Exact Security Analysis of Hash-then-Mask Type Probabilistic MAC Constructions

Avijit Dutta and Ashwin Jha and Mridul Nandi

Indian Statistical Institute, Kolkata

27th September, 2016

Outline of the talk

1. Message Authentication Code.
2. HtM Construction.
3. Contributions.
4. Conclusion

MAC (Stateless and Deterministic): The Popular Story

1. Alice and Bob share a secret key K.

MAC (Stateless and Deterministic): The Popular Story

1. Alice and Bob share a secret key K.
2. Alice sends a message M with a tag $T = MAC_K(M)$ corresponding to the message M to Bob.

MAC (Stateless and Deterministic): The Popular Story

1. Alice and Bob share a secret key K.
2. Alice sends a message M with a tag $T = MAC_K(M)$ corresponding to the message M to Bob.
3. Data Integrity: Bob verifies the sender and the message by computing $VER_K(M, T) = 1$.

MAC (Stateless and Deterministic): The Popular Story

1. Alice and Bob share a secret key K.
2. Alice sends a message M with a tag $T = MAC_K(M)$ corresponding to the message M to Bob.
3. Data Integrity: Bob verifies the sender and the message by computing $VER_K(M, T) = 1$.

Unforgeability

- Adversary asks for tags for queries of his choice.
- Goal is to generate any fresh, valid (message, tag) pair.

Security Requirement: It should be HARD

MAC (Stateful or Probabilistic): The Popular Story

• Alice sends a message M, an auxiliary variable IV with a tag $T = MAC_K(M, IV)$ corresponding to the message M and IV to Bob.

• Data Integrity: Bob verifies the sender and the message by computing $VER_K(M, IV, T) = 1$.

Stateful MAC : When IV is a counter / nonce. (e.g XMACC, PCS)
Probabilistic MAC : When IV is random. (e.g XMACR, EHtM)

Unforgeability

- Adversary asks for T for queries M (Signing Query).
- Adversary asks fresh (M, IV, T) triplet and obtains 1 or 0. Succeed if the response is 1 (Verification Query).

Security: Should be HARD to obtain response 1

Pseudo Random Function (PRF)

PRF

Keyed function which is indistinguishable from a Random Function (RF)

Indistinguishability

- Responses of adversary queries are given either using the function or a RF.
- Goal is to distinguish the function from a RF.

Security Requirement: It should be HARD

Universal and AXU-Hash

Universal Hash

H is a n bit Universal Hash, if for all distinct values, the collision probability of H is negligible.

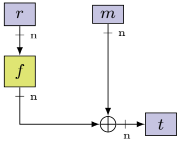
Almost-XOR-Universal Hash

H is a n bit AXU Hash, if for all distinct values x, x' and for all y, $\Pr[H(x) \oplus H(x') = y]$ is negligible.

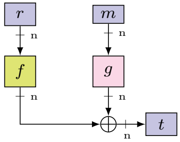
Existing Result on Probablistic MAC

Candidate	Construction	Rand	Eff.	Bound
XMACR[BGR'95]	$(r, H(m) \oplus f(r))$	n	$1H_{\mathrm{xu}}, 1F[n,n]$	$O(\frac{q^2}{2^n} + q_v\epsilon)$
MACRX$_3$[BGK'99]	$(r_1, r_2, r_3, \bigoplus_{i=1}^{3} f(r_i) \oplus H(m))$	$3n$	$1H_{\mathrm{xu}}, 3F[n,n]$	$O(\frac{q^3}{2^{3n}} + q_v\epsilon)$
RMAC[JJV'02]	$(r, f_2^r(\mathrm{CBC}_{f_1}(m))$	n	$(\ell+1)P[n]$	$O(\frac{\ell(q+q_v)}{2^n})$
FRMAC[JJ'04]	$(r, \pi_r(H(m)))$	n	$1H_{\mathrm{u}}, 1P[n,n]$	$O(\ell(q+q_v)\epsilon)$
RWMAC[M'10]	$(r, g(r, H(m)))$	n	$1H_{\mathrm{u}}, 1F[2n,n]$	$O(\frac{q^2\epsilon}{2^n} + q_v\epsilon)$
EHtM[M'10]	$(r, f(r) \oplus g(r \oplus H(m))$	n	$1H_{\mathrm{xu}}, 2F[n,n]$	$O(\frac{q^3\epsilon}{2^n} + q_v\epsilon)$

