Tight PRF-Security of Double-block Hash-then-Sum MACs

Seongkwang Kim, **<u>Byeonghak Lee</u>**, Jooyoung Lee

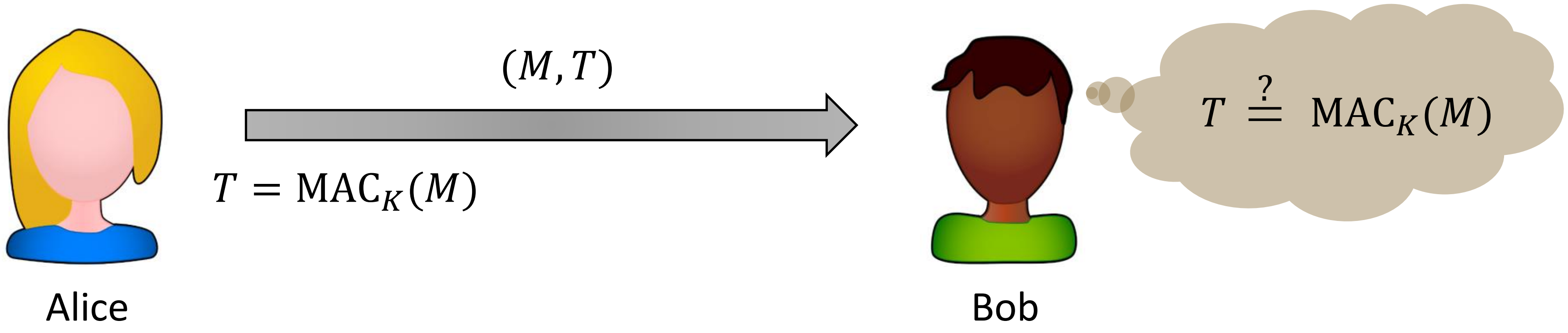
KAIST

Outline

- Introduction
  - Message Authentication Code
  - Double-block Hash-then-Sum paradigm
- Our Contribution
  - Tight security proof of DbHtS MACs
  - Refining Mirror theory
- Conclusion

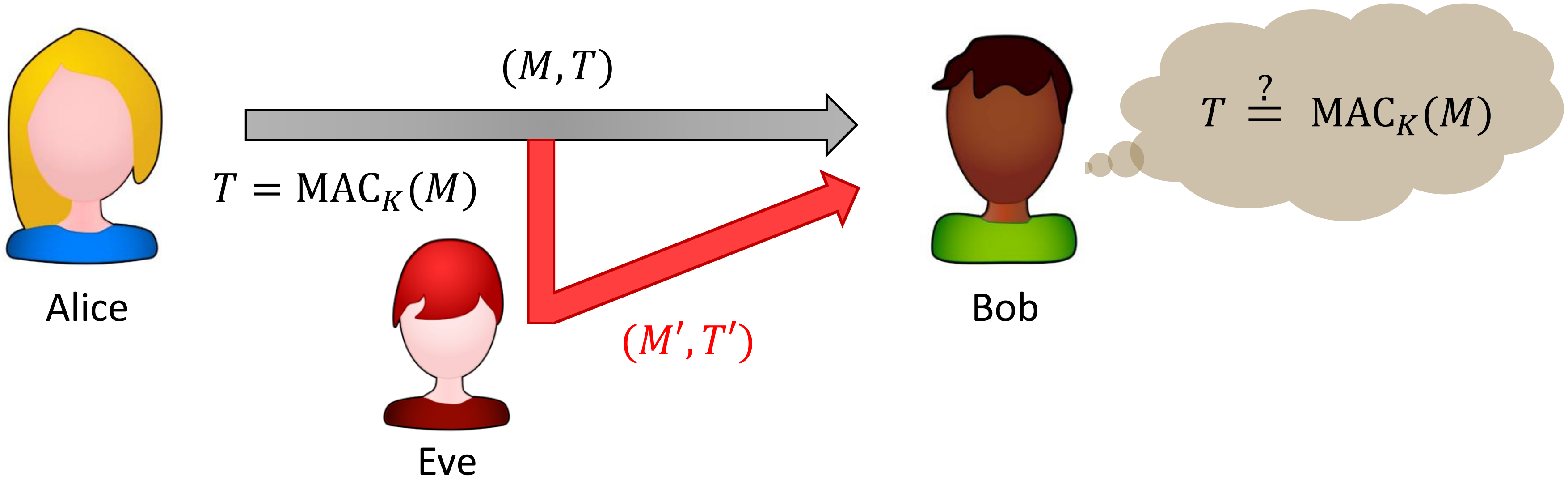
Message Authentication Code (MAC)

- Symmetric key functions to guarantee message integrity
- Alice computes tag $T = \mathrm{MAC}_K(M)$ and sends (M, T) to Bob
- Bob checks whether the tag is valid or not by computing $\mathrm{MAC}_K(M)$



