CIS 5560

Cryptography
Lecture 12

Announcements

- HW 4 is out, due on Friday
- No HW due for next week, we will provide a worksheet to practice problems
- Midterm 1 next Thursday (03/05)
  - If you can't make it, please email me by **tonight** to schedule alternate time.
  -

Recap of last lecture

Collision Resistance

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

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

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

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

Adv$_{CR}$[A,H] = Pr[A outputs collision for H]

is negligible.

Example: SHA-256 (outputs 256 bits)

MACs from Collision Resistance

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

Let $H : M^{\text{big}} \rightarrow M$ be a hash function

Def: $(\text{MAC}^{\text{big}}, \text{Ver}^{\text{big}})$ over (K, M^{big}, T) as:

$$\text{MAC}^{\text{big}}(k, m) = \text{MAC}(k, H(m)); \text{Ver}^{\text{big}}(k, m, t) = V(k, H(m), t)$$

Thm: If MAC is a secure MAC and H is collision resistant
then MAC^{big} is a secure MAC.

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

Generic attack

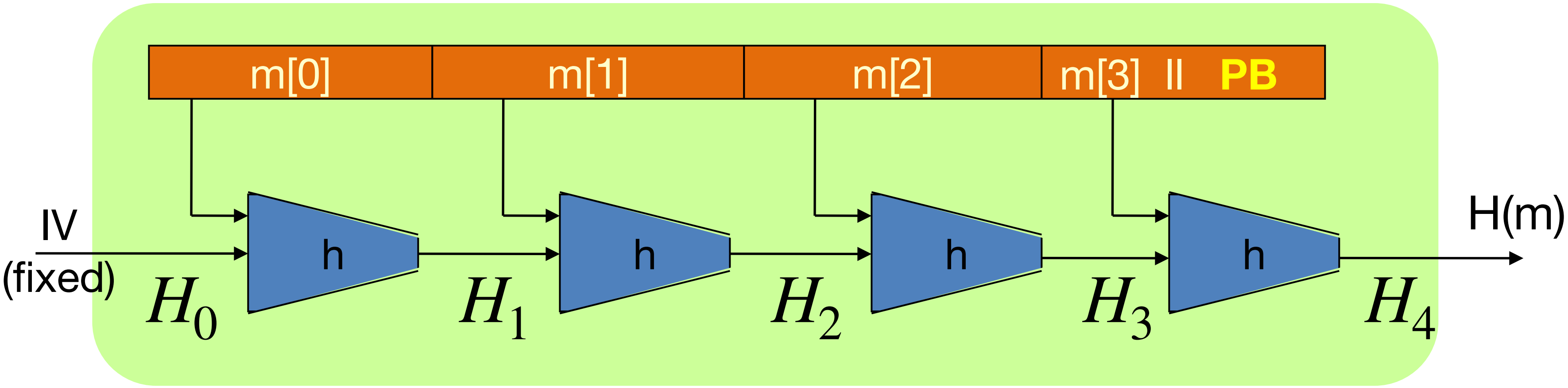
Algorithm:

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

Expected number of iteration ≈ 2

Running time: **O(2^n/2^)** (space O(2^n/2^))

The Merkle-Damgard iterated construction



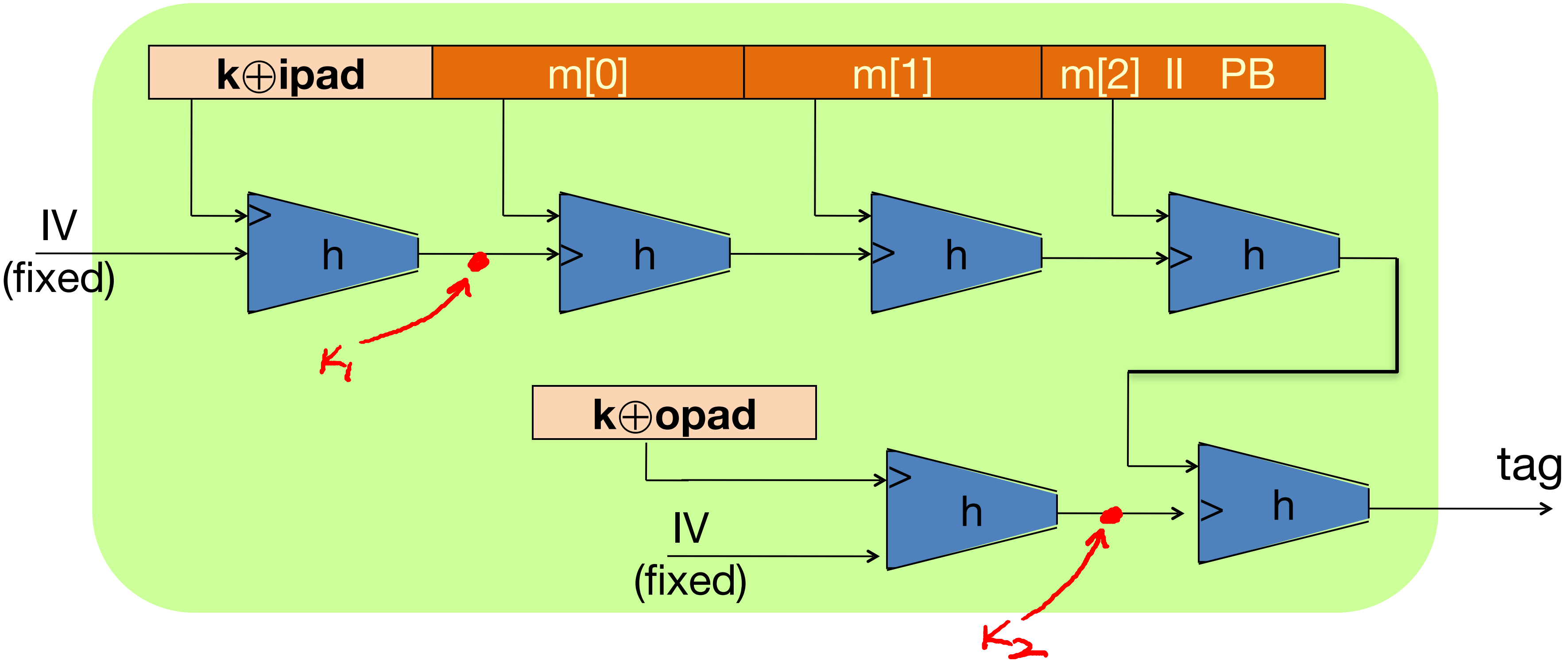
Given $h : T \times X \to T$ (compression function)

