Schemi computazionalmente sicuri

Paolo D'Arco
pdarco@unisa.it

Università di Salerno

Elementi di Crittografia

Contenuti

Schemi di cifratura computazionalmente sicuri

Definizione 3.7. Uno schema di cifratura a chiave privata è una tripla $\Pi = (Gen, Enc, Dec)$ di algoritmi PPT tale che

1. $k \leftarrow Gen(1^n)$, algoritmo probabilistico di generazione della chiave k
   - dove la chiave $k \in K$ è tale che $|k| \geq n$
2. $c \leftarrow Enc_k(m)$, algoritmo probabilistico di cifratura
   - dove il messaggio $m \in \{0,1\}^*$, la chiave $k \in K$ e il cifrato $c \in \{0,1\}^*$
3. $m := Dec_k(c)$, algoritmo deterministico di decifratura
   - dove il cifrato $c \in \{0,1\}^*$, la chiave $k \in K$ e il messaggio $m \in \{0,1\}^*$
   - $Dec(c)$ restituisce $\perp$ in caso di errore

Correttezza. Per ogni n, per ogni k restituito da $Gen(1^n)$ e per ogni $m \in \{0,1\}^*$, risulta

$$Dec_k(Enc_k(m)) = m$$

Schemi di cifratura computazionalmente sicuri

Note e osservazioni:

- se lo spazio dei messaggi è $\{0,1\}^{\ell(n)}$, allora Π è uno schema di cifratura a chiave privata a lunghezza fissa, per messaggi di lunghezza $\ell(n)$.
- solitamente $Gen(1^n)$ restituisce stringhe di n bit scelte uniformemente a caso
- la definizione è senza stato (occasionalmente considereremo schemi con stato)

Definizione di sicurezza di base:

1. **Modello delle minacce.** Adv è PPT. Osserva un singolo cifrato ottenuto usando una certa chiave. Può applicare qualsiasi strategia d'attacco.
2. **Garanzie di sicurezza.** Adv non deve essere in grado di acquisire *alcuna informazione aggiuntiva* sul messaggio in chiaro m a partire dal cifrato c.

Sicurezza semantica ed indistinguibilità

La nozione di *sicurezza semantica* formalizza ciò.

$\Downarrow$

È difficile da maneggiare

$\Downarrow$

Esiste una definizione *equivalente* più semplice

$\Downarrow$

È la nozione di *Indistinguibilità*

Indistinguibilità

Nel contesto della segretezza perfetta abbiamo considerato l'esperimento $PrivK^{eav}_{A,\Pi}$

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

1. $m_0, m_1 \leftarrow A$ (sceglie due messaggi)
2. il challenger calcola $c \leftarrow Enc_k(m_b)$, dove $b \leftarrow \{0,1\}$ e $k \leftarrow Gen(1^n)$
3. A riceve c e dà in output $b' \in \{0,1\}$
4. Se $b = b'$, l'output dell'esperimento è 1 (A vince); altrimenti, 0.

Lo schema Π è sicuro se A vince con probabilità $1/2$, i.e., non c'è strategia migliore per indovinare che *scegliendo a caso*

Indistinguibilità

Nel caso computazionale:

- A è PPT
- A può vincere con probabilità *trascurabilmente* migliore di $1/2$
- L'esperimento dipende da n, il parametro di sicurezza

$PrivK^{eav}_{A,\Pi}(n)$

1. A ottiene 1^n e dà in output m_0, m_1 tali che $|m_0| = |m_1|$
2. il challenger calcola $c \leftarrow Enc_k(m_b)$, dove $b \leftarrow \{0,1\}$ e $k \leftarrow Gen(1^n)$
3. $A(1^n)$ riceve c e dà in output $b' \in \{0,1\}$
4. Se $b = b'$, l'output dell'esperimento è 1 ($A(1^n)$ vince); altrimenti, 0.

