

# Autenticazione

Paolo D'Arco  
pdarco@unisa.it

Università di Salerno

Elementi di Crittografia

- 1 Codici per l'autenticazione di messaggi
- 2 Attacchi di tipo side-channel
- 3 Costruzioni sicure di codici
- 4 Costruzione CBC-MAC
- 5 Costruzioni GMAC e Poly1305

Comunicazione *sicure* su un canale pubblico:

- La cifratura può essere usata per prevenire che ascoltatori non autorizzati cariscano informazioni sul contenuto dei messaggi inviati sul canale sproteetto



Confidenzialit 

- Ogni parte per  dovrebbe essere in grado anche di identificare il mittente del messaggio, esser certo che ha inviato il messaggio e capire se il messaggio   integro o   stato modificato



Autenticit , Integrit 

## Esempi dal mondo reale:

- Utente che comunica con la banca. Richiede il trasferimento di 1000 \$ da  $X$  a  $Y$ 
  - é la richiesta autentica?
  - é stato alterato il totale?
  - le tecniche di rilevamento e correzione degli errori standard non sono utili qui: affrontano un altro problema.
- Cookies usati da un server web
  - *HTTP* é un protocollo senza stato. Aiutano ad implementare una "sessione"
  - l'utente non dovrebbe essere in grado di modificarli
  - il server web dovrebbe essere in grado di verificarne autenticità ed integrità

# Codici per l'autenticazione di messaggi

In generale, l'autenticità di una comunicazione non può essere basata su normali informazioni tipo

- identificatori degli SMS
- ID del chiamante
- indirizzo email
- ...

⇒ non rappresentano una *prova* dell'origine.

Inoltre, non possiamo evitare che le informazioni vengano alterate. Ma possiamo far sí che *ogni* alterazione venga riconosciuta.

Vedremo come ottenere autenticità ed integrità usando tecniche crittografiche.

# Cifratura: strumento per autenticità e integrità?

La cifratura **non** risolve il problema dell'autenticità e dell'integrità.

Esempio. Cifratura attraverso:

$$Enc_k(m) = G(k) \oplus m$$

dove  $G$  è un PRG costruito a partire da uno stream cipher.

Il cifrato è molto facile da manipolare:

- l'*inversione* di un qualsiasi bit nel cifrato risulta nell'inversione dello stesso bit nel messaggio in chiaro



Le implicazioni possono essere nefaste. In un trasferimento bancario:

- inversione del bit meno significativo  $\Rightarrow$  nessun impatto
- inversione dell'11-esimo bit  $\Rightarrow$  impatto significativo!

# Cifratura usando cifrari a blocchi?

Lo stesso attacco appena descritto si applica alle modalità OFB e CTR



Anche usando cifrature CPA-sicure non é sufficiente per prevenire alterazioni dei messaggi. E circa ECB e CBC?

- ECB: modifiche anche di un solo bit del cifrato influenzano ancora il messaggio in chiaro, anche se la posizione *non* é controllabile
- CBC: cambiando il  $j$ -esimo bit di  $IV$  si cambia il  $j$ -esimo bit del primo blocco del messaggio

$$m_1 := F_k^{-1}(c_1) \oplus IV$$

Tutti gli altri blocchi restano inalterati, ma il primo contiene informazioni importanti.

# Cifratura usando cifrari a blocchi?

Nota: tutti gli schemi di cifratura studiati fino ad ora sono tali che:

*ogni cifrato corrisponde ad un qualche messaggio in chiaro*



Adv può scegliere una stringa  $c$  (cifrato arbitrario) e inviarlo ad una parte

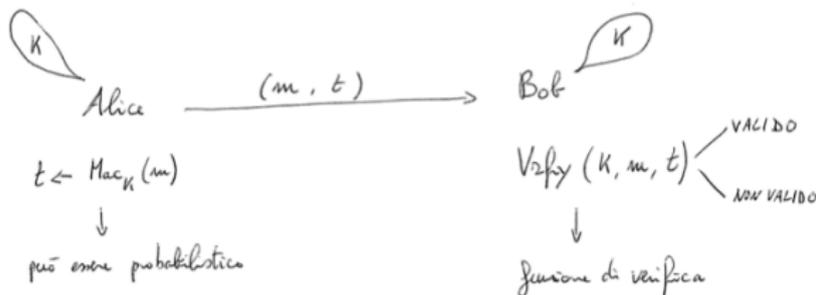


Un meccanismo aggiuntivo é necessario

I codici per l'autenticazione dei messaggi (*message authentication code*)

- permettono di identificare ed autenticare il mittente
- permettono di controllare l'integrità del messaggio

Entrambe le parti devono condividere un segreto che Adv non conosce.



