

Nozioni di Sicurezza per Autenticazione

Autore: Daniele Venturi

Seminario

1 Introduzione

L'autenticazione è uno dei contesti fondamentali in crittografia. Definiamo autenticazione come il processo tramite cui si verifica la genuinità di qualcosa (o qualcuno). Come già abbiamo visto durante il corso, esistono diversi modi per classificare l'autenticazione: in base a *come* essa avviene, in base a *cosa* si autentica ed in base a *dove* il processo ha luogo. In queste note ci occuperemo di *autenticazione di persone* (anche detta *identificazione*) nel contesto della crittografia a *chiave segreta*. In pratica vogliamo indirizzare la seguente domanda.

Supponiamo che Alice e Bob condividano un segreto s (noto solo a loro). Esistono protocolli (ovvero scambi di messaggi possibilmente efficienti) in cui Alice può dimostrare a Bob la sua identità (ovvero autenticarsi) facendo uso del segreto s e che siano sicuri in un qualche senso?

Sicurezza dimostrabile. Il progetto di una qualunque primitiva crittografica è un compito tutt'altro che semplice. Infatti quando definiamo un nuovo schema (ad esempio un protocollo di autenticazione) vorremmo fare in modo che esso resista a *tutti* gli attacchi! Ma come possiamo essere sicuri che questo avvenga? La *sicurezza dimostrabile* è un approccio generale (nato negli anni '80 nel contesto della cifratura, in seguito ad un'idea di Goldwasser e Micali [GM84]) a questo problema: si vuole dare una "prova matematica" che un dato sistema sia sicuro. In realtà, spesso ci si accontenta di dimostrare che un sistema è *computazionalmente sicuro*: nessun attaccante limitato computazionalmente deve essere in grado di violare il sistema con vantaggio non trascurabile. Tecnicamente quando parliamo di attaccante limitato computazionalmente intendiamo in realtà un *algoritmo probabilistico eseguibile in tempo polinomiale*. Diremo che un algoritmo \mathcal{A} è un algoritmo PPT (o anche un avversario PPT) se \mathcal{A} usa un certo grado di randomicità come parte della sua logica (i.e., \mathcal{A} è probabilistico) e se per ogni input $x \in \{0, 1\}^*$ l'esecuzione di $\mathcal{A}(x)$ termina in un numero di passi polinomiale nella lunghezza della stringa x .

La "ricetta" per dimostrare che una data primitiva crittografica è sicura segue più o meno lo schema seguente.

1. Si astrae la realtà, definendo un modello che contenga tutte le caratteristiche salienti presenti nel mondo reale in cui la primitiva è utilizzata.
2. Si definisce cosa significa per la primitiva essere "sicura" in questo modello.
3. Si mostra che nessun attaccante di un certo tipo (ad esempio un attaccante PPT) che "vive" nel modello definito in (1) è in grado di violare la definizione di sicurezza data in (2).

Solitamente quando si parla di sicurezza computazionale, l'ultimo passo consiste nel dimostrare che un attacco al sistema è equivalente ad un tentativo a risolvere un qualche problema computazionale ritenuto difficile: se si ritiene che nessun attaccante PPT sia in grado di risolvere tale problema, allora la nostra primitiva è sicura.¹ Detto in questi termini può sembrare tutto molto astratto, quindi cerchiamo di calare questo paradigma nel caso che ci interessa.

¹Si parla infatti di dimostrazioni di sicurezza su base *riduzione* ovvero di *riduzioni crittografiche*.

Il caso dell'autenticazione. In un protocollo di autenticazione un dimostratore \mathcal{P} tenta di convincere un verificatore \mathcal{V} (entrambi algoritmi PPT) che conosce un certo segreto condiviso s attraverso uno scambio di messaggi. Al termine dell'interazione \mathcal{V} restituisce **accetta** se autentica \mathcal{P} e restituisce **rifiuta** in caso contrario. Diremo che il protocollo ha errore di completezza α se (per ogni segreto condiviso) un'esecuzione *onesta* del protocollo ritorna **accetta** con probabilità $1 - \alpha$. Questo è il nostro modello astratto della realtà.