Indistinguibilità

Definizione 3.8. Uno schema di cifratura a chiave privata $\Pi = (Gen, Enc, Dec)$ ha *cifrature indistinguibili* in presenza di un avversario che ascolta (eavesdropper) o è EAV-sicuro se, per ogni Adv A PPT, esiste una funzione trascurabile *negl* tale che:

$$Pr[PrivK^{eav}_{A,\Pi}(n) = 1] \leq \frac{1}{2} + negl(n),$$

dove la probabilità è calcolata su

- randomness usata da A
- randomness usata nell'esperimento
  - scelta della chiave
  - scelta del bit b
  - random bit usati da $Enc_k(\cdot)$

Indistinguibilità

Nota: qualsiasi schema di cifratura *perfettamente segreto* ha cifrature indistinguibili in presenza di un eavesdropper.

Faremo vedere che esistono schemi con "chiavi più corte"

Esiste una formulazione equivalente: l'idea di fondo è che *ogni Adv PPT si comporta allo stesso modo sia che veda una cifratura di* m_0 *che di* m_1

Definendo $PrivK^{eav}_{A,\Pi}(n,b)$ con $b \in \{0,1\}$ e l'output di A con $out_A(PrivK^{eav}_{A,\Pi}(n,b))$, diamo la seguente

Definizione 3.9. $\Pi = (Gen, Enc, Dec)$ è EAV-sicuro se, per ogni Adv A PPT, esiste una funzione trascurabile *negl* tale che, per tutti gli n si ha:

$$|Pr[out_A(PrivK^{eav}_{A,\Pi}(n,0)) = 1] - Pr[out_A(PrivK^{eav}_{A,\Pi}(n,1)) = 1]| \leq negl(n).$$

Indistinguibilità

Nota: nella definizione *non* richiediamo ad uno schema di nascondere la lunghezza del messaggio da cifrare. Nei casi in cui questa informazione è importante occorre porre rimedio (e.g., estendendo i messaggi ad una lunghezza fissa)

Indistinguibilità e intuizione

Il concetto di indistinguibilità ricorda il concetto di travestimento. E le proprietà che desideriamo sono all'incirca le seguenti:

- Due persone travestite sono indistinguibili: una modella bellissima e l'uomo più brutto del mondo, travestiti, non possono essere distinti
- Ma le due persone devono essere della stessa altezza: altrimenti è immediato distinguere un gigante da un nano

Sicurezza Semantica

Dovrebbe essere impraticabile per un Adv acquisire *alcuna informazione aggiuntiva* sul messaggio in chiaro dal cifrato.

Cominciamo con due nozioni più deboli:

1. il cifrato non rivela alcuna informazione sui singoli bit del messaggio in chiaro
2. il cifrato non aiuta un Adv PPT nel *calcolo* di qualsiasi funzione del messaggio in chiaro

Proveremo che la nozione di indistinguibilità implica 1. e 2.

Notazione: solitamente *Gen* genera chiavi distribuite uniformemente a caso. Quando assumeremo ciò, useremo per semplicità $\Pi = (Enc, Dec)$

Teorema 3.10. Sia $\Pi = (Enc, Dec)$ uno schema di cifratura a chiave privata per messaggi di lunghezza ℓ EAV-sicuro. Allora, per ogni Adv A PPT ed ogni $i \in \{1, \ldots, \ell\}$, esiste una funzione trascurabile *negl* tale che:

$$Pr[A(1^n, Enc_k(m)) = m^i] \leq \frac{1}{2} + negl(n),$$

dove la probabilità è calcolata su

- scelta uniforme di $m \in \{0,1\}^\ell$
- scelta uniforme di $k \in \{0,1\}^n$
- random bit usati da A
- random bit usati da $Enc_k(\cdot)$

Sicurezza Semantica

Dim. Idea: se fosse possibile, con probabilità non trascurabile, calcolare l'i-esimo bit m^i

$$\Downarrow$$

sarebbe anche possibile, con probabilità non trascurabile, distinguere m_0 da m_1 che differiscono nell'i-esimo bit.

Useremo una dimostrazione *per riduzione* (... ci torneremo su a breve).