**Definizione 4.1.** Un codice per l'autenticazione di messaggi (MAC, in breve) consiste di tre algoritmi ppt ( $Gen, Mac, Vrfy$ ) dove

- 1  $k \leftarrow Gen(1^n), \quad |k| \geq n$ , genera le chiavi
- 2  $t \leftarrow Mac_k(m), \quad m \in \{0,1\}^*$ , dove  $t$  é il *tag* associato ad  $m$
- 3  $b := Vrfy_k(m, t), \quad b \in \{0,1\}$

tali che, per ogni  $n$ , per ogni  $k$  dato in output da  $Gen(1^n)$  e per ogni  $m \in \{0,1\}^*$

$$Vrfy_k(m, Mac_k(m)) = 1.$$

Se esiste una funzione  $\ell(n)$  e lo schema é definito solo per  $m \in \{0,1\}^{\ell(n)}$ , allora parleremo di MAC a lunghezza fissa per messaggi di lunghezza  $\ell(n)$ .

Solitamente  $k \leftarrow \text{Gen}(1^n)$  é uniforme.

Una classe importante é quella degli schemi deterministici.

- $\text{Mac}_k(\cdot)$  é deterministica
- La verifica si dice *canonica*



Data la coppia  $(m, t)$ , ricalcola  $\text{Mac}_k(m) = t'$  e verifica che  $t' = t$

In generale, usare una funzione  $\text{Vrfy}(\cdot)$  é utile per distinguere le semantiche di due operazioni diverse: autenticazione e verifica.

*"Nessun Adv PPT dovrebbe essere in grado di generare un tag su qualsiasi messaggio nuovo, che non sia stato autenticato ed inviato in precedenza da una delle parti."*

Potere Adv

- Adv: algoritmo PPT, può influenzare i contenuti da autenticare
- Useremo un oracolo  $Mac_k(\cdot)$ : Adv può chiedere tag di messaggi  $m$  a propria scelta

Cos'è una rottura di un MAC

- Adv dá in output  $(m, t)$  tale che:
  - 1 il tag  $t$  è un tag valido su  $m$ , i.e.,  $Vrfy(m, t) = 1$
  - 2 Adv non ha ottenuto  $t$  per  $m$  con una precedente richiesta all'oracolo

La 1. implica che  $m$  viene accettato come valido.

La 2. che Adv ha generato  $t$  da solo

## Osservazioni:

- Schemi MAC che non permettono ad Adv di avere successo con probabilità piú che trascurabile sono detti *non falsificabili esistenzialmente rispetto ad un attacco adattivo di tipo chosen-message* (existentially unforgeable under an adaptive chosen-message attack)

Sia  $\Pi = (Gen, Mac, Vrfy)$ ,  $A$  un Adv generico ed  $n$  il parametro di sicurezza

$Mac\text{-}forge_{A,\Pi}(n)$

- 1 Il challenger genera  $k \leftarrow Gen(1^n)$  e setta l'oracolo  $Mac_k(\cdot)$
- 2  $A^{Mac_k(\cdot)}(1^n)$  invia all'oracolo un certo numero di query. Sia  $Q$  l'insieme delle query richieste. Alla fine  $A^{Mac_k(\cdot)}(1^n)$  dá in output  $(m, t)$ .
- 3  $A^{Mac_k(\cdot)}(1^n)$  ha successo se e solo se
  - $Vrfy_k(m, t) = 1$
  - $m \notin Q$

In questo caso, l'output dell'esperimento é 1; altrimenti é 0.

**Definizione 4.2.** Un codice per l'autenticazione di messaggi  $\Pi = (Gen, Mac, Vrfy)$  é non falsificabile esistenzialmente rispetto ad un attacco adattivo di tipo chosen-message (existentially unforgeable under an adaptive chosen-message attack) se, per ogni  $A$  PPT, esiste una funzione trascurabile  $negl$ , tale che:

$$Pr[Mac\text{-}forge_{A,\Pi}(n) = 1] \leq negl(n).$$

É la definizione troppo forte?

- Non falsificabile esistenzialmente  $\Rightarrow$  anche un messaggio *senza significato* non é autenticabile da Adv
- Rispetto ad attacchi chosen-message  $\Rightarrow$  Adv nella vita reale potrebbe non avere la possibilità di scegliere ...

Avendo una definizione piú forte possibile, non dobbiamo preoccuparci della specifica applicazione.

Sono una minaccia reale ma *non* sono modellati dalla definizione.

Due soluzioni comuni:

- usare numeri di sequenza nei messaggi
- marche temporali (timestamp) che certificano il tempo

Nota: rispettando la definizione Adv può produrre un *nuovo* Mac su un messaggio per cui un Mac è già stato calcolato. Cioè Adv

dato  $(m, t)$  produce  $(m, t')$  dove  $t \neq t'$ .

In alcune applicazioni può essere un problema

**Definizione 4.2.** Un codice per l'autenticazione di messaggi  $\Pi = (Gen, Mac, Vrfy)$  é fortemente sicuro (o é un Mac forte) se, per ogni  $A$  PPT, esiste una funzione trascurabile  $negl$ , tale che:

$$Pr[Mac-sforge_{A,\Pi}(n) = 1] \leq negl(n),$$

dove  $Mac-sforge_{A,\Pi}(n)$  é come  $Mac-forge_{A,\Pi}(n)$  ma l'insieme delle query  $Q$  contiene coppie  $(m, t)$ .

