Crittografia Moderna
A.A. 2024-25

Segretezza Perfetta: nozioni

Segretezza perfetta

- Ci occuperemo di schemi di cifratura perfettamente sicuri

  confidenzialità / riservatezza

- Avversari di potere computazionale illimitato
- Useremo definizioni matematiche precise e prove
- Nessuna assunzione!

  Inerenti limitazioni → giustificano la necessità di assunzioni

- Useremo esperimenti ed algoritmi randomizzati

Come generare randomness?

- In principio, per pochi bit random, lanciando una moneta!
- Tecniche moderne procedono in due fasi
  - una discreta quantità di dati ad "alta entropia" viene raccolta

    non necessariamente uniformi

    misura impredicibilità
  - i dati vengono elaborati per produrre una sequenza di bit quasi indipendenti e distribuiti quasi uniformemente

- Per la prima fase, occorre una sorgente di dati imprevedibili

Sorgenti

- Input esterni
  - ritardo dei dati in transito in rete
  - tempi di accesso all'hard disk
  - misure sulla pressione dei tasti
  - misure sul movimento del mouse
  - …

dati "lontani" dalla distribuzione uniforme

- Fenomeni fisici
  - misure del rumore nei dispositivi elettronici
  - misure del decadimento radioattivo

Estrazione di bit casuali

- Il processo di “livellamento” dei dati ad alta entropia è non banale. Ne mostriamo uno.
  - Supponiamo di disporre di una sequenza dei risultati del lancio di una moneta difettosa, rappresentata in binario
  - La moneta dà:

    $\boxed{\begin{array}{l} T=1 \text{ con probabilità } p \\ C=0 \text{ con probabilità } (1-p) \end{array}}$

  - I lanci sono indipendenti l’uno dall’altro
  - Consideriamo i bit a coppie
  - Se vediamo:

Se vediamo:	00	01	10	11
scriviamo	/	0	1	/

Estrazione di bit casuali

10	01	11	01	10	00	11	10	10	01
1	0		0	1			1	1	0

- la sequenza binaria calcolata è distribuita uniformemente
- Perchè?
  - La probabilità che una coppia produca 0 è: (1-p) p
  - La probabilità che una coppia produca 1 è: p (1-p)
- La generazione e l'uso di random bit è cruciale per la sicurezza dei protocolli crittografici
  - Generatori di numeri casuali devono essere progettati per uso crittografico.
  - "General-purpouse" non sono idonei.

Definizioni

- Schema di cifratura: (Gen, Enc, Dec), M tale che |M|>1

  spazio dei messaggi

- Gen: algoritmo *probabilistico* di generazione delle chiavi, restituisce $k \in K$ in accordo ad una distribuzione di probabilità

  spazio delle chiavi

- Enc: algoritmo *probabilistico* di cifratura, prende in input $m \in M, k \in K$ e restituisce un cifrato $c \in C$

  spazio dei cifrati

- Dec: algoritmo deterministico di decifratura, prende in input $c \in C, k \in K$ e restituisce $m \in M$

Notazioni

- $k \longleftarrow Gen()$
- $c \longleftarrow Enc_k(m)$

output di alg. probabilistici

- $c := Enc_k(m)$
- $m := Dec_k(c)$

output di alg. deterministici

- Useremo la notazione $x \longleftarrow S$ anche per denotare l'estrazione di un elemento da un insieme S in accordo alla distribuzione uniforme

Condizione di correttezza perfetta

- Per ogni $k \in K$, per ogni $m \in M$, e qualsiasi $c \leftarrow Enc_k(m)$, risulta

$$Dec_k(c) = m, \text{ con probabilità } 1$$

Nei teoremi e nelle definizioni che seguono faremo riferimento alle distribuzioni di probabilità sugli insiemi K, M e C.

La distribuzione su K dipende da Gen()

assumeremo uniforme

La variabile casuale **K** denoterà l'output di Gen()

Variabili casuali

- La variabile casuale **M** denoterà il messaggio "in chiaro"

La Pr[**M** = m] è la probabilità che m debba essere cifrato

$\downarrow$

Non dipende dallo schema di cifratura. Riflette la frequenza dei differenti messaggi che le parti si inviano, così come l'incertezza che l'avversario Adv ha su di essi.

e.g., Adv può sapere che i messaggi sono solo due e che:

Pr[**M** = "attacca oggi"] = 0.7

Pr[**M** = "attacca domani"] = 0.3

- Le variabili casuali **K** ed **M** si assumono **indipendenti**

Variabili casuali

- Fissato uno schema di cifratura ed una distribuzione su M si determina una distribuzione sullo spazio dei cifrati C, ottenuta
  - da k generata tramite Gen()
  - scegliendo m in accordo alla distribuzione su M
  - calcolando $c \leftarrow Enc_k(m)$

scelte casuali

- La variabile casuale **C** denoterà l'output del processo di cifratura
- Nota che $\mathbf{C} = Enc_{\mathbf{K}}(\mathbf{M})$

variabili aleatorie

Esempio

- Consideriamo lo Shift Cipher

  $K = \{0,\ldots,25\}$, $\Pr[\mathbf{K}=k] = 1/26$, per ogni $k \in K$

Sia data su M la distribuzione: $\Pr[\mathbf{M} = a] = 0.7$ e $\Pr[\mathbf{M} = z] = 0.3$

Qual è la probabilità che il cifrato sia B?

Ci sono solo due possibilità: $m = a$ e $k = 1$, oppure $m = z$ e $k = 2$

Poiché le variabili casuali **M** e **K** sono indipendenti, risultano:

$$\Pr[\mathbf{M} = a \wedge \mathbf{K} = 1] = \Pr[\mathbf{M} = a] \bullet \Pr[\mathbf{K} = 1] = 0.7 \bullet 1/26$$

$$\Pr[\mathbf{M} = z \wedge \mathbf{K} = 2] = \Pr[\mathbf{M} = z] \bullet \Pr[\mathbf{K} = 2] = 0.3 \bullet 1/26$$

Pertanto,

$$\Pr[\mathbf{C} = B] = \Pr[\mathbf{M} = a \wedge \mathbf{K} = 1] + \Pr[\mathbf{M} = z \wedge \mathbf{K} = 2]$$

$$= 0.7 \bullet 1/26 + 0.3 \bullet 1/26 = 1/26$$

Esempio

- Consideriamo lo Shift Cipher

Qual è invece la probabilità che il messaggio in chiaro sia “a”, dato il cifrato B?

Usando il teorema di Bayes, risulta:

$$\Pr[\,\mathbf{M} = a \mid \mathbf{C} = B] = \frac{\Pr[\mathbf{C} = B \mid \mathbf{M} = a] \bullet \Pr[\,\mathbf{M} = a]}{\Pr[\mathbf{C} = B\,]}$$

$$= \frac{\Pr[\mathbf{C} = B \mid \mathbf{M} = a] \bullet 0.7}{1/26}$$

Ma $\Pr[\mathbf{C} = B \mid \mathbf{M} = a] = 1/26$ perché, se $\mathbf{M} = a$, l’unico modo per ottenere $\mathbf{C} = B$ è che $\mathbf{K} = 1$ (che accade con prob. 1/26)

Pertanto, $\Pr[\,\mathbf{M} = a \mid \mathbf{C} = B] = \frac{1/26 \bullet 0.7}{1/26} = 0.7$

Ulteriore esempio

- Ancora lo Shift Cipher

- Sia data su M la distribuzione:

$\Pr[\mathbf{M} = \text{kim}] = 0.5,$ $\qquad \Pr[\mathbf{M} = \text{ann}] = 0.2$ e $\qquad \Pr[\mathbf{M} = \text{boo}] = 0.3$

Qual è la probabilità che $\mathbf{C} = \text{DQQ}$?

Ci sono soltanto due possibilità: $\mathbf{M} = \text{ann}$ e $\mathbf{K} = 3$, oppure $\mathbf{M} = \text{boo}$ e $\mathbf{K} = 2$

Pertanto,

$$\begin{aligned}\Pr[\,\mathbf{C} = \text{DQQ}] &= \Pr[\mathbf{M} = \text{ann} \wedge \mathbf{K} = 3] + \Pr[\mathbf{M} = \text{boo} \wedge \mathbf{K} = 2] \\ &= 0.2 \bullet 1/26 + 0.3 \bullet 1/26 = 0.5 \bullet 1/26 \\ &= 1/2 \bullet 1/26 \\ &= 1/52\end{aligned}$$

Ulteriore esempio

Similmente, usando il teorema di Bayes, risulta:

$$\Pr[\,\mathbf{M} = \text{ann} \mid \mathbf{C} = \text{DQQ}] = \frac{\Pr[\mathbf{C} = \text{DQQ} \mid \mathbf{M} = \text{ann}] \bullet \Pr[\,\mathbf{M} = \text{ann}]}{\Pr[\mathbf{C} = \text{DQQ}]}$$

$$= \frac{1/26 \bullet 0.2}{1/52} = 0.4$$

essendo $\Pr[\mathbf{C} = \text{DQQ} \mid \mathbf{M} = \text{ann}] = 1/26$, dato che solo $\mathbf{K} = 3$ produce $\mathbf{C} = \text{DQQ}$.