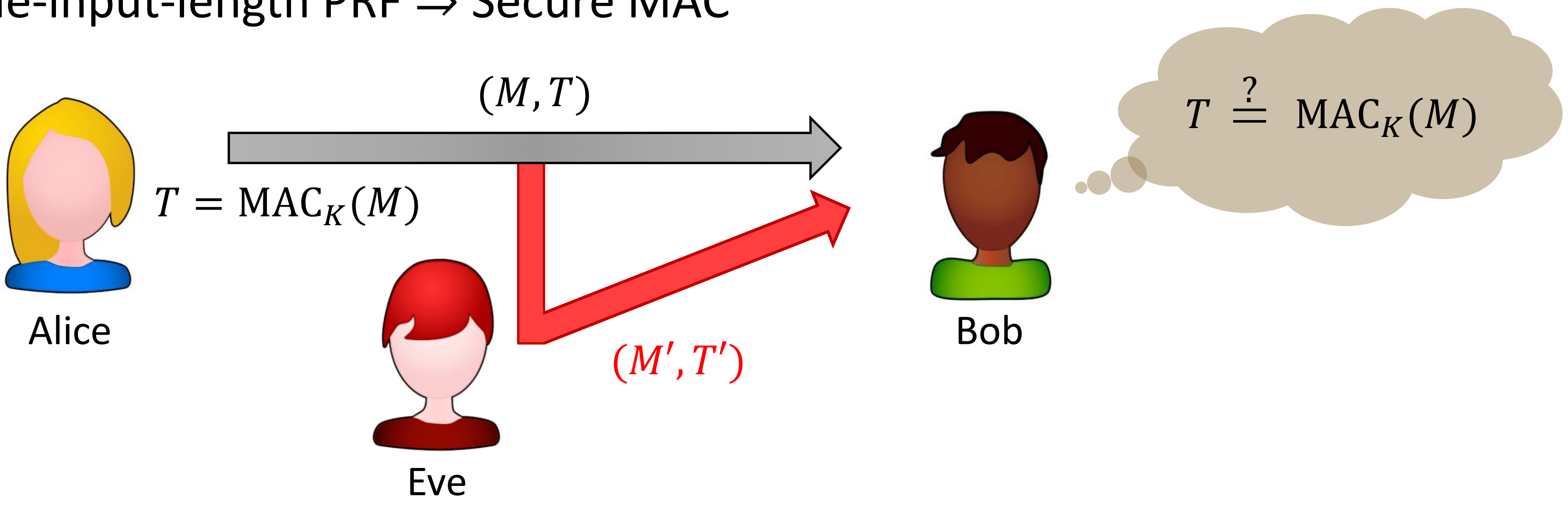
Message Authentication Code (MAC)

- Symmetric key functions to guarantee message integrity
- Alice computes tag $T = \mathrm{MAC}_K(M)$ and sends (M, T) to Bob
- Bob checks whether the tag is valid or not by computing $\mathrm{MAC}_K(M)$

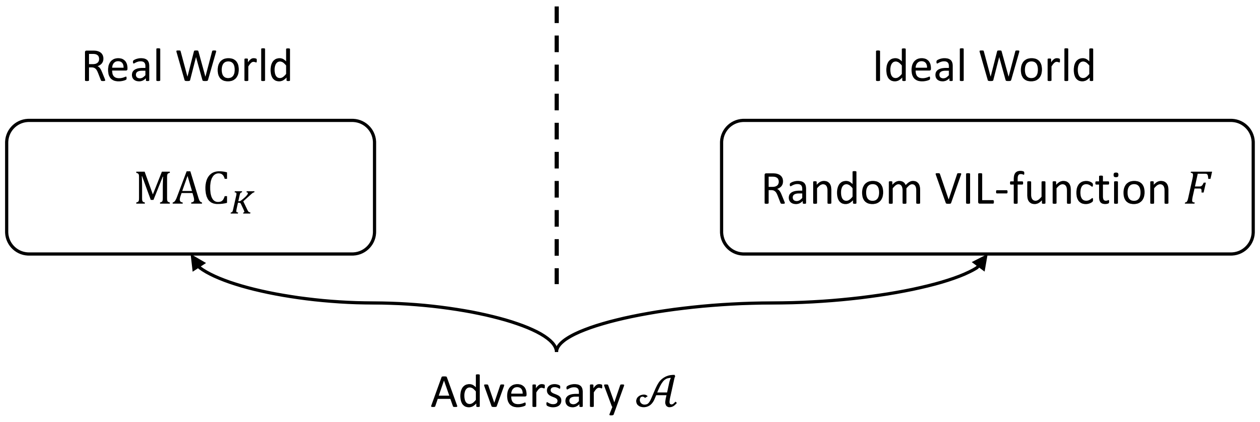


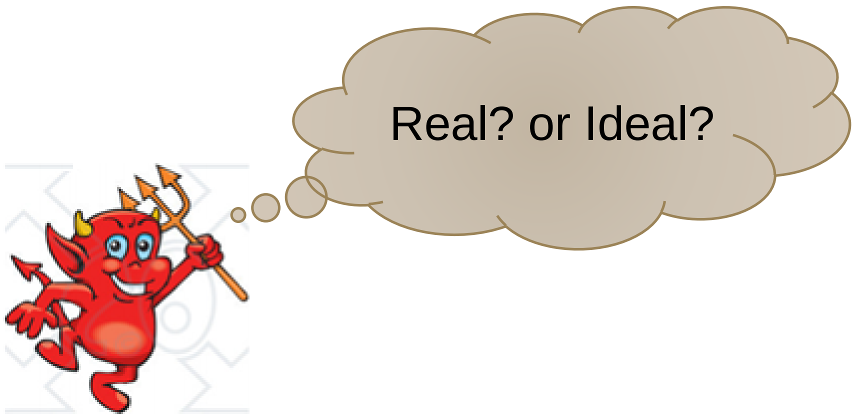
MAC Security

- Unforgeability
  - Infeasible to generate a new valid message/tag pair
- PRF-Security
  - Infeasible to distinguish from a random variable-input-length (VIL) function
  - Secure variable-input-length PRF ⇒ Secure MAC