Ci resta da definire cosa intendiamo per sicurezza in questo modello. Daremo tre diverse definizioni (via via più forti) nel seguito: *sicurezza passiva*, *sicurezza attiva* e *sicurezza contro attacchi "uomo-nel-mezzo"* (Man in the Middle — MiM). Tutte le definizioni sono parametrizzate da un intero κ detto *parametro di sicurezza*.

- Diremo che un protocollo di autenticazione è sicuro contro attaccanti *passivi* se nessun attaccante PPT \mathcal{A} può far sì che \mathcal{V} ritorni **accetta** con probabilità non trascurabile dopo aver osservato passivamente un certo numero di esecuzioni oneste tra \mathcal{P} e \mathcal{V} . Più formalmente diremo che il protocollo è (t, Q, ϵ) -sicuro contro attaccanti passivi se nessun attaccante PPT \mathcal{A} con tempo d'esecuzione $t = t(\kappa)$ e che osserva $Q = Q(\kappa)$ esecuzioni oneste del protocollo può far sì che \mathcal{V} ritorni **accetta** con probabilità migliore di $\epsilon = \epsilon(\kappa)$.
- In un attacco *attivo* \mathcal{A} agisce in due fasi. Nella prima fase può interagire con \mathcal{P} un numero polinomiale di volte impersonando \mathcal{V} anche in modo concorrente.² Nella seconda fase \mathcal{A} interagisce solo con \mathcal{V} e ha un solo tentativo per impersonare \mathcal{P} . Diremo che un protocollo d'autenticazione è (t, Q, ϵ) -sicuro contro attaccanti attivi se nessun attaccante PPT \mathcal{A} con tempo d'esecuzione t che effettua Q richieste a \mathcal{P} nella prima fase, può far sì che \mathcal{V} ritorni **accetta** con probabilità migliore di ϵ .
- In un attacco *MiM* \mathcal{A} può interagire in modo concorrente con un numero polinomiale di istanze sia di \mathcal{P} che \mathcal{V} . Ogni volta l'avversario viene a conoscenza della decisione di \mathcal{V} . (Notare la differenza con il caso attivo, in cui \mathcal{A} ha un solo tentativo per convincere \mathcal{V} .)

Nei paragrafi successivi vedremo alcuni esempi di tecniche (simmetriche) utili a soddisfare queste definizioni.

2 Il Protocollo HB

Si dà il caso che esistono paradigmi generali per soddisfare le definizioni date nel paragrafo precedente. Il *paradigma sfida-risposta* è uno di questi: l'idea di base è che \mathcal{V} invii una sfida a \mathcal{P} chiedendo che quest'ultimo si autentichi legando la sua risposta al segreto s condiviso dalle due parti. Abbiamo già incontrato questo paradigma durante il corso ed in particolare abbiamo visto diversi esempi che usano come scatola nera altre primitive crittografiche di base: schemi MAC e firme digitali oppure cifrari simmetrici/asimmetrici. Seppure non avremo il tempo di dimostrarlo formalmente, questi schemi hanno tutti sicurezza dimostrabile (a volte con qualche semplice modifica) contro attaccanti MiM. (Si vedano [BCK98, BFGM01] per l'introduzione di un modello formale ed alcune estensioni.)

