Message integrity: Confidentiality alone not sufficient, also need message integrity. Otherwise adversary can tamper with the message (e.g., "Send \$100 to Bob" ⟶ "Send \$100 to Eve")

In some cases (e.g., software patches), integrity more important than confidentiality

Idea: Append a "tag" (also called a "signature") to the message to prove integrity (property we want is tags should be hard to forge)

Observation: The tag should be computed using a keyed-function

↳ Example of keyless integrity check: CRC (cyclic redundancy check) [simple example is to set tag to be the parity]

this tolerates a single error (better error-correcting codes can do much better)

↳ this was used in SSH v1 (1995) for data integrity! Fixed in SSH v2 (1996)

↳ also used in WEP (802.11b) protocol for integrity — also broken!

Problem: If there is no key, anyone can compute it! Adversary can tamper with message and compute the new tag.

Definition. A message authentication code (MAC) with key-space $\mathcal{K}$, message space $\mathcal{M}$ and tag space $\mathcal{T}$ is a tuple of algorithms $\Pi_{MAC} = (\mathrm{Sign}, \mathrm{Verify})$:

$$\left.\begin{aligned} &\mathrm{Sign}: \mathcal{K} \times \mathcal{M} \to \mathcal{T} \\ &\mathrm{Verify}: \mathcal{K} \times \mathcal{M} \times \mathcal{T} \to \{0,1\} \end{aligned}\right\} \text{Must be efficiently-computable}$$

Correctness: $\forall k \in \mathcal{K}, \forall m \in \mathcal{M}$:

$$\Pr[\mathrm{Verify}(k, m, \mathrm{Sign}(k,m)) = 1] = 1$$

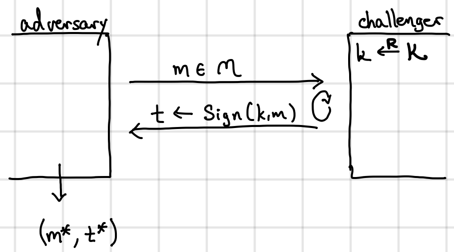
Sign can be a randomized algorithm

Defining security: Intuitively, adversary should not be able to compute a tag on any message without knowledge of the key

↳ Moreover, since adversary might be able to see tags on existing messages (e.g., signed software updates), it should not help towards creating a new MAC

Definition. A MAC $\Pi_{MAC} = (\mathrm{Sign}, \mathrm{Verify})$ satisfies existential unforgeability against chosen message attacks (EUF-CMA) if for all efficient adversaries $\mathcal{A}$, $\mathrm{MACAdv}[\mathcal{A}, \Pi_{MAC}] = \Pr[W=1] = \mathrm{negl}(\lambda)$, where W is the output of the following security game:

adversary gets to choose messages to be signed



As usual, λ denotes the length of the MAC secret key (e.g., $\log |\mathcal{K}| = \mathrm{poly}(\lambda)$)

Note: the key can also be sampled by a special KeyGen algorithm (for simplicity, we just define it to be uniformly random)

Let $m_1, \ldots, m_Q$ be the signing queries the adversary submits to the challenger, and let $t_i \leftarrow \mathrm{Sign}(k, m_i)$ be the challenger's responses. Then, $W = 1$ if and only if:

$$\mathrm{Verify}(k, m^*, t^*) = 1 \quad \text{and} \quad (m^*, t^*) \notin \{(m_1, t_1), \ldots, (m_Q, t_Q)\}$$

MAC security notion says that adversary cannot produce a new tag on any message even if it gets to obtain tags on messages of its choosing.

First, we show that we can directly construct a MAC from any PRF.

MACs from PRFs: Let $F: \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ be a PRF. We construct a MAC Π_{MAC} over $(\mathcal{K}, \mathcal{M}, \mathcal{T})$ as follows:

Sign(k, m): output $t \leftarrow F(k,m)$

Verify(k, m, t): output 1 if $t = F(k,m)$ and 0 otherwise

Theorem. If F is a secure PRF with a sufficiently large range, then Π_{MAC} defined above is a secure MAC. Specifically, for every efficient MAC adversary A, there exists an efficient PRF adversary B such that

$$\text{MACAdv}[A, \Pi_{MAC}] \leq \text{PRFAdv}[B, F] + \frac{1}{|\mathcal{T}|}.$$

Intuition for proof:
1. Output of PRF is computationally indistinguishable from that of a truly random function.
2. If we replace the PRF with a truly random function, adversary wins the MAC game only if it correctly predicts the random function at a new point. Success probability is then exactly $1/|\mathcal{T}|$.