Distinguishing Game





- Adversary $\mathcal{A}$ makes q queries to oracle (MAC_K or F)
- Each query has length at most l blocks
- Transcript $\tau = \left((M_1, T_1), \dots, (M_q, T_q)\right)$
- $\mathrm{Adv}(q, l)$: $\Pr[\mathcal{A}$ correctly determine the interacting world$] - \frac{1}{2}$

Why BBB-Security?

- Most popular MACs provides birthday-bound security
  - With n-bit block cipher, only $2^{n/2}$ security
- In lightweight cryptography, small blocks (64bits / 80bits) are preferred
  - birthday-bound security is insufficient

Construction	key bits	# of allowed queries
ECBC	64	2^{25}
PMAC	128	2^{18}

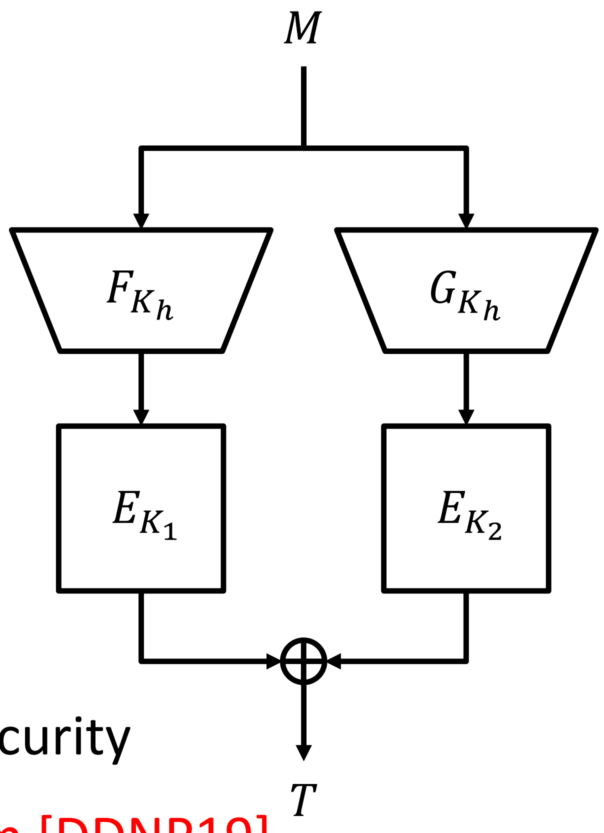
Table*: Data limits of MACs using 64-bit blocks to ensure that the advantage is less than 2^{-10} where each message is shorter than 512KB

- Beyond-Birthday-Bound secure MACs needed!

*Example chosen by Datta et al., in “Double-block Hash-then-Sum: A Paradigm for Constructing BBB Secure PRF”

BBB-Secure MACs

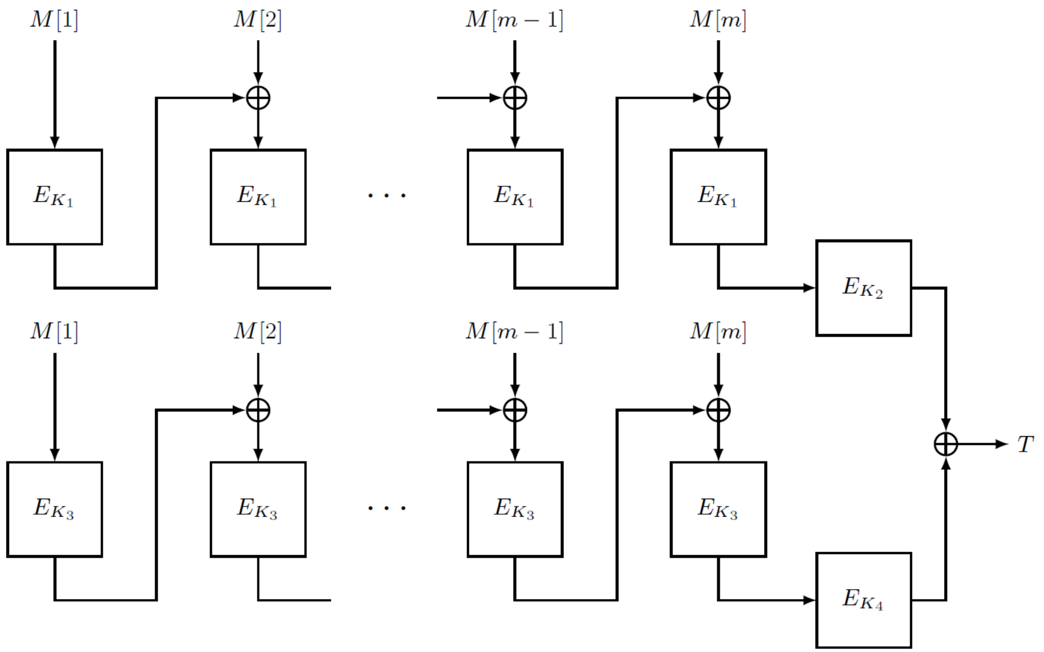
- Ideal cipher / tweakable block cipher based MACs
  - ZMAC[IMPS17], ZMAC+[LN17], HaT, HaK[CLS17]
  - Highly secure MACs from strong primitives