Purtroppo in molti contesti pratici non è possibile affidarsi alle primitive di cui sopra (soprattutto nel caso delle tecniche asimmetriche) perchè troppo dispendiose computazionalmente. Pertanto è interessante ricercare soluzioni alternative che offrano prestazioni migliori in termini d'efficienza. Questa direzione è stata esplorata per la prima volta nel 2001 da Hopper e Blum [HB01], i quali hanno proposto un protocollo d'autenticazione molto elegante basato sulla difficoltà del problema di imparare la parità in presenza di rumore (*Learning Parity in the presence of Noise* — LPN). Il protocollo è così efficiente che, anche per valori realistici dei parametri in gioco, potrebbe essere eseguito da esseri umani senza l'uso di un calcolatore. Più in generale, lo schema di Hopper e Blum è

²Si intende qui che \mathcal{A} può lanciare più istanze di \mathcal{P} anche in parallelo.

potenzialmente utilizzabile in tutti quei contesti in cui le risorse computazionali sono particolarmente limitate, come nel caso dei dispositivi per l'identificazione a radio frequenza (Radio Frequency Identification — RFID).

Il problema LPN. Definiamo innanzitutto il problema LPN. Informalmente, nella sua versione computazionale, il problema LPN richiede di calcolare un *vettore segreto* $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_2^\kappa$ a partire da un certo numero di prodotti scalari tra \mathbf{s} e vettori casuali *sporcati da rumore Bernoulliano*. Più formalmente, sia Bern_τ la distribuzione Bernoulliana con parametro $0 < \tau \leq 1/2$ (i.e., $\mathbb{P}[e = 1] = \tau$ se $e \xleftarrow{\$} \text{Bern}_\tau$). Definiamo con $\Lambda_{\tau,\kappa}(\mathbf{s})$ la distribuzione di probabilità su $\mathbb{Z}_2^\kappa \times \mathbb{Z}_2$ ottenuta estraendo uniformemente $\mathbf{r} \xleftarrow{\$} \mathbb{Z}_2^\kappa$ ed $e \xleftarrow{\$} \text{Bern}_\tau$ e restituendo $(\mathbf{r}, \mathbf{r}^\top \cdot \mathbf{s} \oplus e)$.

Definizione 2.1 (Problema LPN computazionale). Diremo che il problema $\text{LPN}_{\tau,\kappa}$ computazionale è (t, Q, ϵ) -difficile se nessun attaccante PPT \mathcal{A} può determinare \mathbf{s} in tempo t richiedendo Q campioni all'oracolo $\Lambda_{\tau,\kappa}(\mathbf{s})$ con probabilità migliore di ϵ :

$$\mathbb{P} \left[\mathcal{A}^{\Lambda_{\tau,\kappa}(\mathbf{s})}(1^\kappa) = \mathbf{s} : \mathbf{s} \xleftarrow{\$} \mathbb{Z}_2^\kappa \right] \leq \epsilon.$$

La complessità del problema $\text{LPN}_{\tau,\kappa}$ cresce al crescere di τ . Più in generale il problema è stato mostrato essere **NP**-completo [BMvT78]. L'algoritmo più efficiente [BKW03] richiede $t, Q = 2^{\Theta(\kappa \log \kappa)}$.³

Nella sua forma decisionale il problema LPN richiede di distinguere $\Lambda_{\tau,\kappa}$ dalla distribuzione uniforme $U_{\kappa+1}$ su $\mathbb{Z}_2^{\kappa+1}$. Informalmente considereremo un attaccante PPT \mathcal{D} il cui scopo è distinguere l'oracolo $\Lambda_{\tau,\kappa}(\mathbf{s})$ (per un vettore \mathbf{s} casuale) dall'oracolo che restituisce campioni prelevati dalla distribuzione uniforme $U_{\kappa+1}$. Al termine dell'interazione \mathcal{D} restituisce un bit in $\{0, 1\}$ come tentativo d'indovinare l'oracolo con il quale ha dialogato nell'esperimento. Formalmente:

Definizione 2.2 (Problema LPN decisionale). Diremo che il problema $\text{LPN}_{\tau,\kappa}$ decisionale è (t, Q, ϵ) -difficile se per ogni attaccante PPT \mathcal{D} che agisce in tempo t richiedendo Q campioni si ha

$$\left| \mathbb{P} \left[\mathcal{D}^{\Lambda_{\tau,\kappa}(\mathbf{s})}(1^\kappa) = 1 : \mathbf{s} \xleftarrow{\$} \mathbb{Z}_2^\kappa \right] - \mathbb{P} \left[\mathcal{D}^{U_{\kappa+1}}(1^\kappa) = 1 \right] \right| \leq \epsilon.$$

