Relate-Key Almost Universal Hash Functions: Definitions, Constructions and Applications

Peng Wang[1] and Yuling Li[1] and Liting Zhang[2] and Kaiyan Zheng[1]

[1] State Key Laboratory of Information Security
Institute of Information Engineering, Chinese Academy of Sciences
[2] Institution of Software, Chinese Academy of Sciences
wp@is.ac.cn, liyuling@iie.ac.cn, zhangliting@tca.iscas.ac.cn, zhengkaiyan@iie.ac.cn

Abstract. Universal hash functions (UHFs) have been extensively used in the design of cryptographic schemes. But if we consider related-key attack against the schemes, some of them may not be secure, especially when the key of UHF is a part of the key of scheme. In order to solve the issue, we propose a new concept of related-key almost universal hash function, which is a natural extension to almost universal hash function in the related-key scenario. We define related-key almost universal (RK-AU) hash function and related-key almost XOR universal (RK-AXU) hash function. However almost all the existing UHFs do not satisfy the new definitions. We construct fixed-input-length universal hash functions such as RH1 and variable-input-length related-key universal hash functions such as RH2, RH3. We show that RH1 and RH2 are both RK-AXU, and RH3 is RK-AU. Furthermore, RH1, RH2 and RH3 are nearly as efficient as previous similar constructions. RK-AU (AXU) hash functions can be used as components with related-key property in the design of cryptographic schemes. If we replace the universal hash functions in the schemes with our corresponding constructions, the problems about related-key attack can be solved. More specifically, we give four concrete applications of RK-AU and RK-AXU in MACs and TBCs.

Keywords. Almost universal hash function, related-key attack, related-key almost universal hash function, message authentication code, tweakable block cipher.

1 Introduction

AU AND AXU. A universal hash function (UHF) is a family of function indexed by keys. A basic requirement of UHF is that the collision probability of hash values from any two different messages is small when the key is uniformly random. For example we define a polynomial evaluation hash function [5] $Poly : \{0,1\}^n \times \{0,1\}^{nm} \to \{0,1\}^n$,

$$Poly_K(M) = M_1K^m \oplus M_2K^{m-1} \oplus \cdots \oplus M_mK$$

where $M = M_1\|M_2\|\cdots\|M_m \in \{0,1\}^{nm}$, $M_i \in \{0,1\}^n$, $i = 1, 2, \cdots, m$ and all the operations are in the finite field $GF(2^n)$. For any $M \neq M'$, $Poly_K(M) \oplus Poly_K(M')$ is a polynomial in K whose degree is nonzero and no more than m, so there are at most m keys leading to $Poly_K(M) = Poly_K(M')$ and the collision probability is at most $m/2^n$ when K is uniformly random. We say that this function is $m/2^n$-almost-universal (AU). Obviously the probability of $Poly_K(M) \oplus Poly_K(M') = C$ is also at most $m/2^n$ for any $M \neq M'$ and C. That is another commenly used concept: almost XOR universal (AXU) hash functions. $Poly$ is also $m/2^n$-AXU.

Since introduced by Carter and Wegman [11,42], universal hash functions (UHFs) have been adopted in numerous cryptographic schemes. The direct application of UHFs is in c message authentication codes

(MACs). The basic idea is to compress the message using a universal hash function and then encrypt the hash value into a tag. This kind of MACs have been standardized in ISO/IEC 9797-3:2011 [23] which includes UMAC [9], Badger [10], Poly1305-AES [3] and GMAC [28]. UHFs are also used in tweakable block ciphers (TBCs) [27] and tweakable enciphering schemes (TESes), such as XTS-AES [20,31] in IEEE Std 1619-2007 and NIST SP 800-38E, XCB in IEEE Std 1619.2-2010 [21], HCTR [41], HCH [12,13], etc. The third application of UHF is in authenticated encryptions (AEs). The most widely used AE is GCM [28] which has been standardized in ISO/IEC-19772:2009 [22] and NIST SP 800-38D [30]. In the recent CAESAR competition, several UHF based AEs were proposed, such as COBRA [2], Enchilada [19] and POET [1], etc.

In all these schemes, the estimation of the collision probability about the inputs to the underlying primitive such as pseudorandom function (PRF) is a crucial point in the security proofs. The almost universal property of the UHF guarantees that the collision seldom happens.

Related-key attacks. Related-key attack was firstly introduced by Biham et al. [7] for block ciphers [16,8,39] and then extended to other cryptographic algorithms [14] and schemes such as MACs [32], TESes [40], AEs [15], etc. Related-key security has become a basic requirement for cryptographic algorithms. Bellare and Kohno gave a theoretical study of related-key security of block cipher, modeling the concept of related-key pseudorandom permutation (RK-PRP) and related-key pseudorandom function (RK-PRF). Bhattacharyya and Roy [6] gave the related-key security definition of MAC. Compared with the usual single-key scenario, the adversary has more resources to analyze the algorithms. In the related-key setting, the adversary does not know the secret key, but can apply related-key-deriving (RKD) transformations to change the secret key of the algorithm and observe the output under the related keys.

Let Φ be a RKD transformation set. For a function $F : \mathcal{K} \times \mathcal{D} \to \mathcal{R}$ the adversary can make related-key oracle queries $(\phi, M) \in \Phi \times \mathcal{D}$ and is responded with $F_{\phi(K)}(M)$ where K is the secret key. If the adversary can not distinguish F from a uniformly random function from $\mathcal{K} \times \mathcal{D}$ to $\mathcal{R}$, we say that F is a RK-PRF over Φ. There are two commonly used methods to derive related keys, and the corresponding RKD sets are: $\Phi^{\oplus} = \{XOR_{\Delta} : K \mapsto K \oplus \Delta\}$ and $\Phi^{+} = \{ADD_{\delta} : K \mapsto K + \delta \mod p\}$, where p is an integer. In the following, we use $\Phi^{\oplus}$ as the default RKD set unless specified otherwise. For completeness, we give related-key security definitions of MAC, TBC, TES and AE in Appendix A.

Motivations. Although almost all the UHFs based schemes have security proofs under the single-key setting, some of them may have problems in light of the related-key attacks. We show practical attacks against some popular constructions of MAC, TBC, TES and AE using *Poly* as the UHF component.