- Block cipher based MACs?
  - UHF-then-PRF* style MACs with n-bit internal state provides $n/2$-bit security
  - Idea: use $2n$-bit state ⇒ Double-block Hash-then-Sum (DbHtS) paradigm [DDNP19]
    - SUM-ECBC, 3kf9, PMAC-Plus, LightMAC-Plus
    - Their security has been proved up to $O(2^{2n/3})$ queries

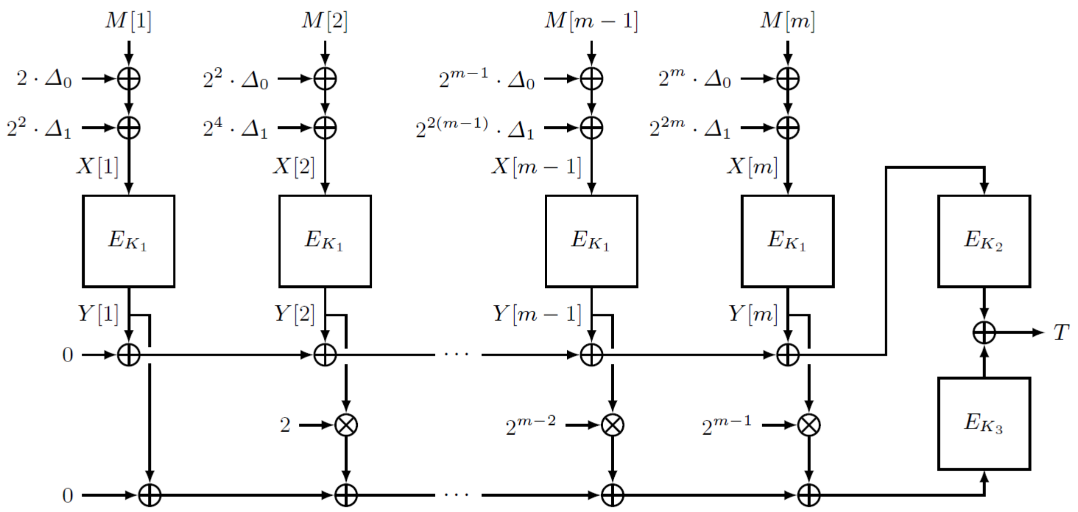
*Universal Hash Function then Pseudorandom Function

Double-Block Hash-then-Sum



SUM-ECBC [Yasuda, CT-RSA 2010]

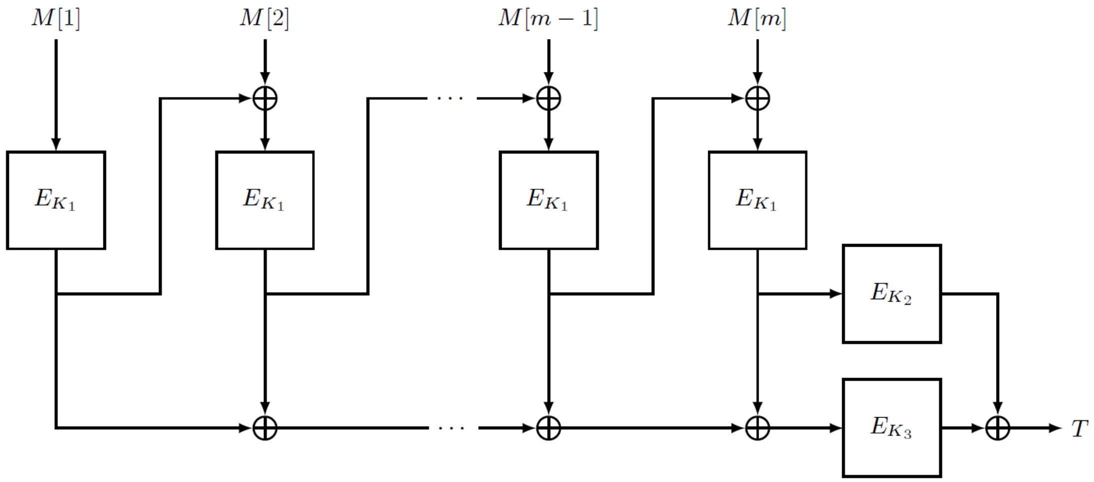
- The first BBB-secure MACs



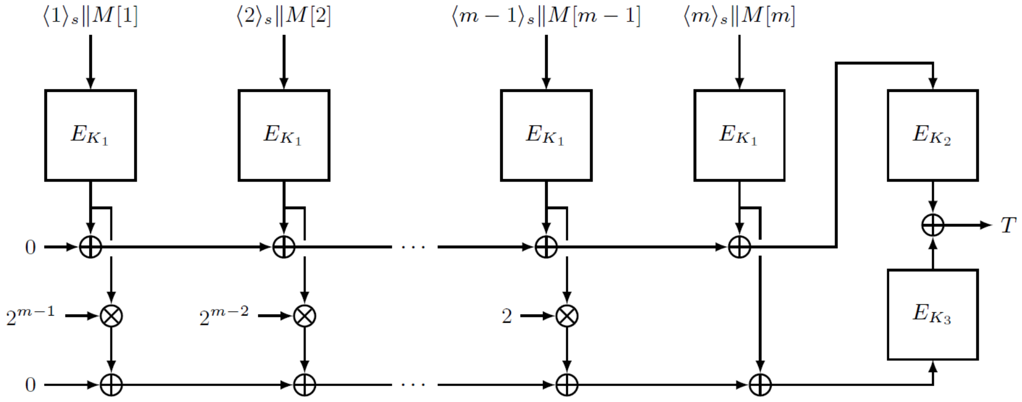
PMAC-Plus [Yasuda, CRYPTO 2011]