Regev [Reg05] ha mostrato che la versione computazionale del problema LPN è equivalente a quella decisionale. Ciò significa che dato un oracolo per il problema LPN decisionale con vettore $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_2^\kappa$ è possibile recuperare \mathbf{s} in tempo polinomiale. Si vedano [Reg05, KS06] per i dettagli.

Descrizione del protocollo. La soluzione proposta da Hopper e Blum ricalca il paradigma sfida e risposta. Le due parti \mathcal{P} e \mathcal{V} condividono un segreto $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_2^\kappa$. Il verificatore \mathcal{V} estrae un vettore $\mathbf{r} \xleftarrow{\$} \mathbb{Z}_2^\kappa$ e lo invia a \mathcal{P} il quale risponde con $z = \mathbf{r}^\top \cdot \mathbf{s} \oplus e$ dove $e \xleftarrow{\$} \text{Bern}_\tau$. Infine \mathcal{V} ritorna *accetta* se e solo se $\mathbf{r}^\top \cdot \mathbf{s} = z$. Notare che il protocollo nella sua versione base ha un *errore di validità* (i.e., la probabilità che \mathcal{A} faccia sì che \mathcal{V} ritorni *accetta* usando un valore casuale di z) pari ad $1/2$ ed *errore di completezza* (i.e., la probabilità che \mathcal{V} ritorni *rifiuta* anche quando \mathcal{P} è onesto) τ . Queste quantità possono essere ridotte ripetendo il protocollo in modo sequenziale (cioè eseguendo un passo base n volte). In questo modo il protocollo è anche provabilmente sicuro contro attaccanti passivi come argomentato dagli stessi Hopper e Blum e mostrato più avanti in modo formale da Jules e Weis [JW05].

³La notazione $f(t) = \Theta(g(t))$ intuitivamente cattura il fatto che la funzione $f(\cdot)$ è dominata asintoticamente (sia dal basso che dall'alto) dalla funzione $g(\cdot)$. In altri termini esistono costanti $c_1, c_2 \in \mathbb{N}$ tali che $\forall t \geq t_0$ risulta

$$c_1 \cdot |g(t)| \leq |f(t)| \leq c_2 \cdot |g(t)|.$$

Per mantenere l'efficienza del protocollo vorremmo però lanciare le diverse istanze *in parallelo* e non sequenzialmente. Il protocollo nella sua versione parallela è indicato come HB ed è definito come segue.

- **Parametri pubblici.** Il protocollo ha i seguenti parametri pubblici, tutti dipendenti dal parametro di sicurezza κ .
 - $0 < \tau < 1/2$: parametro di rumore
 - $\tau < \tau' < 1/2$: soglia di accettazione
 - n : numero di ripetizioni parallele
- **Generazione delle chiavi.** L' algoritmo di generazione delle chiavi restituisce il segreto condiviso \mathbf{s} con $\mathbf{s} \stackrel{\$}{\leftarrow} \mathbb{Z}_2^\kappa$.
- **Protocollo.** Il protocollo è illustrato in Figura 1.

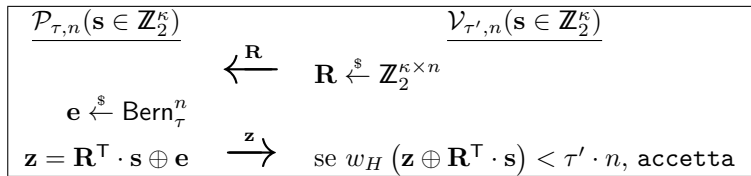


Figura 1: Il protocollo HB, sicuro contro attaccanti passivi.

