zhang,
Shuang Wu, and Bo Liang

State Key Laboratory of Information Security
Institute of Software, Chinese Academy of Sciences, Beijing 100190, P.R. China
Graduate University of Chinese Academy of Sciences, Beijing 100049, P.R. China
{zhangliting,wwl,zhanglei1015,wushuang,liangb}@is.iscas.ac.cn, wp@is.ac.cn

**Abstract.** Almost all current block-cipher-based MACs reduce their security to the pseudorandomness of their underlying block ciphers, except for a few of them to the unpredictability, a strictly weaker security notion than pseudorandomness. However, the latter MACs offer relatively low efficiency. In this paper, we investigate the feasibility of constructing rate-1 MACs from related-key unpredictable block ciphers. First, we show all the existing rate-1 MACs are insecure when instantiated with a special kind of related-key unpredictable block cipher. The attacks on them inspire us to propose an assumption that all the chaining values are available to adversaries for theoretically analyzing such MACs. Under this assumption, we study the security of 64 rate-1 MACs in keyed PGV model, and find that 1) 15 MACs are meaningless; 2) 25 MACs are vulnerable to three kinds of attacks respectively and 3) 24 MACs are provably secure when their underlying block ciphers are related-key unpredictable. Furthermore, we refine these 24 provably secure rate-1 MACs in Compact PGV model by removing a useless parameter away, and find that the resulting 6 provably secure MACs are in fact equivalent to each other. In the aspect of efficiency, however, the low rate of these secure MACs does not necessarily mean they can run faster than none rate-1 one MACs, due to their large number of key schedules.

**Key words:** Message Authentication Code, Block Cipher, Mode of Operation, Provable Security

## 1 Introduction

### 1.1 Background

In cryptography, block ciphers are symmetric-key primitives, and they can only handle fixed-length messages, such as AES [1]. In order to handle variable-length messages and reach different kinds of security targets, modes of operation for them are proposed, such as authentication modes, encryption modes and authenticated encryption modes.

In this paper, we focus on the design of authentication modes, or block-cipher-based Message Authentication Codes. MACs are widely used to protect data integrity and data origin authentication in communications. To use a MAC, the sender and receiver should share a secret key $K$ beforehand. When sending a message $M$, the sender computes $T \leftarrow \mathrm{MAC}(K, M)$ as a tag, and then sends $(M, T)$ out. On receipt of a pair $(M, T)$, the receiver computes $T' \leftarrow \mathrm{MAC}(K, M)$, and deems message $M$ to be valid only if $T = T'$. The security of a MAC algorithm is evaluated by how unpredictable it is. Informally speaking, an adversary $\mathcal{A}$ has access to the MAC algorithm, whose key is randomly selected and kept secret from $\mathcal{A}$. $\mathcal{A}$ can query the MAC with any message in the domain, and receives the corresponding tags; in the end, $\mathcal{A}$ is asked to make a forgery, i.e. to output a pair $(M', T')$ such that 1) $M'$ was never queried to the MAC algorithm by $\mathcal{A}$ and 2) $T'$ is the tag of $M'$. The success probability for $\mathcal{A}$ to do this is called $\mathcal{A}$'s advantage, and the MAC algorithm is deemed to be secure if all the advantages of reasonably restricted adversaries are sufficiently small.

The history of block-cipher-based MACs dates back as early as to CBC-MAC [2]. Although it is secure for fixed-length messages when its underlying block cipher is a PseudoRandom Permutation (PRP) , it is not secure for variable-length messages [3]. Later, several variants of CBC-MAC were proposed to fix this flaw, and usual solutions include different initial and output transformations for CBC-MAC, as suggested in the ISO standard [4]. Furthermore, EMAC [5] and RMAC [6] appends an extra block-cipher invocation at the end of CBC-MAC; XCBC [7] adds secret sub-keys to the last message block; TMAC [8], OMAC [9] and CMAC [10] [1] improve XCBC by taking different sub-key deriving methods. Recently, GCBC [11] was proposed as a generalization of XCBC, TMAC and OMAC, and it avoids length-extension attacks by applying shift operations to chaining values. Besides these, f9 [12] sums the chaining values in CBC structure up and also takes an extra block-cipher invocation in the end, while PMAC [13] takes a parallel structure other than CBC structure, and it adds distinct secret masks to message blocks to ensure the security. All these later-proposed block-cipher-based MACs are highly efficient, and can be classified into rate-1 MACs. [2]

Nevertheless, the provable security of these MACs is based on the assumption that their underlying block ciphers are PseudoRandom Permutations (PRPs) or even Related-Key PseudoRandom Permutations (RK-PRPs). Recall that the security goal for MACs is only unpredictability, and it is strictly weaker than pseudorandomness (we will give an example in Section 3); so, it is desirable to reduce the provable security of MACs to the unpredictability of their underlying block ciphers, other than the pseudorandomness. On the other hand, practical block ciphers seem to be less secure than expected [16, 17], and this depressing fact makes it much more reasonable to reduce MAC security to the unpredictability other than the pseudorandomness of the block cipher.

[1] CMAC belongs to OMAC family; more specifically, it is OMAC1.

[2] Rate is the average number of block-cipher invocations per message block [14, 15].

As far as we know, reducing MAC security to unpredictable primitives is first studied by An and Bellare [18], and later works include [19–24]; however, all those constructions are based on compression functions, while using length-preserving primitives (e.g. block ciphers) to do this initiated by Dodis et al. They proposed enciphered CBC mode [14] and SS-NMAC mode [15] to address this problem. These two MACs are not only provably secure based on unpredictable block ciphers, but also provably secure against Side Channel Attacks (SCAs) as long as their underlying block ciphers are secure against SCAs; unfortunately, their rates are as much as 2 or 3, and this implies they can only offer relatively low efficiency.

Then, there comes a question — *How about the security of rate-1 MACs based on unpredictable block ciphers?*.

### 1.2 Our Work

In this paper, we try to answer this question in two aspects. First, we investigate the security of current rate-1 MACs when they are instantiated with related-key unpredictable block ciphers, and find that they are all insecure by constructing a special related-key unpredictable block cipher. Our attacks on them show that the chaining values of those MACs can hardly be kept secret from adversaries, which is fatal to their security as MACs; then, we propose a natural assumption — *to study the security of MACs based on unpredictable block ciphers, assume all their chaining values are available to adversaries.*

Under this assumption, we try to construct rate-1 MACs in PGV model, which was proposed by Preneel, Govaerts and Vandewalle to study the security of block-cipher-based hash functions [25]. Since MACs can be seen as keyed hash functions, PGV model is naturally suitable to discuss MAC constructions after being equipped with a secret key $K$, as shown in Fig. 1.

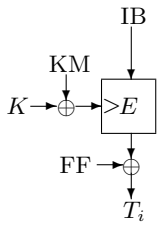


**Fig. 1.** In the keyed PGV model, a basic function $f(K, M_i, T_{i-1})$ is defined as $f(K, M_i, T_{i-1}) = E(K \oplus \mathrm{KM}, \mathrm{IB}) \oplus \mathrm{FF}$, where $K \xleftarrow{\$} \mathcal{K}_E$ and IB, KM, FF$\in \{M_i, T_{i-1}, M_i \oplus T_{i-1}, \mathrm{Cst}\}$.

In the keyed PGV model ($K$-PGV for short), there are three kinds of inputs for a block cipher $E$, i.e. an Input Block IB, a Key Mask KM and a FeedForward

FF, each of which have four choices, i.e. the current message block $M_i$, the last chaining value $T_{i-1}$, their sum $M_i \oplus T_{i-1}$ and a constant Cst. Without loss of generality, we assume $T_0 = \text{Cst}$ and all these four kinds of values and the secret key $K$ have the same length as the block size of $E$. Moreover, we restrict the secret key $K$ to be at the exact position where the block cipher key should be, because it is dangerous to take it as other inputs of block ciphers (IB and FF), even when the block ciphers are assumed to be pseudorandom [26]. $K$-PGV model gives us $4^3 = 64$ rate-1 MACs, among which we find

1) 15 MACs are meaningless, because their inputs are independent of either $M_i$ or $T_{i-1}$;
2) 6 MACs are vulnerable to fixed-$M$ attack;
3) 6 MACs are vulnerable to fixed-$T$ attack;
4) 13 MACs are vulnerable to fixed-$(M \oplus T)$ attack;
5) the remaining 24 MACs are provably secure on the assumption that their underlying block ciphers are independently unpredictable for different keys (or RK-UPs as we will define in Section 2).