1) MAC. A typical UHF based MAC encrypts the hash value into a tag using the one-time-pad encryption. This method was originated from Carter and Wegman [11,42] and dominates the usage of UHF in MACs [23]. Consider a simple example

$$MAC_{K\|K'}(N, M) = Poly_K(M) \oplus E_{K'}(N)$$

where $M = M_1\|M_2 \in \{0,1\}^{2n}$, $Poly_K(M_1\|M_2) = M_1K^2 \oplus M_2K$, E is a block cipher and N is a nonce. It has been proved that [35,4] if E is a pseudorandom function (PRF) and *Poly* is almost XOR universal, MAC is secure. But if we query with $A\|A$ under the derived key $(K \oplus 0^{n-1}1)\|K'$, the answer is $T = (A(K \oplus 0^{n-1}1)^2 \oplus A(K \oplus 0^{n-1}1)) \oplus E_{K'}(N) = (AK^2 \oplus AK)) \oplus E_{K'}(N)$. Therefore we can predict that the tag of $A\|A$ under the original key is also T. Therefore $(N, A\|A, T)$ is a successful forgery which breaks the related-key security of MAC. A similar attack can apply to Poly1305-AES [3] in ISO/IEC 9797-3:2011 [23].

2) TBC. A tweakable block cipher (TBC) is a generalized blockcipher with an extra input called tweak. TBCs were first formalized by Liskov, Rivest and Wanger [27] and found applications largely in the design of mode of operation [33]. In their seminal paper, Liskov et al. gave a construction of TBC from a block cipher:

$$TBC_{K\|K'}(T, M) = E_{K'}(M \oplus H_K(T)) \oplus H_K(T)$$

where E is a block cipher and H is a universal hash function and T is a tweak. They proved that when E is a PRF and H is almost XOR universal, TBC is secure in the sense of indistinguishability from a family of independent random permutation indexed by the tweaks. If we use $Poly_K(T) = TK$ as the underlying UHF, the following is an attack. First we query with (T, M) under the derived key $(K \oplus \Delta)\|K'$, where $\Delta \neq 0$, then the answer is $C = E_{K'}(M \oplus T(K \oplus \Delta)) \oplus T(K \oplus \Delta) = E_{K'}((M \oplus T\Delta) \oplus TK) \oplus TK \oplus T\Delta$. So we can predict that the ciphertext of $(T, (M \oplus T\Delta))$ under the original key is $C \oplus T\Delta$. So it does not resist related-key attack.

3) TES. A tweakable enciphering scheme is a TBC with large or variable input length, especially suitable for disk sector encryption. Recently Sun et al. [40] show that HCTR [41], HCHp and HCHfp [12,13] suffer related-key attacks. All these TESes use the polynomial evaluation hash function as the underlying UHF.

4) AE. An authenticated encryption scheme achieves both privacy and authenticity. One of AE designs, such as OCB [34,33] following from IAPM[24], encrypt the message blocks using independent PRPs into ciphertext blocks and encrypt the XOR of the message blocks into a tag using another independent PRP. Kurosawa [26] proposed a modified IAPM, the encryption of message blocks is

$$C_i = E_{K'}(M_i \oplus Poly_K(IV\|(2i-1))) \oplus Poly_K(IV\|(2i-1))$$

where M_i is the i-th message block, E is a block cipher and the key of the new scheme is $K\|K'$. Kurosawa proved that this modified IAPM is secure even if the underlying block cipher is publicly accessible. But if we query with (IV, M) under the derived key $(K \oplus 0^{n-1}1)\|K'$, the first ciphertext block $C_1 = E_{K'}((M_i \oplus IV \oplus 0^{n-1}1) \oplus (Poly_K(IV\|(0^{n-1}1))) \oplus Poly_K(IV\|(0^{n-1}1)) \oplus IV \oplus 0^{n-1}1$. We can predict that the first ciphertext block of (IV, M') under the original key is $C_1 \oplus IV \oplus 0^{n-1}1$, where M' is changed from M by changing the first block into $M_1 \oplus IV \oplus 0^{n-1}1$. So it does not resist related-key attack.

In the above examples, the key of UHF is a part of the key of scheme, so that the related-key adversary can derive the related key of UHF and get the input collision to the underlying primitive. The essential properties of UHF used in the above attacks are listed in the following.

1) $Poly_{K \oplus 0^{n-1}1}(A\|A) = Poly_K(A\|A)$ used in the MAC example;
2) $Poly_{K \oplus \Delta}(T) \oplus Poly_K(T) = \Delta T$ used in the TBC example;
3) $Poly_{K \oplus \Delta}(A\|B) \oplus Poly_K(A\|B) = A\Delta^2 \oplus B\Delta$ used in the TES and AE examples.

Definitions. In order to solve the above issues, we propose a new concept of related-key almost universal hash function making the above equations seldom happen. The new concept can be seen as a natural extension to almost universal hash function in the related-key scenario. We define related-key almost universal (RK-AU) hash function and related-key almost XOR universal (RK-AXU) hash function. However almost all the existing UHFs do not satisfy the new definitions, such as $Poly$ in the above, MMH [18], Square Hash [17], NMH [18], NH [9], etc. See Appendix B for details.

Constructions. We construct a fixed-input-length universal hash functions RH1 and two variable-input-length related-key universal hash functions RH2 and RH3. We show that RH1 and RH2 are both

RK-AXU, and RH3 is RK-AU, over the RKD set $\Phi^{\oplus}$. Furthermore, RH1, RH2 and RH3 are almost as efficient as previous similar constructions.

Applications. If we replace the universal hash functions in the examples of section 1 with our corresponding constructions, the problems about related-key attacks can be solved. More specifically, we give four concrete applications in MAC and TBC.

2 Definitions

For a finite set X, $|X|$ denotes the size of the set. $x \xleftarrow{\$} X$ means selecting an element x uniformly at random from the set X. $A \Rightarrow b$ denotes that an algorithm A outputs b. For a function $F : \mathcal{K} \times \mathcal{D} \to \mathcal{R}$, when $K \in \mathcal{K}$ is a key, we often write the output $F(K, M)$ as $F_K(M)$, where $(K, M) \in \mathcal{K} \times \mathcal{D}$.