Ora \mathcal{V} genera una matrice $\mathbf{R} \stackrel{\$}{\leftarrow} \mathbb{Z}_2^{\kappa \times n}$ e la risposta prodotta da \mathcal{P} è della forma $\mathbf{z} = \mathbf{R}^\top \cdot \mathbf{s} \oplus \mathbf{e}$ dove $\mathbf{e} \stackrel{\$}{\leftarrow} \text{Bern}_\tau^n$.⁴ In questo caso \mathcal{V} accetta se e solo se al più una frazione τ' delle n istanze parallele fallisce, i.e. se il peso di Hamming⁵ della stringa $\mathbf{z} \oplus \mathbf{R}^\top \cdot \mathbf{s}$ soddisfa $w_H(\mathbf{z} \oplus \mathbf{R}^\top \cdot \mathbf{s}) < \tau' \cdot n$.

Efficienza. Il protocollo richiede semplicemente di calcolare n prodotti scalari in \mathbb{F}_2 (tra stringhe binarie lunghe κ bit) ed un singolo peso di Hamming di una stringa lunga n bit. Valori tipici dei parametri sono $\kappa = 500$ ed $n = 160$, ovvero segreti di lunghezza 500 bit e 160 ripetizioni parallele. Ciò genera un overhead di comunicazione pari a $\approx \kappa \cdot n$ bit ed una complessità computazionale di $\Theta(\kappa \cdot n)$ (espressa come numero di operazioni elementari su \mathbb{F}_2). Come mostrato in [GRS08a] c'è un trade-off naturale tra dimensione della chiave e overhead generato dal protocollo: per ogni costante $1 \leq c \leq n$ possiamo diminuire l'overhead di un fattore c ed aumentare la lunghezza della chiave dello stesso fattore. In particolare, posto $n_s = c$ ed $n_r = n/c$, l'idea è quella di utilizzare una matrice di segreti $\mathbf{S} = (\mathbf{S}[1], \dots, \mathbf{S}[n_s]) \in \mathbb{Z}_2^{\kappa \times n_s}$ (cioè n_s vettori in \mathbb{Z}_2^κ invece che un unico vettore segreto $\mathbf{s} \in \mathbb{Z}_2^\kappa$) ed una matrice casuale $\mathbf{R} \in \mathbb{Z}_2^{\kappa \times n_r}$ più piccola. Si può dimostrare che la sicurezza resta invariata.

Sicurezza. Il protocollo HB è provabilmente resistente ad attacchi passivi, come mostrato per la prima volta da Katz e Shin [KS06]. Mostriamo che gli errori di validità e completezza sono trascurabili. L'errore di completezza è la probabilità che un \mathcal{P} onesto non venga autenticato. Questa non è altro che la probabilità $\alpha_{\tau,n}$ che un vettore $\mathbf{e} \stackrel{\$}{\leftarrow} \text{Bern}_\tau^n$ (cioè n campioni indipendenti prelevati dalla distribuzione Bernoulliana con parametro τ) abbia più di una frazione $\tau' \cdot n$ di valori 1. Invocando

⁴Ovvero ogni elemento del vettore $\mathbf{e} = (\mathbf{e}[1], \dots, \mathbf{e}[n])$ è estratto uniformemente ed indipendentemente dalla distribuzione Bernoulliana con parametro τ , i.e. $\mathbf{e}[i] \stackrel{\$}{\leftarrow} \text{Bern}_\tau$.