Nota che:

- In  $Mac-forge_{A,\Pi}(n)$ :  $A$  produce  $(m, t')$  con  $m \in Q \Rightarrow$  non é un successo per  $A$
- In  $Mac-sforge_{A,\Pi}(n)$ :  $A$  produce  $(m, t')$  con  $(m, t) \in Q \Rightarrow$  é un successo per  $A$

**Proposizione 4.4.** Se  $\Pi = (\text{Gen}, \text{Mac}, \text{Vrfy})$  é un Mac sicuro che usa la verifica canonica, allora  $\Pi$  é anche fortemente sicuro.

Molti Mac usati nelle applicazioni reali usano la verifica canonica.

Osservazione: la definizione fornisce ad Adv accesso all'oracolo per ottenere Mac di messaggi. Perché non diamo ad Adv anche accesso ad un oracolo di verifica?

Per Mac con verifica canonica, non c'è differenza: qualsiasi Mac che soddisfa la Definizione 4.2 soddisferebbe anche una definizione con un oracolo di verifica

Per Mac con verifica generica potrebbe essere diverso.

# Timing Attack

Un potenziale attacco basato sul calcolo del tempo

Adv invia coppie  $(m, t)$  ad un ricevitore, che funziona come un oracolo di verifica, e acquisisce info tipo

- se il ricevitore accetta/rifiuta
- il tempo che spende per decidere

Adv può calcolare un tag corretto per un *nuovo* messaggio.

Per esempio, supponiamo che uno schema Mac usi la verifica canonica.

Per verificare  $t$  su  $m$ , il ricevitore

- calcola  $t' := \text{Mac}_k(m)$  e confronta  $t' = t$ ,
- dá in output 1 (uguale) o 0 (diversi).

# Timing Attack

Supponiamo che il confronto  $t' = t$  sia implementato con la funzione C *strcmp*.

La funzione *strcmp* confronta un byte alla volta e rifiuta appena trova due byte che differiscono. Il tempo di rifiuto dipende dalla posizione del primo byte in cui  $t'$  e  $t$  differiscono.

Supponiamo Adv conosca i primi  $i$  byte di  $t$  per  $m$ .

⇒ Adv può calcolare il prossimo byte del tag corretto  $t$  per  $m$  inviando al ricevitore

$$(m, t_0), \quad (m, t_1), \quad \dots, \quad (m, t_{255})$$

dove in  $t_j$

- i primi  $i$  byte sono corretti (quelli che Adv conosce)
- il byte  $i + 1$  è uguale a  $j$
- i byte rimanenti sono uguali a 0

# Timing Attack

Tutti gli elementi della sequenza saranno rifiutati ma la coppia  $(m, t_j)$  con il  $j$  giusto richiederá piú tempo!

Applicando l'attacco a partire da  $i = 0$  e ripetendo per ogni byte, Adv calcola il tag corretto  $t$ .

Adv ha bisogno al piú di 256 query per calcolare ciascun byte



Per un tag di 16 byte, 4096 query sono sufficienti per il calcolo dell'intero tag



L'attacco é efficiente

L'attacco descritto é un attacco di tipo *side-channel* ed é di tipo fisico.

Attacchi come il precedente che possono essere sferrati nel mondo reale *non* sono modellati dalle definizioni usuali.

Il precedente é stato applicato alla playstation Xbox 360 qualche anno fa.

**Conclusion:** la procedura di verifica del Mac dovrebbe essere implementata usando una funzione time-independent, che confronta *sempre tutti* i byte delle stringhe.

Un attacco di tipo side-channel usa l'osservazione fisica dell'hardware o l'analisi dei messaggi di errore per carpire informazioni circa bit segreti.

L'attacco può

- sfruttare variazioni temporali, il consumo di energia, o la radiazione elettromagnetica che un dispositivo genera durante l'esecuzione di un algoritmo
- misurare le risposte agli errori di formattazione o ai guasti indotti
- sfruttare informazione rilasciata dalla computazione circa i pattern di accesso della CPU alla memoria cache
- sfruttare bug o peculiarità del linguaggio di programmazione usato
- una combinazione dei differenti tipi di informazione rilasciata

Le funzioni pseudocasuali sono uno strumento naturale per la costruzione di codici per l'autenticazione dei messaggi

Intuitivamente  $t = F_k(m) \Rightarrow$  probabilità di individuare il valore  $\approx 1/2^n$

## **CONSTRUCTION 4.5**

Let  $F$  be a pseudorandom function. Define a fixed-length MAC for messages of length  $n$  as follows:

- **Mac**: on input a key  $k \in \{0, 1\}^n$  and a message  $m \in \{0, 1\}^n$ , output the tag  $t := F_k(m)$ . (If  $|m| \neq |k|$  then output nothing.)
- **Vrfy**: on input a key  $k \in \{0, 1\}^n$ , a message  $m \in \{0, 1\}^n$ , and a tag  $t \in \{0, 1\}^n$ , output 1 if and only if  $t \stackrel{?}{=} F_k(m)$ . (If  $|m| \neq |k|$ , then output 0.)

**Teorema 4.6.** Se  $F$  é una funzione pseudocasuale, allora la costruzione 4.5 realizza un Mac sicuro di lunghezza fissata per messaggi di lunghezza  $n$ .

