Capitolo 1

Introduzione

G_0 = **foreach iH** $\leq$ **qH do** OH(x: bitstring):= **return**(hash(x)); |

Ogen():=r $\xleftarrow{R}$ seed;pk $\leftarrow$ pkgen(r);sk $\leftarrow$ skgen(r);**return**(pk);
(**foreach iS** $\leq$ **qS do** OS(m: bitstring):= **return**(invf(sk, hash(m)));) |

find u $\leq$ **qS suchthat** (**defined**(m[u]) $\wedge$ $m' = m[u]$)) **then end**
else event forge

Nel capitolo intoduttivo si introdurranno il modello formale Vs quello computazionale, il concetto di sicurezza perfetta e quella computazionale per poi riprendere quest'ultima nel capitolo 1 parlare di sicurezza TEORICA vs COPMUTAZINOALE parlare delle assunzioni....e delle funzioni one-way.

Da quando la crittografia ha cessato di essere un'arte per assurgere allo stato di scienza[DH76], sono nati almeno due modi profondamente diversi fra loro di vedere la crittografia.
In uno di questi, il modello *formale*, le operazioni crittografiche sono rappresentate da espressioni simboliche, formali. Nell'altro, il modello *computazionale* le operazioni crittografiche sono viste come funzioni su stringhe di bit; le operazioni hanno una semantica probabilistica. Il primo modello è stato teorizzato in []... Il secondo modello trova le basi in lavori di altrettanto illustri studiosi []

In [AR07] gli autori cercano per la primavolta di porre le basi per iniziare a collegare questi due modelli.

Parlare del teorema che unisce i due modelli parlare del concetto di attaccante e di dimostrazioni che sfruttano il fatto che l'attaccante non può fare delle cose

parlare del concetto di funzione one-way.

Capitolo 2

Il Modello Computazionale

2.1 L'Avversario

Il tipico avversario con cui si ha a che fare quando si studiano cifrari o protocolli crittografici nel modello computazionale, è un avversario con risorse di calcolo *limitate*. Limitate nel senso che si sceglie di porre un limite alla potenza di calcolo dell'avversario. Questo significa che non avremo a che fare con un avversario che ha una potenza computazionale infinita o un tempo illimitato a disposizione.
Sebbene siano stati ideati cifrari sicuri anche rispetto ad avversari non limitati[1], questi hanno alcuni difetti come per esempio il fatto che la chiave debba essere lunga quanto il messaggio o che sia utilizzabile una sola volta. Per rappresentare in modo formale un avversario con risorse di calcolo limitate, lo si può pensare come un algoritmo appartenente ad una particolare classe di complessità computazionale[2].
Una linea di pensiero che accomuna ogni campo dell'informatica, considera efficienti gli algoritmi che terminano in un numero di passi polinomiale nella lunghezza dell'input, e inefficienti quelli che hanno una complessità computazionale maggiore (e.g. esponenziale). La scelta di porre un limite alle risorse di calcolo dell'avversario è dettata dal buon senso. È ragionevole infatti pensare che l'attaccante non sia infinitamente potente; è altrettanto ragionevole pensare che un attaccante non sia disposto ad impiegare un tempo *eccessivo* per violare un schema crittografico.
Se un attaccante, infatti, per violare un cifrario, utilizzasse un algoritmo che impega l'età dell'universo (stimata dai fisici intorno ai 13 miliardi di anni), sicuramente non riuscirebbe a sfruttare questa *vittoria*. È logico, quindi,

[1] one-time pad ne è un esempio lampante.
[2] un avversario è, alla fine dei conti, una macchina di Turing che esegue un algoritmo.

pensare che gli avversari vogliano essere *efficienti*. Può sembrare, quindi, naturale immaginare gli avversari come degli algoritmi che terminano in un numero polinomiale di passi rispetto alla lunghezza dell'input.
Come si può notare, non si fa alcuna assunzione particolare sul comportamento dell'avversario. Le uniche cose che sappiamo sono che:

- l'avversario non conosce la chiave, ma conosce l'algoritmo di cifratura utilizzato e i parametri di sicurezza, come per esempio la lunghezza della chiave[3].

- l'avversario vuole essere efficiente, ovvero polinomiale.

Non si fanno ipotesi sull'algoritmo che questo andrà ad eseguire. Per esempio dato un messaggio cifrato $c = E_k(m)$, non ci aspettiamo che l'avversario non decida di utilizzare la stringa c' tale che: $c' = D_{k'}(E_k(m))$ con $k \neq k'$. Ovvero, sarebbe sbagliato supporre che l'avversario non cerchi di decifrare un messaggio mediante una chiave diversa da quella utilizzata per cifrarlo. Nel modello computazionale i messaggi sono stringhe di bit e l'avversario può effettuare qualsiasi operazione su queste. Questa visione è, a differenza di quella che si ha nel modello formale, sicuramente molto più realistica[Dwo06].