we obtain $H : X_{\leq L} \to T$. H_i - chaining variables

PB: padding block 1000…0 ll msg len (64 bits)

If no space for PB add another block

HMAC in pictures



Similar to the NMAC PRF.

main difference: the two keys k_1, k_2 are dependent

Today

- Carter-Wegman MAC
- Other properties of hash functions
- Authenticated encryption

Other properties of (hash) functions

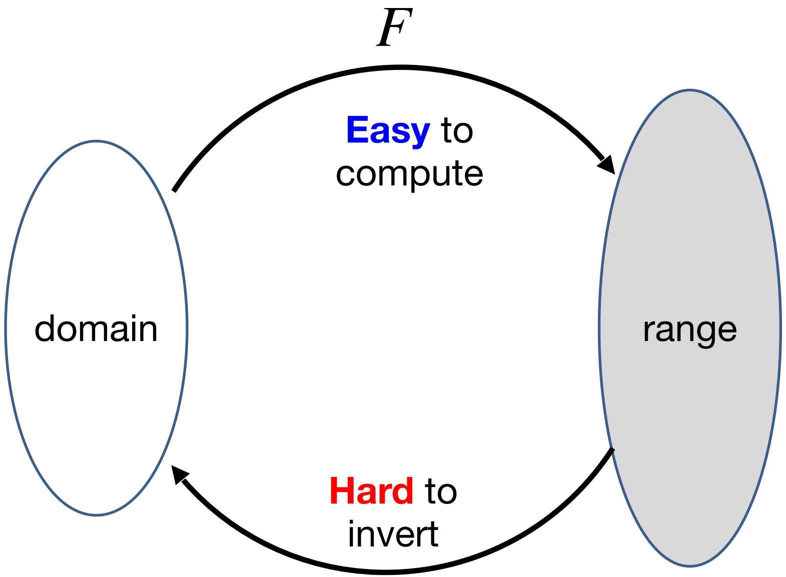
Other properties of (hash) functions

- Collision resistance:
  - Can't find two inputs with same output
  - That is, can't find $x \neq x'$ such that $h(x) = h(x')$
- One-wayness/Preimage resistance:
  - Difficult to find input given an output
  - That is, given $y \in \text{Range}(h)$, can't find x s.t. $h(x) = y$
- 2nd-preimage resistance:
  - Given input x, can't find another input with same output
  - That is, given x, can't find x' s.t. $h(x) = h(x')$

How are these properties related?

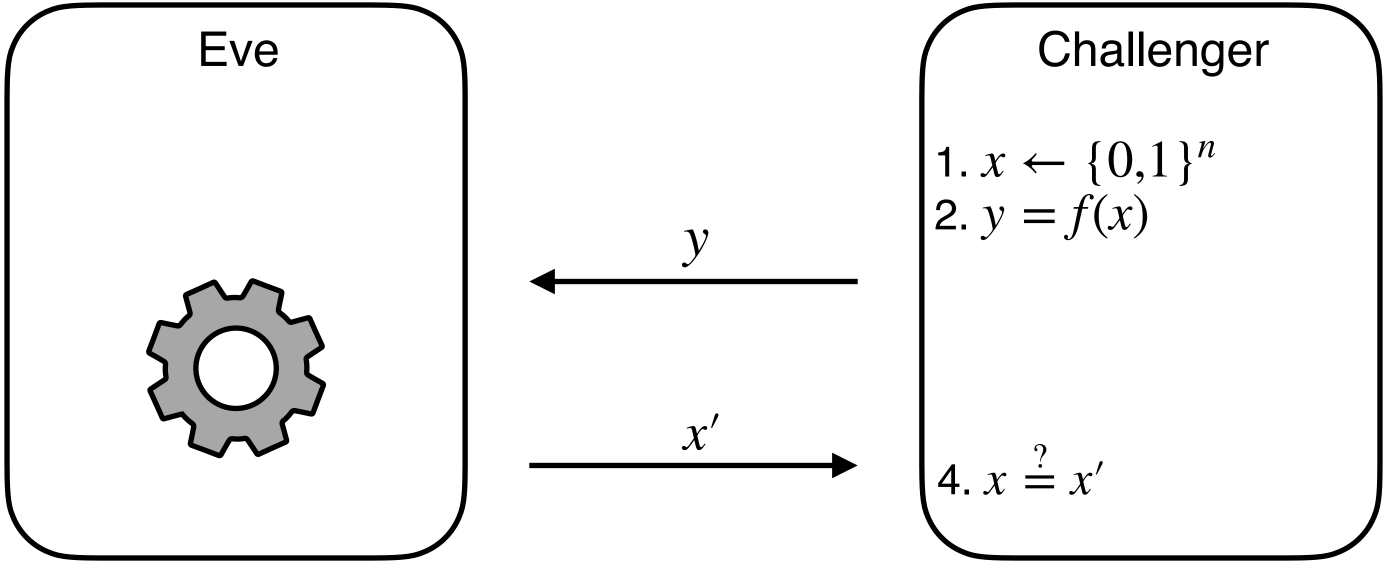
- Q1: If h is collision-resistant, is it also 2nd-preimage resistant?
  - Yes! If you can't find *any* collisions, you also can't find a *specific* collision
- Q2: If h is one-way, is it also collision-resistant?
  - No. E.g.: h outputs 0^n on two inputs.
- Q2: If h is collision-resistant, is it also one-way?
  - Not necessarily! E.g.: let h be CRH. Then construct f such that if first bit of input x is 0, then output rest of input, otherwise, output $h(x)$.