Implication: Any PRF with large output space can be used as a MAC.

↳ AES has 128-bit output space, so can be used as a MAC

Drawback: Domain of AES is 128-bits, so can only sign 128-bit (16-byte) messages

How do we sign longer messages? We will look at two types of constructions:

1. Constructing a large-domain PRF from a small-domain PRF (i.e., AES)
2. Hash-based constructions

So far, we have focused on constructing a large-domain PRF from a small-domain PRF in order to construct a MAC on long messages

↳ Alternative approach: "compress" the message itself (e.g., "hash" the message) and MAC the compressed representation

Still require unforgeability: two messages should not hash to the same value [otherwise trivial attack: if $H(m_1) = H(m_2)$, then MAC on m_1 is also MAC on m_2]

↳ counter-intuitive: if hash value is shorter than messages, collisions always exist — so we can only require that they are hard to find

Definition. A hash function $H: \mathcal{M} \to \mathcal{T}$ is collision-resistant if for efficient adversaries $\mathcal{A}$,

$$\text{CRHFAdv}[\mathcal{A}, H] = \Pr[(m_0, m_1) \leftarrow \mathcal{A} : H(m_0) = H(m_1)] = \text{negl.}$$

As stated, definition is problematic: if $|\mathcal{M}| > |\mathcal{T}|$, then there always exists a collision m_0^*, m_1^* so consider the adversary that has m_0^*, m_1^* hard coded and outputs m_0^*, m_1^*

↳ Thus, some adversary always exists (even if we may not be able to write it down explicitly)

↳ Formally, we model the hash function as being parameterized by an additional parameter (e.g., a "system parameter" or a "key") so adversary cannot output a hard-coded collision

↳ In practice, we have a concrete function (e.g., SHA-256) that does not include security or system parameters

↳ believed to be hard to find a collision even though there are infinitely-many (SHA-256 can take inputs of arbitrary length)

MAC from CRHFs: Suppose we have the following

- A MAC (Sign, Verify) with key-space $\mathcal{K}$, message space $\mathcal{M}_0$ and tag space $\mathcal{T}$
- A collision-resistant hash function $H: \mathcal{M}_1 \to \mathcal{M}_0$

[e.g., $\mathcal{M}_0 = \{0,1\}^{256}$, $\mathcal{M}_1 = \{0,1\}^*$]

Define $S'(k,m) = S(k, H(m))$ and

$V'(k,m,t) = V(k, H(m), t)$

Theorem. Suppose $\Pi_{MAC} = (\text{Sign}, \text{Verify})$ is a secure MAC and H is a CRHF. Then, Π'_{MAC} is a secure MAC. Specifically, for every efficient adversary $\mathcal{A}$, there exist efficient adversaries $\mathcal{B}_0$ and $\mathcal{B}_1$ such that

$$\text{MACAdv}[\mathcal{A}, \Pi'_{MAC}] \le \text{MACAdv}[\mathcal{B}_0, \Pi_{MAC}] + \text{CRHFAdv}[\mathcal{B}_1, H]$$

Proof Idea. Suppose $\mathcal{A}$ manages to produce a valid forgery t on a message m. Then, it must be the case that

- t is a valid MAC on $H(m)$ under Π_{MAC}
- If $\mathcal{A}$ queries the signing oracle on $m' \neq m$ where $H(m') = H(m)$, then $\mathcal{A}$ breaks collision-resistance of H
- If $\mathcal{A}$ never queries signing oracle on m' where $H(m') = H(m)$, then it has never seen a MAC on $H(m)$ under Π_{MAC}. Thus, $\mathcal{A}$ breaks security of Π_{MAC}.

[See Boneh-Shoup for formal argument — very similar to above: just introduce event for collision occurring vs. not occurring]