Non bisogna però dimenticare che un avversario può sempre *indovinare* il segreto che cerchiamo di nascondere, o che cifriamo. Per esempio: se il segreto che si cerca di nascondere ha una lunghezza di n bit, l'avversario può sempre lanciare una moneta n volte e associare, via via, la testa della moneta al valore 1 e la croce al valore 0. La probabilità che l'avversario ottenga una stringa uguale al segreto è ovviamente di $\frac{1}{2^n}$. Questa probabilità tende a 0 in modo esponenziale al crescere della lunghezza del segreto, ma per valori finiti di n questa probabilità non sarà mai 0. È quindi più realistico cercare di rappresentare l'avversario come un algoritmo che, oltre a terminare in tempo polinomiale, ha anche la possibilità di effettuare scelte random. La classe dei problemi risolti da questo tipo di algoritmi è indicata con la sigla *BPP* (i.e. *Bounded-Probability Polynomial Time*).
Un modo più formale di vedere questo tipo di algoritmi è il seguente: si suppone che la macchina di Turing che esegue l'algoritmo, oltre a ricevere l'input, diciamo x, riceva anche un input ausiliario r. Questa stringa di bit r, rappresenta una possibile sequenza di lanci di moneta. Quando la macchina dovrà effettuare una scelta random, non dovrà far altro che prendere

[3]Un famoso principio della crittografia afferma che l'algoritmo di cifratura non deve essere segreto e deve poter cadere nelle mani del nemico senza inconvenienti.

il successivo bit dalla stringa r, e prendere una decisione in base ad esso (è, in effetti, come se avesse preso una decisione lanciando una moneta). Ecco quindi che il nostro tipico avversario si configura come un algoritmo polinomiale probabilistico. È inoltre giustificato cercare di rendere sicuri[4] gli schemi crittografici rispetto, principalmente, a questo tipo di avversario. Con questa scelta si cerca di rispettare il più possibile un famoso principio di Kerckhoffs[5] che afferma:

> *Un cifrario deve essere, se non matematicamente, almeno praticamente indecifrabile.*

Non è quindi necessario dimostrare che un particolare schema crittografico sia inviolabile, ma basta dimostrare che:

- in tempi ragionevoli lo si può violare solo con scarsissima probabilità.
- lo si può violare con alta probabilità, ma solo in tempi non ragionevoli.

Sappiamo che il concetto di *tempo ragionevole* è catturato dalla classe degli algoritmi polinomiali probabilistici. Vediamo ora di catturare il concetto di *scarsa probabilità*.

2.2 Funzioni Trascurabili e non ...

In crittografia i concetti di *scarsa probabilità* e di evento *raro* vengono formalizzati attraverso la nozione di funzione trascurabile.

Definizione 2.1 *Funzione Trascurabile (negligible). Sia $\mu : \mathbb{N} \to \mathbb{R}^+$ una funzione. Si dice che μ è trascurabile se e solo se per ogni polinomio p, esiste $C \in \mathbb{N}$ tale che $\forall n > C$: $\mu(n) < \frac{1}{p(n)}$.*

Una funzione trascurabile, quindi, è una funzione che tende a 0 più velocemente dell'inverso di qualsiasi polinomio. Un'altra definizione utile è la seguente:

Definizione 2.2 *Funzione Distinguibile (noticeable). Sia $\mu : \mathbb{N} \to \mathbb{R}^+$ una funzione. Si dice che μ è distinguibile se e solo se esiste un polinomio p, tale per cui esiste $C \in \mathbb{N}$ tale che $\forall n > C$: $\mu(n) > \frac{1}{p(n)}$.*

[4] In qualsiasi modo si possa intendere il concetto di sicurezza. Vedremo che in seguito si daranno delle definizioni rigorose di questo concetto.

[5] Auguste Kerckhoffs (19 Gennaio 1835 – 9 Agosto 1903) fu un linguista olandese e un famoso crittografo.

Per esempio la funzione $n \to 2^{-\sqrt{n}}$ è una funzione trascurabile, mentre la funzione $n \to \frac{1}{n^2}$ non lo è. Ovviamente esistono anche funzioni che non sono né trascurabili né distinguibili. Per esempio, la seguente funzione definita per casi: $f(n) = \begin{cases} 1, & \text{se } n \text{ è pari} \\ 0, & \text{se } n \text{ è dispari} \end{cases}$ non è né trascurabile né distinguibile. Questo perché le definizioni precedenti, pur essendo molto legate, non sono l'una la negazione dell'altra.