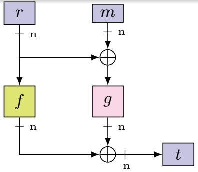
HtM: Probabilistic MAC



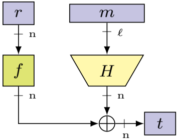
C1 : $t := f(r) \oplus m$



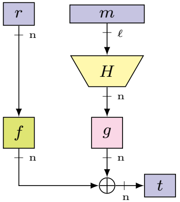
C3 : $t := f(r) \oplus g(m)$



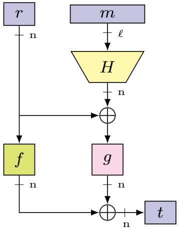
C5 : $t := f(r) \oplus g(r \oplus m)$



C2 : $t := f(r) \oplus H(m)$



C4 : $t := f(r) \oplus g(H(m))$



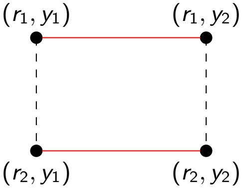
C6 : $t := f(r) \oplus g(r \oplus H(m))$

Our Contribution

1. Tight PRF, pPRF and MAC Security Analysis of Different Types of HtM Constructions.
2. An Impossibility Result on Probabilistic MAC:
   Unlike deterministic MAC, in probabilistic MAC, there is no such ideal system, indistinguishable to which, ensures forging advantage.

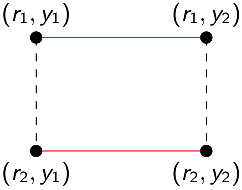
	C1	C2	C3	C4	C5	C6
PRF	X	X	X	X	X	$\Theta(2^{n/2})$
pPRF	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{3n}{4}})$	$\Theta(2^{\frac{3n}{4}})$
MAC	X	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{n}{2}})$	$\Theta(2^{\frac{2n}{3}})$	$\Theta(2^{\frac{3n}{4}})$

PRF Attack Idea of C1,C2,C3,C4



$$\boxed{\mathrm{SUM}_{\mathrm{f,g}}(r, y) = f(r) \oplus g(y)}$$

PRF Attack Idea of C1,C2,C3,C4



$$\text{SUM}_{\mathrm{f,g}}(r, y) = f(r) \oplus g(y)$$

Alternating Cycle (Alt-Cycle)

- For an Alt-Cycle C, $\sum_{i=1}^{4} \text{SUM}^{C}_{\mathrm{f,g}}(r_i, y_i) = 0$ (distinguishing event)
- For C1, C2 : g is identity function.
- For C1, C3 : y is m; For C2, C4 : y is $H(m)$; For C5 : y is $r + m$

PRF Attack Idea of C5 and C6

Attack Algorithm C5 : $f(r) \oplus g(r \oplus m)$

- Choose $(r_1, m_1), (r_2, m_2)$ s.t $r_1 + m_1 = r_2 + m_2$
- Query Phase :
  $t_1 \leftarrow (r_1, m_1), t_2 \leftarrow (r_2, m_2), t_3 \leftarrow (r_1, m_2), t_4 \leftarrow (r_2, m_1)$
- Distinguishing Event : If $\bigoplus_{i=1}^{4} t_i = 0$, return 1.