⁵Data una stringa $\mathbf{v} \in \mathbb{Z}_2^\kappa$ il suo peso di Hamming $0 \leq w_H(\mathbf{v}) \leq \kappa$ è il numero di valori 1 in \mathbf{v} , i.e. $w_H(\mathbf{v}) \stackrel{\text{def}}{=} \#\{v_i : v_i = 1, 1 \leq i \leq \kappa\}$.

il limite di Chernoff⁶ è facile vedere che

$$\alpha_{\tau,n} \stackrel{\text{def}}{=} \mathbb{P} \left[w_H(\mathbf{e}) > \tau' \cdot n : \mathbf{e} \stackrel{\$}{\leftarrow} \text{Bern}_{\tau}^n \right] \leq 2^{-\Theta(n)}.$$

L'errore di validità è la probabilità che \mathcal{A} faccia sì che \mathcal{V} ritorni *accetta* inviando una risposta casuale. Questa non è altro che la probabilità $\alpha'_{\tau',n}$ che una stringa random $\mathbf{v} \stackrel{\$}{\leftarrow} \mathbb{Z}_2^{\kappa}$ abbia peso di Hamming $\leq \tau' \cdot n$, dove $0 < \tau < \tau' < 1/2$. Ancora una volta il limite di Chernoff implica che

$$\alpha'_{\tau',n} \stackrel{\text{def}}{=} 2^{-n} \cdot \sum_{i=0}^{\lfloor \tau' \cdot n \rfloor} \binom{n}{i} \leq 2^{-\Theta(n)}.$$

Teorema 2.1 (Sicurezza del protocollo HB). *Se il problema $\text{LPN}_{\tau,\kappa}$ è $(t, n \cdot (Q+1), \epsilon)$ -difficile con $0 < \tau < \tau' < 1/4$, il protocollo HB è (t', Q, ϵ') -sicuro contro attaccanti passivi dove $t' = O(t)$ ed*

$$\epsilon' = \epsilon + 2^{-\Theta(n)}.$$

Dimostrazione. Mostriamo che se esiste un attaccante PPT \mathcal{A} in grado di violare la sicurezza del protocollo HB attraverso un attacco passivo con probabilità ϵ' , allora possiamo usare \mathcal{A} per costruire un attaccante PPT \mathcal{D} in grado di risolvere la versione decisionale del problema $\text{LPN}_{\tau,\kappa}$ con probabilità ϵ come nell'enunciato del teorema. L'attaccante \mathcal{D} ha accesso ad un oracolo che restituisce stringhe in $\mathbb{Z}_2^{\kappa+1}$ e deve distinguere se tale oracolo è $\Lambda_{\tau,\kappa}(\mathbf{s})$ (per un qualche $\mathbf{s} \stackrel{\$}{\leftarrow} \mathbb{Z}_2^{\kappa}$) oppure la distribuzione uniforme $U_{\kappa+1}$. Per fare ciò \mathcal{D} usa \mathcal{A} come segue:

1. Ogni volta in cui \mathcal{A} richiede di osservare passivamente un'esecuzione onesta del protocollo HB, interroga n volte l'oracolo ottenendo coppie di elementi $(\mathbf{R}[i], z_i)$ con $i = 1, \dots, n$. Posto $\mathbf{R} = (\mathbf{R}[1], \dots, \mathbf{R}[n])$ e $\mathbf{z} = (z_1, \dots, z_n)$, ritorna ad \mathcal{A} la coppia (\mathbf{R}, \mathbf{z}) .
2. Quando \mathcal{A} tenta di impersonare \mathcal{P} , interroga nuovamente l'oracolo per n volte ottenendo coppie di elementi $(\mathbf{R}'[i], z'_i)$ con $i = 1, \dots, n$. Sfida quindi \mathcal{A} con $\mathbf{R}' = (\mathbf{R}'[1], \dots, \mathbf{R}'[n])$ ed ottieni in cambio la risposta \mathbf{z}'' .
3. Sia $\mathbf{z}' = (z'_1, \dots, z'_n)$. Ritorna 1 se e solo se $w_H(\mathbf{z}' \oplus \mathbf{z}'') \leq 2\tau' \cdot n$.

Dobbiamo distinguere due casi.