Se sappiamo che, in un esperimento, un evento avviene con una probabiltà trascurabile, quest'evento si verificherà con una probabilità trascurabile anche se l'esperimento viene ripetuto molte volte (ma sempre un numero polinomiale di volte), e quindi per la legge dei grandi numeri, con una frequenza anch'essa trascurabile[6]. Le funzioni trascurabili, infatti, godono di due particolari proprietà di chiusura, enunciate nella seguente:

Proposizione 2.1 *Siano μ_1, μ_2 due funzioni trascurabili e sia p un polinomio. Se $\mu_3 = \mu_1 + \mu_2$, e $\mu_4 = p \cdot \mu_1$, allora μ_3, μ_4 sono funzioni trascurabili.*

Se quindi, in un esperimento, un evento avviene solo con probabilità trascurabile, ci aspettiamo che, anche se ripetiamo l'esperimento un numero polinomiale di volte, questa probabilità rimanga comunque trascurabile.
Per esempio: supponiamo di avere un dado truccato in modo che la probabilità di ottenere 1 sia trascurabile. Allora se lanciamo il dado un numero polinomiale di volte, la probabilità che esca 1 rimane comunque trascurabile.
È ora importantissimo notare che: **gli eventi che avvengono con una probabilità trascurabile possono essere ignorati per fini pratici**. In [KL07], infatti leggiamo:

> *Events that occur with negligible probability are so unlikely to occur that can be ignored for all practical purposes. Therefore, a break of a cryptographic scheme that occurs with negligible probability is not significant.*

Potrebbe sembrare pericoloso utilizzare degli schemi crittografici che ammettono di essere violati con probabilità trascurabile, ma questa possibilità è cosi remota, che se ci preoccupassimo, per amor di coerenza, dovremmo anche essere ragionevolmente sicuri di fare sei all'enalotto giocando una schedina semplice. Finora abbiamo parlato sempre di funzioni che prendono in input un argomento non meglio specificato. Al crescere di questo parametro, le funzioni si comportano in modo diverso, a seconda che siano trascurabili, oppure no. Ma cosa rappresenta nella realtà questo input? Di solito, questo

[6] In modo informale, la legge debole dei grandi numeri afferma che: per un numero grande di prove, la frequenza approssima la probabilità di un evento.

valore rappresenta un generico parametro di sicurezza, indipendente dal segreto. È comune immaginarlo come la lunghezza in bit delle chiavi.
D'ora in poi con affermazioni del tipo «ł'algoritmo è polinomiale, o esponenziale», si intenderanno algoritmi polinomiali o esponenziali nella lunghezza (in bit) del parametro di sicurezza (indicato con n). Si utilizzerà questa assunzione anche quando si faranno affermazioni su funzioni trascurabili o meno. Quelle funzioni saranno trascurabili o meno nel parametro n. Tutte le definizioni di sicurezza che vengono date nel modello computazionale e che utilizzano le probabilità trascurabili, sono di tipo *asintotico*. Un template di definizione di sicurezza è il seguente [KL07]:

> *A scheme is secure if for every probablistic polynomial-time adversary* ***A*** *[...], the probability that* ***A*** *succeds in this attack [...] is negligible*

Essendo questo schema di definizione asintotico (nel parametro di sicurezza n), è ovvio che non considera valori piccoli di n. Quindi se si dimostra che un particolare schema crittografico è sicuro secondo una definizione di questo tipo, può benissimo capitare che per valori piccoli di n lo schema sia violabile con alta probabilità e in tempi ragionevoli.

2.3 Indistinguibilità Computazionale

Se due oggetti, sebbene profondamente diversi fra loro, non possono essere distinti, allora sono da un certo punto di vista equivalenti. Nel caso della crittografia computazionale, due oggetti sono computazionalmente equivalenti se nessun algoritmo efficiente li può distinguere. Possiamo immaginare che un algoritmo riesca a distinguere due oggetti, se quando gli si da in input il primo, lui da in output una costante c, mentre se gli si fornisce come input il secondo da in output una costante c' e ovviamente $c \neq c'$. La definizione tipica di indistinguibilità computazionale è data prendendo come oggetti da distinguere alcune particolari distribuzioni statistiche detti *ensembles*.

Definizione 2.3 *Ensemble. Sia I un insieme numerabile infinito.*
$X = \{X_i\}_{i \in I}$ è un ensemble su I se e solo se è una sequenza di variabili statistiche.

Un *ensemble* è quindi una sequenza infinita di distribuzioni di probabilità[7]. Tipicamente le variabili dell'ensemble sono stringhe di lunghezza i. X_i è quindi una distribuzione di probabilità su stringhe di lunghezza i.

[7] siccome si parla di distribuzioni su stringhe di bit con lunghezza finita, in crittografia computazionale si considerano ensemble che sono una sequenza inifinta di distribuzioni finite di stringhe di bit.