**Dim.** Sia  $A$  ppt e sia  $\tilde{\Pi} = (\tilde{Gen}, \tilde{Mac}, \tilde{Vrfy})$  (ipotetico,  $f \in Func_n$  invece di  $F_k$ ).

Risulta

$$Pr[\text{Mac-forge}_{A, \tilde{\Pi}}(n) = 1] = 1/2^n.$$

Infatti, per ogni  $m \notin Q$ , il valore  $t = f(m)$  é uniformemente distribuito in  $\{0, 1\}^n$  dal punto di vista di  $A$ . Mostreremo che  $\exists$  una funzione  $negl()$  tale che:

$$|Pr[\text{Mac-forge}_{A, \tilde{\Pi}}(n) = 1] - Pr[\text{Mac-forge}_{A, \Pi}(n) = 1]| \leq negl(n)$$

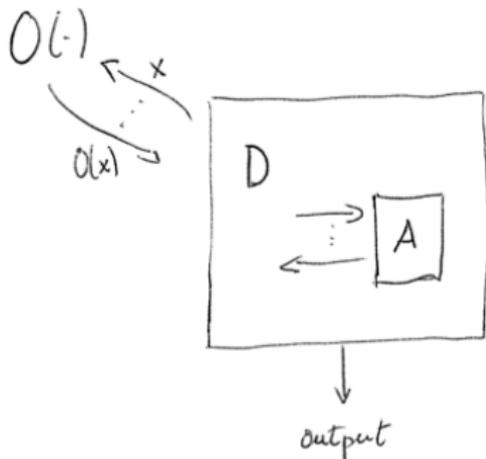
che, combinato con il risultato precedente, implica

$$Pr[\text{Mac-forge}_{A, \Pi}(n) = 1] \leq 1/2^n + negl(n), \quad \text{provando il teorema.}$$

# Costruzioni sicure di codici

Per provare l'affermazione, costruiremo un distinguisher  $D$  con accesso oracolare ad una funzione, il cui obiettivo é stabilire se la funzione é casuale o pseudocasuale.

$D$  emula l'esperimento  $\text{Mac-forge}_{A,?}(n)$  e osserva se  $A$  vince o perde.



*A pensa di star eseguendo l'esperimento reale. In realt  e  $D$  che lo sta "emulando".*

## Distinguisher $D$ :

$D$  is given input  $1^n$  and access to an oracle  $\mathcal{O} : \{0, 1\}^n \rightarrow \{0, 1\}^n$ , and works as follows:

1. Run  $\mathcal{A}(1^n)$ . Whenever  $\mathcal{A}$  queries its MAC oracle on a message  $m$  (i.e., whenever  $\mathcal{A}$  requests a tag on a message  $m$ ), answer this query in the following way:

Query  $\mathcal{O}$  with  $m$  and obtain response  $t$ ; return  $t$  to  $\mathcal{A}$ .

2. When  $\mathcal{A}$  outputs  $(m, t)$  at the end of its execution, do:
  - (a) Query  $\mathcal{O}$  with  $m$  and obtain response  $\hat{t}$ .
  - (b) If (1)  $\hat{t} = t$  and (2)  $\mathcal{A}$  never queried its MAC oracle on  $m$ , then output 1; otherwise, output 0.

$D$  é ppt se  $A$  é ppt. Inoltre:

- Se  $O(\cdot)$  é pseudocasuale  
 $\Rightarrow$  la vista di  $A$  che esegue come subroutine di  $D$  é identica a quella che avrebbe nell'esperimento  $\text{Mac-forge}_{A,\Pi}(n)$
- Se  $O(\cdot)$  é casuale  
 $\Rightarrow$  la vista di  $A$  che esegue come subroutine di  $D$  é identica a quella che avrebbe nell'esperimento  $\text{Mac-forge}_{A,\tilde{\Pi}}(n)$

Pertanto, si ha:

$$\Pr[D^{F_k(\cdot)}(n) = 1] = \Pr[\text{Mac-forge}_{A,\Pi}(n) = 1]$$

e

$$\Pr[D^{f(\cdot)}(n) = 1] = \Pr[\text{Mac-forge}_{A,\tilde{\Pi}}(n) = 1].$$

Ma poiché  $F$  é pseudocasuale,  $\exists$  una funzione  $negl(n)$  tale che

$$|Pr[D^{F_k(\cdot)}(n) = 1] - Pr[D^{f(\cdot)}(n) = 1]| \leq negl(n)$$

$\Downarrow$

$$|Pr[\text{Mac-forge}_{A, \tilde{\Pi}}(n) = 1] - Pr[\text{Mac-forge}_{A, \Pi}(n) = 1]| \leq negl(n)$$

□

Come possiamo estendere la costruzione a messaggi *di lunghezza arbitraria*?

Sia  $\Pi' = (Mac', Vrfy')$  un mac sicuro a lunghezza fissa per messaggi di lunghezza  $n$ .