Furthermore, we find that FF in fact has no influence over the security of these MACs, so we propose the Compact PGV model in which FF is removed from $K$-PGV model away. In the new model, we have six provably secure MACs, all of which are equivalent to each other in the sense that their basic functions can be transformed into one another by some invertible $2 \times 2$ matrix over GF(2). Unfortunately, this equivalence implies the security of the six MACs affects each other. That is, if one MAC is used with a secret key $K$, adversaries can easily make forgeries against all the other five MACs with the same key $K$, although the other five may never be used with $K$ before. This can be seen as a related-mode attack introduced by Phan and Siddiqi [27]. To avoid this attack, we break these six MACs into three groups, in each of which the two MACs can take distinct-and-fixed initial value $T_0$ to ensure the security with each other. As we will prove, by taking distinct-and-fixed $T_0$, the two MACs in the same group are in fact independent of each other.

### 1.3 Related Works

In PGV model, Preneel et al study the security of 64 block-cipher-based hash functions from the attackers' point of view, and conclude that 4 schemes are secure and 8 more are less secure, while other schemes are vulnerable to different kinds of attacks [25]. Then, Black et al review these hash functions by provable security techniques [28], and show that the Preneel's 12 schemes are really secure, and 8 more are provably secure with larger security bounds, while other schemes are not. Interestingly, the 24 secure MAC constructions found in $K$-PGV model include the previous 20 secure hash constructions (after being equipped with a secret key $K$), and 4 more schemes are also provably secure as MACs, because here adversaries are not allowed to make inverse queries to block ciphers, different from that of [28]. More clear relationships are illustrated in Table 1 of Section 4.

The rest of this paper is organized as follows: section 2 introduces the symbols and security notions we will use in this paper; section 3 gives detailed attacks on current rate-1 MACs by constructing a special unpredictable block cipher; section 4 lists the results we obtain from $K$-PGV model and section 5 investigates MAC security and their relationships in Compact PGV model. At last, section 6 concludes the full paper.

## 2 Preliminaries

**Symbols.** Suppose $A$ is a set, then $\#A$ denotes the size of set $A$, and $x \xleftarrow{\$} A$ denotes that $x$ is chosen from set $A$ uniformly at random. If $a, b \in \{0,1\}^*$ are strings of equal length then $a \oplus b$ is their bitwise XOR. If $a, b \in \{0,1\}^*$ are strings, then $a||b$ denotes their concatenation. Sometimes, we write $ab$ for $a||b$ if there is no confusion. Furthermore, $\mathrm{msb}_i(a)$ stands for the most significant $i$ bits of $a$, and $\mathrm{lsb}_i(a)$ stands for the least significant $i$ bits of $a$. If $M \in \{0,1\}^*$ is a string then $|M|$ stands for its length in bits, and we let $\mathbf{pad}(M) = M10^{n-1-(|M| \bmod n)} = M_1 M_2 \cdots M_l$, where $|M_i| = n$ for $1 \le i \le l$.

**Security Definitions.** Denote $\mathrm{Perm}(n)$ as the set containing all the permutations over $\{0,1\}^n$. An adversary $\mathcal{A}$ is an algorithm with an oracle. $\mathcal{A}$ can query the oracle with any message in the domain, but should not repeat a query. For a block cipher $E : \mathcal{K}_E \times \{0,1\}^n \to \{0,1\}^n$, and a function family $F : \mathcal{K}_F \times \{0,1\}^* \to \{0,1\}^n$, the security notions of prp and mac are listed below, where the maximum is taken over computation time at most $t$, oracle queries at most $q$, and the aggregate length of queries at most $\sigma$ blocks. In the mac security notions, the event adversary $\mathcal{A}^{F(K,\cdot)}$ forges means $\mathcal{A}$ outputs a pair $(M', T')$ such that $F(K, M') = T'$ and $M'$ was never queried to $F(K, \cdot)$ by $\mathcal{A}$.

$$\begin{cases} \mathbf{Adv}_E^{\mathrm{prp}}(\mathcal{A}) \stackrel{\mathrm{def}}{=} |\Pr[K \xleftarrow{\$} \mathcal{K}_E : \mathcal{A}^{E(K,\cdot)} = 1] - \Pr[P \xleftarrow{\$} \mathrm{Perm}(n) : \mathcal{A}^{P(\cdot)} = 1]|, \\ \mathbf{Adv}_E^{\mathrm{prp}}(t,q,\sigma) \stackrel{\mathrm{def}}{=} \max_{\mathcal{A}}\{\mathbf{Adv}_E^{\mathrm{prp}}(\mathcal{A})\}. \end{cases}$$

$$\begin{cases} \mathbf{Adv}_F^{\mathrm{mac}}(\mathcal{A}) \stackrel{\mathrm{def}}{=} \Pr[K \xleftarrow{\$} \mathcal{K}_F : \mathcal{A}^{F(K,\cdot)} \text{ forges}], \\ \mathbf{Adv}_F^{\mathrm{mac}}(t,q,\sigma) \stackrel{\mathrm{def}}{=} \max_{\mathcal{A}}\{\mathbf{Adv}_F^{\mathrm{mac}}(\mathcal{A})\}. \end{cases}$$

More details about these two security notions can be found in [29, 3].

Next, we define the unpredictability of a block cipher $E : \mathcal{K}_E \times \{0,1\}^n \to \{0,1\}^n$ under related-key chosen message attack. A Related-Key-Deriving (RKD) function $\phi \in \Phi$ is a map $\phi : \mathcal{K}_E \to \mathcal{K}_E$, where $\Phi$ is a set of functions mapping $\mathcal{K}_E$ to $\mathcal{K}_E$. Then, for a block cipher $E : \mathcal{K}_E \times \{0,1\}^n \to \{0,1\}^n$ and a RKD function family $\Phi : \mathcal{K}_E \to \mathcal{K}_E$, consider the following experiment:

Experiment $\mathbf{Exp}_{E,\mathcal{A}}^{\text{rk-up}}$
$K \xleftarrow{\$} \mathcal{K}_E$;
while $\mathcal{A}$ makes a query $(\phi, M)$ to $E(K, \cdot)$, do
    $T \leftarrow E(\phi(K), M)$; return $T$ to $\mathcal{A}$;
until $\mathcal{A}$ stops and outputs $(\phi', M', T')$ such that
    1) $E(\phi'(K), M') = T'$;
    2) $(\phi', M')$ was never queried to $E(K, \cdot)$;
then return 1 else return 0.

Define

$$\begin{cases} \mathbf{Adv}_E^{\text{rk-up}}(\mathcal{A}) \stackrel{\text{def}}{=} \Pr[\mathbf{Exp}_{E,\mathcal{A}}^{\text{rk-up}} = 1], \\ \mathbf{Adv}_E^{\text{rk-up}}(t, q, \mu) \stackrel{\text{def}}{=} \max_{\mathcal{A}}\{\mathbf{Adv}_E^{\text{rk-up}}(\mathcal{A})\}, \end{cases}$$

where the maximum is taken over computation time at most $t$, oracle queries at most $q$, whose total length is at most $\mu$. If $\mathbf{Adv}_E^{\text{rk-up}}(t, q, \mu)$ is sufficiently small, we say block cipher $E : \mathcal{K}_E \times \{0,1\}^n \to \{0,1\}^n$ is secure against $\Phi$-restricted related-key chosen message attack.

*Remark 1.* The way to define rk-up is similar to that of prp-rka, which was proposed by Bellare et al to theoretically study the pseudorandomness of block ciphers under related-key attacks [30]. Nevertheless, rk-up is strictly weaker than prp-rka since unpredictability is strictly weaker than pseudorandomness.

*Remark 2.* The RKD function family $\Phi$ plays an important role in rk-up security. If $\Phi$ is not properly restricted, there may be no rk-up secure block ciphers. For example, we consider a special RKD function family $\Phi_K^{\text{Cst}} = \{\phi | \phi : \mathcal{K}_E \to \text{Cst}\}$. That is, for any $K \xleftarrow{\$} \mathcal{K}_E$ and $\phi \in \Phi_K^{\text{Cst}}$, we have $\phi(K) = \text{Cst}$. Obviously, any block cipher under such a $\Phi_K^{\text{Cst}}$-restricted related-key attack is easy to predict, not to mention rk-up security. For more discussions about $\Phi$, refer to [30, 31].

Almost all current block-cipher-based MACs take only one secret key for their underlying block ciphers, except for a few of them who aim to get higher security by taking more than one block-cipher keys, e.g. RMAC [6], f9 [12] and some in the ISO standards [4]. In RMAC, the authors suggest the second block-cipher key can be obtained by $K_2 \oplus R$, where $K_2$ is a secret key for all messages and $R$ is a random value for only one message; while f9 just lets $K_2 = K_1 \oplus \text{Cst}$ to obtain the second block-cipher key. In this paper, we only consider such a commonly used RKD function family $\Phi_K^{\oplus} = \{\text{XOR}_{\text{KM}} | \text{XOR}_{\text{KM}} : K \to K \oplus \text{KM}, \text{KM} \in \{0,1\}^n\}$. Then, any $\Phi_K^{\oplus}$-restricted adversary $\mathcal{A}$ attacking the rk-up security of $E$ has access to an oracle $E(K \oplus \cdot, \cdot)$ with $K \xleftarrow{\$} \mathcal{K}_E$, who will accept queries (KM,$M$) $\in \{0,1\}^n \times \{0,1\}^n$ and returns the tag $T \leftarrow E(K \oplus \text{KM}, M)$ to $\mathcal{A}$. At last, $\mathcal{A}$ is asked to output a three-tuple (KM′,$M'$, $T'$) such that 1) (KM′,$M'$) was never queried to $E(K \oplus \cdot, \cdot)$ by $\mathcal{A}$ and 2) $T' = E(K \oplus \text{KM}', M')$. If all reasonably restricted adversaries can do this within sufficiently small probability, we say block cipher $E$ is secure against $\Phi_K^{\oplus}$-restricted related-key chosen message attack.