One-way Functions (Informally)



Source of all hard problems in cryptography!

OWF Security Attempt #1



One-way Functions (Take 1)

A function (family) $\{F_n\}_{n \in \mathbb{N}}$ where $F(\cdot): \{0,1\}^n \rightarrow \{0,1\}^{m(n)}$ is **one-way** if for every PPT adversary A, the following holds:

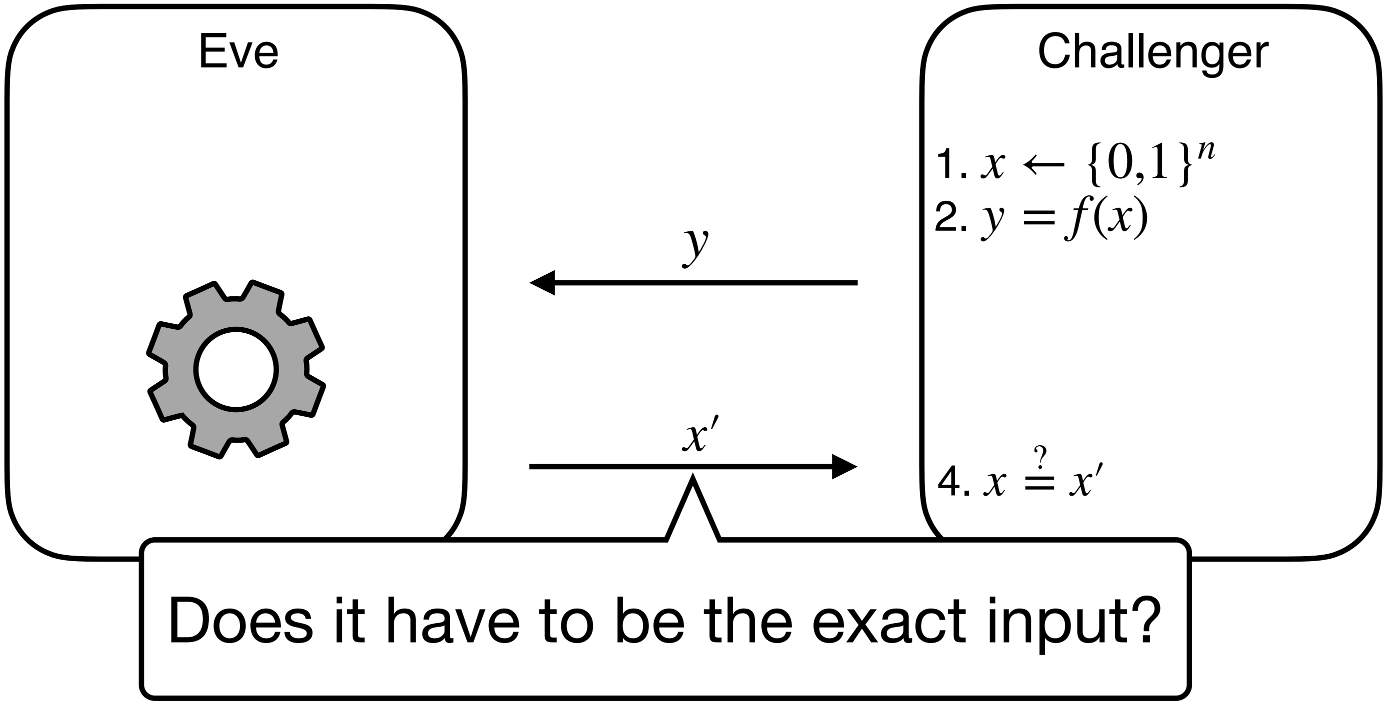
$$\Pr\left[A(1^n, y) = x \;\middle|\; \begin{array}{c} x \leftarrow \{0,1\}^n \\ y := F_n(x) \end{array}\right] = \mathrm{negl}(n)$$

Consider $\boldsymbol{F_n(x) = 0}$ for all x.

This is one-way according to the above definition.
In fact, impossible to find *the* inverse even if A has unbounded time.

Conclusion: not a useful/meaningful definition.