PRF Attack Idea of C5 and C6

Attack Algorithm C5 : $f(r) \oplus g(r \oplus m)$

- Choose $(r_1, m_1), (r_2, m_2)$ s.t $r_1 + m_1 = r_2 + m_2$
- Query Phase :
  $t_1 \leftarrow (r_1, m_1), t_2 \leftarrow (r_2, m_2), t_3 \leftarrow (r_1, m_2), t_4 \leftarrow (r_2, m_1)$
- Distinguishing Event : If $\bigoplus_{i=1}^{4} t_i = 0$, return 1.

Attack Algorithm C6 : $f(r) \oplus g(r \oplus H(m))$

- Query Phase :
  $t_1 \leftarrow (r, m_1), t_2 \leftarrow (r, m_2), \ldots, t_{2^{n/2}} \leftarrow (r, m_{2^{n/2}})$
- If $H(m_i) = H(m_j)$, query $t_i' \leftarrow (r', m_i), t_j' \leftarrow (r', m_j)$, output 1 if $t_i' = t_j'$
- Else, collision in g.

Probabilistic PRF (pPRF)

Definition and Security Game

Keyed function that takes two inputs (r, M) is indistinguishable from RF

- Adversary can only query the oracle with M.
- Goal is to distinguish the function from a RF; secure if it is hard

Probabilistic PRF (pPRF)

Definition and Security Game

Keyed function that takes two inputs (r, M) is indistinguishable from RF

- Adversary can only query the oracle with M.
- Goal is to distinguish the function from a RF; secure if it is hard

pPRF Attack Algorithm of C1 : $f(r) \oplus m$

- Query Phase : $t_1 \leftarrow m_1, t_2 \leftarrow m_1, \ldots, t_{2^{n/2}} \leftarrow m_1$
- W.h.p $\exists i, j \in \{1, 2, \ldots, 2^{n/2}\}$ s.t $r_i = r_j$
- If $t_i = t_j$, return 1.

Probabilistic PRF (pPRF)

Definition and Security Game

Keyed function that takes two inputs (r, M) is indistinguishable from RF

- Adversary can only query the oracle with M.
- Goal is to distinguish the function from a RF; secure if it is hard

pPRF Attack Algorithm of C1 : $f(r) \oplus m$

- Query Phase : $t_1 \leftarrow m_1, t_2 \leftarrow m_1, \ldots, t_{2^{n/2}} \leftarrow m_1$
- W.h.p $\exists i, j \in \{1, 2, \ldots, 2^{n/2}\}$ s.t $r_i = r_j$
- If $t_i = t_j$, return 1.

pPRF Attack for C2, C3, C4 is same as that of C1

pPRF Attack Idea of C5

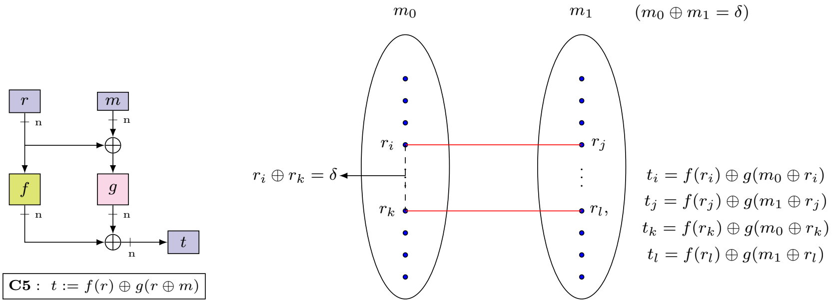


Figure 0.1: Distinguishing Event : If $t_i \oplus t_j \oplus t_k \oplus t_l = 0$, output 1.

pPRF Attack Idea of C6

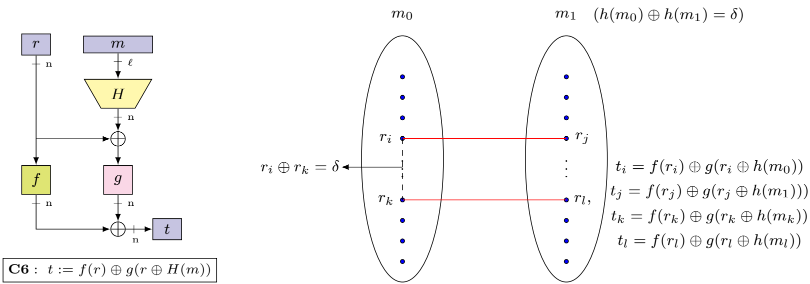


Figure 0.1: Distinguishing Event : If $t_i \oplus t_j \oplus t_k \oplus t_l = 0$, output 1.