Fissiamo un Adv arbitrario A PPT ed $i \in \{1, \ldots, \ell\}$.

Vogliamo usare A (e la sua capacità di calcolare m^i con prob. non trascurabile) per costruire un Adv A' che *usa* A per distinguere con prob. non trascurabile m_0 da m_1 che differiscono nell'i-esimo bit.

Sicurezza Semantica

Siano:

$I_0 \subset \{0,1\}^\ell$ insieme di stringhe con i-esimo bit uguale a 0

$I_1 \subset \{0,1\}^\ell$ insieme di stringhe con i-esimo bit uguale a 1

Essendo $|I_0| = |I_1| = 2^{\ell-1}$ ed m scelto in modo uniforme, risulta

$$\begin{aligned} Pr[A(1^n, Enc_k(m)) = m^i] &= Pr[m \in I_0] \cdot Pr[A(1^n, Enc_k(m)) = 0 | m \in I_0] \\ &\quad + Pr[m \in I_1] \cdot Pr[A(1^n, Enc_k(m)) = 1 | m \in I_1] \end{aligned}$$

$$= \frac{1}{2} \cdot Pr[A(1^n, Enc_k(m_0)) = 0] + \frac{1}{2} \cdot Pr[A(1^n, Enc_k(m_1)) = 1]$$

Costruiamo A' come segue:

Adv A'

1. sceglie uniformemente $m_0 \in I_0$ e $m_1 \in I_1$ e li passa al challenger
2. dopo aver ricevuto c dal challenger, invoca $A(1^n, c)$
3. Se A dà in output 0, allora dà in output $b' = 0$; altrimenti, $b' = 1$.

A' gioca nell'esperimento $PrivK^{eav}_{A',\Pi}(n)$ e usa A, che calcola m^i, come subroutine.

A' è PPT poichè A è PPT e fa poco più che invocare A.

Dalla definizione di $PrivK^{eav}_{A',\Pi}(n)$, A' ha successo se e solo se A restituisce b dopo aver ricevuto $Enc_k(m_b)$. Pertanto, risulta:

Sicurezza Semantica

$$
\begin{aligned}
Pr[PrivK^{eav}_{A',\Pi}(n) = 1] &= Pr[A(1^n, Enc_k(m_b)) = b] \\
&\quad \text{(dato che } b \text{ viene scelto uniform. nell'esperimento)} \\
&= \frac{1}{2} \cdot Pr[A(1^n, Enc_k(m_0)) = 0] + \frac{1}{2} \cdot Pr[A(1^n, Enc_k(m_1)) = 1] \\
&= Pr[A(1^n, Enc_k(m)) = m^i]
\end{aligned}
$$

Poichè abbiamo assunto che $\Pi = (Enc, Dec)$ è EAV-sicuro, esiste una funzione trascurabile *negl* tale che

$$Pr[PrivK^{eav}_{A',\Pi}(n) = 1] \leq 1/2 + negl(n)$$

$$\Downarrow$$

$$Pr[A(1^n, Enc_k(m)) = m^i] \leq 1/2 + negl(n).$$

Sicurezza Semantica

Circa il punto 2., mostreremo che:

un Adv A che calcola $f(m)$ con una certa probabilità quando riceve $Enc_k(m)$

$$\Downarrow$$

un Adv A' che calcola $f(m)$ con la *stessa* probabilità, *senza* conoscere $Enc_k(m)$.

Sicurezza Semantica

Teorema 3.11. Sia $\Pi = (Enc, Dec)$ uno schema di cifratura a chiave privata per messaggi di lunghezza ℓ EAV-sicuro. Allora, per ogni Adv A PPT, esiste un Adv A' PPT tale che, per ogni distr. di prob. $\mathcal{D}$ su $\{0,1\}^\ell$ ed ogni $f : \{0,1\}^\ell \to \{0,1\}$ esiste una funzione trascurabile *negl* tale che:

$$|Pr[A(1^n, Enc_k(m)) = f(m)] - Pr[A'(1^n) = f(m)]| \leq negl(n),$$

dove la prima probabilità è calcolata su

- scelta di m in accordo a $\mathcal{D}$ e uniforme di $k \in \{0,1\}^n$
- random bit usati da A
- random bit usati da $Enc_k(\cdot)$

e la seconda su

- scelta di m in accordo a $\mathcal{D}$ e uniforme dei random bit usati da A'