Ora supponiamo di avere due ensemble X e Y. Intuitivamente queste distribuzioni sono indistinguibili se nessun algoritmo (efficiente) può accettare infiniti elementi di X_n (per esempio stampando 1 su input preso da X_n) e scartare infiniti elementi di Y_n (per esempio stampare 0 su input preso da Y_n). È importante notare che sarebbe facile distinguere due *singole* distribuzioni usando un approccio esaustivo, ecco perché si considerano sequenze infinite di ditribuzioni finite. In poche parole questi ensemble sono indistinguibili se ogni algoritmo (efficiente) accetta $x \in X_n$ se e solo se accetta $y \in Y_n$. Ovviamente il *se e solo se* non può e non deve essere inteso in senso *classico*, ma deve essere inteso in senso statistico. Poiché in crittografia si è soliti indicare con U_m una variabile uniformemente distribuita sull'insieme delle stringhe di lunghezza m, chiameremo $U = \{U_n\}_{n\in\mathbb{N}}$ l'ensemble uniforme. Dopo questa breve introduzione all'indistinguibilità siamo pronti per dare una definizione rigorosa:

Definizione 2.4 *Indistinguibilità computazionale. Due ensemble $X = \{X_n\}$, $Y = \{Y_n\}$ sono computazionalmente indistinguibili se e solo se per ogni algoritmo $D \in BPP$ (detto distinguitore) esiste μ trascurabile tale che:*

$$|Pr[D(1^n, X_n) = 1] - Pr[D(1^n, Y_n) = 1]| \leq \mu(n).$$

Nella definizione precedente: $Pr[D(1^n, X_n) = 1]$ è la probabilità che, scegliendo x secondo la distribuzione X_n e fornendo questo valore al distinguitore insieme al valore 1^n, il distinguitore stampi 1. Il fatto che al distinguitore si fornisca anche il valore del parametro di sicurezza in base unaria, serve ad esser sicuri che in ogni caso il distinguitore impieghi un tempo polinomiale nel parametro di sicurezza. Infatti, il distinguitore quando si troverà a dover leggere il primo parametro, necessariamente impiegherà un tempo polinomiale nel parametro di sicurezza, visto che questo è stato fornito in base unaria[8].

La definizione di indistinguibilità computazionale cattura quindi il seguente concetto: se due ensemble sono computazionalmente indistinguibili, allora la probabilità che un distinguitore riesca a discernere i valori provenienti da un insieme rispetto all'altro è trascurabile; di conseguenza agli occhi del distinguitore gli ensemble non sono differenti e quindi sono per lui equivalenti (o meglio computazionalmente equivalenti o ancora, indistinguibili in tempo polinomiale). Non è raro, nell'ambito scientifico in particolare, basarsi sul concetto generale di indistinguibilità al fine di creare nuove classi di equivalenza di oggetti.

[8] ignoreremo, d'ora in poi, questo cavillo formale.

> *The concept of efficient computation leads naturally to a new kind of equivalence between objects: Objects are considered to be computationally equivalent if they cannot be differentiated by any efficient procedure.* We note that considering indistinguishable objects as equivalent is one of the basics paradigms of both science and real-life situations. Hence, we believe that the notion of computational indistinguishability is a very natural one.[Gol00]

2.4 Pseudocasualità e Generatori Pseudocasuali

Argomento centrale di questa sezione è il concetto di *pseudocasualità* (pseudorandomness), applicato a stringhe di bit di lunghezza finita. Parlare di pseudocasualità applicata ad una *singola* stringa, ha poco senso quanto ne ha poco parlare di singola stringa casuale (random). Il concetto di casualità (come quello di pseudocasualità) si applica, infatti, a distribuzioni di oggetti (stringhe di bit nel nostro caso) e non a singoli oggetti.
La nozione di casualità è fortemente legata a quella di distribuzione uniforme. Un insieme di oggetti è caratterizzato da una distribuzione uniforme se la probabilità è equamente distribuita su tutti gli oggetti. Quindi *l'estrazione* di un elemento è del tutto casuale, perché non ci sono elementi più probabili di altri.

Il concetto di pseudorandomness è un caso particolare di indistinuguibilità, infatti una distribuzione è *pseudorandom* se nessuna procedura efficiente, può distinguerla dalla distribuzione uniforme.

Definizione 2.5 *Pseudorandomness. L'ensemble* $X = \{X_n\}_{n\in\mathbb{N}}$ *è detto pseudorandom se e solo se* $\exists l : \mathbb{N} \to \mathbb{N}$ *tale che:* X *è computazionalmente indistinguibile da* $U = \{U_{l(n)}\}_{n\in\mathbb{N}}$.

Data questa definizione, possiamo finalmente definire formalmente cosa sia un generatore pseudorandom.