OWF Security Attempt #2



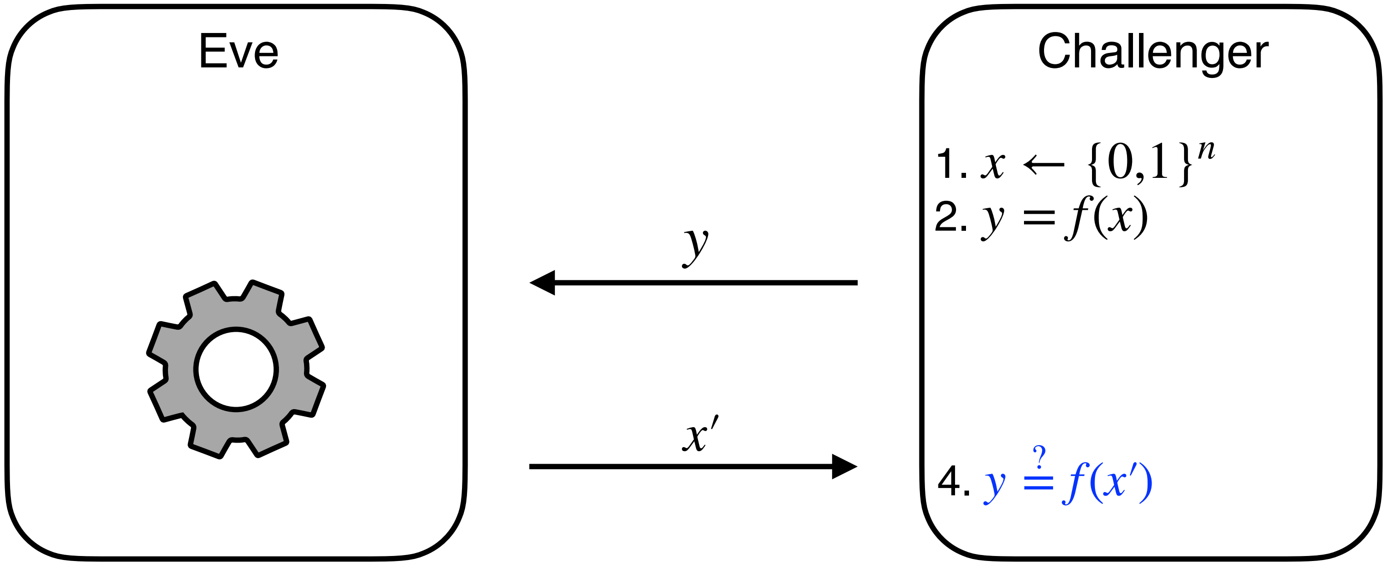
One-way Functions (Take 1)

A function (family) $\{F_n\}_{n\in\mathbb{N}}$ where $F(\cdot) : \{0,1\}^n \rightarrow \{0,1\}^{m(n)}$ is **one-way** if for every p.p.t. adversary A, the following holds:

$$\Pr\left[A(1^n, y) = x \,\middle|\, \begin{array}{r} x \leftarrow \{0,1\}^n \\ y := F_n(x) \end{array}\right] = \text{negl}(n)$$

The Right Definition: Impossible to find ***an*** inverse efficiently.

OWF Security Attempt #2



One-way Functions: The Definition

A function (family) $\{F_n\}_{n\in\mathbb{N}}$ where $F(\cdot):\{0,1\}^n \to \{0,1\}^{m(n)}$ is **one-way** if for every p.p.t. adversary A, the following holds:

$$\Pr\left[F_n(x') = y \;\middle|\; \begin{array}{r} x \leftarrow \{0,1\}^n \\ y := F_n(x) \\ x' \leftarrow A(1^n, y) \end{array}\right] = \mathrm{negl}(n)$$

- Can always find *an* inverse with unbounded time
- … but should be hard with probabilistic polynomial time

One-way Permutations:

One-to-one one-way functions with $m(n) = n$.

Story so far

Confidentiality: semantic security against a CPA attack

- Encryption secure against **eavesdropping only**

Integrity:

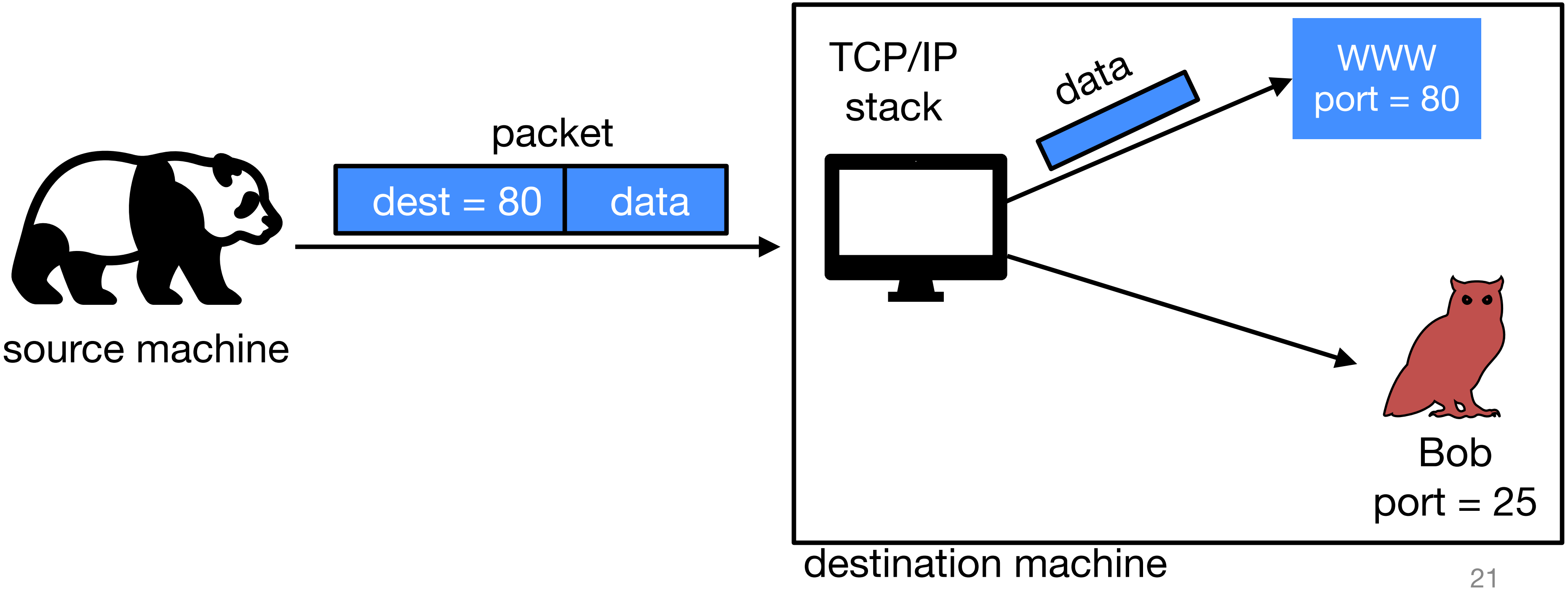
- Existential unforgeability under a chosen message attack
- CBC-MAC, HMAC, CMAC