Let $H : \mathcal{K} \times \mathcal{D} \to \mathcal{R}$ be a function. Let Φ be a RKD transformation set.

Definition 1 (AU [37]). *H is an ϵ-almost-universal (ϵ-AU) hash function, if for any $M, M' \in \mathcal{D}$, $M \neq M'$,*

$$\Pr[K \xleftarrow{\$} \mathcal{K} : H_K(M) = H_K(M')] \leq \epsilon.$$

Definition 2 (AXU [25]). *Let $(\mathcal{R}, \oplus)$ be an abelian group[3]. H is an ϵ-almost-XOR-universal (ϵ-AXU), if for any $M, M' \in \mathcal{D}$, $M \neq M'$, and $C \in \mathcal{R}$,*

$$\Pr[K \xleftarrow{\$} \mathcal{K} : H_K(M) \oplus H_K(M') = C] \leq \epsilon.$$

Clearly, if H is ϵ-AXU, it is also ϵ-AU, for ϵ-AU is a special case of ϵ-AXU when $C = 0$.

We give new definitions to extend the above definitions in the related-key scenario.

Definition 3 (RK-AU). *H is an ϵ-related-key-almost-universal (ϵ-RK-AU) hash function over the RKD set Φ, if for any $\phi, \phi' \in \Phi$, $M, M' \in \mathcal{D}$, $(\phi, M) \neq (\phi', M')$,*

$$\Pr[K \xleftarrow{\$} \mathcal{K} : H_{\phi(K)}(M) = H_{\phi'(K)}(M')] \leq \epsilon.$$

Definition 4 (RK-AXU). *Let $(\mathcal{R}, \oplus)$ be an abelian group. H is an ϵ-related-key-almost-universal (ϵ-RK-AXU) hash function over the RKD set Φ, if for any $\phi, \phi' \in \Phi$, $M, M' \in \mathcal{D}$, $(\phi, M) \neq (\phi', M')$, and $C \in \mathcal{R}$,*

$$\Pr[K \xleftarrow{\$} \mathcal{K} : H_{\phi(K)}(M) \oplus H_{\phi'(K)}(M') = C] \leq \epsilon.$$

Obviously *Poly* is neither RK-AXU nor RK-AU. Unfortunately almost all the existing UHFs do not satisfy the new definitions, such as MMH [18], Square Hash [17], NMH [18], NH [9], etc. See Appendix B for more details.

3 Constructions

We construct two types of related-key almost universal hash functions: fixed-input-length universal hash functions such as RH1 and variable-input-length universal hash functions such as RH2 and RH3. We show that RH1 and RH2 are both RK-AXU, and RH3 is RK-AU, over the RKD set $\Phi^{\oplus}$.

Fil-rk-axu (au). *Poly* is a fixed-input-length universal hash function, but it is neither RK-AXU nor RK-AU. Our constructions are based on one-block-input-length polynomial evaluation function: $Poly_K(M) = MK$.

[3] For arbitrary abelian groups the generalized notion is almost Delta universal (AΔU) hash functions [38]. In the following when we say AXU we may sometimes refer to AΔU.

Construction 1 RH1 : $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$,

$$\mathrm{RH1}_K(M) = MK \oplus K^3.$$

Theorem 1. RH1 *is* $2/2^n$*-RK-AXU over the RKD set* $\Phi^{\oplus}$.

Proof. We prove that for any $M, M', \Delta_1, \Delta_2 \in \{0,1\}^n$, $(\Delta_1, M) \neq (\Delta_2, M')$, and $C \in GF(2^n)$, $\Pr[K \xleftarrow{\$} \{0,1\}^n : F(K) = C] \leq \epsilon$, where $F(K) = \mathrm{RH1}_{K\oplus\Delta_1}(M) \oplus \mathrm{RH1}_{K\oplus\Delta_2}(M')$. We have

$$F(K) = (\Delta_1 \oplus \Delta_2)K^2 \oplus (\Delta_1^2 \oplus \Delta_2^2 \oplus M \oplus M')K \oplus (\Delta_1^3 \oplus \Delta_2^3 \oplus M\Delta_1 \oplus M'\Delta_2).$$

If $\Delta_2 \neq \Delta_1$, $F(K) = C$ has two roots at most. If $\Delta_1 = \Delta_2$, then $M \neq M'$. The degree of $F(K)$ is 1 and $F(K) = C$ has one root. Therefore RH1 is $2/2^n$-RK-AXU. □

More generally we consider the polynomial $H_K^{i,j}(M) = MK^i + K^j$ over the finite field $GF(2^n)$ or $GF(p)$ where i, j are integers and p is a large prime. We have the following results when $1 \leq i, j \leq 4$.

(i,j)	(1,1)	(1,2)	(1,3)	(1,4)	(2,1)	(2,2)	(2,3)	(2,4)	(3,1)	(3,2)	(3,3)	(3,4)	(4,1)	(4,2)	(4,3)	(4,4)
$GF(2^n)$	-	-	1	-	-	-	1	-	0	0	-	0	-	-	1	-
$GF(p)$	-	-	1	1	0	-	0	1	0	1	-	1	0	1	1	-

Table 1. RK-AXU or RK-AU. Here "1" means it is RK-AXU, "0" means it is RK-AU, and "-" means it is neither RK-AU nor RK-AXU.

VIL-RK-AXU (AU). *Poly* does not support variable input length. For any message $M \in \{0,1\}^*$, a general padding method as in [28] is to firstly pad minimum zeroes to make the length multiple of the block length and then pad the bit length of M as the last block:

$$pad(M) = M\|0^i\||M|$$

where $i = n - (|M| \mod n)$. Then $Poly_K(pad(M))$ is AXU but not RK-AU (RK-AXU). In the follow construction we add some monomial K^i to get the RK-AXU property.

