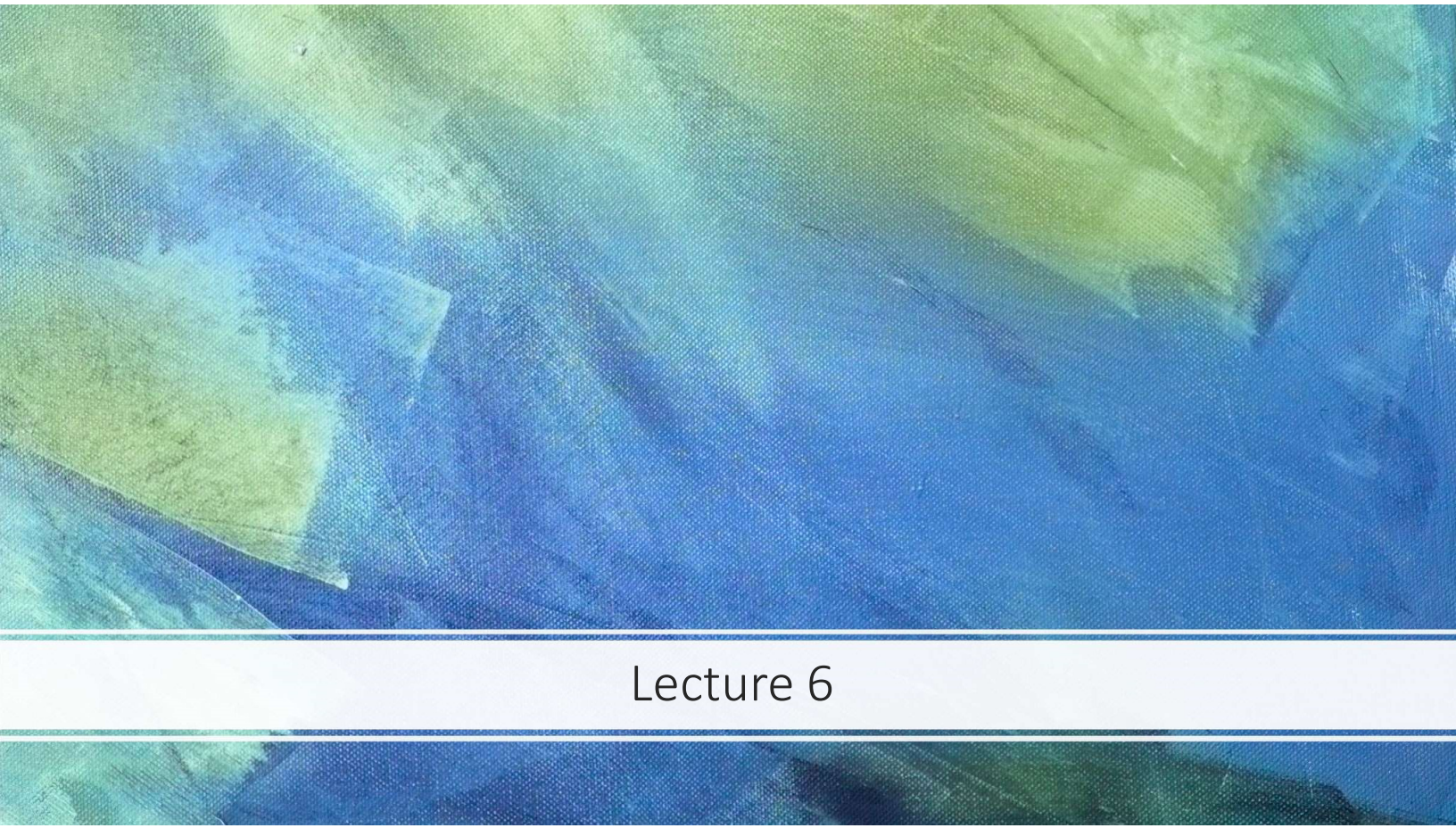
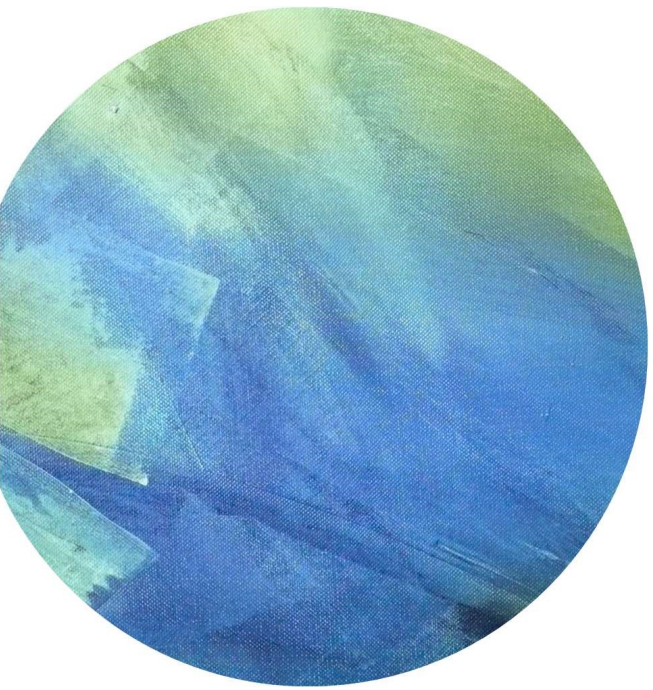