Forging Idea of C1-C6

Forging C1 : $f(r) \oplus m$

- Query Phase : $t \leftarrow (r, m)$.
- Forge : $(r, m \oplus \delta, t \oplus \delta)$.

Forging Idea of C1-C6

Forging C1 : $f(r) \oplus m$

- Query Phase : $t \leftarrow (r, m)$.
- Forge : $(r, m \oplus \delta, t \oplus \delta)$.

Forging C2 : $f(r) \oplus H(m)$

- Query Phase :
  $t_1 \leftarrow (r_1, m_1), t_2 \leftarrow (r_2, m_2), \ldots, t_{2^{n/2}} \leftarrow (r_{2^{n/2}}, m_{2^{n/2}})$.
- W.h.p $i, j \in \{1, 2, \ldots, 2^{n/2}\}$ such that $r_i = r_j$. It leaks $H(m_i) \oplus H(m_j) = \delta$.
- Query $t \leftarrow (r, m_i)$.
- Forge : $(r, m_j, t \oplus \delta)$.

Forging Idea of C1-C6

Forging C1 : $f(r) \oplus m$

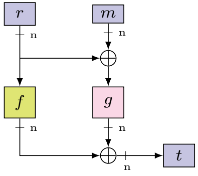
- Query Phase : $t \leftarrow (r, m)$.
- Forge : $(r, m \oplus \delta, t \oplus \delta)$.

Forging C2 : $f(r) \oplus H(m)$

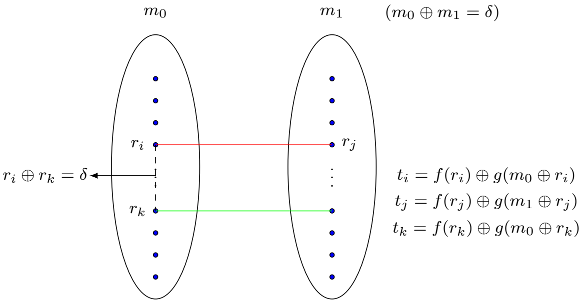
- Query Phase :
  $t_1 \leftarrow (r_1, m_1), t_2 \leftarrow (r_2, m_2), \ldots, t_{2^{n/2}} \leftarrow (r_{2^{n/2}}, m_{2^{n/2}})$.
- W.h.p $i, j \in \{1, 2, \ldots, 2^{n/2}\}$ such that $r_i = r_j$. It leaks $H(m_i) \oplus H(m_j) = \delta$.
- Query $t \leftarrow (r, m_i)$.
- Forge : $(r, m_j, t \oplus \delta)$.

Forging attack of C3, C4 is same as that of C2

Forging Idea of C5



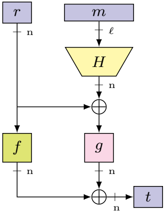
C5 : $t := f(r) \oplus g(r \oplus m)$



Forge: $(r_k, m_1, t_i \oplus t_j \oplus t_k)$

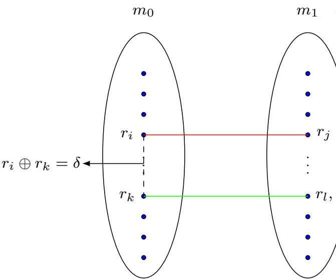
$(\because t_i \oplus t_j \oplus t_k = f(r_k) \oplus g(m_1 \oplus r_k))$