Construction 2 RH2 : $\{0,1\}^n \times \{0,1\}^* \to \{0,1\}^n$,

$$\mathrm{RH2}_K(M) = \begin{cases} K^{l+2} \oplus Poly_K(pad(M)), & l \text{ is odd}, \\ K^{l+3} \oplus Poly_K(pad(M))K, & l \text{ is even}, \end{cases}$$

where $l = \lceil |M|/n \rceil + 1$ *is the number of blocks in* $pad(M)$.

Theorem 2. RH2 *is* $(l_{max}+3)/2^n$*-RK-AXU over the RKD set* $\Phi^{\oplus}$*, where* l_{max} *is the maximum block number of messages after padding.*

Proof. For any message M, suppose $pad(M) = M_1\|M_2\|\cdots\|M_l$. When l is odd

$$\mathrm{RH2}_K(M) = K^{l+2} \oplus M_1K^l \oplus \cdots \oplus M_lK.$$

When l is even

$$\mathrm{RH2}_K(M) = K^{l+3} \oplus M_1K^{l+1} \oplus \cdots \oplus M_lK^2.$$

We prove that for any $M, M' \in \{0,1\}^*$, $\Delta_1, \Delta_2, C \in \{0,1\}^n$, $(\Delta_1, M) \neq (\Delta_2, M')$, $\Pr[F(K) = C] \leq \epsilon$, where $F(K) = \mathrm{RH2}_{K\oplus\Delta_1}(M) \oplus \mathrm{RH2}_{K\oplus\Delta_2}(M')$. We only need to show the degree of $F(K)$ is nonzero. Suppose $pad(M) = M_1\|M_2\|\cdots\|M_l$ and $pad(M') = M'_1\|M'_2\|\cdots\|M'_{l'}$. Consider $F(K)$ in the following two cases.

Case 1. $\Delta_1 \neq \Delta_2$. Suppose the degrees of $\mathrm{RH2}_{K\oplus\Delta_1}(M)$ and $\mathrm{RH2}_{K\oplus\Delta_2}(M')$ are d and d' respectively, which are both odd.

When $d = d'$, the coefficient of K^{d-1} in $F(K)$ is $\Delta_1 \oplus \Delta_2$ which is nonzero.

When $d \neq d'$, suppose $d > d'$ w.l.o.g. The coefficient of K^d in $F(K)$ is 1.

Case 2. $\Delta_1 = \Delta_2$. If we look at $K \oplus \Delta_1$ as a whole, that is the single-key situation. So we only consider $\Delta_1 = \Delta_2 = 0$ in the following.

When $l = l'$, there exists $1 \leq j \leq l$ s.t. $M_j \neq M'_j$. So the coefficient of K^{l+1-j} (if l is odd) or K^{l+2-j} (if l is even) in $F(K)$ is $M_j \oplus M'_j$ which is nonzero.

When $l' \neq l$ and are both odd, the coefficient of K is $|M| \oplus |M'|$ which is nonzero.

When $l' \neq l$ and are both even, the coefficient of K^2 is $|M| \oplus |M'|$ which is nonzero.

When $l' \neq l$, one is odd and one is even, the coefficient of K is $|M|$ or $|M'|$ which are both nonzero.

Therefore the degree of $F(K)$ is nonzero. □

RH2 is RK-AXU, so it is also RK-AU. But we still can improve the efficiency of RK-AU construction if replace $Poly$ in RH2 with the following $Poly'$:

$$Poly'_K(M) = M_1K^{m-1} \oplus M_2K^{m-2} \oplus \cdots \oplus M_m$$

where $M = M_1\|M_2\|\cdots\|M_m \in \{0,1\}^{nm}$. $Poly'$ is AU but not AXU and more efficient than $Poly$. We have the following construction and the proof is similar to that of theorem 2.

Construction 3 $\mathrm{RH3} : \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^n$,

$$\mathrm{RH3}_K(M) = \begin{cases} K^{l+2} \oplus Poly'_K(pad(M)), & l \text{ is odd}, \\ K^{l+3} \oplus Poly'_K(pad(M))K, & l \text{ is even}, \end{cases}$$

where $l = \lceil |M|/n \rceil + 1$ *is the number of blocks in* $pad(M)$.

Theorem 3. RH3 *is* $(l_{max}+3)/2^n$*-RK-AU over the RKD set* $\Phi^{\oplus}$*, where* l_{max} *is the maximum number of blocks in messages after padding.*

Efficiency of constructions. We analyze the efficiency of RH1, RH2 and RH3 compared with previous similar constructions.

1) RH1. Compared with $Poly_K(M) = MK$, in $\mathrm{RH1}_K(M) = MK \oplus K^3$ the monomial K^3 can be pre-computed. So RH1 need extra one pre-computation and one XOR operation.

2) RH2. The polynomial $T = M_1K^m \oplus M_2K^{m-1} \oplus \cdots \oplus M_mK$ is often computed as: $T \leftarrow 0$, $T \leftarrow (T \oplus M_i)K$ for $1 \leq i \leq m$. Assume that $pad(M) = M_1\|M_2\|\cdots\|M_l$, the following are the computation processes of $\mathrm{RH2}_K(M)$and $Poly_K(pad(M))$ respectively.

$\mathrm{RH2}_K(M)$:	$\mathrm{Poly}_K(pad(M))$:
$T \leftarrow K^2$	$T \leftarrow 0$
for $i = 1$ **to** l	**for** $i = 1$ **to** l
$T \leftarrow (T \oplus M_i)K$	$T \leftarrow (T \oplus M_i)K$
if l is even	
$T \leftarrow TK$	
return T	**return** T

We can see that compared with $Poly_K(pad(M))$, RH2 needs additional one pre-computation of K^2, and one multiplication if l is even.

3) RH3. Similar to the analysis of RH2, RH3 needs additional one pre-computation of K^2, and one multiplication if l is even, compared with $Poly'_K(pad(M))$.

In brief, RH1, RH2 and RH3 are almost as efficient as previous similar constructions.

4 Applications

RK-AU (AXU) hash functions can be used as components with related-key property in the design of cryptographic schemes. If we replace the UHFs in section 1 with our corresponding constructions, the issues about related-key attacks can be solved. Informally speaking, if the UHF is RK-AU or RK-AXU over the RKD set Φ_1 and the underlying primitive is RK-PRP or RK-PRF over the RKD set Φ_2, even if the key of UHF is a part of the key of scheme, the scheme is secure in the related-key scenario over the RKD set $\Phi_1 \times \Phi_2$.