This module: encryption secure against **tampering**

- Ensuring both confidentiality and integrity

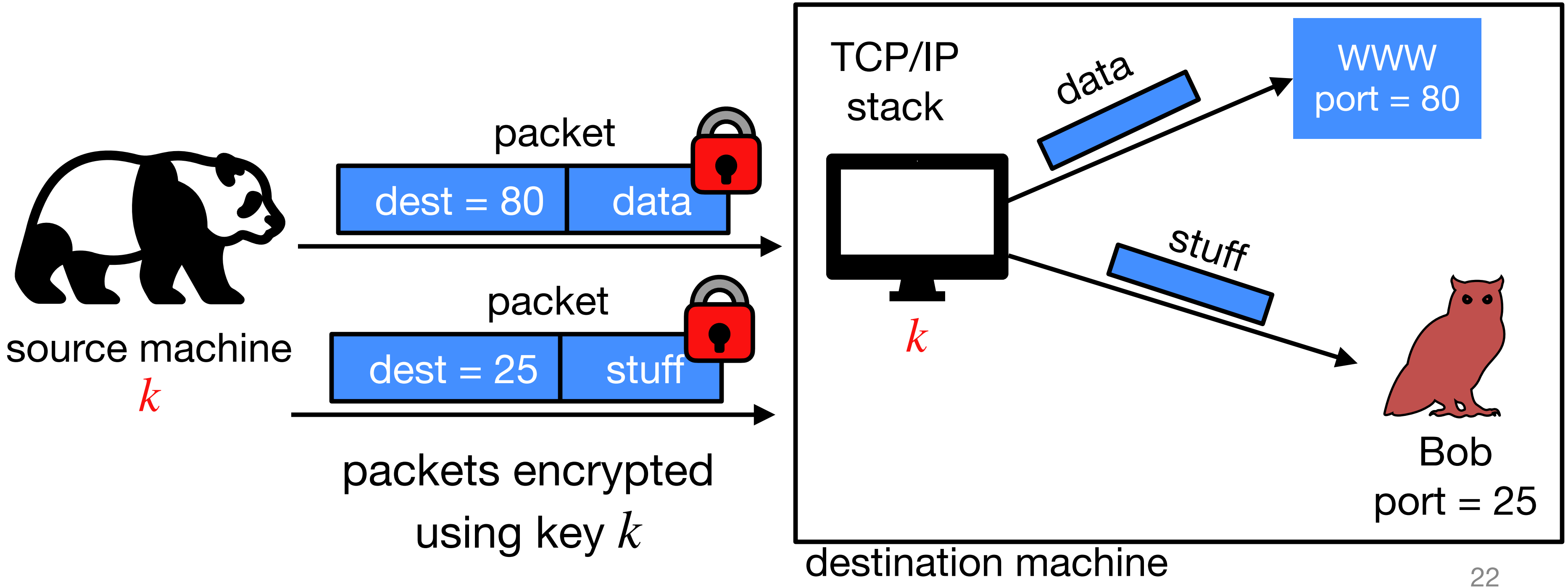
Why is integrity important for privacy?

TCP/IP: (highly abstracted)



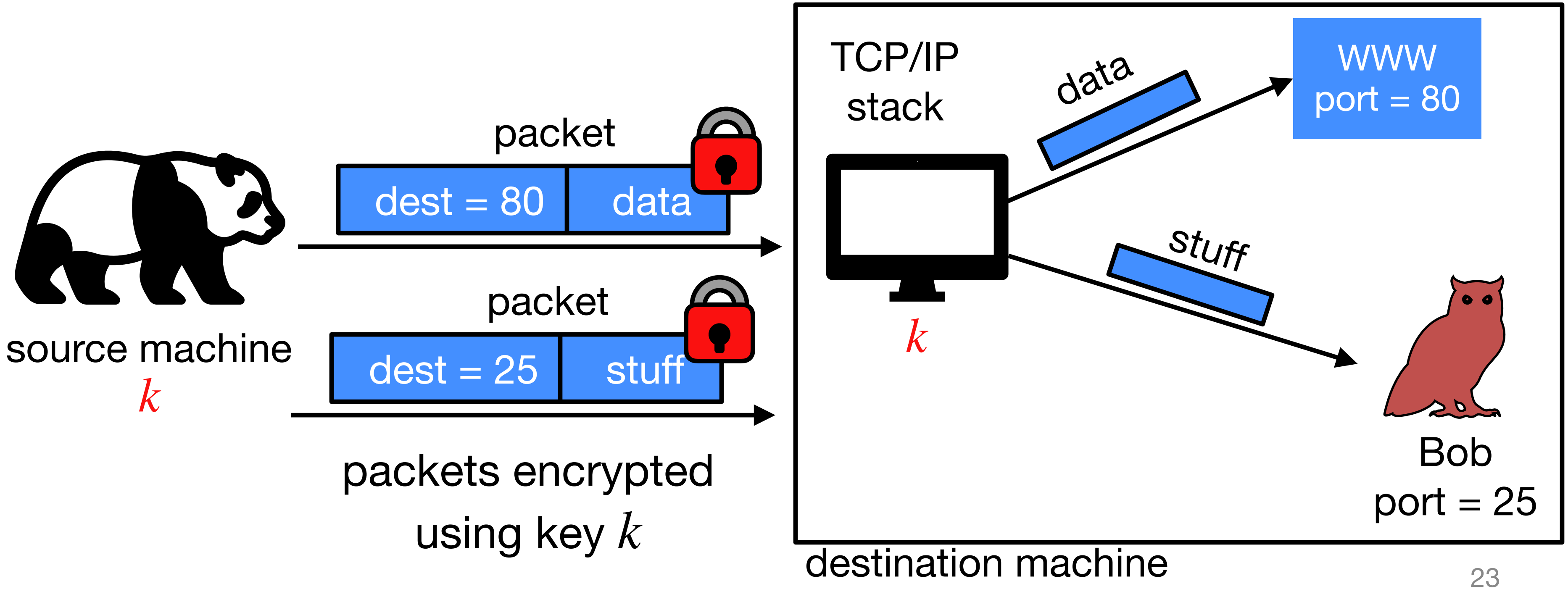
Why is integrity important for privacy?