Sia  $m = m_1 m_2 \dots m_d$ , messaggio di  $d$  blocchi di lunghezza  $n$

Come usare  $\Pi'$  per costruire  $\Pi$ ? Alcune idee

- 1  $m_1 m_2 \dots m_d \rightarrow t_1 t_2 \dots t_d$ , dove  $t_i = Mac'_k(m_i)$ 
  - previene l'invio di blocchi non autenticati
  - non previene attacchi di riordino  $m_2 m_1 \rightarrow t_2 t_1$
- 2 per evitare il riordino, usiamo un indice di blocco
  - non previene attacchi "per troncamento" (A butta via la fine del msg)

3. aggiungiamo in ogni blocco la lunghezza del messaggio

$$t_i = \text{Mac}'_k(\ell || i || m_i)$$

- non previene attacchi del tipo "mix-and-match"

$$m_1 \dots m_d \rightarrow t_1 \dots t_d \quad \text{e} \quad m'_1 \dots m'_d \rightarrow t'_1 \dots t'_d$$

↓

$$m_1 m'_2 \dots m_d \rightarrow t_1 t'_2 \dots t_d$$

4. aggiungiamo un "identificatore casuale del messaggio" in ogni blocco.

## CONSTRUCTION 4.7

Let  $\Pi' = (\text{Mac}', \text{Vrfy}')$  be a fixed-length MAC for messages of length  $n$ . Define a MAC as follows:

- **Mac**: on input a key  $k \in \{0, 1\}^n$  and a message  $m \in \{0, 1\}^*$  of (nonzero) length  $\ell < 2^{n/4}$ , parse  $m$  into  $d$  blocks  $m_1, \dots, m_d$ , each of length  $n/4$ . (The final block is padded with 0s if necessary.) Choose a uniform identifier  $r \in \{0, 1\}^{n/4}$ . For  $i = 1, \dots, d$ , compute  $t_i \leftarrow \text{Mac}'_k(r \parallel \ell \parallel i \parallel m_i)$ , where  $i, \ell$  are encoded as strings of length  $n/4$ .<sup>†</sup> Output the tag  $t := \langle r, t_1, \dots, t_d \rangle$ .
- **Vrfy**: on input a key  $k \in \{0, 1\}^n$ , a message  $m \in \{0, 1\}^*$  of length  $\ell < 2^{n/4}$ , and a tag  $t = \langle r, t_1, \dots, t_{d'} \rangle$ , parse  $m$  into  $d$  blocks  $m_1, \dots, m_d$ , each of length  $n/4$ . (The final block is padded with 0s if necessary.) Output 1 if and only if  $d' = d$  and  $\text{Vrfy}'_k(r \parallel \ell \parallel i \parallel m_i, t_i) = 1$  for  $1 \leq i \leq d$ .

<sup>†</sup> Note that  $i$  and  $\ell$  can be encoded using  $n/4$  bits because  $i, \ell < 2^{n/4}$ .

**Teorema 4.8** Se  $\Pi'$  é un Mac a lunghezza fissa per messaggi di lunghezza  $n$ , la Costruzione 4.7 produce un Mac sicuro per messaggi di lunghezza arbitraria.

**Dim.** Intuizione. Nell'ipotesi in cui  $\Pi'$  é sicuro:

- $A$  non può introdurre un nuovo blocco con un tag valido
- le informazioni aggiunte ad ogni blocco prevengono il riuso

Sia  $\Pi$  il Mac della Costruzione 4.7.

Mostreremo che:

$$\Pr[\text{Mac-forge}_{A,\Pi}(n) = 1] \leq \text{negl}(n).$$

# Lunghezza arbitraria

Indichiamo con *Repeat* l'evento "lo stesso identificatore é presente in due dei tag restituiti dall'oracolo nell'esperimento".

Sia  $(m, t = \langle r, t_1, \dots \rangle)$  l'output finale di *A*.

Indichiamo con *Newblock* l'evento "almeno uno dei blocchi  $r||\ell||i||m_i$  non é stato autenticato in precedenza tramite query all'oracolo."

Risulta  $Pr[\text{Mac-forge}_{A,\Pi}(n) = 1]$  uguale a:

$$Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \text{Repeat}] + Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}}]$$

Inoltre,  $Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}}]$  risulta uguale a:

$$\begin{aligned} &Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}} \wedge \text{Newblock}] \\ &+ Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}} \wedge \overline{\text{Newblock}}]. \end{aligned}$$

↓

Dalle precedenti risulta:

$$\begin{aligned} & Pr[\text{Mac-forge}_{A,\Pi}(n) = 1] \\ & \leq Pr[\text{Repeat}] + Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \text{Newblock}] \\ & \quad + Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}} \wedge \overline{\text{Newblock}}]. \end{aligned}$$

Consideriamo i tre termini separatamente.