For simplicity, we directly say $E$ is rk-up secure in the rest of this paper, without pointing out that all adversaries attacking $E$ are $\Phi_K^{\oplus}$-restricted, and we denote $E$ as RK-UP (Related-Key Unpredictable Permutation).

## 3 Attacks on Current Rate-1 MACs

The provable security of current rate-1 MACs relies on the assumption that their underlying block ciphers are PRPs or RK-PRPs. In this section, we give detailed attacks to show that their provable security can no longer exist if their underlying block ciphers are only RK-UPs.

The idea comes from [18], in which An and Bellare show the basic CBC-MAC does not hold unpredictability. In our attacks, we first construct a special block cipher $E' : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ that is rk-up secure, but not pseudorandom, and then give attacks against the unpredictability of current rate-1 MACs instantiated with $E'$,

$$E'(K, M) = \begin{cases} m_1||m_2||m_3||c, & \text{if } \mathrm{msb}_1(m_1) = 0, \\ m_1||c||m_3||m_4, & \text{if } \mathrm{msb}_1(m_1) = 1, \end{cases}$$

where $M = m_1||m_2||m_3||m_4$, $|m_i| = n/4$ for $1 \le i \le 4$, $c = \mathrm{CBC}[Q_K](m_1m_2m_3m_4)$ and $Q : \mathcal{K} \times \{0,1\}^{n/4} \to \{0,1\}^{n/4}$ is a block cipher with RK-PRP security. Notice that $c$ is obtained by applying a RK-PRP $Q_K$ to $m_1m_2m_3m_4$ in Cipher-Block-Chaining mode, which has been proved to hold pseudorandomness when its inputs are of fixed-length [3]. So, $c$ is pseudorandom, and this indicates $E'(K, M)$ is rk-up secure; however, it is absolutely not pseudorandom since parts of its inputs are listed in the ciphertext directly.

Next, we give an attack on the unpredictability of XCBC [7] instantiated with $E'$. Notice that, to authenticate messages of length $ln$ bits, XCBC first deal with its first $l-1$ blocks by $\mathrm{CBC}[E'_{K_1}]$, and then XORs a secret sub-key $K_2$ and the last message block to the output of $\mathrm{CBC}[E'_{K_1}]$. Finally, it encrypts the sum by $E'_{K_1}$. The attack on $\mathrm{XCBC}_{E'}(\cdot)$ is as follows,

1) Adversary $\mathcal{A}$ queries $\mathrm{XCBC}_{E'}(\cdot)$ with $0^n$, obtains the tag $T^1 = t_1^1 t_2^1 t_3^1 t_4^1$;
2) $\mathcal{A}$ queries $\mathrm{XCBC}_{E'}(\cdot)$ with $10^{n-1}$, obtains the tag $T^2 = t_1^2 t_2^2 t_3^2 t_4^2$;
3) $\mathcal{A}$ makes a forgery $(M', T^1)$, where

$$\begin{cases} M' = (t_1^1 t_2^1 t_3^1 t_4^2)||T^1, & \text{if } \mathrm{msb}_1(t_1^1) = 0, \\ M' = (t_1^2 t_2^2 t_3^2 t_4^1)||T^1, & \text{if } \mathrm{msb}_1(t_1^1) = 1. \end{cases}$$

By the definitions of $E'$ and XCBC, it is easy to get that the secret sub-key $K_2$ in XCBC is $(t_1^1 t_2^1 t_3^1 t_4^2)$ if $\mathrm{msb}_1(t_1^1) = 0$ or $(t_1^2 t_2^2 t_3^2 t_4^1)$ if $\mathrm{msb}_1(t_1^1) = 1$ after the above attack. Then, the validity of the forgery is obvious. Since TMAC [8], OMAC [9] and CMAC (OMAC1) [10] are variants of XCBC by taking different sub-key deriving methods, the same attack applies to them as well. What is more, the other existing rate-1 MACs are also vulnerable when instantiated with $E'$, and we describe the attacks on them in Appendix A.

The reason behind the insecurity of these MACs instantiated with $E'$ is that the secrecy of their chaining values can no longer be kept; so, we propose the following assumption,

**Assumption:** ***To study the security of MACs based on unpredictable block ciphers, assume all their chaining values are available to adversaries.***

This assumption gives much more power to the attackers than that in the usual black-box model [3], and it may even overkill the current rate-1 MACs; however, it indeed explain why the existing rate-1 MACs are no longer secure when their underlying block ciphers are only RK-UPs, and also it helps to understand why SS-NMAC is provably secure against SCAs as long as its underlying block ciphers are [15].

Moreover, this assumption affects the security definition of MACs a little. That is, under such an assumption adversaries should not forge with a message which after being padded is a prefix of a queried message, although the forgery message may never be queried to the MACs before. This seems to bring trouble into MAC security; however, we can apply prefix-free encoding to messages and it is easy to achieve by simply prepending each message with a block denoting its length in bits, as suggested in [32].

## 4 Rate-1 MACs from $K$-PGV Model

In this section, we consider the feasibility of constructing rate-1 MACs from RK-UPs in $K$-PGV model. As shown in Fig. 1, $K$-PGV model gives us 64 basic functions $f_s(K, M_i, T_{i-1})$ $(s = 1, 2, \cdots, 64)$, all of which can be used in an iterative way to construct MACs $F_s(K, M)$ who can authenticate arbitrary-length messages. Without loss of generality, we assume $\mathbf{pad}(M) = M_1 M_2 \cdots M_l$; then, $F_s(K, M)$ is defined as follows,

MAC $F_s(K, M)$
$K \xleftarrow{\$} \mathcal{K}_E$;
**for** $i = 1$ **to** $l$ **do** $T_i \leftarrow f_s(K, M_i, T_{i-1})$ **end for**
**return** $T_l$.

Next, we study the security of $F_s(K, M)$ as MACs and find the main results as follows, while the details are listed in Table. 1.

1) 15 MACs are meaningless, because their inputs are independent of either $M_i$ or $T_{i-1}$;
2) 6 MACs are vulnerable to attack 1 — fixed-$M$ attack, who can make a forgery for $M^2$ by simply choosing any queried message $M^1$, where $\mathbf{pad}(M^1) = M_1^1 M_2^1 \cdots M_l^1$, and let $\mathbf{pad}(M^2) = \mathbf{pad}(M^1)||M_l^1$;
3) 6 MACs are vulnerable to attack 2 — fixed-$T$ attack, who can forge with $M^2$, where $\mathbf{pad}(M^2) = M_1^1 M_2^1 \cdots M_{l-1}^1||(M_l^1 \oplus \Delta)$ and $\Delta$ can be any non-zero value in $\{0,1\}^n$;

**Table 1.** The security of 64 MACs from $K$-PGV model. "–" means the MAC is meaningless because its inputs are independent of either $M_i$ or $T_{i-1}$; a number $i$ ($i = 1, 2, 3$) means the MAC is vulnerable to attack $i$; the MACs marked with $f_i$ ($i = 1, 2, \cdots, 24$) are provably secure.

| | | choice of IB | | | |
|---|---|---|---|---|---|
| choice of KM | choice of FF | $M_i$ | $T_{i-1}$ | $M_i \oplus T_{i-1}$ | Cst |
| $M_i$ | $M_i$ | – | $f_{17}$ | $f_{20}$ | – |
| | $T_{i-1}$ | 1 | $f_5$ | $f_8$ | 1 |
| | $M_i \oplus T_{i-1}$ | 1 | $f_7$ | $f_6$ | 1 |
| | Cst | – | $f_{15}$ | $f_{19}$ | – |
| $T_{i-1}$ | $M_i$ | $f_1$ | 2 | $f_4$ | 2 |
| | $T_{i-1}$ | $\boxed{f_{21}}$ | – | $\boxed{f_{24}}$ | – |
| | $M_i \oplus T_{i-1}$ | $f_3$ | 2 | $f_2$ | 2 |
| | Cst | $\boxed{f_{23}}$ | – | $\boxed{f_{22}}$ | – |
| $M_i \oplus T_{i-1}$ | $M_i$ | $f_9$ | $f_{12}$ | 3 | 3 |
| | $T_{i-1}$ | $f_{11}$ | $f_{10}$ | 3 | 3 |
| | $M_i \oplus T_{i-1}$ | $f_{14}$ | $f_{18}$ | 3 | 3 |
| | Cst | $f_{13}$ | $f_{16}$ | 3 | 3 |
| Cst | $M_i$ | – | 2 | 3 | – |
| | $T_{i-1}$ | 1 | – | 3 | – |
| | $M_i \oplus T_{i-1}$ | 1 | 2 | 3 | 3 |
| | Cst | – | – | 3 | – |

4) 13 MACs are vulnerable to attack 3 — fixed-$(M \oplus T)$ attack, who can forge with $M^2$, where $\mathbf{pad}(M^2) = \mathbf{pad}(M^1)||(M_l^1 \oplus T_l^1 \oplus T_{l-1}^1)$;
5) 24 MACs are provably secure, on the assumption that their underlying block cipher is rk-up secure. We will prove this in Theorem 1.