Definizione 2.6 *Generatore Pseudorandom. Sia* $l : \mathbb{N} \to \mathbb{N}$ *un polinomio detto fattore d'espansione. Sia* G *un algoritmo polinomiale deterministico tale che:* $\forall s \in \{0,1\}^n \; G(s) \in \{0,1\}^{l(n)}$. *Allora* G *è un generatore pseudorandom se e solo se valgono le seguenti condizioni:*

- *Espansione:* $\forall n : l(n) > n$
- *Pseudocasualità:* $\forall D \in BPP, \exists \mu$ *trascurabile tale che*

$$|Pr[D(r) = 1] - Pr[D(G(s)) = 1]$$

con $r \in U_{l(n)}$ *e* $s \in U_n$

Quindi: se data una stringa di bit $s \in U_n$, nessun distinguitore efficiente riesce a distinguere (con una probabilità non trascurabile) $G(s)$ da una stringa $r \in U_{l(n)}$, allora G è un generatore pseudorandom. Il suo output, infatti, non è distinguibile dalla distribuzione effettivamente uniforme.

È importante però notare, che la stringa in output di un generatore pseudorandom è fortemente differente da una stringa effettivamente random. Per rendere più chiara questa distinzione procederemo con un importante esempio. Supponiamo di avere un generatore pseudorandom G con fattore d'espansione $l(n) = 2n$. L'insieme $A = \{0,1\}^{2n}$ ha, ovviamente, una cardinalità pari a 2^{2n}. Fissando quindi una certa stringa $x \in A$, questa ha una probabilità di esser scelta in maniera random pari a: $\frac{1}{|A|} = \frac{1}{2^{2n}}$.

Ragioniamo adesso sull'output del generatore G. Questo prende un input appartenente al dominio: $B = \{0,1\}^n$. Anche considerando il caso *migliore* di un generatore iniettivo[9], il codominio di G avrà una cardinalità pari a quella del dominio ovvero 2^n. La maggior parte degli elementi dell'insieme A non ricadrà nell'output di G; questo a causa dell'abissale differenza di cardinalità fra gli insiemi $G(B)$ e A. Quindi la probabilità che una stringa scelta in maniera uniforme dall'insieme A ricada nel codominio di G è di $\frac{2^n}{2^{2n}}$, cioè 2^{-n}. In teoria, quindi, è facile immaginare un distinguitore D che riesca a discernere l'output di G dalla distribuzione uniforme con probabilità non trascurabile. Supponiamo che D prenda in input $y \in A$. Tutto ciò che D deve fare è ricercare in modo esaustivo un $w \in B$ tale che $G(w) = y$. Se $y \in G(B)$ allora D se ne accorgerà con probabilità 1, mentre se y è stato scelto in maniera uniforme dall'insieme A, D stamperà 1 con probabilità 2^{-n}.

[9] una generica funzione f è iniettiva se e solo se $\forall x_1, x_2 : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.

Quindi abbiamo che:

$$|Pr[D(r) = 1] - Pr[D(G(s)) = 1]| \geq 1 - 2^{-n} \text{ con } r \xleftarrow{R} A \text{ e } s \xleftarrow{R} B^{10}.$$

Il membro a destra della disequazione è una funzione distinguibile. Sembrerebbe quindi che G non sia un generatore pseudorandom. C'è un' importante constatazione da fare però. Il distinguitore D non è efficiente! Infatti impiega un tempo esponenziale nel parametro n, e non polinomiale. La distribuzione generata da G dunque, è si ben lontana dall'essere uniforme, ma questo non è importante dal momento che nessun distinguitore che viaggia in tempo polinomiale può accorgersene.

Nella pratica lo scopo di G è prendere in input un *seed* random, e da quello generare una variabile pseudocasuale molto più lunga. Si inutisce da questo la grandissima importanza che hanno i generatori pseudorandom in crittografia. Per esempio il seed potrebbe corrispondere alla chiave di un cifrario, mentre l'output di G di lunghezza k potrebbe essere il valore con cui viene fatto lo XOR del messaggio (anch'esso di lunghezza k); otteniamo così una versione del one-time pad basato su una chiave più corta del messaggio. Siccome una stringa pseudorandom appare, ad un distinguitore efficiente D, come una stringa random, D non ottiene un vantaggio sensibile nel passaggio dal vero one-time pad al one-time pad che usa una chiave pseudorandom. In generale i generatori pseudorandom sono molto utili in crittografia per creare schemi crittografici simmetrici.