**Claim 4.9**  $Pr[\text{Repeat}]$  é trascurabile.

**Dim.** Sia  $q(n)$  il numero (polinomiale) di query all'oracolo effettuate da  $A$ . L'oracolo usa  $r_i$  tali che  $|r_i| = n/4$ , scelti uniformemente a caso in  $\{0, 1\}^{n/4}$ .

$$\Rightarrow Pr[\text{Repeat}] = Pr[r_i = r_j \text{ per } i \neq j] \leq \frac{q(n)^2}{2^{n/4}}.$$

□

# Lunghezza arbitraria

Soffermiamoci ora sul termine

$$Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}} \wedge \overline{\text{Newblock}}].$$

Faremo vedere che se  $\text{Mac-forge}_{A,\Pi}(n) = 1$  ma *Repeat* non si verifica, allora **deve** verificarsi *Newblock*. Pertanto, risulta:

$$Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \overline{\text{Repeat}} \wedge \overline{\text{Newblock}}] = 0.$$

Sia  $q(n)$  il num. di query che  $A$  fa all'oracolo ed  $r_i$  l'identificatore della  $i$ -esima query.

Se *Repeat* non si verifica, allora  $r_1, \dots, r_{q(n)}$  sono *tutti distinti*.

Se  $r \notin \{r_1, \dots, r_{q(n)}\}$  allora *Newblock* si verifica! Altrimenti,  $r = r_j$  per qualche  $j$  ed i blocchi

$$r \parallel \ell \parallel 1 \parallel m_1, \quad r \parallel \ell \parallel 2 \parallel m_2, \quad \dots \text{ di } (m, t = \langle r, t_1, \dots \rangle)$$

possono essere stati autenticati soltanto durante la  $j$ -esima query all'oracolo.

Sia  $m_j$  il messaggio usato da  $A$  per la sua  $j$ -esima query all'oracolo ed  $\ell_j$  la sua lunghezza.

Occorre considerare due casi:

- 1  $\ell \neq \ell_j \Rightarrow$  i blocchi autenticati durante la  $j$ -esima query hanno *tutti*  $\ell_j \neq \ell$  in seconda posizione. Quindi, nessun blocco di  $m$  può essere stato autenticato durante la  $j$ -esima query. Per esempio,  $r||\ell||1||m_1$  **non** è stato autenticato durante la  $j$ -esima query  $\Rightarrow$  *Newblock* si verifica
- 2  $\ell = \ell_j \Rightarrow$  poiché  $m \neq m_j$  (per ipotesi  $\text{Mac-forge}_{A,\Pi}(n) = 1$ )  $\exists$  un indice  $i$  per cui  $m_i \neq m_j^{(i)}$ , dove  $m_j^{(i)}$  denota l' $i$ -esimo blocco di  $m_j$ . Il blocco  $r||\ell||i||m_i$  **non** è stato autenticato durante la  $j$ -esima query  $\Rightarrow$  *Newblock* si verifica.



**Claim 4.10**  $Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \text{Newblock}]$  é trascurabile.

Il claim poggia sulla sicurezza di Mac'.

Costruiamo, infatti, un avversario  $A'$  ppt che attacca lo schema  $\Pi'$  (utilizzato da  $\Pi$ )

$A'$  ha successo nel produrre un tag valido su un messaggio non autenticato in precedenza con

$$Pr[\text{Mac-forge}_{A',\Pi'}(n) = 1] \geq Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \text{Newblock}]$$

Poiché  $\Pi'$  é per ipotesi sicuro, allora la prima probabilità deve essere trascurabile



$Pr[\text{Mac-forge}_{A,\Pi}(n) = 1 \wedge \text{Newblock}]$  deve essere altrettanto.

$A'$  esegue  $A$  come subroutine e risponde alle query di  $A$  per tag di specifici messaggi come segue

- sceglie  $r \leftarrow \{0, 1\}^{n/4}$
- divide il messaggio in modo opportuno
- effettua le query necessarie al proprio oracolo  $Mac'(\cdot)$

$A'$  tiene traccia di tutte le query che ha rivolto a  $Mac'(\cdot)$

Quando  $A$  dá in output  $(m, t = \langle r, t_1, \dots \rangle)$ ,  $A'$  controlla se l'evento *Newblock* si verifica.

Se ciò accade,  $A'$  trova il primo blocco  $r||\ell||i||m_i$  che non é stato usato in una query per  $Mac'(\cdot)$  e dá in output  $(r||\ell||i||m_i, t_i)$ .

Altrimenti, non dá nulla in output.

La "vista" che  $A$  ha quando eseguito come subroutine da  $A'$  é distribuita esattamente come in  $\text{Mac-forge}_{A,\Pi}(n)$



Le probabilità di  $\text{Mac-forge}_{A,\Pi}(n)$  e *Newblock* **non** cambiano.



Se  $\text{Mac-forge}_{A,\Pi}(n) = 1$ , il tag restituito da  $A$  é corretto su ciascun blocco del messaggio  $m$ , ed in particolare sul blocco che non é stato mai usato in una query da  $A'$  verso l'oracolo  $\text{Mac}'$  e che  $A'$  dá in output.

Quindi, ogni volta che  $\text{Mac-forge}_{A,\Pi}(n) = 1$  e *Newblock* occorre,  $\text{Mac-forge}_{A',\Pi'}(n) = 1 \Rightarrow$  il Claim 4.10 vale.

Pertanto, mettendo assieme i risultati provati,  $\Pr[\text{Mac-forge}_{A,\Pi}(n) = 1] \leq \text{negl}(n)$ .