L'oracolo di \mathcal{D} è $U_{\kappa+1}$. In questo caso \mathcal{A} vede solamente stringhe uniformemente random. Quindi la stringa $\mathbf{z}' \oplus \mathbf{z}''$ è a distribuzione uniforme in \mathbb{Z}_2^{κ} . Invocando il limite di Chernoff ne segue che la probabilità con cui \mathcal{D} restituisce 1 è esattamente $2^{-n} \cdot \sum_{i=0}^{\lfloor 2\tau' \cdot n \rfloor} \binom{n}{i} \leq 2^{-\Theta(n)}$.

L'oracolo di \mathcal{D} è $\Lambda_{\tau,\kappa}(\mathbf{s})$. In questo caso la simulazione eseguita da \mathcal{D} nella prima fase dell'attacco è perfetta. Indichiamo con $\mathbf{z}^* = (\mathbf{R}')^T \cdot \mathbf{s}$ il vettore contenente i valori dei prodotti scalari (non affetti da rumore) $z_i^* = \mathbf{R}'[i]^T \cdot \mathbf{s}$. Siccome per ipotesi \mathcal{A} impersona \mathcal{P} con probabilità ϵ' , ne segue che con probabilità almeno ϵ' i vettori \mathbf{z}'' e \mathbf{z}^* differiscono in al più $\tau' \cdot n$ valori. D'altra parte, siccome il vettore \mathbf{z}' in questo caso soddisfa $\mathbf{z}' = (\mathbf{R}')^T \cdot \mathbf{s} \oplus \mathbf{e}$ con $\mathbf{e} \stackrel{\$}{\leftarrow} \text{Bern}_{\tau}^n$, il vettore \mathbf{z}' è distribuito come in una risposta onesta di \mathcal{P} nel protocollo e quindi (poiché il protocollo ha errore di completezza

⁶Il limite di Chernoff è un risultato classico in teoria delle probabilità (si veda ad esempio [MU07, Teorema 4.4], per una dimostrazione). Siano X_1, X_2, \dots, X_n variabili aleatorie mutuamente indipendenti su $\{0, 1\}$ e sia $\mu = \sum_{i=1}^n \mathbb{E}[X_i]$. Allora $\forall 0 < \delta \leq 1$ si ha

$$\mathbb{P} \left[\sum_{i=1}^n X_i \geq (1 + \delta)\mu \right] \leq e^{-\mu\delta^2/3} = 2^{-\Theta(n)}.$$

$\alpha_{\tau,n}$) \mathbf{z}' e \mathbf{z}^* differiscono in al più $\tau' \cdot n$ valori tranne che con probabilità al più $\alpha_{\tau,n}$. Ma allora la probabilità che \mathbf{z}' e \mathbf{z}'' differiscano in al più $2\tau' \cdot n$ valori⁷ è almeno $\epsilon' - \alpha_{\tau,n}$.

Mettendo tutto insieme abbiamo trovato

$$\left| \mathbb{P} \left[\mathcal{D}^{\Lambda_{\tau,\kappa}(\mathbf{s})}(1^\kappa) = 1 \right] - \mathbb{P} \left[\mathcal{D}^{U_{\kappa+1}}(1^\kappa) = 1 \right] \right| \geq \epsilon' - \alpha_{\tau,n} - 2^{-\Theta(n)},$$

e quindi \mathcal{D} può risolvere il problema $\text{LPN}_{\tau,\kappa}$ decisionale almeno con probabilità $\epsilon = \epsilon' - \alpha_{\tau,n} - 2^{-\Theta(n)} = \epsilon' - 2^{-\Theta(n)}$, come richiesto. \square

La dimostrazione porta ad un risultato utile solo quando $\tau < \tau' < 1/4$, poichè quando $\tau \geq 1/4$ risulta $2\tau' \cdot n \geq 2\tau \cdot n \geq n/2$ e quindi $2^{-n} \sum_{i=0}^{2\lceil \tau' \cdot n \rceil} \binom{n}{i} \geq 1/2$. In altri termini in questo caso anche quando l'oracolo di \mathcal{D} è $U_{\kappa+1}$ con alta probabilità \mathcal{D} ritorna 1. Si veda [KSS10] per una dimostrazione nel caso in cui $1/4 \leq \tau' < 1/2$.