Forging Idea of C6



C6 : $t := f(r) \oplus g(r \oplus H(m))$

$(h(m_0) \oplus h(m_1) = \delta)$



$t_i = f(r_i) \oplus g(h(m_0) \oplus r_i)$
$t_j = f(r_j) \oplus g(h(m_1) \oplus r_j)$
$t_k = f(r_k) \oplus g(h(m_0) \oplus r_k)$

Forge: $(r_k, m_1, t_i \oplus t_j \oplus t_k)$

$(\because t_i \oplus t_j \oplus t_k = f(r_k) \oplus g(h(m_1) \oplus r_k))$

Alternating Cycle

A transcript $\tau := \{(x_1, y_1), (x_2, y_2), \ldots, (x_q, y_q)\}$ has an alternating-cycle in τ of length k (k is even and ≥ 2), if we have k pairwise distinct indices $i_1, i_2, \ldots, i_k$ such that

$x_{i_1} = x_{i_2}, y_{i_2} = y_{i_3}, x_{i_3} = x_{i_4}, \ldots, x_{i_{k1}} = x_{i_k}, y_{i_k} = y_{i_1}$.

Alternating Cycle

A transcript $\tau := \{(x_1, y_1), (x_2, y_2), \ldots, (x_q, y_q)\}$ has an alternating-cycle in τ of length k (k is even and ≥ 2), if we have k pairwise distinct indices $i_1, i_2, \ldots, i_k$ such that

$x_{i_1} = x_{i_2}, y_{i_2} = y_{i_3}, x_{i_3} = x_{i_4}, \ldots, x_{i_{k1}} = x_{i_k}, y_{i_k} = y_{i_1}$.

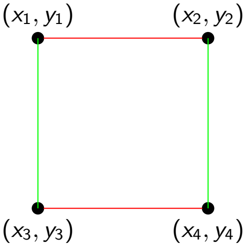


Figure: Alternating Cycle of length 4. Red line indicates first coordinate matches. Green line indicates second coordinates matches

Benes Butterfly Result

Theorem (Benes-Butterfly (AV'96))

Let f and g be two n-bit independent and uniformly distributed random functions. Let us consider a transcript $\tau = \{(x_i, y_i, t_i)_{1 \le i \le q}\}$ which does not contain any alternating cycle. Then

$$\Pr[f(x_i) \oplus g(y_i) = t_i, 1 \le i \le q] = \frac{1}{2^{nq}}.$$

Proof Sketch : If there is no alternating cycle in $\tau = \{(x_i, y_i)_{1 \le i \le q}\}$ then from each of q many equations, we get at least one uniform random variable

pPRF Advantage of C5 and C6

Theorem

$$\mathbf{Adv}^{\mathrm{pprf}}_{\mathrm{C5/C6}}(q,t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q,t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q,t) + \frac{q^4}{2^{3n}}.$$

pPRF Advantage of C5 and C6

Theorem

$$\mathbf{Adv}^{\mathrm{pprf}}_{\mathrm{C5/C6}}(q,t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q,t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q,t) + \frac{q^4}{2^{3n}}.$$

- Bad Transcript : Alternating cycle on $(r, r \oplus m)/(r, r \oplus h(m))$.

pPRF Advantage of C5 and C6

Theorem

$$\mathbf{Adv}^{\mathrm{pprf}}_{\mathrm{C5/C6}}(q,t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q,t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q,t) + \frac{q^4}{2^{3n}}.$$

- Bad Transcript : Alternating cycle on $(r, r \oplus m)/(r, r \oplus h(m))$.
- No bad event $\Rightarrow$ No alternating cycle in the transcript.

pPRF Advantage of C5 and C6

Theorem

$$\mathbf{Adv}^{\mathrm{pprf}}_{\mathrm{C5/C6}}(q,t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q,t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q,t) + \frac{q^4}{2^{3n}}.$$

- Bad Transcript : Alternating cycle on $(r, r \oplus m)/(r, r \oplus h(m))$.
- No bad event $\Rightarrow$ No alternating cycle in the transcript.
- Probability of bad event : $\frac{q^4}{2^{3n}}$