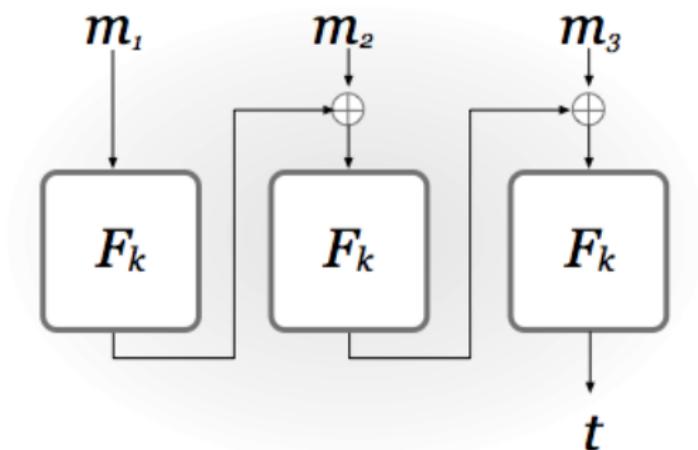
Abbiamo visto che usando funzioni pseudocasuali che prendono input di lunghezza fissa  $n$ , é possibile costruire codici di autenticazione per messaggi di lunghezza arbitraria

La costruzione **non** é efficiente.

$$m \quad : \quad |m| = dn \quad \Rightarrow \quad \text{tag } t \quad : \quad |t| = 4dn$$

Inoltre,  $F_k$  viene valutata  $4d$  volte.

Possiamo far meglio usando CBC-MAC, un codice per l'autenticazione di messaggi standardizzato.



**CONSTRUCTION 4.11**

Let  $F$  be a pseudorandom function, and fix a length function  $\ell > 0$ . The basic CBC-MAC construction is as follows:

- **Mac**: on input a key  $k \in \{0, 1\}^n$  and a message  $m$  of length  $\ell(n) \cdot n$ , do the following (we set  $\ell = \ell(n)$  in what follows):
  1. Parse  $m$  as  $m = m_1, \dots, m_\ell$  where each  $m_i$  is of length  $n$ .
  2. Set  $t_0 := 0^n$ . Then, for  $i = 1$  to  $\ell$ :  
Set  $t_i := F_k(t_{i-1} \oplus m_i)$ .

Output  $t_\ell$  as the tag.

- **Vrfy**: on input a key  $k \in \{0, 1\}^n$ , a message  $m$ , and a tag  $t$ , do: If  $m$  is not of length  $\ell(n) \cdot n$  then output 0. Otherwise, output 1 if and only if  $t \stackrel{?}{=} \text{Mac}_k(m)$ .

**Teorema 4.12** Sia  $\ell(\cdot)$  un polinomio. Se  $F$  é una PRF, allora la Costruzione 4.11 é un Mac sicuro per messaggi di lunghezza  $\ell(n) \cdot n$ .

Attenzione: sicura **soltanto** quando la lunghezza dei messaggi é fissata e convenuta in precedenza.

Vantaggi: fornisce un Mac a lunghezza fissa molto piú efficiente rispetto alla costruzione precedente.

Confronto con la modalitá di cifratura CBC

- 1 CBC-MAC **non** usa un vettore di inizializzazione IV. Questo aspetto é cruciale. CBC-MAC che usa un IV casuale **non** é sicuro.
- 2 La modalitá di cifratura CBC dá in output **tutti** i valori  $t_i$  intermedi. CBC-MAC solo  $t_\ell$ . Se viene modificata per dare in output tutti gli  $\{t_i\}_i$ , CBC-MAC **non** é piú sicuro.

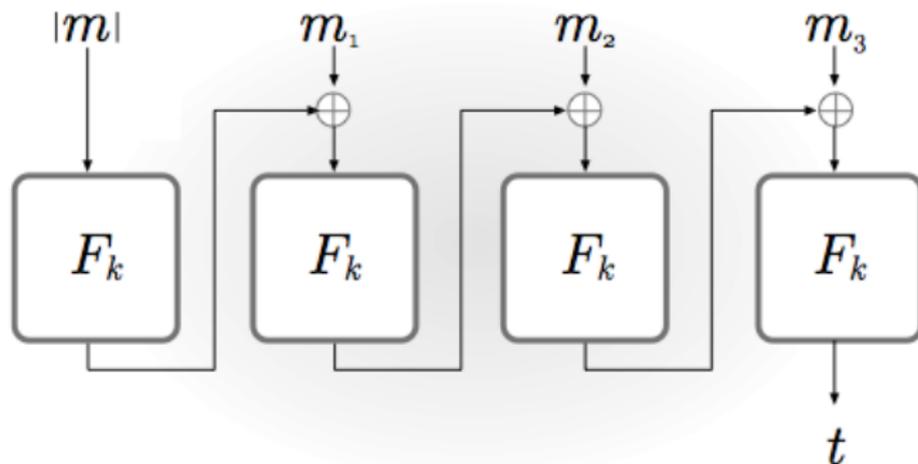
Conclusione: nuovamente, anche piccole ed apparentemente innocue modifiche ad una primitiva crittografica possono renderla insicura.