- Parallelizable, Rate-1 with BBB-security

Double-Block Hash-then-Sum



3kf9 [Zhang et al., ASIACRYPT 2012]



LightMAC-Plus [Naito, ASIACRYPT 2017]

- 3GPP-MAC + ECBC
- Rate-1 without field operation

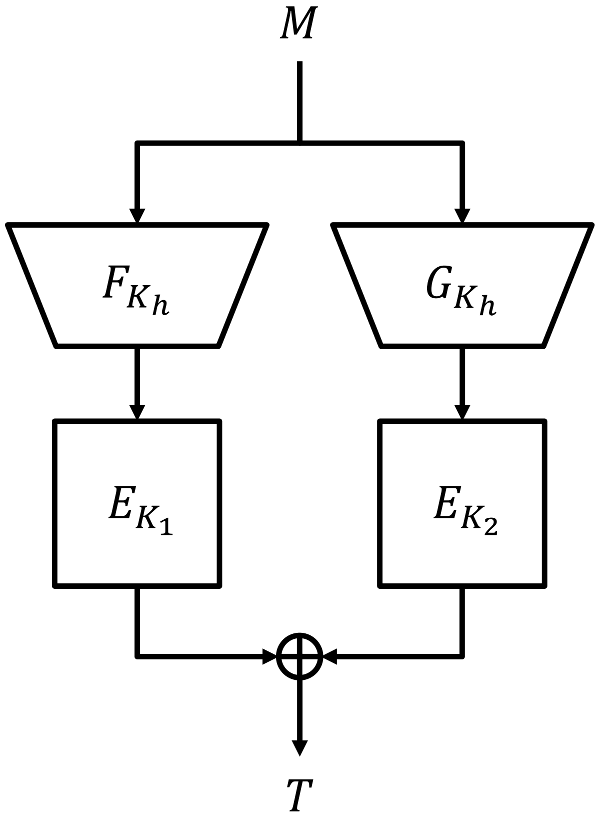
- Message-length-independent security

Generic Attacks on DbHtS MACs

- Generic attacks with $O(2^{3n/4})$ queries [LNS18]
  - Exploited the difference between Xor of Permutations (XoP) and the ideal $2n$-to-n bit function

$$E_{K_1}(F(M_1)) \oplus E_{K_2}(G(M_1)) = T_1$$
$$E_{K_1}(F(M_2)) \oplus E_{K_2}(G(M_2)) = T_2$$
$$E_{K_1}(F(M_3)) \oplus E_{K_2}(G(M_3)) = T_3$$
$$E_{K_1}(F(M_4)) \oplus E_{K_2}(G(M_4)) = T_4$$

$$\Rightarrow \quad T_1 \oplus T_2 \oplus T_3 \oplus T_4 = 0$$



Gap exists between the best known attacks and their provable security!

Outline

- Introduction
  - Message Authentication Code
  - Double-block Hash-then-Sum paradigm
- Our Contribution
  - Tight security proof of DbHtS MACs
  - Refining Mirror theory
- Conclusion