We also find that all the MACs with a fixed key $K \oplus \text{Cst}$ (KM = Cst) are either insecure or meaningless, and this implies within this model, it is impossible to construct a rate-1 MAC from only unpredictable block ciphers.

The basic functions in the 24 provably secure MACs are marked as $f_i$ ($i = 1, 2, \cdots, 24$), of which the first 20 (being removed the secret key $K$ away) are the exact compression functions of the 20 provably secure hash functions [28]. The extra 4 ($f_{21}, f_{22}, f_{23}, f_{24}$) can be used to construct provably secure MACs (with $K$), but not hash functions (without $K$). The reason is that, in attacks on MACs adversaries are not allowed to make inverse queries to block cipher $E$, but they can do this in attacks on hash functions, since the latter is considered within the ideal cipher model [33, 28].

**Theorem 1.** *Suppose the underlying block cipher $E : \mathcal{K}_E \times \{0,1\}^n \times \{0,1\}^n$ is rk-up secure, then $F_s[E]$ ($s = 1, 2, \cdots, 24$) is provably secure for prefix-free messages. More concretely, we have*

$$\mathbf{Adv}^{\text{mac}}_{F_s[E]}(t, q, \mu) \leq (\sigma^2 - \sigma + 1)\mathbf{Adv}^{\text{rk-up}}_{E}(t', q', \mu'),$$

*where $\sigma$ is the total block length of all $q$ queried messages plus the block length of the forgery message, $t' = t + O(\sigma)$, $q' = \sigma - 1$, $\mu' = \mu + O(\sigma)$.*

*Proof.* To upper bound the success probability for any adversary $\mathcal{A}$ attacking the mac security of $F_s[E]$, we construct an adversary $\mathcal{B}$ attacking the rk-up security of $E$. $\mathcal{B}$ will simulate $\mathcal{A}$'s oracle $F_s[E](\cdot)$ with its own oracle $\mathcal{O}_{\mathcal{B}}(\cdot,\cdot) = E(\cdot,\cdot)$ and the definition of $F_s$, as in Fig. 2. In either Game 0 or Game 1, $\mathcal{A}$ can make any prefix-free queries, get not only the corresponding tags but also the chaining values; at last, he is asked to make a forgery. However, the forgery message should not be a prefix of a queried message by the arguments in the end of Section 3.

| Game 0 | Game 1 |
|---|---|
| Range$\leftarrow \{T_0\}$; Collision$_w \leftarrow$False, for $w \geq 1$; $z \leftarrow 1$. | |
| when $\mathcal{A}$ makes a query $M^j$, where $\mathbf{Pad}(M^j) = M_1^j M_2^j \cdots M_{l_j}^j$, $j = 1, 2, \cdots, q$ | |
| 01. **for** $i = 1$ **to** $l_j$ **do** | |
| 02. renew KM, IB, FF with $(M_i^j, T_{i-1}^j, M_i^j \oplus T_{i-1}^j, \mathrm{Cst})$ by the definition of $f_s$; | |
| **03.** $T_i^j = \mathcal{O}_{\mathcal{B}}(K \oplus \mathrm{KM}, \mathrm{IB}) \oplus \mathrm{FF}$; | |
| **04.** **if** $T_i^j \in$Range and $\nexists j1 < j$ s.t. $M_1^{j1} M_2^{j1} \cdots M_{i-1}^{j1} = M_1^j M_2^j \cdots M_{i-1}^j$ | |
| **05.** **then** { Collision$_z \leftarrow$True; Stop. } | |
| **06.** **end if** | |
| 07. Range$\leftarrow$Range$\cup\{T_i^j\}$; $z \leftarrow z + 1$; return $T_i^j$ to $\mathcal{A}$; | |
| 08. **end for** | |
| when $\mathcal{A}$ makes a forgery $(M', T')$, where $\mathbf{Pad}(M') = M_1' M_2' \cdots M_{l'}'$ | |
| 11. **for** $i = 1$ **to** $l' - 1$ **do** | |
| 12. renew KM, IB, FF with $(M_i', T_{i-1}', M_i' \oplus T_{i-1}', \mathrm{Cst})$ by the definition of $f_s$; | |
| **13.** $T_i' = \mathcal{O}_{\mathcal{B}}(K \oplus \mathrm{KM}, \mathrm{IB}) \oplus \mathrm{FF}$; | |
| **14.** **if** $T_i' \in$Range and $\nexists j1 \in \{1, 2, \cdots, q\}$ s.t. $M_1^{j1} M_2^{j1} \cdots M_{i-1}^{j1} = M_1' M_2' \cdots M_{i-1}'$ | |
| **15.** **then** { Collision$_z \leftarrow$True; Stop. } | |
| **16.** **end if** | |
| 17. Range$\leftarrow$Range$\cup\{T_i'\}$; $z \leftarrow z + 1$; return $T_i'$ to $\mathcal{A}$; | |
| 18. **end for** | |
| **19.** renew KM, IB, FF with $(M_{l'}', T_{l'-1}', M_{l'}' \oplus T_{l'-1}', \mathrm{Cst})$ by the definition of $f_s$; | |
| 20. **if** $T' = \mathcal{O}_{\mathcal{B}}(K \oplus \mathrm{KM}, \mathrm{IB}) \oplus \mathrm{FF}$ return 1 **else** return 0 **end if** | |

**Fig. 2.** Definitions for Game 0 (excluding the boxed codes) and Game 1 (including the boxed codes), in which adversary $\mathcal{B}$ simulates $\mathcal{A}$'s oracle $F_s[E]$ with its own oracle $\mathcal{O}_{\mathcal{B}}(\cdot,\cdot) = E(\cdot,\cdot)$ and the definition of $F_s$, for $s = 1, 2, \cdots, 24$.

The only differences between Game 0 and Game 1 are the boxed codes in lines 05 and 15, where the flag Collision$_z$ would be true and then $\mathcal{B}$ would stop the simulation. We denote such an event by $\mathsf{Coll}_z$.

Let event $\mathsf{Coll}$ be $\mathsf{Coll}_1 \vee \mathsf{Coll}_2 \vee \cdots \vee \mathsf{Coll}_{\sigma-1}$. So, we get

$$
\begin{aligned}
&|\Pr[\mathcal{A} \text{ forges in Game } 0] - \Pr[\mathcal{A} \text{ forges in Game } 1]| \\
=\ &|\Pr[\mathcal{A} \text{ forges in Game } 0 \wedge \mathsf{Coll}] + \Pr[\mathcal{A} \text{ forges in Game } 0 \wedge \overline{\mathsf{Coll}}] \\
&- \Pr[\mathcal{A} \text{ forges in Game } 1 \wedge \mathsf{Coll}] - \Pr[\mathcal{A} \text{ forges in Game } 1 \wedge \overline{\mathsf{Coll}}]| \\
=\ &|\Pr[\mathcal{A} \text{ forges in Game } 0 \wedge \mathsf{Coll}] - \Pr[\mathcal{A} \text{ forges in Game } 1 \wedge \mathsf{Coll}]| \\
=\ &|\Pr[\mathcal{A} \text{ forges in Game } 0 | \mathsf{Coll}] - \Pr[\mathcal{A} \text{ forges in Game } 1 | \mathsf{Coll}]| \times \Pr[\mathsf{Coll}] \\
\leq\ &\Pr[\mathsf{Coll}]. \qquad (1)
\end{aligned}
$$

Furthermore, noticing that

$$
\begin{aligned}
&\Pr[\mathcal{A} \text{ forges in Game } 1] \\
=\ &\Pr[\mathcal{A} \text{ forges in Game } 1 | \mathsf{Coll}] \times \Pr[\mathsf{Coll}] \\
&+ \Pr[\mathcal{A} \text{ forges in Game } 1 | \overline{\mathsf{Coll}}] \times \Pr[\overline{\mathsf{Coll}}] \\
\leq\ &\Pr[\mathsf{Coll}] + \Pr[\mathcal{A} \text{ forges in Game } 1 | \overline{\mathsf{Coll}}], \qquad (2)
\end{aligned}
$$

by inequalities (1) and (2), we have

$$
\Pr[\mathcal{A} \text{ forges in Game } 0] \leq 2\Pr[\mathsf{Coll}] + \Pr[\mathcal{A} \text{ forges in Game } 1 | \overline{\mathsf{Coll}}] \qquad (3)
$$

Next, we show that the two items in the right side of inequality (3) are both sufficiently small, because the occurrence of either $\mathsf{Coll}$ or $[\mathcal{A}$ forges in Game $1|\overline{\mathsf{Coll}}]$ implies $\mathcal{B}$ can make a successful forgery against the rk-up security of $E$.