Lecture 6



Outline

MACS: Continued

Collision Resistance

Administrative Details

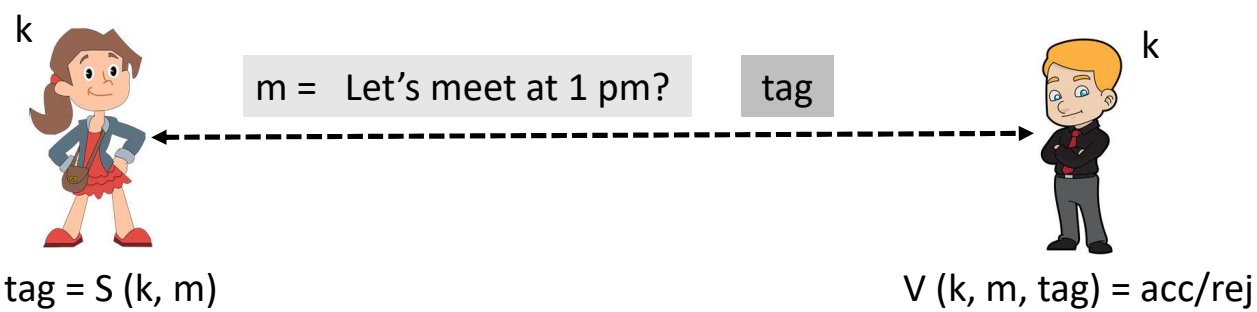
- Course website: https://courses.grainger.illinois.edu/cs498ac3/fa2020/
- Has syllabus, instructor and TA info, office hours
- IMPORTANT: **Join Piazza!** piazza.com/illinois/fall2020/ececs498ac/home

I strongly encourage class participation.
If you don't understand something in class, please interrupt me and ask questions.
Please make abundant use of office hours.

HOMEWORK 1 IS OUT.
DUE IN A WEEK.

Message Integrity: MACs

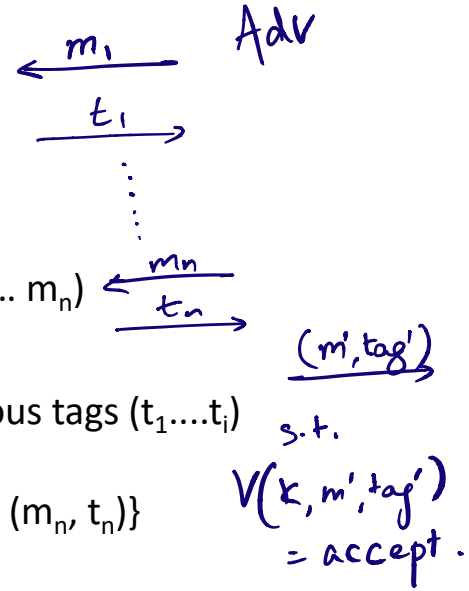
Recap: MACs



MAC = (S,V) is a pair of algorithms that satisfy:

1. CORRECTNESS. V(k, m, tag) = 1 when tag = S (k, m)

Recap: Security of MACs

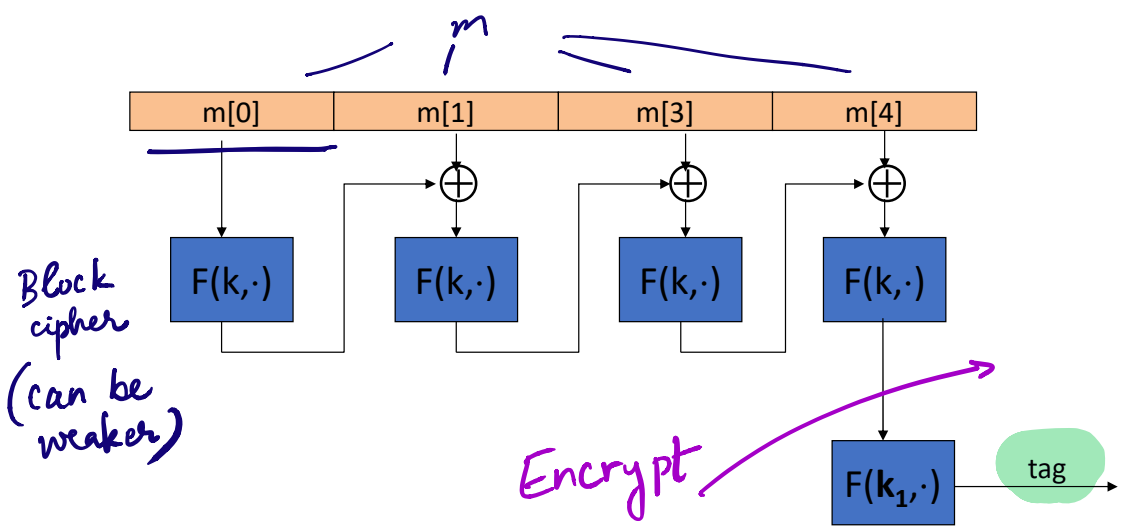


- Attacker can demand tags (t_1, t_2, …. t_n) for messages (m_1, m_2, … m_n)
- Message m_{i+1} can be chosen adaptively as a function of previous tags (t_1….t_i)
- Attacker wins if it outputs (m', tag') not in {(m_1, t_1), (m_2, t_2), …. (m_n, t_n)} such that V (k, m', tag') = 1

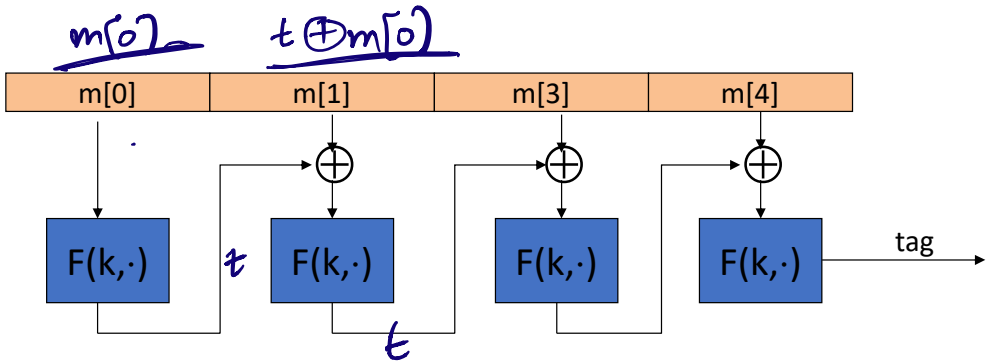
Constructing MACs

- S(k,m) = PRF(k,m), V(k, m, t) = accept if and only if t = PRF(k,m)
- What about long messages, larger than the input size of PRF?
- For larger messages, we use:
  - CBC-MAC
  - HMAC

Encrypted CBC-MAC



What happens if we remove encryption step?