IPsec: (highly abstracted)



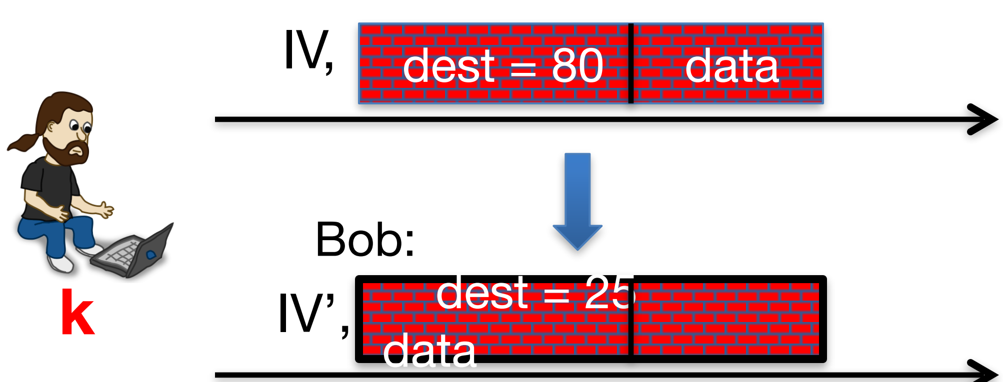
Why is integrity important for privacy?

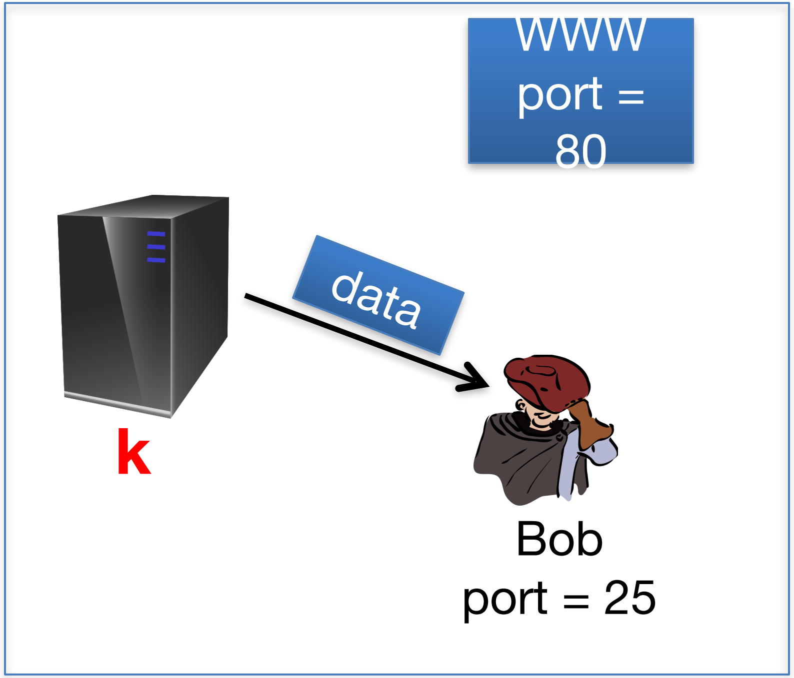
IPsec: (highly abstracted)



Reading someone else's data

Note: attacker obtains decryption of any ciphertext beginning with "dest=25"





Easy to do for CBC with rand. IV

(only IV is changed)

Encryption is done with CBC with a random IV.

What should IV' be?

$m[0] = D(k, c[0]) \oplus IV = \text{"dest=80..."}$

- ○ $IV' = IV \oplus (...25...)$
- ○ $IV' = IV \oplus (...80...)$
- ○ $IV' = IV \oplus (...80...) \oplus (...25...)$
- ○ It can't be done

The lesson

CPA security cannot guarantee secrecy under active attacks.