We give four concrete applications of related-key almost universal hash function in constructions of MAC and TBC with respect to related-key securities.

Beside the Carter-Wegman scheme [42] $MAC1_{K\|K'}(N, M) = E_{K'}(N) \oplus H_K(M)$, the other method [36] to construct MAC is $MAC2_{K\|K'}(M) = E_{K'}(H_K(M))$, where E is a block cipher, H is a universal hash function, M is the message and N is a nonce. We show that the two methods are all related-key secure under some conditions by the following two theorems.

Theorem 4. *$H : \mathcal{K}_1 \times \mathcal{D} \to \{0,1\}^n$ is ϵ-RK-AXU over the RKD set Φ_1 and $F : \mathcal{K}_2 \times \{0,1\}^n \to \{0,1\}^n$ is a RK-PRF over the RKD set Φ_2, then $MAC2_{K\|K'}(N, M) = F_{K'}(N) \oplus (H_K(M))$ is RK-UF over the RKD set $\Phi_1 \times \Phi_2$. More specifically,*

$$\mathbf{Adv}^{rk-uf}_{MAC1}(q, t) \leq \mathbf{Adv}^{rk-prf}_{F}(q, t') + \epsilon$$

where where the adversary makes q queries to $MAC2$ and $t' = t + O(q)$.

The proof is similar to that in [25].

Theorem 5. *$H : \mathcal{K}_1 \times \mathcal{D} \to \mathcal{R}$ is ϵ-RK-AU over the RKD set Φ_1 and $F : \mathcal{K}_2 \times \mathcal{R} \to \mathcal{R}$ is a RK-PRF over the RKD set Φ_2, then $MAC2_{K\|K'}(M) = F_{K'}(H_K(M))$, where N is a nonce, is a RK-PRF over the RKD set $\Phi_1 \times \Phi_2$. More specifically,*

$$\mathbf{Adv}^{rk-prf}_{MAC1}(q, t) \leq \mathbf{Adv}^{rk-prf}_{F}(q, t') + \epsilon q^2/2$$

where where the adversary makes q queries to $MAC1$ and $t' = t + O(q)$.

The proof is similar to that in [36].

The other two applications are in constructions of related-key strongly tweakable pseudorandom permutation permutation (RK-STPRP).

Theorem 6. *$H : \mathcal{K}_1 \times \mathcal{D} \to \{0,1\}^n$ is ϵ-RK-AXU over the RKD set Φ_1 and $E : \mathcal{K}_2 \times \{0,1\}^n \to \{0,1\}^n$ is RK-PRP over the RKD set Φ_2, then $TBC1_{K\|K'}(T,M) = E_{K'}(M \oplus H_K(T)) \oplus H_K(T)$ is a RK-STPRP over the RKD set $\Phi_1 \times \Phi_2$. More specifically,*

$$\mathbf{Adv}^{rk-stprp}_{TBC1}(q,t) \leq \mathbf{Adv}^{rk-sprp}_{E}(q,t') + 3\epsilon q^2$$

where the adversary makes q queries to $TBC1$ or $TBC1^{-1}$ and $t' = t + O(q)$.

The proof is similar to that in [27]. If we replace the block cipher with a complex permutation, TBC is still can be proved to be related-key secure in the ideal permutation model.

Theorem 7. *$H : \mathcal{K}_1 \times \mathcal{D} \to \{0,1\}^n$ is (ϵ,δ)-RK-AXU [4] over the RKD set Φ and π is a public random permutation from $\{0,1\}^m$ to $\{0,1\}^m$, $n \leq m$, then $TBC2_K(T,M) = \pi(M \oplus H_K(T)) \oplus H_K(T)$ [5] is a RK-TPRP over the RKD set Φ. More specifically,*

$$\mathbf{Adv}^{rk-stprp}_{TBC2}(q_0,q_1) \leq q_0^2\epsilon + 2q_0q_1\delta + 2^{-m}(q_0^2 + 2q_0q_1)$$

where the adversary makes q_0 queries to $TBC2$ or $TBC2^{-1}$ and q_1 queries to π or π^{-1}.

The proof is similar to that in [26]. Compared with the construction in [29] we only need one permutation invocation (two in [29]).

5 Conclusions

In order to solve the related-key attack problem again some of the UHF based schemes, we define a new concept of related-key almost universal hash function, which is a natural extension to almost universal hash function in the related-key scenario. We also give several efficient constructions such as RH1, RH2 and RH3 which are nearly as efficient as previous similar constructions. Finally we give four concrete applications of RK-AU and RK-AXU in MACs and TBCs.

References

1. Abed, F., Fluhrer, S., Foley, J., Forler, C., List, E., Lucks, S., McGrew, D., , Wenzel, J.: The POET family of on-line authenticated encryption schemes (2014), http://competitions.cr.yp.to/caesar-submissions.html 2
2. Andreeva, E., Bogdanov, A., Lauridsen, M.M., Luykx, A., Mennink, B., Tischhauser, E., Yasuda, K.: AES-COBRA (2014), http://competitions.cr.yp.to/caesar-submissions.html 2
3. Bernstein, D.J.: The poly1305-aes message-authentication code. In: Gilbert, H., Handschuh, H. (eds.) Fast Software Encryption: 12th International Workshop, FSE 2005, Paris, France, February 21-23, 2005, Revised Selected Papers. Lecture Notes in Computer Science, vol. 3557, pp. 32–49. Springer (2005), http://dx.doi.org/10.1007/11502760_3 2

[4] H is (ϵ,δ)-RK-AXU means H is ϵ-RK-AXU and for any $\phi \in \Phi$, $M \in \mathcal{D}$ and $C \in \{0,1\}^n$, $\Pr[K \xleftarrow{\$} \mathcal{K} : H_{\phi(K)}(M) = C] \leq \delta$.

[5] When $n < m$, $A \in \{0,1\}^n$, $B \in \{0,1\}^m$, $A \oplus B$ is defined as $(A\|0^*) \oplus B$.