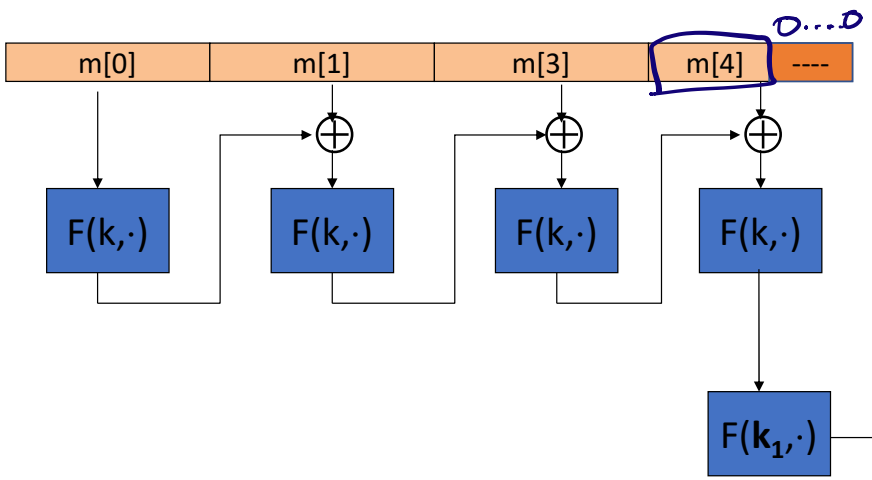
Obtain tag for m, then get t = F (k, m)

Output forgery (m', t') where m' = (m||(t \xor m)) and t' = t.

$t = F(k, m[0])$ $m = m[0]$ ask for tag. Get t.

Generate forgery on $(m[0] \| t \oplus m[0])$. Output $((m[0] \| t \oplus m[0]), t)$

What if message length is not a multiple of block size?

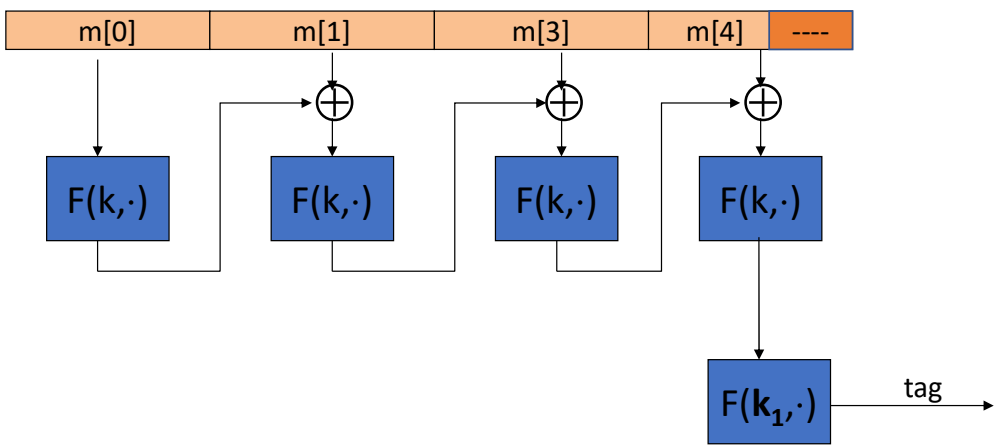


m[4] = 0…0 (128 bits)

OR

m[4] = 00 2… (2 bits)

First idea?



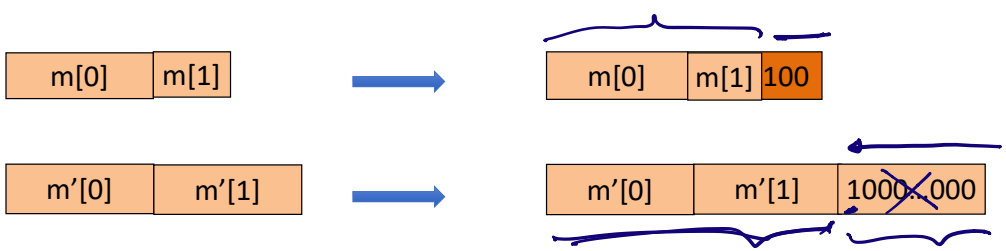
Need invertible padding!

For security, padding must be invertible !

$$m_0 \neq m_1 \quad \Rightarrow m_0 \,||\, pad(m_0) \neq m_1 \,||\, pad(m_1)$$

Pad with "1000…00". Add new dummy block if needed.

- The "1" indicates beginning of pad.



HMAC

- Hash MAC
- Apply a *hash* function H to your original message
- What properties should H satisfy?