Only use one of two modes:

- If message needs integrity but no confidentiality:

  use a **MAC**

- If message needs both integrity and confidentiality:

  use **authenticated encryption** modes

Goals

An **authenticated encryption** system (Gen, Enc, Dec) is a cipher where

As usual: $\text{Enc} : \mathscr{K} \times \mathscr{M} \rightarrow \mathscr{C}$

but $\text{Dec} : \mathscr{K} \times \mathscr{C} \rightarrow \mathscr{M} \cup \{\text{Error} / \perp\}$

ciphertext is rejected

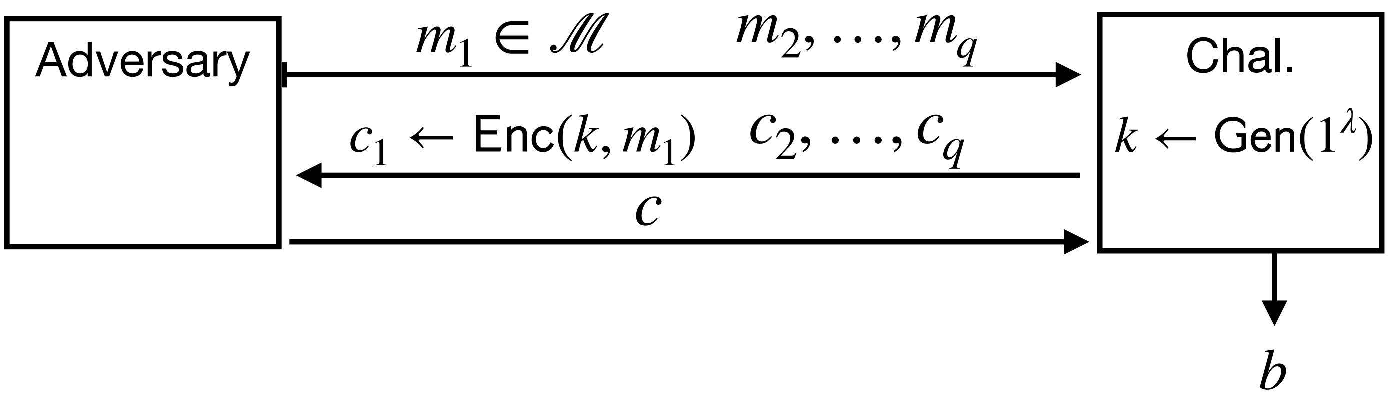
<u>Security</u>: the system must provide

- IND-CPA, and
- **ciphertext integrity:**

  attacker cannot create new ciphertexts that decrypt properly

Ciphertext integrity

Let (Gen, Enc, Dec) be a cipher with message space $\mathscr{M}$.



$b = 1$ if $\text{Dec}(k, c) \neq \perp$ and $c \notin \{c_1, \ldots, c_q\}$

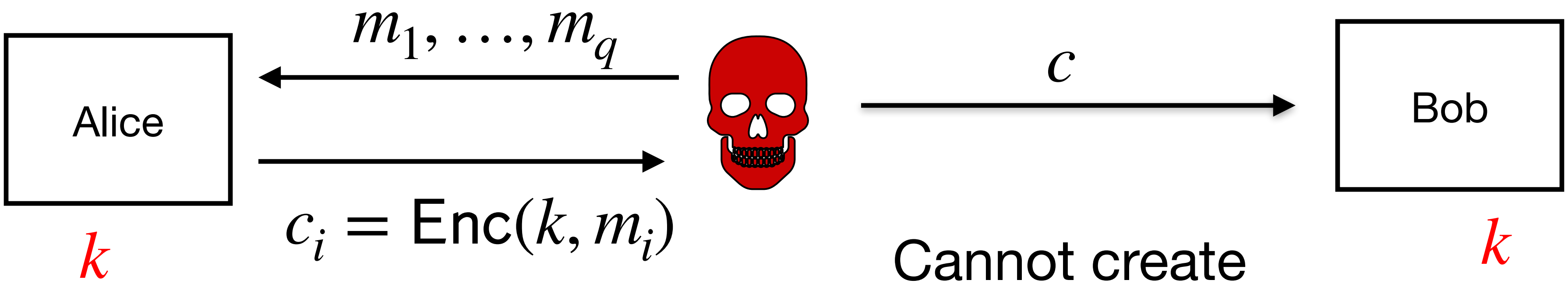
$b = 0$ otherwise

Def: (Gen, Enc, Dec) has **ciphertext integrity** if for all PPT A:

$$\text{Adv}_{\text{CI}}[A] = \Pr[b = 1] = \text{negl}(\lambda)$$

Implication 1: authenticity

Attacker cannot fool Bob into thinking a message was sent from Alice



Cannot create
valid $c \notin \{c_1, \ldots, c_q\}$

$\Rightarrow$ if $\mathsf{Dec}(k, c) \neq \bot$ Bob knows message is from someone who knows k

(but message could be a replay)

Implication 2

Authenticated encryption

↓

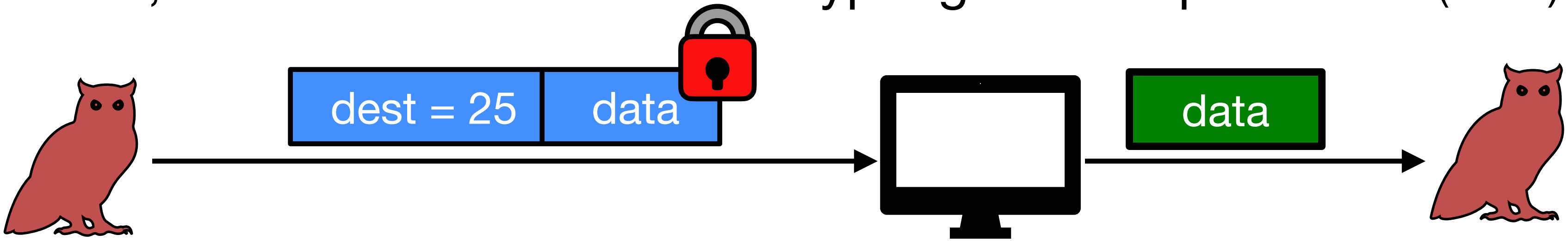
Security against **chosen ciphertext attacks**