Segretezza perfetta

- Scenario: Adv conosce
  - la distribuzione di probabilità su M
  - lo schema di cifratura
  - **non** conosce la chiave k
  - può ascoltare/vedere la comunicazione

Per ottenere segretezza perfetta, l'osservazione del cifrato non dovrebbe fornire alcuna informazione aggiuntiva ad Adv sul messaggio in chiaro. Cioè:

Il cifrato non dovrebbe rivelare ad Adv nulla in più al di là di ciò che già sa

Segretezza perfetta

Definizione 2.3. Uno schema di cifratura (Gen, Enc, Dec) con spazio dei messaggi M è perfettamente segreto se

- per ogni distribuzione di probabilità su M
- per ogni messaggio $m \in M$
- e per ogni cifrato $c \in C$ per cui risulta Pr[**C**=c]>0, si ha:

$$\Pr[\,\mathbf{M} = m \mid \mathbf{C} = c\,] = \Pr[\,\mathbf{M} = m\,]$$

Definizione equivalente

- Possiamo fornire una definizione equivalente richiedendo che la distribuzione di probabilità dei cifrati **non dipenda** dal messaggio in chiaro

- Formalmente, per ogni $m, m' \in M$ e per ogni $c \in C$

$$\Pr[\mathrm{Enc}_{\mathbf{K}}(m) = c] = \Pr[\mathrm{Enc}_{\mathbf{K}}(m') = c]$$

 dove le prob sono calcolate sulle possibili scelte
  - della chiave k
  - dei bit casuali utilizzati da $\mathrm{Enc}_{\mathbf{k}}()$

Definizione equivalente

- La precedente condizione implica
  - è impossibile distinguere una cifratura di m da una di m', poiché le distribuzioni sui cifrati sono le stesse
  - un cifrato non rivela alcuna informazione sul messaggio in chiaro

Lemma 2.4 Uno schema di cifratura (Gen, Enc, Dec) con spazio dei messaggi M è perfettamente segreto **se e solo se** per ogni $m, m' \in M$ ed ogni $c \in C$ risulta:

$$\Pr[\mathrm{Enc}_{\mathbf{K}}(m) = c] = \Pr[\mathrm{Enc}_{\mathbf{K}}(m') = c]$$

Dimostrazione

- Proviamo che, se la condizione vale, **allora** lo schema è perfettamente segreto. Fissiamo una distribuzione su M e sia c tale che $\Pr[\mathbf{C} = c] > 0$.
  - se $\Pr[\mathbf{M} = m] = 0$, allora banalmente

$$\Pr[\,\mathbf{M} = m \mid \mathbf{C} = c] = \frac{\Pr[\,\mathbf{C} = c \mid \mathbf{M} = m] \cdot \Pr[\mathbf{M} = m]}{\Pr[\mathbf{C} = c]}$$

0

$$= 0$$

$$\Rightarrow \quad \Pr[\,\mathbf{M} = m \mid \mathbf{C} = c] = \Pr[\mathbf{M} = m]$$

  - se $\Pr[\mathbf{M} = m] > 0$, invece, notiamo che

$$\Pr[\mathbf{C} = c \mid \mathbf{M} = m] = \Pr[\mathrm{Enc}_{\mathbf{K}}(\mathbf{M}) = c \mid \mathbf{M} = m] = \Pr[\mathrm{Enc}_{\mathbf{K}}(m) = c]$$

Sia $\delta_c = \Pr[\mathrm{Enc}_{\mathbf{K}}(m) = c]$. Se la condizione del lemma vale, per ogni m’

Dimostrazione

risulta:

$$\Pr[\mathrm{Enc}_{\mathbf{K}}(m') = c] = \Pr[\mathbf{C} = c \mid \mathbf{M} = m'] = \delta_c$$

Applicando il teorema di Bayes,

$$\Pr[\, \mathbf{M} = m \mid \mathbf{C} = c] = \frac{\Pr[\, \mathbf{C} = c \mid \mathbf{M} = m] \bullet \Pr[\mathbf{M} = m]}{\Pr[\, \mathbf{C} = c]}$$

$$= \frac{\Pr[\, \mathbf{C} = c \mid \mathbf{M} = m] \bullet \Pr[\mathbf{M} = m]}{\sum_{m' \in M} \Pr[C = c \mid M = m'] \cdot \Pr[M = m']}$$