4. Bernstein, D.J.: Stronger security bounds for wegman-carter-shoup authenticators. In: Cramer, R. (ed.) Advances in Cryptology - EUROCRYPT 2005. Lecture Notes in Computer Science, vol. 3494, pp. 164–180. Springer (2005), http://dx.doi.org/10.1007/11426639_10 2
5. Bernstein, D.J.: Polynomial evaluation and message authentication (2011), http://cr.yp.to/papers.html#pema 1
6. Bhattacharyya, R., Roy, A.: Secure message authentication against related-key attack. In: Moriai, S. (ed.) Fast Software Encryption - 20th International Workshop, FSE 2013. Lecture Notes in Computer Science, vol. 8424, pp. 305–324. Springer (2013), http://dx.doi.org/10.1007/978-3-662-43933-3_16 2
7. Biham, E.: New types of cryptoanalytic attacks using related keys (extended abstract). In: Helleseth, T. (ed.) EUROCRYPT. Lecture Notes in Computer Science, vol. 765, pp. 398–409. Springer (1993) 2
8. Biryukov, A., Khovratovich, D.: Related-key cryptanalysis of the full AES-192 and AES-256. In: Matsui, M. (ed.) Advances in Cryptology - ASIACRYPT 2009. Lecture Notes in Computer Science, vol. 5912, pp. 1–18. Springer (2009), http://dx.doi.org/10.1007/978-3-642-10366-7_1 2
9. Black, J., Halevi, S., Krawczyk, H., Krovetz, T., Rogaway, P.: UMAC: fast and secure message authentication. In: Wiener [43], pp. 216–233, http://dx.doi.org/10.1007/3-540-48405-1_14 2, 3, 4, 12
10. Boesgaard, M., Christensen, T., Zenner, E.: Badger - A fast and provably secure MAC. In: Ioannidis, J., Keromytis, A.D., Yung, M. (eds.) Applied Cryptography and Network Security, Third International Conference, ACNS 2005, New York, NY, USA, June 7-10, 2005, Proceedings. Lecture Notes in Computer Science, vol. 3531, pp. 176–191 (2005), http://dx.doi.org/10.1007/11496137_13 2
11. Carter, L., Wegman, M.N.: Universal classes of hash functions. J. Comput. Syst. Sci. 18(2), 143–154 (1979) 1, 2
12. Chakraborty, D., Sarkar, P.: HCH: A new tweakable enciphering scheme using the hash-encrypt-hash approach. In: Barua, R., Lange, T. (eds.) Progress in Cryptology - INDOCRYPT 2006. Lecture Notes in Computer Science, vol. 4329, pp. 287–302. Springer (2006), http://dx.doi.org/10.1007/11941378_21 2, 3
13. Chakraborty, D., Sarkar, P.: HCH: A new tweakable enciphering scheme using the hash-counter-hash approach. IEEE Transactions on Information Theory 54(4), 1683–1699 (2008), http://dx.doi.org/10.1109/TIT.2008.917623 2, 3
14. Chen, J., Miyaji, A.: A new practical key recovery attack on the stream cipher RC4 under related-key model. In: Lai, X., Yung, M., Lin, D. (eds.) Inscrypt 2010. Lecture Notes in Computer Science, vol. 6584, pp. 62–76. Springer (2010), http://dx.doi.org/10.1007/978-3-642-21518-6_5 2
15. Dobraunig, C., Eichlseder, M., Mendel, F.: Related-key forgeries for prøst-otr. IACR Cryptology ePrint Archive, to appear in FSE 2015 (2015), http://eprint.iacr.org/2015/091 2
16. Dunkelman, O., Keller, N., Shamir, A.: A practical-time related-key attack on the KASUMI cryptosystem used in GSM and 3g telephony. J. Cryptology 27(4), 824–849 (2014), http://dx.doi.org/10.1007/s00145-013-9154-9 2
17. Etzel, M., Patel, S., Ramzan, Z.: SQUARE HASH: fast message authenication via optimized universal hash functions. In: Wiener [43], pp. 234–251, http://dx.doi.org/10.1007/3-540-48405-1_15 3, 4, 12
18. Halevi, S., Krawczyk, H.: MMH: software message authentication in the gbit/second rates. In: Biham, E. (ed.) Fast Software Encryption 1997. Lecture Notes in Computer Science, vol. 1267, pp. 172–189. Springer (1997), http://dx.doi.org/10.1007/BFb0052345 3, 4, 12
19. Harris, S.: AES-COBRA (2014), http://competitions.cr.yp.to/caesar-submissions.html 2
20. IEEE Std 1619-2007: IEEE standard for cryptographic protection of data on block-oriented storage devices (2008) 2
21. IEEE Std 1619.2-2010: IEEE standard for wide-block encryption for shared storage media (2011) 2
22. ISO/IEC 19772:2009: Information technology – security techniques – authenticated encryption (2009) 2
23. ISO/IEC9797-3:2011: Information technology – security techniques – message authentication codes (MACs) – part 3: Mechanisms using a universal hash-function (2011) 2