2.5 Dimostrazioni Basate su Games

In questo paragrafo si cercherà di spiegare cosa sono nell'ambito della crittografia i *games*[11] e come sono strutturate la maggior parte delle dimostrazioni che utilizzano sequenze di games.
Si possono trovare approfondimenti riguardo a questi concetti nel lavoro di Shoup [Sho]. Quella basata sul concetto di game[12] è una tecnica molto utilizzata per provare la sicurezza di primitive crittografiche o di protocolli crittografici. Questi games sono giocati da un' ipotetica entità maligna, l'attaccante, e da un'ipotetica parte benigna di solito chiamato sfidante[13]. È difficile dare una definizione formale di game, infatti il concetto di game cambia, sebbene in maniera non sensibile, da situazione a situazione da ambiente ad ambiente e da dimostrazione a dimostrazione (sia che queste siano fatte a

[10] con la notazione $s \xleftarrow{R} O$, si intende la scelta dell'elemento $s \in O$ in maniera random.
[11] che intenderemo letteralmente come giochi
[12] game hopping tecnique
[13] perchè *sfida* l'attaccante a vincere questo gioco.

mano sia che queste siano automatiche e quindi dipendenti dal framework in cui vengono costruite). Intuitivamente però, i game possono essere immaginati come un insieme di azioni, modellate in una particolare algebra di processi, che servono a specificare il comportamento dei partecipanti al gioco, ovvero le entità che partecipano come *principals* al protocollo crittografico. Il lettore troverà utile pensarli, almeno nell'ambito di questa tesi, come insiemi di processi che, fra le altre cose, forniscono un'interfaccia all'attaccante attraverso degli *oracoli* che possono resitituire dei valori all'attaccante. Questi oracoli, prima di ritornare il valore, possono effettuare calcoli, dichiarare ed utilizzare variabili che non saranno visibili all'esterno.
Si deve pensare agli oracoli come delle scatole nere inaccessibili dall'esterno. Questi oracoli una volta interrogati forniscono una risposta, e questo è il massimo livello di interazione che dall'esterno si può avere con queste entità.

Il principio che seguono le dimostrazioni di sicurezza basate su sequenze di game è quello di partire da un game iniziale G_0 che modella il protocollo crittografico reale. In G_0 esiste la probabilità non nulla che un evento *negativo* possa accadere (immaginiamolo come una sorta di vittoria da parte dell'attaccante). Ora, in generale, si procede effettuando delle modifiche al game G_i ottenendo un game G_{i+1} tale che G_i e G_{i+1} sono computazionalmente indistinguibili. Le modifiche che si effettuano fra un game e un altro devono introdurre delle differenze computazionalmente irrilevanti. Queste modifiche possono essere viste come regole di riscrittura delle distribuzioni di probabilità delle variabili in gioco nei game. Se per esempio in un game G_i un processo ha a che fare con un variabile random, e noi nel game G_{i+1} la sostituiamo con una variabile che invece è pseudorandom, non introduciamo modifiche computazionalmente rilevanti, e quindi il passaggio è lecito visto che i due game non sono distinguibili se non con probabilità trascurabile. Alla fine si deve arrivare ad un game G_f in cui *l'evento* non può accadere. Se, quindi, nel game finale l'attaccante non ha possibilità di vincere, e il game finale è computazionalmente indistinguibile dal penultimo, e il penultimo lo è dal terzultimo e il terzultimo lo è dal quartultimo e cosi via fino al primo, allora il game finale è computazionalmente indistinguibile dal primo[14]. Adesso, quindi, se nel game iniziale esiste la possibilità che un evento avvenga, e nel game finale no, possiamo dare un limite superiore alla probabilità che l'evento avvenga nel game iniziale. Questo limite è la somma di tutte le probabilità[15] con cui un attaccante riesce a distinguere un game

[14] Ricordiamo infatti che la somma di due probabilità trascurabili rimane trascurabile.

[15] ovviamente tutte trascurabili, visto che i due game sono computazionalmente indistinguibili.

dal successivo.

È importante dire che, anche nel modello formale si possono utilizzare le sequenze di game, la differenza è che due game successivi non sono computazionalmente indistinguibili ma sono perfettamente indistinguibili. In particolare quello che si vuole sottolineare è che, se in una sequenza di game, G_b e G_a sono l'uno il successore dell'altro, allora i due game devono appartenere ad una stessa classe di equivalenza indotta dalla particolare relazione di equivalenza che il modello in cui si sta costruendo la sequenza sfrutta. Nel modello formale questa relazione sarà l'equivalenza osservazionale mentre in quello computazionale sarà l'indistinguibilità computazionale.

Capitolo 3

CryptoVerif

3.1 Introduzione al Tool

CryptoVerif è un *dimostratore* automatico che lavora direttamente nel modello computazionale. È un tool molto recente e in continuo sviluppo.
CryptoVerif è stato scritto da Bruno Blanchet[1] in ML. Si possono trovare più informazioni nella home page di Blanchet[2]. Questo tool è utilizzato per dimostrare proprietà di segretezza e autenticazione Queste prove si basano sulla tecnica delle sequenze di game. Il tool è liberamente scaricabile[3] sotto la licenza CeCill[4]. CryptVerif è stato già utilizzato per dimostrare la correttezza di alcuni protocolli crittografici, come per esempio: FDH [BP06], Kerberos [BJST08].