δ_c

$$= \frac{\delta_c \bullet \Pr[\mathbf{M} = m]}{\delta_c \bullet \sum_{m' \in M} \Pr[M = m']}$$

1

$$= \Pr[\mathbf{M} = m]$$

Pertanto lo schema è perfettamente segreto

Dimostrazione

- Viceversa, sia lo schema perfettamente segreto. Sia M **uniformemente distribuita**. Per il teorema di Bayes, per ogni c tale che Pr[**C** = c] > 0 e per ogni m, risulta

$$\Pr[\,\mathbf{M} = m \mid \mathbf{C} = c] \bullet \Pr[\,\mathbf{C} = c] = \Pr[\,\mathbf{C} = c \mid \mathbf{M} = m] \bullet \Pr[\mathbf{M} = m]$$

- La segretezza perfetta $\Pr[\,\mathbf{M} = m \mid \mathbf{C} = c] = \Pr[\mathbf{M} = m]$ implica

$$\cancel{\Pr[\,\mathbf{M} = m]} \bullet \Pr[\,\mathbf{C} = c] = \Pr[\,\mathbf{C} = c \mid \mathbf{M} = m] \bullet \cancel{\Pr[\mathbf{M} = m]}$$

ovvero

$$\Pr[\,\mathbf{C} = c] = \Pr[\,\mathbf{C} = c \mid \mathbf{M} = m]$$

- Pertanto, per ogni m, m' e per ogni c tale che Pr[**C** = c] > 0

$$\Pr[\mathrm{Enc}_{\mathbf{K}}(m) = c] = \Pr[\mathbf{C} = c] = \Pr[\mathrm{Enc}_{\mathbf{K}}(m') = c].$$

Perfetta indistinguibilità

- Possiamo dare ancora un'altra definizione di segretezza perfetta equivalente alle precedenti
  - basata su un esperimento che coinvolge Adv e un Challenger C
- Informalmente
  - Adv sceglie due messaggi in chiaro, m ed m'
  - C sceglie uniformemente a caso uno dei due e lo cifra usando una chiave k casuale
  - Il cifrato ottenuto c viene dato ad Adv
  - Adv dà in output una sua ipotesi su quale dei due è stato cifrato
  - Adv ha successo se la sua ipotesi è corretta

Perfetta indistinguibilità

- Uno schema di cifratura è perfettamente indistinguibile se nessun Adv può avere successo con probabilità > ½

- Osservazioni:
  - un Adv che ipotizza a caso ha successo con probabilità ½ , richiediamo che Adv non disponga di una strategia migliore
  - non poniamo alcun limite al potere computazionale di Adv

Esperimento $PrivK^{eav}_{A,\Pi}$

Sia Π = (Gen, Enc, Dec) uno schema di cifratura con spazio dei messaggi M.
Sia A un Adv (i.e., un algoritmo che mantiene uno stato)

$PrivK^{eav}_{A,\Pi}$

- A dà in output una coppia di messaggi m_0, m_1
- C genera k tramite Gen() e sceglie uniformemente a caso il bit b in {0,1}
- C calcola $c \leftarrow Enc_K(m_b)$, detto **cifrato di sfida**, e lo passa ad A
- A dà in output un bit b'
- L'output dell'esperimento è 1 se b' = b, 0 altrimenti. Se l'output è 1, A ha successo.

Definizione

Definizione 2.5. Uno schema di cifratura (Gen, Enc, Dec) con spazio dei messaggi M è perfettamente indistinguibile se, per ogni A, risulta:

$$\Pr[\ \underbrace{\mathrm{Pr}ivK^{eav}_{A,\Pi}}_{\text{Risultato dell'esperimento}} = 1] = 1/2$$

È possibile dimostrare che vale il seguente

Lemma 2.6. Uno schema di cifratura Π è perfettamente segreto **se e solo se** è perfettamente indistinguibile.

Conclusione

- Disponiamo di tre definizioni di segretezza perfetta equivalenti!

per ogni distribuzione su M,
per ogni $m \in M$ e per ogni c
tale che $\Pr[\mathbf{C} = c] > 0$
$\Pr[\,\mathbf{M} = m \mid \mathbf{C} = c] = \Pr[\,\mathbf{M} = m]$

per ogni $m, m' \in M$ ed ogni $c \in C$

$\Pr[\mathrm{Enc}_{\mathbf{K}}(m) = c] = \Pr[\mathrm{Enc}_{\mathbf{K}}(m') = c]$

Per ogni Adv A,

$\Pr[\,\mathrm{Priv}K^{eav}_{A,\Pi} = 1] = 1/2$