Collision-resistance

Let H: $M \rightarrow T$ be a hash function ($|M| >> |T|$)

A **collision** for H is a pair $m_0, m_1 \in M$ such that:

$$H(m_0) = H(m_1) \quad \text{and} \quad m_0 \neq m_1$$

A function H is **collision resistant** if for all PPT algs. A:

$$Adv_{CR}[A,H] = \Pr[\text{A outputs collision for H}] = \text{negl}$$

Example: SHA-256 (outputs 256 bits)

MAC from Collision-resistant Hash Functions

(Not HMAC)

Let (S,V) be a MAC for short messages over (K,M,T) (e.g. AES)

Let H: $M^{big} \rightarrow M$

Def: (S^{big}, V^{big}) over (K, M^{big}, T) as:

$S^{big}(k,m) = S(k,H(m))$; $V^{big}(k,m,t) = V(k,H(m),t)$

Thm: If (S,V) is a secure MAC and H is collision resistant

then I^{big} $= (S^{big}, V^{big})$ is a secure MAC.

m', t' s.t. $V^{big}(k, m', t') = 1$

$\Rightarrow m_0, m_1$ s.t.

$H(m_0) = H(m_1)$

$(t_{m_0} = t_{m_1})$

OR Adv. broke (S,V)

Example: $S(k,m) = AES_{2\text{-block-cbc}}(k, \text{SHA-256}(m))$ is a secure MAC.

MAC from Collision-resistant Hash Functions

$S^{big}(k, m) = S(k, H(m))$; $V^{big}(k, m, t) = V(k, H(m), t)$

Collision resistance is necessary for security:

Suppose adversary can find $m_0 \neq m_1$ s.t. $H(m_0) = H(m_1)$.

Then: S^{big} is insecure under a 1-chosen msg attack

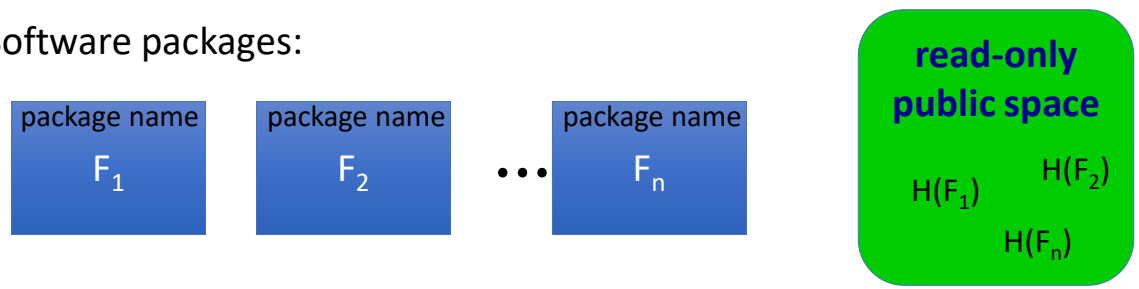
step 1: adversary asks for $t \leftarrow S(k, m_0)$

step 2: output (m_1, t) as forgery

If $H(m_0) = H(m_1)$ then $tag_{m_0} = S(k, H(m_0))$
$= S(k, H(m_1))$
$= tag_{m_1}$

Protecting File Integrity

Software packages:



The birthday attack

Let $H: M \rightarrow \{0,1\}^n$ be a hash function ($|M| >> 2^n$)

Want: 2^n

Generic alg. to find a collision **in time $O(2^{n/2})$** hashes

Space $O(2^{n/2})$

The birthday attack

Let $H: M \rightarrow \{0,1\}^n$ be a hash function ($|M| >> 2^n$)

Generic alg. to find a collision **in time $O(2^{n/2})$** hashes

Algorithm:

1. Choose $\mathbf{2^{n/2}}$ random messages in M: $m_1, \ldots, m_{2^{n/2}}$ (distinct w.h.p)
2. For i = 1, …, $2^{n/2}$ compute $t_i = H(m_i) \in \{0,1\}^n$
3. Look for a collision ($t_i = t_j$). If not found, got back to step 1.

How well will this work?

The birthday attack

Let $r_1, \ldots, r_n \in \{1,\ldots,B\}$ be indep. identically distributed integers.