Dim. (Sketch) Poichè Π è EAV-sicuro, per ogni distribuzione $\mathcal{D}$, nessun Adv PPT può distinguere tra $Enc_k(m)$ ed $Enc_k(1^\ell)$

Consideriamo la probabilità con cui A calcola $f(m)$ data $Enc_k(m)$.

A dovrebbe calcolare $f(m)$ data $Enc_k(1^\ell)$ con $\approx$ la stessa probabilità.

Altrimenti A potrebbe essere usato per distinguere tra $Enc_k(m)$ ed $Enc_k(1^\ell)$.

Distinguisher

1. sceglie m in accordo a $\mathcal{D}$ e passa al challenger $m_0 = m$ e $m_1 = 1^\ell$
2. dopo aver ricevuto c dal challenger, invoca $A(1^n, c)$
3. Se A dà in output $f(m)$, allora dà in output $b' = 0$; altrimenti, $b' = 1$.

Sicurezza Semantica

Se A dà in output $f(m)$ con una probabilità *significativamente* migliore nel caso in cui riceve $Enc_k(m)$ rispetto a quando riceve $Enc_k(1^\ell)$, allora l'algoritmo Distinguisher viola la Definizione 3.8.

Detto ciò, possiamo costruire A' come segue

Adv $A'(1^n)$

1. sceglie uniformemente $k \in \{0,1\}^n$
2. invoca $A(1^n, Enc_k(1^\ell))$
3. dà in output qualsiasi cosa A dà in output

A dà in output $f(m)$ quando viene eseguito come subroutine di A' con $\approx$ la stessa probabilità di quando riceve $Enc_k(m)$. Pertanto A' ha i requisiti richiesti dal teorema.

Sicurezza Semantica

La garanzia offerta dalla sicurezza semantica è più forte della garanzia offerta dal Teorema 3.11

- la lunghezza dei messaggi dipende dal parametro di sicurezza n
- la distribuzione di probabilità su M è arbitraria
  - unica condizione: sia efficientemente campionabile (samplable). Cioè, esiste $Samp(1^n)$, algoritmo PPT, che dà in output messaggi in accordo alla distribuzione di probabilità definita su M
- inoltre, la definizione tiene anche conto di eventuali informazioni aggiuntive $h(m)$ sul messaggio m che l'avversario può ottenere attraverso altri mezzi

Sicurezza Semantica

Definizione 3.12. Uno schema di cifratura a chiave privata $\Pi = (Gen, Enc, Dec)$ è semanticamente sicuro in presenza di un eavesdropper se, per ogni Adv A PPT, esiste un Adv A' PPT tale che, per qualsiasi $Samp(1^n)$ PPT e per ogni coppia di funzioni f ed h, calcolabili in tempo polinomiale, esiste una funzione trascurabile $negl$ per cui si ha:

$$|Pr[A(1^n, Enc_k(m), h(m)) = f(m)] - Pr[A'(1^n, |m|, h(m)) = f(m)]| \leq negl(n),$$

dove la prima probabilità è calcolata su

- scelta uniforme di $k \in \{0,1\}^n$
- random bit usati da $Samp(1^n)$
- random bit usati da A
- random bit usati da $Enc_k(\cdot)$

e la seconda su

- random bit usati da $Samp(1^n)$ e random bit usati da A'

Sicurezza Semantica

Teorema 3.13. $\Pi = (Enc, Dec)$ ha cifrature indistinguibili in presenza di un eavesdropper *se e solo se* è semanticamente sicuro in presenza di un eavesdropper.

$$\Downarrow$$

Possiamo usare la definizione più semplice di indistinguibilità come definizione di lavoro!