3.2 Un Esempio: FDH

Full Domain Hash è uno schema di firma che segue il paradigma *hash-and-sign*[5]. Quella che a breve seguirà è una sequenza di game costruita dal CryptoVerif che dà un limite alla probabilità che un attaccante riesca a forgiare una firma valida per un messaggio. L'input che si fornisce al CryptoVerif consta di un game iniziale chiamato G_0 in cui si modella la sicurezza dello

[1] Ricercatore al LIENS (Computer Science Laboratory of Ecole Normale Supérieure)
[2] `http://www.di.ens.fr/~blanchet/index-eng.html`
[3] `http://www.cryptoverif.ens.fr/cryptoverif.html`
[4] http://www.cecill.info/licences/
[5] Questo paradigma vuole che, dato un messaggio m se ne ritorni la firma di $hash(m)$ e non la firma di m, dove la funzione hash è può essere istanziata con qualsiasi funzione hash collision resistant. Si ottiene così una firma di lunghezza fissa e non dipendente dalla lunghezza del messaggio

schema e delle equivalenze necessarie a CryptoVerif per effettuare le modifiche ai game. La descrizione dello schema di firma *FDH* attraverso il seguente game:

$$G0 = \textbf{foreach } iH \leq qH \textbf{ do } OH(x : bitstring) := \textbf{return}(hash(x)) \mid$$
$$Ogen() := r \xleftarrow{R} seed; pk \leftarrow pkgen(r); sk \leftarrow skgen(r); \textbf{return}(pk) \mid$$
$$(\textbf{foreach } iS \leq qS \textbf{ do } OS(m : bitstring) := \textbf{return}(invf(sk, hash(m))) \mid$$
$$OT(m : bitstring, s : D) := \textbf{if} f(pk, s) = hash(m) \textbf{then find } u \leq qS$$
$$\textbf{suchthat } (\textbf{defined}(m[u]) \wedge m = m[u]) \textbf{ then end}$$
$$\textbf{else event } forge$$

Possiamo vedere come in questo game si forniscano qH copie dell' oracolo OH che ritorna l'hash della stringa che gli si fornisce in input: x. Abbiamo poi il processo $Ogen()$ che ritorna al contesto una chiave pubblica dopo aver creato, partendo da un seme random, una coppia (chiave privata, chiave pubblica). Abbiamo poi qS oracoli OS che si occupano di fornire una firma del messaggio che gli viene dato in input. Infine abbiamo un singolo oracolo OT che si occupa di verificare se la firma s è valida per il messaggio $m^{'}$. In particolare se $hash(m^{'}) \neq f(pk, s)$ allora la firma non è valida per il messaggio. Se invece $hash(m^{'}) = f(pk, s)$ l'oracolo si occupa di verificare se, per il messaggio $m^{'}$, è stata mai rilasciata una firma dall'oracolo OS, in caso affermativo il processo termina, altrimenti significa che il contesto, ovvero l'attaccante, è stato in grado di forgiare una firma valida per il messaggio, e quindi è avvenuto l'evento *forge*. Questo game modella la sicurezza dello schema di firma FDH. Notiamo come l'attaccante non sia modellato esplicitamente, infatti non possiamo fare nessuna assunzione su questo. Possiamo immaginare un attaccante come un altro processo messo in parallelo con questo game. In gergo l'attaccante viene anche detto anche *contesto*. Per poter trasformare un game in un altro CryptoVerif ha bisogno di alcune equivalenze da poter utilizzare. Le definizioni in CryptoVerif vengono fornite attraverso delle equivalenze che possono essere viste come regole di riscrittura delle distribuzioni di probabilità delle variabili in gioco. Queste regole di riscrittura di un generico elemento L in un generico elemento R possono valere incondizionatamente, oppure possono valere a meno di una certa probabilità. Nel secondo caso la riscrittura di un termine L nell'equivalente R comporta l'introduzione di una differenza non nulla. Per esempio la definizione di *one-wayness* viene fornita a CryptoVerif mediante questa equivalenza:

```
                                R
     foreach ik ≤ nk do r ← seed ; (Opk() := return(pkgen(r))
                                    R
          | foreach if ≤ nf do x ← D; (Oy() := return(f(pkgen(r), x))
               | foreach i1 ≤ n1 do Oeq(x : D) := return(x = x)
               | Ox() := return(x)))

                                R
≈
pow foreach ik ≤ nk do r ← seed ; (Opk() := return(pkgen (r))
                                    R
          | foreach if ≤ nf do x ← D; (Oy() := return(f (pkgen (r), x))
               | foreach i1 ≤ n1 do Oeq(x : D) :=
                    if defined(k) then return(x = x) else return(false)
               | Ox() := k ← mark; return(x)))
```

Questa equivalenza cattura e definisce il concetto di funzione one-way. Supponiamo infatti che la funzione f non sia one-way, esisterà dunque un avversario efficiente A che riuscirà ad invertire la funzione. Questo avversario sarà dunque in grado di distinguere i due membri dell'equivalenza L<=>R. Infatti L'avversario non dovrà far altro che chiamare Oy(), e invertire la funzione. Quando poi chiamerà Oeq fornendo la preimmagine di y da lui calcolata se Oeq appartiene al primo membro allora Oeq tironerà sempre true, mentre se appartiene al secondo ritornerà sempre false.