Thm: when $n = 1.2 \times B^{1/2}$ then $\Pr[\exists i \neq j: r_i = r_j] \geq \frac{1}{2}$

Proof: (for uniform indep. $r_1, \ldots, r_n$)

$$\Pr[\exists i \neq j \mid r_i = r_j] = 1 - \Pr[\forall i \neq j, r_i \neq r_j]$$

$$= 1 - \frac{(B-1)}{B} \cdot \frac{(B-2)}{B} \cdots \cdots \frac{(B-n+1)}{B}$$

$$= 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{B}\right)$$

true because $\left(1 - \frac{i}{B} \leq e^{-i/B}\right)$

$$\geq 1 - \prod_{i=1}^{n-1} e^{-i/B} \geq 1 - e^{-n^2/2B} = 0.53 > \frac{1}{2}.$$

The birthday attack

Take any $\hookrightarrow 2^{n/2}$

$H: M \to \{0,1\}^n$. Collision finding algorithm:

comparisons $O(2^{n/2})^2$

1. Choose $\mathbf{2^{n/2}}$ random elements in M: $m_1, \ldots, m_{2^{n/2}}$
2. For i = 1, …, $2^{n/2}$ compute $t_i = H(m_i) \in \{0,1\}^n$
3. Look for a collision ($t_i = t_j$). If not found, got back to step 1.

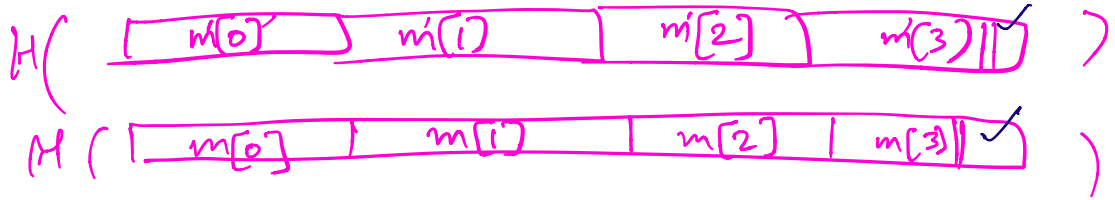
$\exists\ i, j$ s.t. $m_i \neq m_j$ but $H(m_i) = H(m_j)$

Expected number of iteration ≈ 2 (by previous Thm)

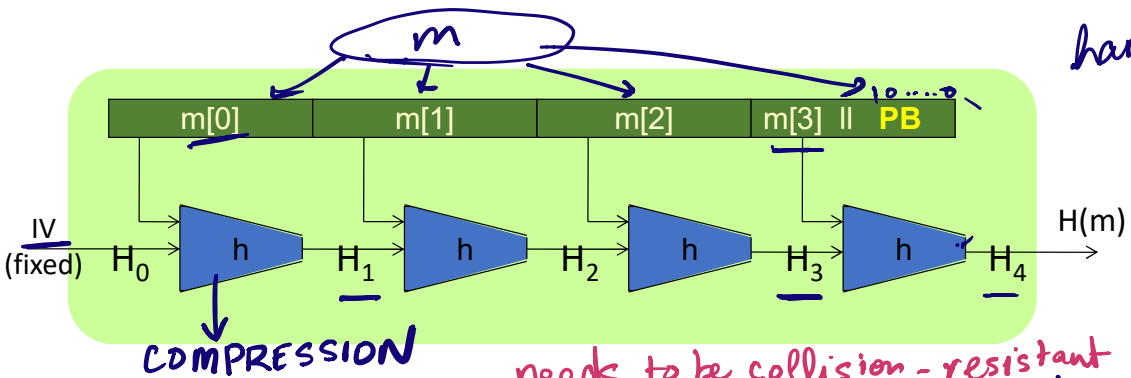
Running time: $\mathbf{O(2^{n/2})}$ (space $O(2^{n/2})$)

Example: SHA1 has output size 160 bits. Birthday attack: 2^{80}. Best attack: 2^{51}

(Do not use SHA1
Use SHA256 instead!)



Merkle-Damgard (Domain extension)



hard to find m_0, m_1 s.t. $H(m_0) = H(m_1)$

Given **h: T × X ⟶ T** (compression function)

needs to be collision-resistant

we obtain **H: $X^{\leq L}$ ⟶ T**

H_i - chaining variables

PB: padding block

-- If no space for PB add another block

First step: $h(IV \| m[0])$

2nd step: $h(H_1 \| m[1])$

3rd step: $h(H_2 \| m[2])$

⋮

Merkle-Damgard

H.W.