Chosen ciphertext attacks

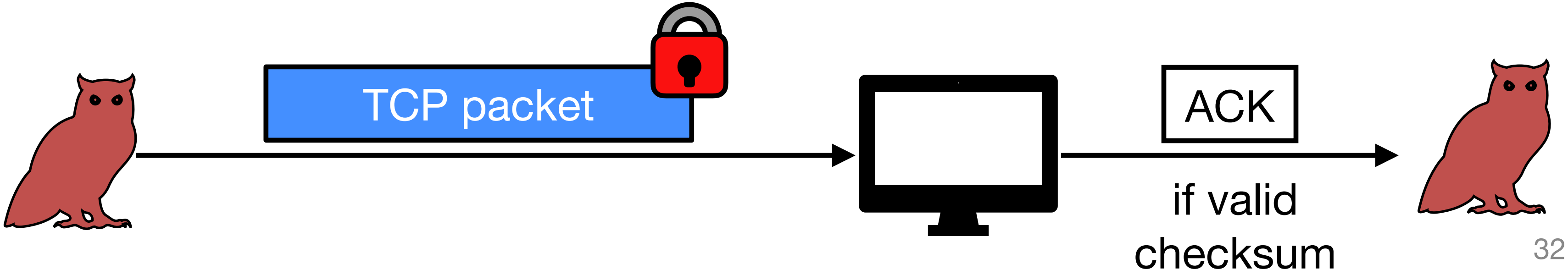
Example chosen ciphertext attacks

Adversary A has ciphertext c that it wants to decrypt

- Often, A can fool server into decrypting ***other*** ciphertexts (not c)



- Often, adversary can learn partial information about plaintext



Chosen ciphertext security

Adversary's power: both CPA and CCA

- Can obtain the encryption of arbitrary messages of his choice
- Can decrypt any ciphertext of his choice, other than challenge

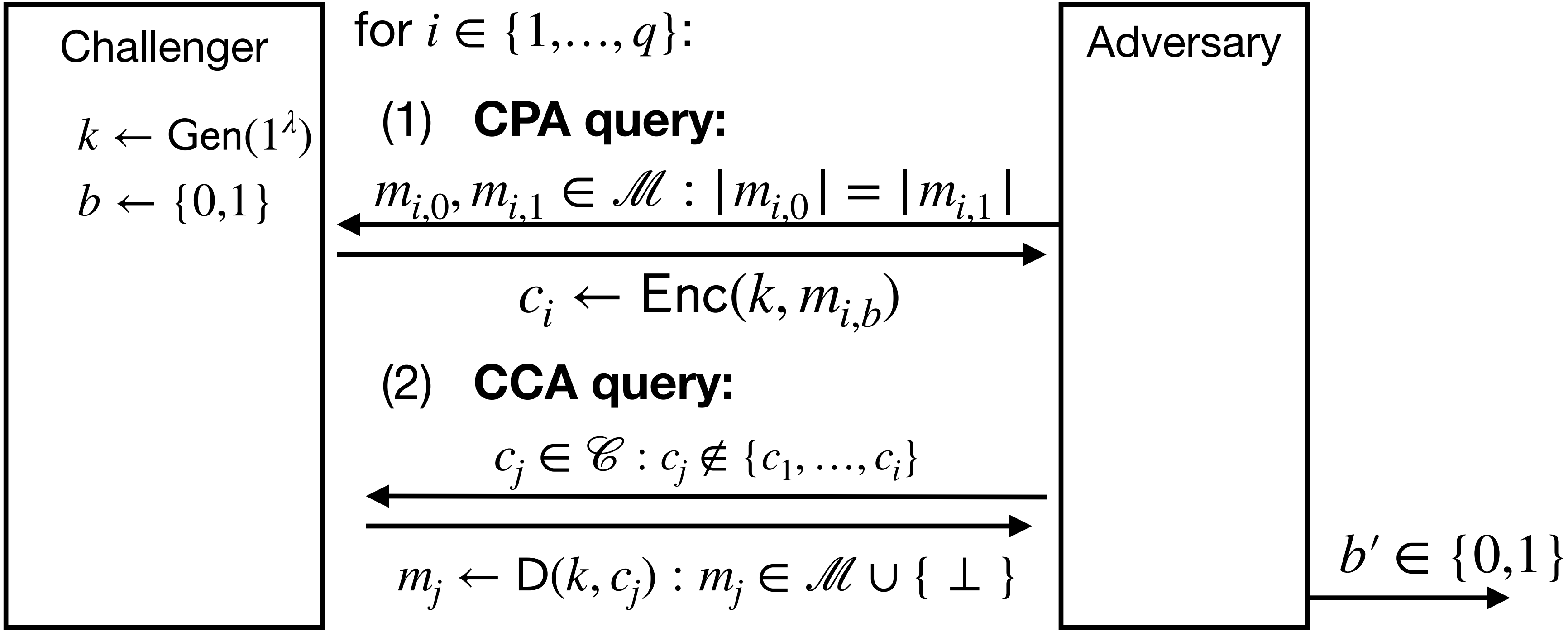
(conservative modeling of real life)

Adversary's goal:

Learn partial information about challenge plaintext

Chosen ciphertext security: definition

Let (Gen, Enc, Dec) be a cipher with message space $\mathscr{M}$



Chosen ciphertext security: definition

E is CCA secure if for all "efficient" A: $\Pr[b = b'] = 1/2 + \mu(\lambda)$

Question: Is CBC with rand. IV CCA-secure?

Authenticated enc. ⇒ CCA security

<u>Thm</u>: Let (E,D) be a cipher that provides AE.

Then (E,D) is CCA secure !

In particular, for any q-query eff. A there exist eff. B_1, B_2 s.t.

$$Adv_{CCA}[A,E] \leq 2q \cdot Adv_{CI}[B_1,E] + Adv_{CPA}[B_2,E]$$

Proof by pictures