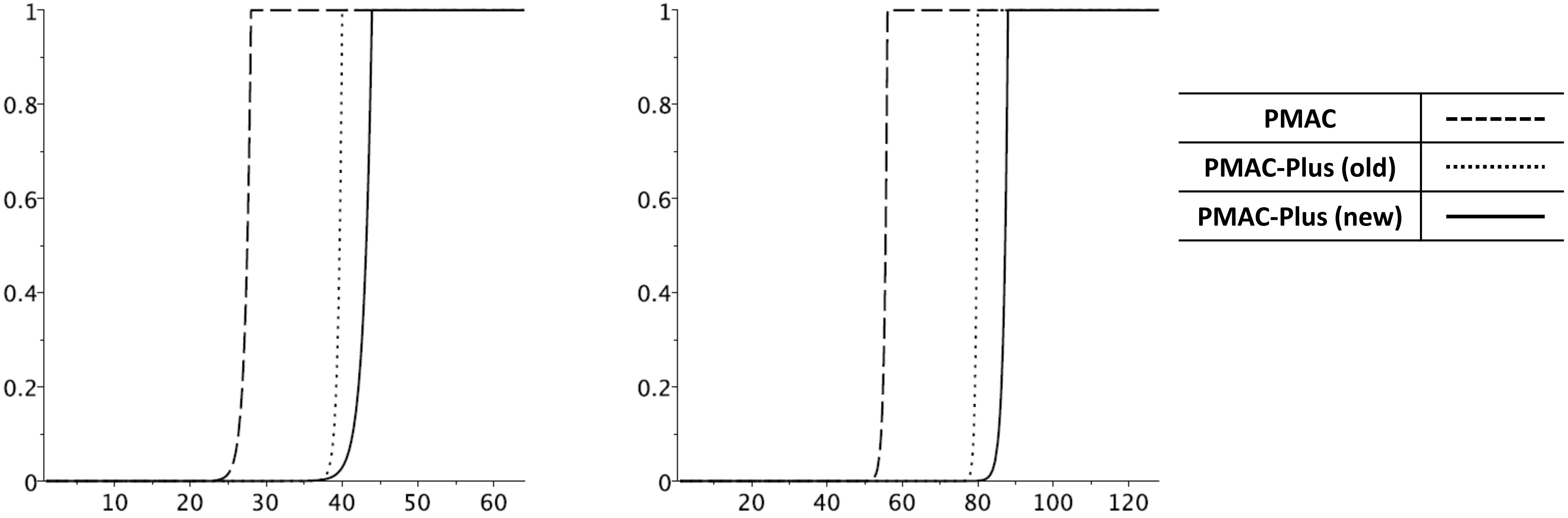
Tight Security of DbHtS MACs

- Proved $3n/4$-bit security of DbHtS MACs
  - Closed the gap between generic attacks and provable security bounds
  - Identify the required properties of the underlying hash functions

Construction	# Keys	Rate	Old Bound	New Bound
PolyMAC	4	-	$l^2q^3/2^{2n}$	$l^3q^4/2^{3n}$
SUM-ECBC	4	$1/2$	$l^2q/2^n + q^3/2^{2n}$	$l^3q^4/2^{3n}$
PMAC-Plus	3	1	$lq^3/2^{2n}$	$l^2q^4/2^{3n} + l^2q/2^n$
3kf9	3	1	$l^4q^3/2^{2n}$	$l^6q^4/2^{3n}$
LightMAC-Plus	3	$1 - s/n$	$q^3/2^{2n}$	$q^4/2^{3n}$

Table: Security bound of DbHtS MACs. q denotes the number of queries, l denotes maximum block length, and s denotes the length of prefix for LightMAC-Plus

Comparison of Security Bounds for PMAC-Plus

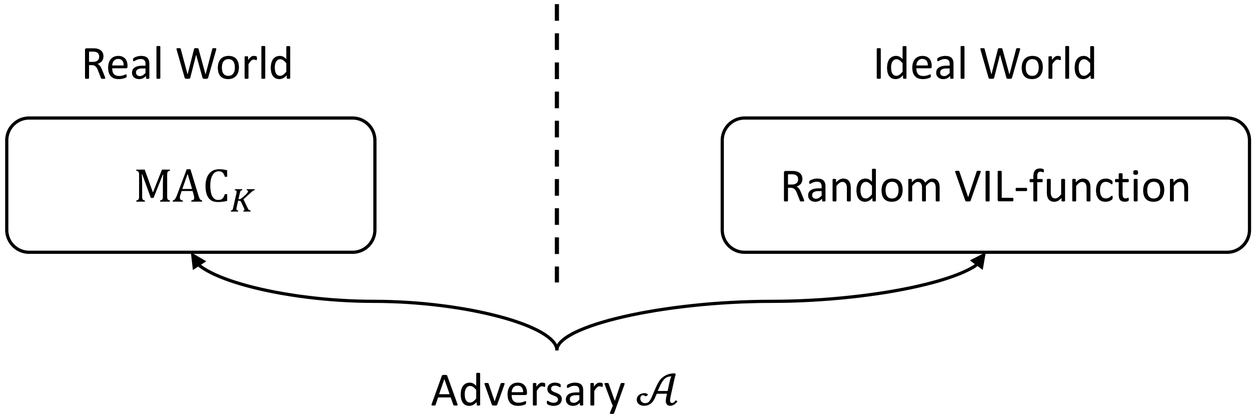


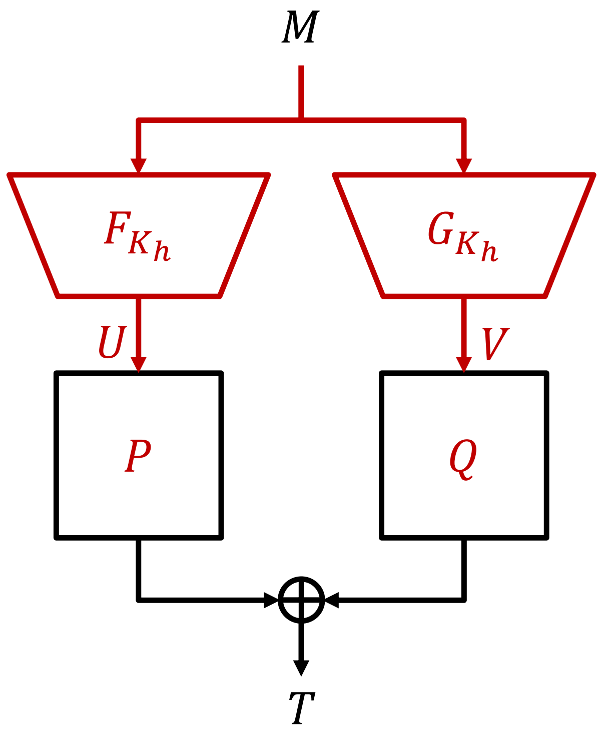
(a) $n = 64$ and $\ell = 2^8$

(b) $n = 128$ and $\ell = 2^{16}$

Figure: Upper bounds on distinguishing advantage for PMAC and PMAC-Plus.
x-axis gives the log of number of queries, and y-axis gives the security bounds.