SUF Advantage of C5 and C6

Theorem (SUF Advantage of C5)

$$\mathbf{Adv}^{\mathrm{suf}}_{C_5}(q, q', t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q + q', t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q + q', t) + \frac{q^3}{2^{2n}} + \frac{q'}{2^n}.$$

SUF Advantage of C5 and C6

Theorem (SUF Advantage of C5)

$$\mathbf{Adv}^{\mathrm{suf}}_{C_5}(q, q', t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q + q', t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q + q', t) + \frac{q^3}{2^{2n}} + \frac{q'}{2^n}.$$

• Bad Transcript : Alternating cycle on $(r, r \oplus m)$ after making signing and verfication queries.

SUF Advantage of C5 and C6

Theorem (SUF Advantage of C5)

$$\mathbf{Adv}^{\mathrm{suf}}_{C_5}(q, q', t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q + q', t) + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q + q', t) + \frac{q^3}{2^{2n}} + \frac{q'}{2^n}.$$

- Bad Transcript : Alternating cycle on $(r, r \oplus m)$ after making signing and verfication queries.
- Good Transcript $\Rightarrow$ No Alternating Cycle in the transcript.
- Probability of Bad Transcript : $\frac{q^3}{2^{2n}}$.

Proof Idea of SUF Advantage of C5 and C6

Theorem (SUF Advantage of C6)

$$\mathbf{Adv}^{\mathrm{suf}}_{\mathrm{C6}}(q, q', \ell, t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q+q', t') + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q+q', t') + \frac{q^4}{2^{3n}} + \frac{10q'}{2^n},$$

where $t = t' + O(qT_h)$

Proof Idea of SUF Advantage of C5 and C6

Theorem (SUF Advantage of C6)

$\mathbf{Adv}^{\mathrm{suf}}_{\mathrm{C6}}(q, q', \ell, t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q+q', t') + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q+q', t') + \frac{q^4}{2^{3n}} + \frac{10q'}{2^n}$, *where* $t = t' + O(qT_h)$

- Bad Transcript : Alternating cycle on $(r, r \oplus h(m))$ after making signing and verfication queries.

Proof Idea of SUF Advantage of C5 and C6

Theorem (SUF Advantage of C6)

$$\mathbf{Adv}^{\mathrm{suf}}_{\mathrm{C6}}(q, q', \ell, t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q+q', t') + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q+q', t') + \frac{q^4}{2^{3n}} + \frac{10q'}{2^n},$$

where $t = t' + O(qT_h)$

- Bad Transcript : Alternating cycle on $(r, r \oplus h(m))$ after making signing and verfication queries.
- Good Transcript $\Rightarrow$ No Alternating cycle.

Proof Idea of SUF Advantage of C5 and C6

Theorem (SUF Advantage of C6)

$\mathbf{Adv}^{\mathrm{suf}}_{\mathrm{C6}}(q, q', \ell, t) \leq \mathbf{Adv}^{\mathrm{prf}}_{f_{k_1}}(q+q', t') + \mathbf{Adv}^{\mathrm{prf}}_{f_{k_2}}(q+q', t') + \frac{q^4}{2^{3n}} + \frac{10q'}{2^n}$, *where* $t = t' + O(qT_h)$

- Bad Transcript : Alternating cycle on $(r, r \oplus h(m))$ after making signing and verfication queries.
- Good Transcript $\Rightarrow$ No Alternating cycle.
- Probability of Bad Transcript : $\frac{q^4}{2^{3n}}$ as (we need one more point)

Summary

- Tight Security Analysis of HtM Probabilistic MAC.
- Tight Security Analysis of EHtM.
- Impossibility result on Probabilistic MAC.

Summary

- Tight Security Analysis of HtM Probabilistic MAC.
- Tight Security Analysis of EHtM.
- Impossibility result on Probabilistic MAC.

Thank You