Messaggi di lunghezza arbitraria: esistono due modi per estendere la costruzione

- 1 Anteporre ad  $m$  la sua lunghezza  $|m|$ , codificata come una stringa di  $n$  bit e calcolare il CBC-MAC di base
- 2 L'algoritmo di generazione della chiave sceglie due chiavi e poi
  - calcola il CBC-MAC di base usando  $k_1$ , sia esso  $t$
  - dá in output  $t^* := F_{k_2}(t)$ .

Le due chiavi possono essere anche generate al volo usando una PRF, i.e.,  $k_1 = F_k(1)$  e  $k_2 = F_k(2)$ , memorizzando una sola chiave  $k$ . Ma é ancora costoso.

Solitamente la prima strategia viene preferita.



Nuove soluzioni efficienti, adottate da diversi standard.

- Schema generale
  - 1 funzione  $\epsilon(n)$ -difference universal  $h_{k_h}(m)$  e PRF  $F_{k_F}(\cdot)$
  - 2 MAC  $t = \langle r, s \rangle$  (randomizzato,  $r$  uniforme), dove
$$s = h_{k_h}(m) + F_{k_F}(r)$$
- GMAC e Poly1305: block cipher +  $h_{k_h}(m)$  basata su polinomi
- Usano campi finiti diversi

La PRF viene valutata una sola volta

Riduzioni di sicurezza *piú strette* rispetto a CBC-MAC

Al solito  $n$  indica il parametro di sicurezza.  $K_n, M_n, T_n$  insiemi di chiavi, messaggi e tag ( $T_n$  deve essere un gruppo per ogni  $n$ , ci torneremo ...)

**Definizione 4.14.** Una funzione con chiave  $h : K_n \times M_n \rightarrow T_n$  è  $\epsilon(n)$ -difference universal se, per tutti gli  $n$ , per ogni  $m, m' \in M_n$ , ed ogni  $\Delta \in T_n$ , risulta

$$\Pr[h_k(m) - h_k(m') = \Delta] \leq \epsilon(n),$$

per  $k$  scelta uniformemente in  $K_n$ .

Nota che deve essere  $\epsilon(n) \geq 1/|T_n|$ . Inoltre  $h$  deve essere efficientemente calcolabile.

## Schema Generico

- 1 Gen: su input  $1^n$ , sceglie uniformemente  $k_h \in K_n$  e  $k_F \in \{0, 1\}^n$
- 2 Mac: su input  $(k_h, k_F)$  ed  $m \in M_n$ , sceglie  $r$  uniformemente in  $\{0, 1\}^n$  e calcola il tag

$$t = \langle r, h_{k_h}(m) + F_{k_F}(r) \rangle .$$

- 3 Vrfy: su input  $(k_h, k_F)$ ,  $m \in M_n$  e  $t = \langle r, s \rangle$ , dá in output 1 se

$$s = h_{k_h}(m) + F_{k_F}(r).$$

Tecnicalità che ci servono (ci torneremo ...).

Un campo finito  $F_q$  contenente  $q$  elementi esiste per ogni potenza di un numero primo.

Un polinomio non nullo di grado  $\ell$  su un campo finito ha al più  $\ell$  zeri.

Sia  $F$  un campo finito. Sia  $F^{<\ell}$  l'insieme dei vettori con meno di  $\ell$  elementi. Sia  $M = F^{<\ell}$ . Sia  $m = m_1, \dots, m_{t-1} \in M$ , con  $t \leq \ell$ , e sia  $m_t$  una codifica della lunghezza di  $m$ . Sia  $m(X)$  il polinomio

$$m(X) = m_1 \cdot X^t + m_2 \cdot X^{t-1} + \dots + m_t \cdot X.$$

La funzione  $h : F \times F^{<\ell} \rightarrow F$  con chiave definita come

$$h_k(m) = m(k)$$

è una funzione  $\ell/|F|$ -difference universal.

**Prova.** Siano  $m, m' \in F^{<\ell}$  e  $\Delta \in F$ . Consideriamo il polinomio

$$P(X) = m(X) - m'(X) - \Delta.$$

$P$  é un polinomio non nullo di grado al piú  $\ell$ . Risulta

$$\Pr_{k \in F}[h_k(m) - h_k(m') = \Delta] = \Pr_{k \in F}[P(k) = 0] \leq \ell/|F|,$$

poiché  $P$  ha al piú  $\ell$  radici (zeri).

La costruzione GMAC usa un cifrario a blocchi per instanziare la PRF  $F(\cdot, \cdot)$  e la funzione  $\ell/|F|$ -difference universal basata sui polinomi definita sul campo  $F_{2^{128}}$ , contenente  $2^{128}$  elementi.

La costruzione Poly1305 é realizzata in modo simile, ma usa il campo  $F_p$  definito per il primo  $p = 2^{130} - 5$ .

Il tag finale é calcolato come

$$t = \langle r, [h_{k_h}(m) + F_{k_F}(r) \bmod 2^{128}] \rangle .$$