If $\mathsf{Coll}$ occurs, then at least one of $[\mathsf{Coll}_i|\overline{\mathsf{Coll}_{i-1}} \wedge \cdots \wedge \overline{\mathsf{Coll}_0}]$ for $i = 1, 2, \cdots, \sigma - 1$ occurs, where $\mathsf{Coll}_0$ is the null event. By this we let $\mathcal{B}$ select $T \in \text{Range}$ uniformly at random, and make a forgery $(\text{KM}, \text{IB}, T \oplus \text{FF})$ against the rk-up security of $E$, where $\text{KM}, \text{IB}, \text{FF}$ are from lines 03 and 13 of Fig. 2 at the moment $z = i$. Notice that event $[\mathsf{Coll}_i|\overline{\mathsf{Coll}_{i-1}} \wedge \cdots \wedge \overline{\mathsf{Coll}_0}]$ occurs implies $\mathcal{B}$'s forgery is valid; however, $E$ is rk-up secure by assumption, so we get $\Pr[\mathsf{Coll}_i|\overline{\mathsf{Coll}_{i-1}} \wedge \cdots \wedge \overline{\mathsf{Coll}_0}] \times \frac{1}{i} \leq \mathbf{Adv}_E^{\text{rk-up}}(t_{i-1}, i-1, \mu_{i-1})$, which implies $\Pr[\mathsf{Coll}_i|\overline{\mathsf{Coll}_{i-1}} \wedge \cdots \wedge \overline{\mathsf{Coll}_0}] \leq i \times \mathbf{Adv}_E^{\text{rk-up}}(t_{i-1}, i-1, \mu_{i-1})$. Then, we get

$$
\begin{aligned}
\Pr[\mathsf{Coll}] &= \Pr[\mathsf{Coll}_1 \vee \mathsf{Coll}_2 \vee \cdots \vee \mathsf{Coll}_{\sigma-1}] \\
&\leq \textstyle\sum_{i=1}^{\sigma-1} \Pr[\mathsf{Coll}_i|\overline{\mathsf{Coll}_{i-1}} \wedge \cdots \wedge \overline{\mathsf{Coll}_0}] \\
&\leq \textstyle\sum_{i=1}^{\sigma-1} (i \times \mathbf{Adv}_E^{\text{rk-up}}(t_{i-1}, i-1, \mu_{i-1})) \\
&\leq \textstyle\sum_{i=1}^{\sigma-1} i \times \mathbf{Adv}_E^{\text{rk-up}}(t_{\sigma-2}, \sigma-2, \mu_{\sigma-2}) \\
&= \frac{\sigma(\sigma-1)}{2} \mathbf{Adv}_E^{\text{rk-up}}(t_{\sigma-2}, \sigma-2, \mu_{\sigma-2}) \qquad (4)
\end{aligned}
$$

Similarly, if event $[\mathcal{A}$ forges in Game $1|\overline{\mathsf{Coll}}]$ occurs, we let $\mathcal{B}$ directly make a forgery $(\text{KM}, \text{IB}, T' \oplus \text{FF})$ against the rk-up security of $E$, where $\text{KM}, \text{IB}, \text{FF}$ are from line 19 in Fig. 2 and $T'$ is from $\mathcal{A}$'s forgery. Also, by assumption $E$ is rk-up secure, we have

$$
\Pr[\mathcal{A} \text{ forges in Game } 1|\overline{\mathsf{Coll}}] \leq \mathbf{Adv}_E^{\text{rk-up}}(t_{\sigma-1}, \sigma-1, \mu_{\sigma-1}) \qquad (5)
$$

Combining inequalities (3), (4) and (5), we know that for any adversary $\mathcal{A}$ attacking the mac security of $F_s[E]$, the following holds,

$$\begin{aligned}
&\mathbf{Adv}^{\mathrm{mac}}_{F_s[E]}(\mathcal{A}) \\
&= \Pr[\mathcal{A} \text{ forges in Game 0}] \\
&\leq 2 \times \frac{\sigma(\sigma-1)}{2}\mathbf{Adv}^{\mathrm{rk-up}}_{E}(t_{\sigma-2}, \sigma-2, \mu_{\sigma-2}) + \mathbf{Adv}^{\mathrm{rk-up}}_{E}(t_{\sigma-1}, \sigma-1, \mu_{\sigma-1}) \\
&\leq (\sigma^2-\sigma+1)\mathbf{Adv}^{\mathrm{rk-up}}_{E}(t_{\sigma-1}, \sigma-1, \mu_{\sigma-1}).
\end{aligned}$$

Finally, we get $\mathbf{Adv}^{\mathrm{mac}}_{F_s[E]}(t, q, \mu) \leq (\sigma^2-\sigma+1)\mathbf{Adv}^{\mathrm{rk-up}}_{E}(t', q', \mu')$, where $\sigma$ is the total block length of all $q$ queried messages plus the block length of the forgery message, $t' = t + O(\sigma)$, $q' = \sigma - 1$, $\mu' = \mu + O(\sigma)$. □

In Theorem 1, we reduce the mac security of $F_s[E]$ to the rk-up security of $E$, under the assumption that adversaries of $F_s[E]$ can observe all its chaining values. This implies in practical implementations for $F_s[E]$, engineers do not have to protect the secrecy of its chaining values, so $F_s[E]$ is provably secure against SCAs as long as $E$ is. In this sense, the security of $F_s[E]$ is more reliable than those from PRF (PseudoRandom Function) to PRF reductions since the latter MACs are treated as black boxes in the analysis [3, 5–13].

Furthermore, notice that unpredictability requires block ciphers much less than pseudorandomness does; on the other hand, related-key attacks (especially the kind we consider here, $\Phi^{\oplus}_{K}$-restricted as in Section 2) have become common analysis methods for block ciphers, and block ciphers are expected to be secure against such attacks in their designs. So, the security level that $F_s[E]$ asks for $E$ is not hard for practical block ciphers to reach.

Nevertheless, we note that rk-up and PRF are two separate security notions from theory, although unpredictability is strictly weaker than PRF, because related key is a notion independent of unpredictability and pseudorandomness.

In the aspect of efficiency, the 24 secure rate-1 MACs may not run faster than none rate-1 MACs, due to their large number of key schedules. Since for many practical block ciphers, the time for key schedule is no shorter than that for an encryption. However, notice that there are 8 MACs out of the 24 ($F_i$ for $i = 5, 6, 7, 8, 15, 17, 19.20$) whose KM are independent of the chaining values $T_{i-1}$, and they may pre-compute the key-schedules once having obtained $M_i$, so these 8 MACs may offer relatively high efficiency.

## 5 Compact PGV Model

In the proof for the 24 secure MACs from $K$-PGV model, it is easy to find that FF in fact has no influence over their security, and this observation can also be gotten from Table 1. So, it is natural to remove FF from $K$-PGV model away, and we call the remaining Compact PGV model, as illustrated in Fig. 3.

In Compact PGV model, we have 16 basic functions $g_s(K, M_i, T_{i-1}) = E(K \oplus \mathrm{KM}, \mathrm{IB})$ $(s = 0, 1, \cdots, 15)$ to construct rate-1 MACs $G_s(K, M)$ in the same way

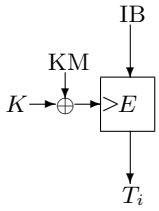


**Fig. 3.** In Compact PGV model with secret key $K \xleftarrow{\$} \mathcal{K}_E$, a block cipher $E$ has two inputs: an input block IB and a key mask KM. A basic function $g(K, M_i, T_{i-1})$ is defined as $g(K, M_i, T_{i-1}) = E(K \oplus \text{KM}, \text{IB})$, where IB, KM$\in \{M_i, T_{i-1}, M_i \oplus T_{i-1}, \text{Cst}\}$.

as we define $F_s$ $(s = 1, 2, \cdots, 64)$ by $f_s$. To be concrete, $G_s(K, M)$ is defined as follows,

MAC $G_s(K, M)$
$K \xleftarrow{\$} \mathcal{K}_E$;
**for** $i = 1$ **to** $l$ **do** $T_i \leftarrow g_s(K, M_i, T_{i-1})$ **end for**
**return** $T_l$.

where $\mathbf{pad}(M) = M_1 M_2 \cdots M_l$ and $T_0 = \text{Cst}$.