24. Jutla, C.S.: Encryption modes with almost free message integrity. In: Pfitzmann, B. (ed.) EUROCRYPT. Lecture Notes in Computer Science, vol. 2045, pp. 529–544. Springer (2001), http://dx.doi.org/10.1007/3-540-44987-6_32 3
25. Krawczyk, H.: LFSR-based hashing and authentication. In: Desmedt, Y. (ed.) Advances in Cryptology - CRYPTO '94. Lecture Notes in Computer Science, vol. 839, pp. 129–139. Springer (1994), http://dx.doi.org/10.1007/3-540-48658-5_15 4, 7
26. Kurosawa, K.: Power of a public random permutation and its application to authenticated encryption. IEEE Transactions on Information Theory 56(10), 5366–5374 (2010), http://dx.doi.org/10.1109/TIT.2010.2059636 3, 8
27. Liskov, M., Rivest, R.L., Wagner, D.: Tweakable block ciphers. In: Yung, M. (ed.) CRYPTO. Lecture Notes in Computer Science, vol. 2442, pp. 31–46. Springer (2002) 2, 3, 8
28. McGrew, D.A., Viega, J.: The Galois/Counter mode of operation (GCM) (2004), http://csrc.nist.gov/groups/ST/toolkit/BCM/ 2, 5
29. Mennink, B.: XPX: generalized tweakable even-mansour with improved security guarantees. IACR Cryptology ePrint Archive 2015, 476 (2015), http://eprint.iacr.org/2015/476 8
30. NIST SP 800-38D: Recommendations for block cipher modes of operation: Galois/counter mode (GCM) and GMAC (November 2007) 2
31. NIST SP 800-38E: Recommendation for block cipher modes of operation: The XTS-AES mode for confidentiality on storage devices (2010) 2
32. Peyrin, T., Sasaki, Y., Wang, L.: Generic related-key attacks for HMAC. In: Wang, X., Sako, K. (eds.) Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7658, pp. 580–597. Springer (2012), http://dx.doi.org/10.1007/978-3-642-34961-4_35 2
33. Rogaway, P.: Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC. In: Lee, P.J. (ed.) Advances in Cryptology - ASIACRYPT 2004. Lecture Notes in Computer Science, vol. 3329, pp. 16–31. Springer (2004), http://dx.doi.org/10.1007/978-3-540-30539-2_2 3
34. Rogaway, P., Bellare, M., Black, J., Krovetz, T.: OCB: a block-cipher mode of operation for efficient authenticated encryption. In: Reiter, M.K., Samarati, P. (eds.) ACM Conference on Computer and Communications Security. pp. 196–205. ACM (2001) 3
35. Shoup, V.: On fast and provably secure message authentication based on universal hashing. In: Koblitz, N. (ed.) Advances in Cryptology - CRYPTO '96. Lecture Notes in Computer Science, vol. 1109, pp. 313–328. Springer (1996), http://dx.doi.org/10.1007/3-540-68697-5_24 2
36. Shoup, V.: Sequences of games: a tool for taming complexity in security proofs. IACR Cryptology ePrint Archive 2004, 332 (2004), http://eprint.iacr.org/2004/332 7, 8
37. Stinson, D.R.: Universal hashing and authentication codes. In: Feigenbaum, J. (ed.) Advances in Cryptology - CRYPTO '91, 11th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1991, Proceedings. Lecture Notes in Computer Science, vol. 576, pp. 74–85. Springer (1991), http://dx.doi.org/10.1007/3-540-46766-1_5 4
38. Stinson, D.R.: On the connections between universal hashing, combinatorial designs and error-correcting codes. Electronic Colloquium on Computational Complexity (ECCC) 2(52) (1995), http://eccc.hpi-web.de/eccc-reports/1995/TR95-052/index.html 4
39. Sun, S., Hu, L., Wang, P., Qiao, K., Ma, X., Song, L.: Automatic security evaluation and (related-key) differential characteristic search: Application to SIMON, PRESENT, LBlock, DES(L) and other bit-oriented block ciphers. In: Sarkar, P., Iwata, T. (eds.) Advances in Cryptology - ASIACRYPT 2014. Lecture Notes in Computer Science, vol. 8873, pp. 158–178. Springer (2014), http://dx.doi.org/10.1007/978-3-662-45611-8_9 2

40. Sun, Z., Wang, P., Zhang, L.: Weak-key and related-key analysis of hash-counter-hash tweakable enciphering schemes. In: Foo, E., Stebila, D. (eds.) Information Security and Privacy - 20th Australasian Conference, ACISP 2015. Lecture Notes in Computer Science, vol. 9144, pp. 3–19. Springer (2015), http://dx.doi.org/10.1007/978-3-319-19962-7_1 2, 3
41. Wang, P., Feng, D., Wu, W.: HCTR: A variable-input-length enciphering mode. In: Feng, D., Lin, D., Yung, M. (eds.) Information Security and Cryptology, First SKLOIS Conference, CISC 2005, Beijing, China, December 15-17, 2005, Proceedings. Lecture Notes in Computer Science, vol. 3822, pp. 175–188. Springer (2005), http://dx.doi.org/10.1007/11599548_15 2, 3
42. Wegman, M.N., Carter, L.: New hash functions and their use in authentication and set equality. J. Comput. Syst. Sci. 22(3), 265–279 (1981) 1, 2, 7
43. Wiener, M.J. (ed.): Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings, Lecture Notes in Computer Science, vol. 1666. Springer (1999) 9

A Related-key security of MAC, TBC, TES and AE

1) MAC. A message authentication code (MAC) is a function $F : \mathcal{K} \times \mathcal{N} \times \mathcal{M} \to \{0,1\}^n$, where $\mathcal{K}$, $\mathcal{N}$, $\mathcal{M}$ and $\{0,1\}^n$ are spaces of key, nonce, message and tag respectively. The nonce space can be an empty set $\mathcal{N} = \emptyset$. For a RKD set Φ, an adversary $\mathbb{A}$ queries the MAC algorithm with $(\phi, N, M) \in \Phi \times \mathcal{N} \times \mathcal{M}$ but never repeats N, and gets $T = F_{\phi(K)}(N, M)$. After several queries $\mathbb{A}$ returns a quadruple (ϕ', N', M', T') which never appear before in the queries. We define the probability of $T' = F_{\phi'(K)}(N', M')$ as the advantage of $\mathbb{A}$ and write it as:

$$\mathbf{Adv}_F^{rk-uf}(\mathbb{A}) = \Pr[\mathbb{A}^{F_{\cdot(K)}(\cdot,\cdot)} \text{ forges}].$$

For all adversaries with computation time at most t, oracle queries at most q, we denote $\mathbf{Adv}_F^{rk-uf}(q,t) = max_{\mathbb{A}}\mathbf{Adv}_F^{rk-uf}(\mathbb{A})$. When the advantage is negligible, we say that F is related-key unforgeable (RK-UF) or related-key unpredictable.