What if $IV = h(IV \| M_1)$? Does this contradict collision resistance of h?

Theorem: If h is collision resistant, then so is H. (h: compression function)

Proof: collision on H ⇒ collision on h

Suppose H(M) = H(M'). We build collision for h.

$IV = H_0, H_1, \ldots, H_t, H_{t+1} = H(M)$

$IV = H_0', H_1', \ldots, H'_r, H'_{r+1} = H(M')$

$h(H_t, M_t \| PB) = H_{t+1} = H'_{r+1} = h(H'_r, M'_r \| PB')$

$H_{t+1} = H'_{r+1}$

$\Rightarrow h(H_t, M_t \| PB) = h(H'_r, M'_r \| PB')$

$\Rightarrow H_t = H'_r$ and $M_t \| PB = M'_r \| PB'$

$\Rightarrow h(H_{t-1}, M_{t-1}) = h(H'_{r-1}, M'_{r-1})$

$\Rightarrow H_{t-1} = H'_{r-1}$ and $M_{t-1} = M'_{r-1}$

$\Rightarrow H_{t-2} = H'_{r-2} \ldots$

BASE CASE.

$IV = H_1 = H'_{r-t}$ and $M_1 = M'_{r-t}$ ...

Merkle-Damgard

Theorem: If h is collision resistant, then so is H.

Proof: collision on H $\Rightarrow$ collision on h

Suppose H(M) = H(M'). We build collision for h.

$IV = H_0, H_1, \ldots, H_t, H_{t+1} = H(M)$

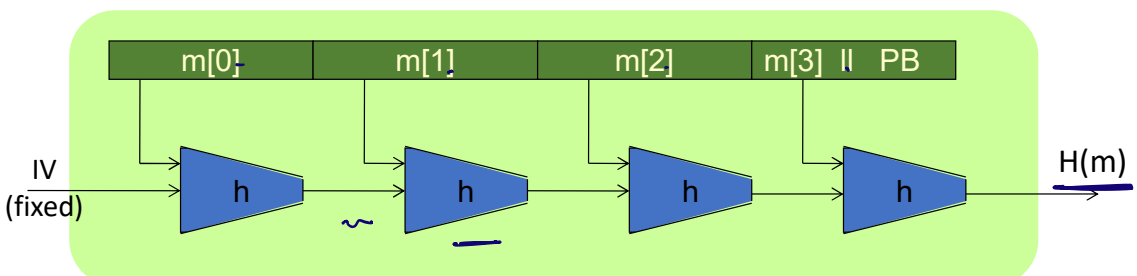
$IV = H_0', H_1', \ldots, H'_r, H'_{r+1} = H(M')$

$h(H_t, M_t \,||\, PB) = H_{t+1} = H'_{r+1} = h(H'_r, M'_r \,||\, PB')$

Otherwise suppose $H_t = H'_r$ and $M_t = M'_r$ and $PB = PB'$

Then: $h(H_{t-1}, M_{t-1}) = H_t = H'_t = h(H'_{t-1}, M'_{t-1})$

Merkle-Damgard



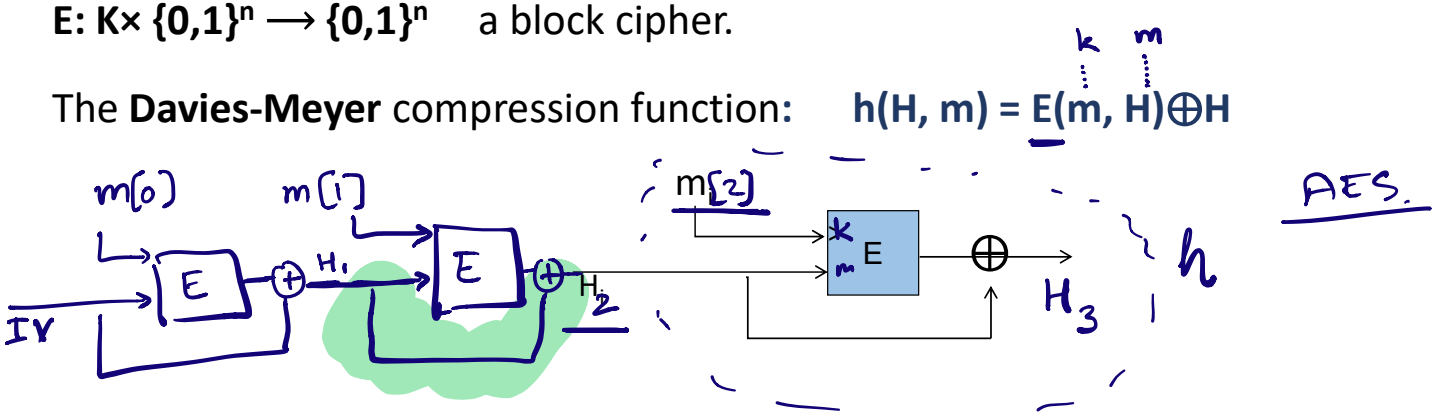
Thm: h collision resistant ⇒ H collision resistant