H-Coefficient Technique





- SPRP switch
  - Replace E_{K_1} and E_{K_2} by random permutations P and Q up to the to the pseudorandomness of E
- Transcript $\tau = \left((M_1, T_1), \dots, (M_q, T_q), K_h\right) \Rightarrow \tau = \left((U_1, V_1, T_1), \dots, (U_q, V_q, T_q)\right)$
  - T_{id} : Probability distribution of τ in the ideal world
  - T_{re} : Probability distribution of τ in the real world

$$U_i = F_{K_h}(M_i)$$
$$V_i = G_{K_h}(M_i)$$

H-Coefficient Technique

H-coefficient lemma (informal)

If there exists $\epsilon_{bad}, \epsilon_{ratio}$ *such that*

1) *for a set of bad transcripts* $\mathcal{T}_{bad}$, $\Pr[\mathrm{T_{id}} \in \mathcal{T}_{bad}] \leq \epsilon_{bad}$

2) *with* $\tau \notin \mathcal{T}_{bad}$, $\frac{\Pr[\mathrm{T_{re}}=\tau]}{\Pr[\mathrm{T_{id}}=\tau]} \geq 1 - \epsilon_{ratio}$

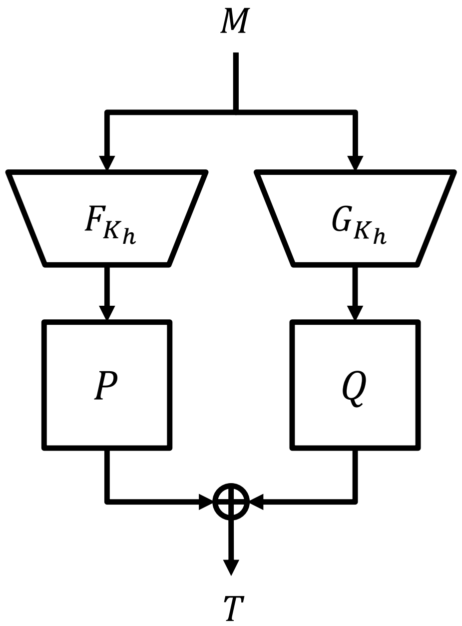
Then,

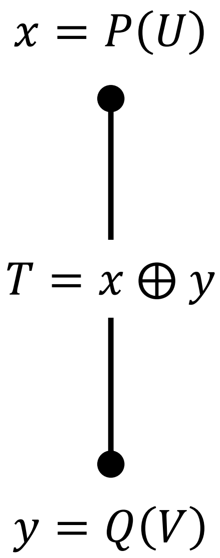
$$\mathrm{Adv} \leq \epsilon_{bad} + \epsilon_{ratio}$$

- Define a proper set of bad transcripts then upper bound ϵ_{bad} and ϵ_{ratio}
- $\Pr[\mathrm{T_{id}} = \tau]$ is easy to compute, while $\Pr[\mathrm{T_{re}} = \tau]$ is challenging

Proof Sketch

- Step 1: Represent the transcript by a graph





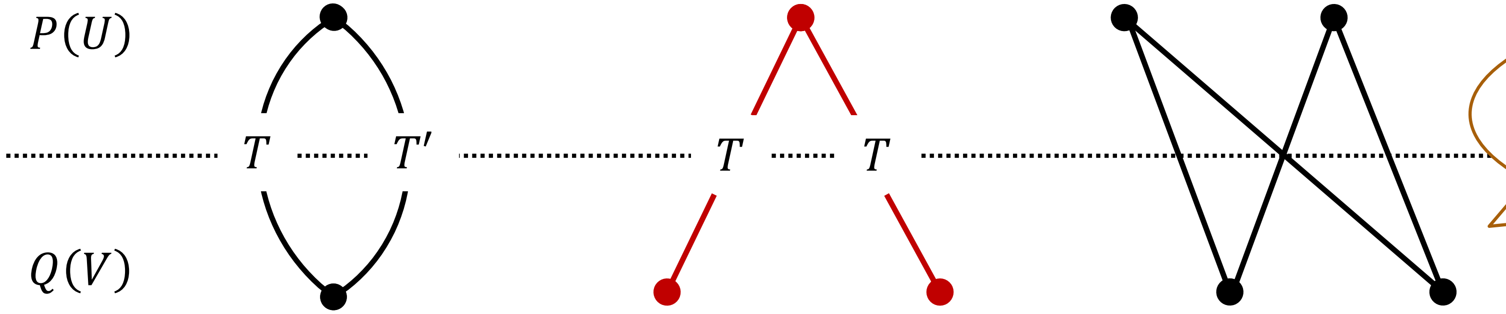
$$U = F_{K_h}(M)$$
$$V = G_{K_h}(M)$$

  - Each query makes an affine equation between two variables
  - Since we target BBB-security, hash collisions are allowed

    $\Rightarrow$ edges might be connected each other

Proof Sketch

- Step 2: Identify bad graphs
  - Some transcript graphs might lead to a contradiction!
    - When the graph contains a cycle
    - When the graph contains a path of even length whose tag sum is 0 (degeneracy)