2) TBC. A tweakable block cipher consists of two algorithms $(\mathbf{E}, \mathbf{D})$. The encryption algorithm $\mathbf{E} : \mathcal{K} \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$, where $\mathcal{K}$, $\mathcal{T}$ and $\{0,1\}^n$ are spaces of key, tweak, plaintext/ciphertext respectively. For input $(K, T, P) \in \mathcal{K} \times \mathcal{T} \times \{0,1\}^n$, we write the result as $C = \mathbf{E}_K^T(P)$ The decryption algorithm $\mathbf{D} : \mathcal{K} \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$. We require that for any $(K, T) \in \mathcal{K} \times \mathcal{T}$, $\mathbf{E}_K^T(\cdot)$ and $\mathbf{D}_K^T(\cdot)$ are permutations, and $\mathbf{D}_K^T(\mathbf{E}_K^T(P)) = P$. For a RKD set Φ, an adversary $\mathbb{A}$ queries $\mathbf{E}$ with $(\phi, T, P) \in \Phi \times \mathcal{T} \times \{0,1\}^n$ or queries $\mathbf{D}$ with $(\phi, T, C) \in \Phi \times \mathcal{T} \times \{0,1\}^n$. $\mathbb{A}$ tries to distinguish $\mathcal{S}$ from an ideal TBF, where for any $(K, T) \in \mathcal{K} \times \mathcal{T}$ π_K^T is independent uniformly random permutation. We define the advantage as

$$\mathbf{Adv}_{\mathbf{E}}^{rk-stprp}(\mathbb{A}) = \Pr[\mathbb{A}^{\mathbf{E}_{\cdot(T)}^{\cdot}(\cdot),\mathbf{D}_{\cdot(T)}^{\cdot}(\cdot)} \Rightarrow 1] - \Pr[\mathbb{A}^{\pi_{\cdot(T)}^{\cdot}(\cdot),\pi^{-1}_{\cdot(T)}(\cdot)} \Rightarrow 1].$$

For all adversaries with computation time at most t, oracle queries at most q, we denote $\mathbf{Adv}_F^{rk-stprp}(q,t) = max_{\mathbb{A}}\mathbf{Adv}_F^{rk-stprp}(\mathbb{A})$. When the advantage is negligible, we say that $\mathbf{E}$ is a related-key relate-key strongly pseudorandom permutation (RK-STPRP).

3) TES. A tweakable enciphering schemes are TBCs with large or variable input length. The definition is the same as that of TBC.

4) AE. An authenticated encryptions consists of two algorithms $\mathcal{SE} = (\mathbf{E}, \mathbf{D})$. The encryption $\mathbf{E} : \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{P} \to \mathcal{C}$, where $\mathcal{K}$, $\mathcal{N}$, $\mathcal{A}$, $\mathcal{P}$ and $\mathcal{C}$ are spaces of key, nonce, associated data, plaintext and

ciphertext respectively. For input $(K, N, A, P) \in \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{P}$, we write the result as $C = \mathbf{E}_K(N, A, P)$. The decryption algorithm $\mathbf{D} : \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{C} \to \mathcal{P} \cup \{\perp\}$. We require that $\mathbf{D}_K(N, A, \mathbf{E}_K(N, A, P)) = P$. For a RKD set Φ, an adversary $\mathbb{A}$ queries the $\mathbf{E}$ with $(\phi, N, A, P) \in \Phi \times \mathcal{N} \times \mathcal{A} \times \mathcal{P}$ but never repeats (ϕ, N), or queries the $\mathbf{D}$ with (ϕ, N, A, C). $\mathbb{A}$ tries to distinguish $\mathcal{SE}$ from an ideal AE $(\$, \perp)$, where for any query $\$$ returns a random string and $\perp$ always returns $\perp$. We define the advantage as

$$\mathbf{Adv}_{\mathbf{E}}^{rk-ae}(\mathbb{A}) = \Pr[\mathbb{A}^{\mathbf{E}_{\cdot(K)}(\cdot,\cdot,\cdot),\mathbf{D}_{\cdot(K)}(\cdot,\cdot,\cdot)} \Rightarrow 1] - \Pr[\mathbb{A}^{\$(\cdot,\cdot,\cdot,\cdot),\perp(\cdot,\cdot,\cdot,\cdot)} \Rightarrow 1].$$

For all adversaries with computation time at most t, oracle queries at most q, we denote $\mathbf{Adv}_F^{rk-ae}(q, t) = max_{\mathbb{A}}\mathbf{Adv}_F^{rk-ae}(\mathbb{A})$. When the advantage is negligible, we say that $\mathcal{SE}$ is related-key secure.

B Existing UHFs are not RK-AXU (RK-AU)

The following universal hash functions are proved to be AXU (AΔU).

1) MMH [18]: $H_K(M) = (((\sum_{i=1}^{t} M_i K_i) \mod 2^{64}) \mod p) \mod 2^{32}$, $M_i, K_i \in \mathbf{Z}_{2^{32}}$ and $p = 2^{32} + 15$;
2) Square Hash [17]: $H_K(M) = \sum_{i=1}^{t}(M_i + K_i)^2 \mod p$, $M_i, K_i \in \mathbf{Z}_p$;
3) NMH [18]: $H_K(M) = (\sum_{i=1}^{t/2}(M_{2i-1} + K_{2i-1})(M_{2i} + K_{2i})) \mod p$, $M_i, K_i \in \mathbf{Z}_{2^{32}}$, $p = 2^{32} + 15$;
4) NH [9]: $H_K(M) = (\sum_{i=1}^{t/2}((M_{2i-1}+K_{2i-1}) \mod 2^w)((M_{2i}+K_{2i}) \mod 2^w)) \mod 2^{2w}$, $M_i, K_i \in \mathbf{Z}_{2^w}$.

In 1) we set $t = 1$, then $H_K(M) = (MK \mod 2^{32} + 15) \mod 2^{32}$. If $M = M' = \Delta' = 1$, $\Delta = 0$, then $H_K(M) = K$, $H_{K+\Delta'}(M') = K + 1 \mod 2^{32}$, therefore $H_K(M) + 1 = H_{K+\Delta'}(M')$, MMH is not RK-A$\Delta$U. 2), 3) and 4) all have the term $M_1 + K_1$. From $M_1 + K_1 = (M_1 - 1) + (K_1 + 1)$ we know that they are all not RK-AU.