### 5.1 Rate-1 MACs from Compact PGV Model

The security evaluations for the 16 MACs are shown in Table 2, and they can also be got from Table 1 directly.

**Table 2.** The security of 16 MACs from Compact PGV model. "–" means the MAC is meaningless because its inputs are independent of $M_i$ or $T_{i-1}$; the number 3 means the MAC is vulnerable to attack 3; the MACs marked with $g_s$ $(s = 0, 1, \cdots, 5)$ are provably secure.

| | choice of IB | | | |
|---|---|---|---|---|
| choice of KM | $M_i$ | $T_{i-1}$ | $M_i \oplus T_{i-1}$ | Cst |
| $M_i$ | – | $g_0$ | $g_5$ | – |
| $T_{i-1}$ | $g_1$ | – | $g_4$ | – |
| $M_i \oplus T_{i-1}$ | $g_2$ | $g_3$ | 3 | 3 |
| Cst | – | – | 3 | – |

Compact PGV model gives us a more clear view on the MACs, whose security is related with the independence of KM and IB. More specifically, the MACs with independent KM and IB are provably secure, while others are not.

### 5.2 Equivalence of the Six Secure MACs

Next, we study the relationships among these six secure MACs from Compact PGV model, and find that they are in fact equivalent to each other, in the sense that $\forall 0 \leq s1 \leq s2 \leq 5$ there exists an invertible $2 \times 2$ matrix $A_i$ $(1 \leq i \leq 6)$ over GF(2) who can transform $g_{s1}$ into $g_{s2}$. That is, $(\mathrm{KM}_{s1}, \mathrm{IB}_{s1}) \times A_i = (\mathrm{KM}_{s2}, \mathrm{IB}_{s2})$, where $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $A_5 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $A_6 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. In other words, the basic functions in these six MACs can be generated from any one of them by invertible $2 \times 2$ matrixes over GF(2), whose total number is exactly six.

The goodness of this equivalence includes not only the convenience for us to get a better understanding on these six secure MACs, but also the fact that any one of the six is as secure as the others; however, this equivalence also implies if one of them is used with secret key $K$, then adversaries can easily make forgeries against the other five with the same $K$, although the other five may never be used with $K$ before. As an example, suppose an adversary has access to $G_0$ under a secret key $K$, and he queries $G_0$ with message $M = T_0||T_1||\cdots||T_l$, obtaining the tag $T_{l+1}$ (He can make such a query because he can observe the chaining values by our assumption and he can decide $M_{i+1} = T_i$ once he has obtained $T_i$). Then, he outputs a forgery $(M, T_{l+1})$ against $G_1$ under the same $K$. The forgery is valid because the adversary never queried $(M, T_{l+1})$ to $G_1$ under $K$. What is more, this adversary can also make similar forgeries against $G_s$ under $K$ (for $s = 2, 3, 4, 5$) by querying $G_0$ with a carefully selected message. This can be seen as a related-mode attack introduced by Phan and Siddiqi [27], which is dangerous since in many practical protocols, such as those of IPSec [34], there are several comparable algorithms for the users to choose. Some lazy users may take the same key for different algorithms, and in such a case $G_s$ $(s = 0, 1, \cdots, 5)$ can not be used in the same protocol together.

### 5.3 Independence Classes

The equivalence of the six secure MACs makes it inconvenient to use them in practice; luckily, for parts of the six, it is easy to break their equivalence. That is, we let $G_0$ and $G_3$ take distinct-and-fixed $T_0$, then we can prove they are independent of each other. The same technique also applies to break the independence of $G_1$ and $G_4$, $G_2$ and $G_5$; thus, we have three independent classes.

**Theorem 2.** *Suppose $G_{s1}$ and $G_{s2}$ ($s1 = (s2 + 3) \bmod 6$) have the same secret key $K \xleftarrow{\$} \mathcal{K}_E$ but distinct-and-fixed $T_0$, and their underlying block cipher $E : \mathcal{K}_E \times \{0,1\}^n \times \{0,1\}^n$ is rk-up secure, then they are independent of each other.*

The proof idea is that, for any adversary $\mathcal{A}$ attacking the independence of $G_{s1}$ and $G_{s2}$, we let it have access to two oracles $G_{s1}[E]$ and $G_{s2}[E]$. $\mathcal{A}$ can query these two oracles with any prefix-free messages and obtain not only the

corresponding tags but also the chaining values; at last, $\mathcal{A}$ is asked to make a forgery against either $G_{s1}[E]$ or $G_{s2}[E]$. If $\mathcal{A}$ can do this with a non-trivial probability, we say that $G_{s1}$ and $G_{s2}$ are not independent of each other. However, as we will prove, the success probability for $\mathcal{A}$ to forge against either $G_{s1}[E]$ or $G_{s2}[E]$ is sufficiently small, and it can be reduced to the rk-up security of $E$. Thus, $G_{s1}[E]$ and $G_{s2}[E]$ are independent of each other. The detailed proof is given in Appendix B.

However, taking distinct-and-fixed $T_0$ can not guarantee the independence of $G_{s1}$ and $G_{s2}$, where $s1 \neq (s2 + 3) \bmod 6$.

## 6 Conclusions and Future Work

To sum up, we study the provable security of MACs based on related-key unpredictable block ciphers in this paper, and obtain both good news and bad news. The bad news are mainly two folds: firstly, all current rate-1 MACs may not guarantee their provable security when instantiated with related-key unpredictable block ciphers; secondly, in the keyed PGV model 25 MACs are vulnerable to three kinds of attacks respectively. The good news is that 24 provably secure rate-1 MACs are found in the keyed PGV model, whose provable security relies on the related-key unpredictability of their underlying block ciphers. Furthermore, we study the 16 rate-1 MACs in Compact PGV model, and find that the six provably secure MACs are equivalent to each other, which implies related-mode attacks on them. Then, we give a suggestion for parts of the six to avoid such attacks by taking distinct-and-fixed initial values. In the aspect of efficiency, these provably secure rate-1 MACs may not run faster than none rate-1 MACs due to their large number of key schedules.

Furthermore, we find that in the keyed PGV model all the MACs with a fixed key $K \oplus \mathrm{Cst}$ ($\mathrm{KM} = \mathrm{Cst}$) are either insecure or meaningless. This implies *within* this model, it is impossible to construct a rate-1 MAC from only unpredictable block ciphers. However, it is still unknown whether it is possible to do this *beyond* the keyed PGV model, and we leave this as an open question.

**Acknowledgments.** The authors would like to thank the anonymous referees for their valuable comments. Special thanks to Kan Yasuda for his help to revise this paper. Furthermore, this work is supported by the National High-Tech Research and Development 863 Plan of China (No. 2007AA01Z470), the National Natural Science Foundation of China (No. 60873259, and No. 60903219) and the Knowledge Innovation Project of The Chinese Academy of Sciences.

## References

1. FIPS 197. Advanced Encryption Standard (AES). National Institute of Standards and Technology. (2001)
2. FIPS 113. Computer Data Authentiaction. National Institute of Standards and Technology. (1985)