This event was used to break DbHtS in [LNS18]

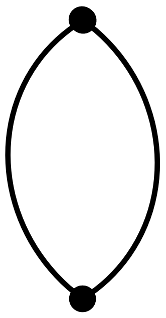
$$P(U) \oplus Q(V) = T$$
$$P(U) \oplus Q(V) = T'$$

$$P(U) \oplus Q(V) = T$$
$$P(U) \oplus Q(V') = T$$

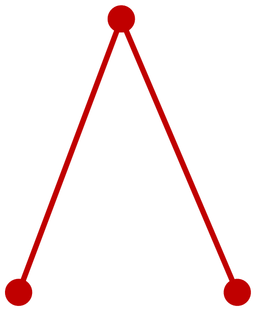
Proof Sketch

- Step 3: Upper bound the probability of obtaining bad graphs ($= \epsilon_{bad}$)

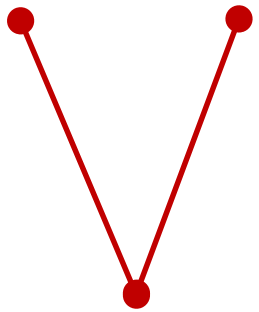
Bad1 : $U_i = U_j$ & $V_i = V_j$



Bad2 : $U_i = U_j$ & $T_i = T_j$



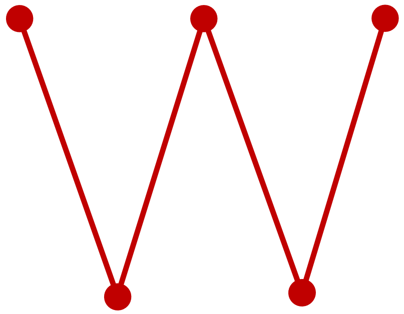
Bad3 : $V_i = V_j$ & $T_i = T_j$



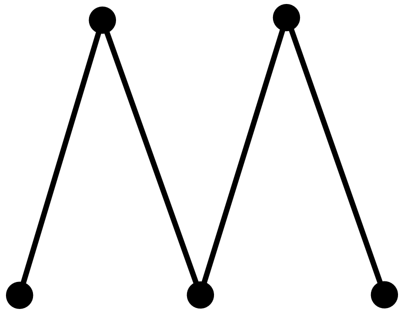
No Bad1 & Bad5 $\Rightarrow$ No cycle

No Bad2 - Bad5
$\Rightarrow$ No even length trail of zero tag sum

Bad4 : $V_i = V_j$ & $U_j = U_k$ & $V_k = V_l$ & $\sum T = 0$



Bad5 : $U_i = U_j$ & $V_j = V_k$ & $U_k = U_l$



Proof Sketch

- Step 4: Apply Patarin's Mirror theory to upper bound ϵ_{ratio}
  - Mirror theory: evaluates the number of solutions of affine systems ⇒ evaluates $\Pr[\mathrm{T_{re}} = \tau]$

- Mirror theory should be extended!
  - The original Mirror theory can be used when the maximum component size is bounded
    - This is not the case for DbHtS
  - We relaxed the constraints to allow a component of an arbitrary size
  - Instead, the ratio of the number of connected edges to the number of all the edges should be bounded

Refined Mirror Theory

- Patarin's Mirror theory

Authors	Publication	Application	Max Comp Size	Security
Patarin	eprint 2010/287	XoP	2	n
Patarin	eprint 2010/293	Feistel	$2^n/q$	n
Mennink, Neves	Crypto 17	EWCDM	2	n
Datta, Dutta, Nandi, Yasuda	Crypto 18	DWCDM	3	2n/3
Dutta, Nandi, Talnikar	EC 19	CWC+	$2^n/q$	2n/3
Mennink	TCC 18	CLRW2	4	3n/4
Jha, Nandi	JoC 20	CLRW2	Any[1)]	3n/4
This work	EC 20	DbHtS	Any[2)]	3n/4

- The first refinement allows a component of an arbitrary size up to 3n/4-bit security (concurrent work with [JN20])

1) Without path of length 3 2) With bounded number of connected edges

Result

- Security of DbHtS MACs with two independent δ-universal hash functions F and G

$$\mathbf{Adv}_{\mathsf{DbHtS}[F,G]}(q) \leq 4q^{\frac{4}{3}}\delta + \frac{22q^{\frac{4}{3}}}{2^n} + \epsilon(q,\delta)$$

- Security of PMAC-Plus

$$\mathbf{Adv}_{\mathsf{PMAC\text{-}Plus}}(q,\ell) \leq \frac{53\ell^{\frac{2}{3}}q^{\frac{4}{3}}}{2^n} + \frac{\ell^2 q}{2^{n+1}} + \epsilon(q,\ell)$$

Conclusion

- Proved tight security bounds for DbHtS MACs
  - PolyMAC, SUM-ECBC, 3kf9, PMAC-Plus, LightMAC-Plus are PRF up to $2^{3n/4}$ queries
  - All the security bounds are tight in terms of the threshold number of queries

- Future Works
  - Find better security bounds considering the influence of message length ℓ
  - Find tight security of key-reduced variants of DbHtS MACs

Thank you

Q&A : lbh0307@kaist.ac.kr