Goal: construct compression function **h: T × X ⟶ T**

Compression function from block cipher

E: K× {0,1}^{n} ⟶ {0,1}^{n} a block cipher.

The **Davies-Meyer** compression function: **h(H, m) = E(m, H)⊕H**



Thm: Suppose E is an ideal cipher (collection of |K| random perms.). Finding a collision **h(H,m)=h(H',m')** takes **O(2$^{n/2}$)** evaluations of (E,D).

Goal: find H, m, H', m' s.t. $h(H, m) = h(H', m')$!

What about a simpler construction?

Suppose we define **h(H, m) = E(m, H)** (without XOR step)

Then the resulting h(.,.) is not collision resistant:

to find a collision (H,m) and (H',m')

choose random (H,m,m') and construct H' as follows:

H'=D(m', E(m,H))

$E(m', H') = E(m, H)$

$h(H', m') = h(H, m)$

We saw: $MAC^{big} = PRF(k, H(m^{big}))$

Standardized Method: HMAC

↳ Doesn't ever invoke a PRF.

Most widely used MAC on the Internet.

H: hash function.
example: SHA-256 ; output is 256 bits

Can we build a MAC directly out of a hash function?

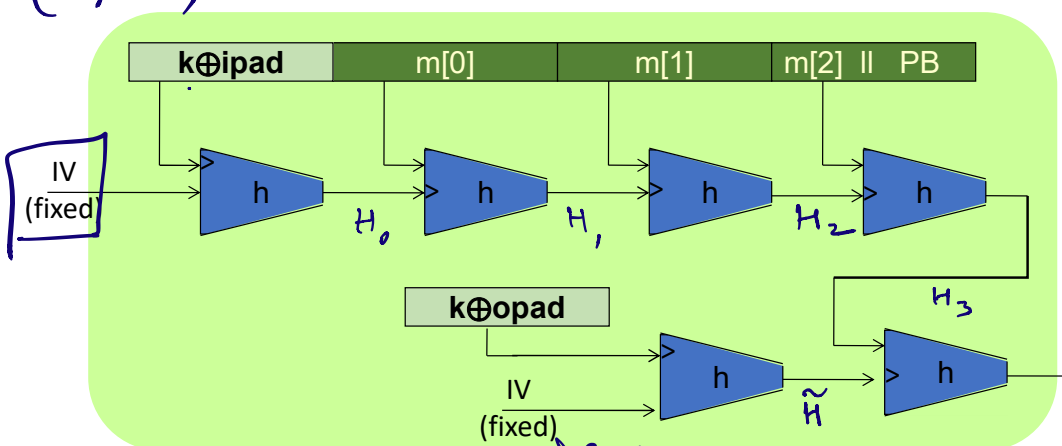
HMAC: $S(k, m) = H(k \oplus opad \;||\; \mathbf{H(k \oplus ipad \;||\; m)})$

Fixed domain: h : bounded input length

Variable domain: H : unbounded input length.

The HMAC Construction

MAC(k, m)



Variable domain

Directly build MAC { from h. ← (fixed domain)

Previously: H (c.r.h.f) } from h. ← (fixed domain)

HMAC: Features

Built from a black-box implementation of SHA-256.

HMAC is assumed to be a secure PRF

- Can be proven under certain PRF assumptions about h(.,.)
- Can even be truncated, to say the first 80 bits of output

This is used in TLS

Summary

- Message Authentication Codes (MACs)
- Hash Functions
- HMAC