Supponiamo invece che f si effettivamente one-way. Allora non esisterà nessun attaccante efficiente che riesca ad invertire f. L'unico modo che ha quindi un attaccante per invertire la funzione è chiamare oltre Ox() e ottenere così una preimmagine valida. Se però l'attaccante chiama Ox() allora L e R sono indistinguibili (a meno di probabilità trascurabili). Infatti se l'Ox() chiamato dall'attaccante apaprtiene ad L allora viene semplicemente ritornrata x. Se invece Ox() appartiene ad R l'oracolo prima di ritornare x segna k con una costante mark. Quando poi L'attaccante chiama Oeq() nel caso Oeq sia in R allora Oeq effettua un test e se k è stato definito ritorna quello che avrebbe ritornato L.Oeq. SE invece k non è stato definito allora siginifica che l'attaccante non ha richiamato Oeq e quindi è probabilissimo che non sia riuscto ad invertire la f e quindi Oeq ritorna false.

Un'ulteriore definizione viene fornita al CryptoVerif, quella di funzione hash. La funzione hash viene fornita attraverso il seguente codice:

```
726 equiv foreach iH <= nH do OH(x:hashinput) := return(hash(x)) [all]
727       <=(0)=>
```

```
728       foreach iH <= nH do OH(x:hashinput) :=
729     find[unique] u <= nH suchthat defined(x[u],r[u])
     && otheruses(r[u]) && x= x[u] then return(r[u])
     else r <-R hashoutput; return(r).
730
731 let hashoracle =
732         foreach iH <= qH do
733     OH(x:hashinput) :=
734     return(hash(x)).
```

La funzione hash è intesa implementata nel modello dell'oracolo random. Ovvero la funzione se non è mai stata richiamata su un particolare valore x_0 allora ritorna un valore random, altrimenti ritorna il valore che ha ritornato precedentemente. Questo viene fatto salvando il valore ritornato per un particolare input in un array, epoi al momento della chiamata effettuando un look up nell'array.

925 forall r:seed, x:D, x':D; 926 (x' = invf(skgen(r),x)) = (f(pkgen(r),x') = x).

forall k:skey, x:D, x':D; (invf(k,x) = invf(k,x')) = (x = x').

915 forall k:pkey, x:D, x':D; (f(k,x) = f(k,x')) = (x = x'). 916 forall k:pkey, x:D, x':D; (f'(k,x) = f'(k,x')) = (x = x').

911 forall r:seed, x:D; f(pkgen(r), invf(skgen(r), x)) = x.

forall r:seed, x:D; invf(skgen(r), f(pkgen(r), x)) = x.

Capitolo 4

Risultati Raggiunti

Capitolo 5

Conclusioni

Bibliografia

[AR07] Martín Abadi and Phillip Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). *J. Cryptology*, 20(3):395, 2007.

[BJST08] Bruno Blanchet, Aaron D. Jaggard, Andre Scedrov, and Joe-Kai Tsay. Computationally sound mechanized proofs for basic and public-key Kerberos. In *ACM Symposium on Information, Computer and Communications Security (ASIACCS'08)*, pages 87–99, Tokyo, Japan, March 2008. ACM.

[BP06] Bruno Blanchet and David Pointcheval. Automated security proofs with sequences of games. In Cynthia Dwork, editor, *CRYPTO'06*, volume 4117 of *Lecture Notes on Computer Science*, pages 537–554, Santa Barbara, CA, August 2006. Springer Verlag.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, November 1976.

[Dwo06] Cynthia Dwork, editor. *Advances in Cryptology - CRYPTO 2006, 26th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2006, Proceedings*, volume 4117 of *Lecture Notes in Computer Science*. Springer, 2006.

[Gol00] Oded Goldreich. *Foundations of Cryptography: Basic Tools*. Cambridge University Press, New York, NY, USA, 2000.

[KL07] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography (Chapman & Hall/Crc Cryptography and Network Security Series)*. Chapman & Hall/CRC, 2007.

[Sho] Victor Shoup. Sequences of games: A tool for taming complexity in security proofs. `http://eprint.iacr.org/2004/332.pdf`.