3. Bellare, M., Kilian, J., Rogaway, P.: The Security of Cipher Block Chaining. In: Desmedt, Y., (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 341–358. Springer, Heidelberg (1994)
4. ISO/IEC 9797 – 1, Information Technology – Security Techniques – Message Authentication Codes (MACs) – Part 1: Mechanisms Using A Block Cipher. International Organization for Standardization. (1999)
5. Petrank, E., Rackoff, C.: CBC MAC for Real-Time Data Sources. J. Cryptology **13**(3) 315–338 (2000)
6. Jaulmes, É., Joux, A., Valette, F.: On the Security of Randomized CBC-MAC Beyond the Birthday Paradox Limit: A New Construction. In: Daemen, J., Rijmen, V., (eds.) FSE 2002. LNCS, vol. 2365, pp. 237–251. Springer, Heidelberg (2002)
7. Black, J., Rogaway, P.: CBC MACs for Arbitrary-Length Messages: The Three-Key Constructions. In: Bellare, M., (ed.) CRYPTO 2000. LNCS, vol. 1880, pp. 197–215. Springer, Heidelberg (2000)
8. Kurosawa, K., Iwata, T.: TMAC: Two-Key CBC MAC. In: Joye, M., (ed.) CT-RSA 2003. LNCS, vol. 2612, pp. 33–49. Springer, Heidelberg (2003)
9. Iwata, T., Kurosawa, K.: OMAC: One-Key CBC MAC. In: Johansson, T., (ed.) FSE 2003. LNCS, vol. 2887, pp. 129–153. Springer, Heidelberg (2003)
10. Special Publication 800-38B. Recommendation for Block Cipher Modes of Operation: The CMAC Mode for Authentication. National Institute of Standards and Technology Available at: `http://csrc.nist.gov/groups/ST/toolkit/BCM/current_modes.html`.
11. Nandi, M.: Fast and Secure CBC-Type MAC Algorithms. In: Dunkelman, O., (ed.) FSE 2009. LNCS, vol. 5665, pp. 375–393. Springer, Heidelberg (2009)
12. Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 1: f8 and f9 Specifications Available at: `http://www.3gpp.org/ftp/Specs/archive/35_series/35.201/`.
13. Black, J., Rogaway, P.: A Block-Cipher Mode of Operation for Parallelizable Message Authentication. In: Knudsen, L.R., (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 384–397. Springer, Heidelberg (2002)
14. Dodis, Y., Pietrzak, K., Puniya, P.: A New Mode of Operation for Block Ciphers and Length-Preserving MACs. In: Smart, N.P., (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 198–219. Springer, Heidelberg (2008)
15. Dodis, Y., Steinberger, J.: Message Authentication Codes from Unpredictable Block Ciphers. In: Halevi, S., (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 267–285. Springer, Heidelberg (2009)
16. Biham, E., Dunkelman, O., Keller, N.: A Related-Key Rectangle Attack on the Full KASUMI. In: Roy, B.K., (ed.) ASIACRYPT 2005. LNCS, vol. 3788, pp. 443–461. Springer, Heidelberg (2005)
17. Biryukov, A., Khovratovich, D., Nikolić, I.: Distinguisher and Related-Key Attack on the Full AES-256. In: Halevi, S., (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 231–249. Springer, Heidelberg (2009)
18. An, J.H., Bellare, M.: Constructing VIL-MACs from FIL-MACs: Message Authentication under Weakened Assumptions. In: Wiener, M.J., (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 252–269. Springer, Heidelberg (1999)
19. Maurer, U.M., Sjödin, J.: Single-Key AIL-MACs from Any FIL-MAC. In: Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M., (eds.) ICALP 2005. LNCS, vol. 3580, pp. 472–484. Springer, Heidelberg (2005)
20. Bellare, M.: New Proofs for NMAC and HMAC Security Without Collision-Resistance. In: Dwork, C., (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 602–619. Springer, Heidelberg (2006)

21. Bellare, M., Ristenpart, T.: Hash Functions in the Dedicated-Key Setting: Design Choices and MPP Transforms. In: Arge, L., Cachin, C., Jurdzinski, T., Tarlecki, A., (eds.) ICALP 2007. LNCS, vol. 4596, pp. 399–410. Springer, Heidelberg (2007)
22. Hirose, S., Park, J.H., Yun, A.: A Simple Variant of the Merkle-Damgård Scheme with a Permutation. In: Kurosawa, K., (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 113–129. Springer, Heidelberg (2007)
23. Yasuda, K.: A Single-Key Domain Extender for Privacy-Preserving MACs and PRFs. In: Lee, P.J., Cheon, J.H., (eds.) ICISC 2008. LNCS, vol. 5461, pp. 268–285. Springer, Heidelberg (2008)
24. Yasuda, K.: A Double-Piped Mode of Operation for MACs, PRFs and PROs: Security beyond the Birthday Barrier. In: Joux, A., (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 242–259. Springer, Heidelberg (2009)
25. Preneel, B., Govaerts, R., Vandewalle, J.: Hash Functions Based on Block Ciphers: A Synthetic Approach. In: Stinson, D.R., (ed.) CRYPTO 1992. LNCS, vol. 773, pp. 368–378. Springer, Heidelberg (1993)
26. Wang, P., Feng, D., Wu, W., Zhang, L.: On the Unprovable Security of 2-Key XCBC. In: Mu, Y., Susilo, W., Seberry, J., (eds.) ACISP 2008. LNCS, vol. 5107, pp. 230–238. Springer, Heidelberg (2008)
27. Phan, R.C.W., Siddiqi, M.U.: Related-Mode Attacks on Block Cipher Modes of Operation. In: Gervasi, O., Gavrilova, M.L., Kumar, V., Laganà, A., Lee, H.P., Mun, Y., Taniar, D., Tan, C.J.K., (eds.) ICCSA 2005. LNCS, vol. 3482, pp. 661–671. Springer, Heidelberg (2005)
28. Black, J., Rogaway, P., Shrimpton, T.: Black-Box Analysis of the Block-Cipher-Based Hash-Function Constructions from PGV. In: Yung, M., (ed.) CRYPTO 2002. LNCS, vol. 2442, pp. 320–335. Springer, Heidelberg (2002)
29. Luby, M., Rackoff, C.: How to Construct Pseudo-Random Permutations from Pseudo-Random Functions (Abstract). In: Williams, H.C., (ed.) CRYPTO 1985. LNCS, vol. 218, pp. 447. Springer, Heidelberg (1985)
30. Bellare, M., Kohno, T.: A Theoretical Treatment of Related-Key Attacks: RKA-PRPs, RKA-PRFs, and Applications. In: Biham, E., (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 491–506. Springer, Heidelberg (2003)
31. Lucks, S.: Ciphers Secure against Related-Key Attacks. In: Roy, B.K., Meier, W., (eds.) FSE 2004. LNCS, vol. 3017, pp. 359–370. Springer, Heidelberg (2004)
32. Coron, J.S., Dodis, Y., Malinaud, C., Puniya, P.: Merkle-Damgård Revisited: How to Construct a Hash Function. In: Shoup, V., (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 430–448. Springer, Heidelberg (2005)
33. Shannon, C.E.: Communication Theory of Secrecy Systems. Bell Systems Technical Journal **28**(4) 656–715 (1949)
34. Kent, S., Atkinson, R.: Security Architecture for the Internet Protocol. RFC 2401, standards track, the Internet Society. (1998)

## A Attacks on Some Current Rate-1 MACs

Here, we give the attacks on the unpredictability of other existing rate-1 MACs instantiated with a special kind of rk-up block cipher $E'$, as defined in Section 3. Our attacks can be seen as extensions of An and Bellare's attack on the basic CBC-MAC [18], and they show that all the existing rate-1 MACs may not hold their unpredictability when their underlying block ciphers are only rk-up secure;

however, these attacks do not necessarily mean the non-existence of secure rate-1 MACs based on only unpredictable block ciphers.

Due to limitation of pages, we describe the attacks without introducing the corresponding MAC algorithms.

**Attack on RMAC [6]** Adversary $\mathcal{A}$ does as follows,

1) queries $\mathrm{RMAC}_{E'}(\cdot)$ with $0^n||10^{n-2}$, obtains the tag $T = t_1t_2t_3t_4$ and a random value $R$;

2) makes a forgery $(R, M', T)$, where $M' = 0^n||(0^{3n/4}||(t_4 \oplus 0^{n/4-1}1))||10^{n-2}$.

This attack also applies to EMAC [5].

**Attack on GCBC1 [11]** Adversary $\mathcal{A}$ does as follows,

1) queries $\mathrm{GCBC1}_{E'}(\cdot)$ with $0^n||10^{n-1}$, obtains the tag $T = t_1t_2t_3t_4$;

2) makes a forgery $(M', T)$, where $M' = 0^n||(0^{3n/4}||x)||10^{n-1}$ and $x = \mathrm{lsb}_1(t_3)||\mathrm{msb}_{n/4-1}(t_4)$.

**Attack on GCBC2 [11]** Adversary $\mathcal{A}$ does as follows,

1) queries $\mathrm{GCBC2}_{E'}(\cdot)$ with $0^n||10^{n-1}$, obtains the tag $T^1 = t_1^1t_2^1t_3^1t_4^1$;

2) queries $\mathrm{GCBC2}_{E'}(\cdot)$ with $(01^{n-4}000)||10^{n-1}$, obtains the tag $T^2 = t_1^2t_2^2t_3^2t_4^2$;

3) makes a forgery $(M', T^2)$, where $M' = 0^n||((0^{3n/4}||x^1) \oplus (0^{3n/4}||x^2) \oplus 10^{n-1})$, $x^1 = \mathrm{lsb}_1(t_3^1)||\mathrm{msb}_{n/4-1}(t_4^1)$ and $x^2 = \mathrm{lsb}_1(t_3^2)||\mathrm{msb}_{n/4-1}(t_4^2)$.

**Attack on f9 [12]** If $\mathrm{msb}_1(N) = 0$ ($N = \mathrm{COUNT}||\mathrm{FRESH}$), adversary $\mathcal{A}$ does as follows,

1) queries $\mathrm{f9}_{E'}(\cdot)$ with $M = 10^{n-2}$, obtains the tag $T = t_1t_2t_3t_4$ and a nonce $N = n_1n_2n_3n_4$;

2) makes a forgery $(N, M', T)$, where $M' = M_1||M_1||10^{n-2}$ and $M_1 = 0^{3n/4}||(n_4 \oplus n_3 \oplus t_3)$.

If $\mathrm{msb}_1(N) = 1$, we can define $E''$ as follows to attack f9,

$$E''(K, M) = \begin{cases} m_1||m_2||m_3||c, & \text{if } \mathrm{msb}_1(m_1) = 1, \\ m_1||c||m_4||m_3, & \text{if } \mathrm{msb}_1(m_1) = 0, \end{cases}$$

where $c$ is the same as in $E'$.

**Attack on PMAC [13]** Adversary $\mathcal{A}$ does as follows,

1) queries $\mathrm{PMAC}_{E'}(\cdot)$ with $0^n$, obtains the tag $T^1 = t_1^1t_2^1t_3^1t_4^1$;

2) queries $\mathrm{PMAC}_{E'}(\cdot)$ with $10^{n-1}$, obtains the tag $T^2 = t_1^2t_2^2t_3^2t_4^2$;

3) makes a forgery $(M', T')$, where

$$\begin{cases} M' = t_1^1t_2^1t_3^1t_4^2, & \text{if } \mathrm{msb}_1(t_1^1) = 0, \\ M' = t_1^2t_2^2t_3^2t_4^1, & \text{if } \mathrm{msb}_1(t_1^1) = 1, \\ T' = M' \cdot x. \end{cases}$$

## B Proof for the Independence of $G_{s1}$ and $G_{s2}$, where $s1 = (s2 + 3) \bmod 6$

*Proof.* If there exists an adversary $\mathcal{A}$ who can attack the independence of $G_{s1}[E]$ and $G_{s2}[E]$, it implies $\mathcal{A}$ is able to attack the mac security of either $G_{s1}[E]$ or $G_{s2}[E]$. Now we show that the success probability for $\mathcal{A}$ to do the latter is upper bounded since $E$ is rk-up secure by assumption.

| Game 0 | Game 1 |
|---|---|
| Range$\leftarrow \{T_{s1,0}, T_{s2,0}\}$, Collision$_w \leftarrow$False, for $w \geq 1$; $z \leftarrow 1$. | |
| when $\mathcal{A}$ makes a query $M_s^j$ to $G_s$, where $s \in \{s1, s2\}$ and | |
| $\mathbf{Pad}(M_s^j) = M_{s,1}^j M_{s,2}^j \cdots M_{s,l_{s,j}}^j$, $j = 1, 2, \cdots, q_s$ | |
| 01. **for** $i = 1$ **to** $l_{s,j}$ **do** | |
| 02. renew KM, IB with $(M_{s,i}^j, T_{s,i-1}^j, M_{s,i}^j \oplus T_{s,i-1}^j, \text{Cst})$ by the definition of $g_s$; | |
| **03.** $T_{s,i}^j = \mathcal{O}_{\mathcal{B}}(K \oplus \text{KM}, \text{IB})$; | |
| **04.** **if** $T_{s,i}^j \in$Range and $\nexists j1 < j$ s.t. $M_{s,1}^{j1} M_{s,2}^{j1} \cdots M_{s,i-1}^{j1} = M_{s,1}^j M_{s,2}^j \cdots M_{s,i-1}^j$ | |
| **05.** **then** {$\boxed{\text{Collision}_z \leftarrow \text{True; Stop.}}$} | |
| **06.** **end if** | |
| 07. Range$\leftarrow$Range$\cup\{T_{s,i}^j\}$; $z \leftarrow z + 1$; return $T_{s,i}^j$ to $\mathcal{A}$; | |
| 08. **end for** | |
| 10. when $\mathcal{A}$ makes a forgery $(M', T')$ to $G_s$, where $s \in \{s1, s2\}$ and $\mathbf{Pad}(M') = M'_1 M'_2 \cdots M'_{l'}$ | |
| 11. **for** $i = 1$ **to** $l' - 1$ **do** | |
| 12. renew KM, IB with $(M'_i, T'_{i-1}, M'_i \oplus T'_{i-1}, \text{Cst})$ by the definition of $g_s$; | |
| **13.** $T'_i = \mathcal{O}_{\mathcal{B}}(K \oplus \text{KM}, \text{IB})$; | |
| **14.** **if** $T'_i \in$Range and $\nexists j1 \in \{1, 2, \cdots, q_s\}$ s.t. $M_{s,1}^{j1} M_{s,2}^{j1} \cdots M_{s,i-1}^{j1} = M'_1 M'_2 \cdots M'_{i-1}$ | |
| **15.** **then** {$\boxed{\text{Collision}_z \leftarrow \text{True; Stop.}}$} | |
| **16.** **end if** | |
| 17. Range$\leftarrow$Range$\cup\{T'_i\}$; $z \leftarrow z + 1$; return $T'_i$ to $\mathcal{A}$; | |
| 18. **end for** | |
| **19.** renew KM, IB with $(M'_{l'}, T'_{l'-1}, M'_{l'} \oplus T'_{l'-1}, \text{Cst})$ by the definition of $g_s$; | |
| 20. **if** $T' = \mathcal{O}_{\mathcal{B}}(K \oplus \text{KM}, \text{IB})$ return 1 **else** return 0 **end if** | |

**Fig. 4.** Definitions for Game 0 (excluding the boxed codes) and Game 1 (including the boxed codes), in which adversary $\mathcal{B}$ simulates adversary $\mathcal{A}$'s oracles $G_{s1}[E]$ and $G_{s2}[E]$ with its own oracle $\mathcal{O}_{\mathcal{B}}(\cdot, \cdot) = E(\cdot, \cdot)$ combining the definitions of $G_{s1}$ and $G_{s2}$, where $s1 = (s2 + 3) \bmod 6$.

The following proof is much similar to that for Theorem 1. We define two Games in Fig. 4, where an adversary $\mathcal{B}$ will simulate $\mathcal{A}$'s oracles $G_{s1}[E]$ and $G_{s2}[E]$ with its own oracle $\mathcal{O}_{\mathcal{B}}(\cdot, \cdot) = E(\cdot, \cdot)$ combining the definitions of $G_{s1}$ and $G_{s2}$. Finally, $\mathcal{B}$ will attack the rk-up security of $E$. In either Game 0 or Game 1, $\mathcal{A}$ can make any prefix-free queries, get not only the corresponding tags

but also the chaining values; at last, he is asked to make a forgery against either $G_{s1}[E](\cdot)$ or $G_{s2}[E](\cdot)$. Also, the forgery message should not be a prefix of a queried message. Unlike that in Fig. 2, the Range in Fig. 4 is defined as the set containing the outputs of $E$ when dealing with both $G_{s1}[E](\cdot)$ and $G_{s2}[E](\cdot)$.

By similar discussions as in the proof for Theorem 1, we get

$$
\begin{aligned}
&\Pr[\mathcal{A} \text{ breaks the independence of } G_{s1}[E] \text{ and } G_{s2}[E]] \\
&\leq \max\{\mathbf{Adv}^{\mathrm{mac}}_{G_{s1}[E]}(\mathcal{A}), \mathbf{Adv}^{\mathrm{mac}}_{G_{s2}[E]}(\mathcal{A})\} \\
&\leq \Pr[\mathcal{A} \text{ forges in Game 0}] \\
&\leq \Pr[\mathsf{Coll}] + \Pr[\mathcal{A} \text{ forges in Game 1}] \\
&\leq 2\Pr[\mathsf{Coll}] + \Pr[\mathcal{A} \text{ forges in Game 1} | \overline{\mathsf{Coll}}] \\
&\leq 2 \times \frac{\sigma'(\sigma'-1)}{2} \mathbf{Adv}^{\mathrm{rk-up}}_{E}(t_{\sigma'-2}, \sigma'-2, \mu_{\sigma'-2}) + \mathbf{Adv}^{\mathrm{rk-up}}_{E}(t_{\sigma'-1}, \sigma'-1, \mu_{\sigma'-1}) \\
&\leq (\sigma'^2 - \sigma' + 1)\mathbf{Adv}^{\mathrm{rk-up}}_{E}(t_{\sigma'-1}, \sigma'-1, \mu_{\sigma'-1}),
\end{aligned}
$$

where the event $\mathsf{Coll}$ is the same as that in the proof for Theorem 1 and $\sigma'$ is the total block length of all queried messages (to both $G_{s1}[E]$ and $G_{s2}[E]$) plus the block length of the forgery message (to either $G_{s1}[E]$ or $G_{s2}[E]$).

Thus, any adversary $\mathcal{A}$ in fact has a sufficiently small probability to make a forgery against either $G_{s1}[E](\cdot)$ or $G_{s2}[E](\cdot)$, after having queried $G_{s1}[E](\cdot)$ and $G_{s2}[E](\cdot)$ for some time (this is measured by the total block length $\sigma'$). Finally, we conclude that $G_{s1}[E](\cdot)$ and $G_{s2}[E](\cdot)$ are independent of each other. $\square$