3 Il protocollo HB^+

Il protocollo HB è insicuro contro attaccanti attivi. Ad esempio \mathcal{A} potrebbe sostituirsi a \mathcal{V} e quindi sfidare \mathcal{P} ripetutamente con lo stesso $\mathbf{R} = (\mathbf{R}[1], \dots, \mathbf{R}[n])$. Prendendo la maggioranza delle risposte relative ad ogni valore $\mathbf{R}[i]^\top \cdot \mathbf{s}$, se i vettori in \mathbf{R} sono linearmente indipendenti, è possibile recuperare \mathbf{s} con probabilità alta. Per contrastare questo tipo di attacchi, Jules e Weis [JW05] hanno proposto la seguente estensione.

Le due parti condividono due segreti $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{Z}_2^\kappa$. L'idea di base è che stavolta \mathcal{P} cominci l'interazione inviando a \mathcal{V} un vettore $\mathbf{r}_1 \xleftarrow{\$} \mathbb{Z}_2^\kappa$; quindi \mathcal{V} risponde con una sfida random $\mathbf{r}_2 \xleftarrow{\$} \mathbb{Z}_2^\kappa$. Infine \mathcal{P} risponde con $z = \mathbf{r}_1^\top \cdot \mathbf{s}_1 \oplus \mathbf{r}_2^\top \cdot \mathbf{s}_2 \oplus e$, dove $e \leftarrow \text{Bern}_\tau$. In questo caso \mathcal{V} ritorna *accetta* se e solo se $z = \mathbf{r}_1^\top \cdot \mathbf{s}_1 \oplus \mathbf{r}_2^\top \cdot \mathbf{s}_2$. Ancora una volta gli errori di validità e completezza possono essere ridotti ripetendo l'interazione base in modo sequenziale o parallelo. La versione sequenziale è stata analizzata da Juel e Weils [JW05] i quali hanno mostrato che il protocollo è sicuro contro attaccanti attivi. La versione parallela, indicata nel seguito con HB^+ , è stata analizzata sempre in [KS06, KSS10].

- **Parametri pubblici.** Il protocollo ha i seguenti parametri pubblici, tutti dipendenti dal parametro di sicurezza κ .
 $0 < \tau < 1/2$: parametro di rumore
 $\tau < \tau' < 1/2$: soglia di accettazione
 n : numero di ripetizioni parallele
- **Generazione delle chiavi.** L'algoritmo di generazione delle chiavi restituisce i segreti condivisi $\mathbf{s}_1, \mathbf{s}_2$ con $\mathbf{s}_1, \mathbf{s}_2 \xleftarrow{\$} \mathbb{Z}_2^\kappa$.
- **Protocollo.** Il protocollo è illustrato in Figura 2.

Gli errori di validità e completezza sono gli stessi del protocollo HB.

Teorema 3.1 (Sicurezza del protocollo HB^+). *Se il problema $\text{LPN}_{\tau,\kappa}$ è $(t, n \cdot Q, \epsilon)$ -difficile con $0 < \tau < \tau' < 1/2$, il protocollo HB^+ è (t', Q, ϵ') -sicuro contro attaccanti attivi dove $t' = O(t)$ ed*

$$(\epsilon')^2 = \epsilon + 2^{-\Theta(n)}.$$

⁷Questo viene dal fatto che calcolare $w_H(\mathbf{v}_1 \oplus \mathbf{v}_2)$ è identico a calcolare la distanza di Hamming $d_H(\mathbf{v}_1, \mathbf{v}_2)$. Siccome la distanza di Hamming è una metrica, essa soddisfa la disuguaglianza triangolare

$$d_H(\mathbf{v}_1, \mathbf{v}_2) \leq d_H(\mathbf{v}_1, \mathbf{v}^*) + d_H(\mathbf{v}_2, \mathbf{v}^*).$$

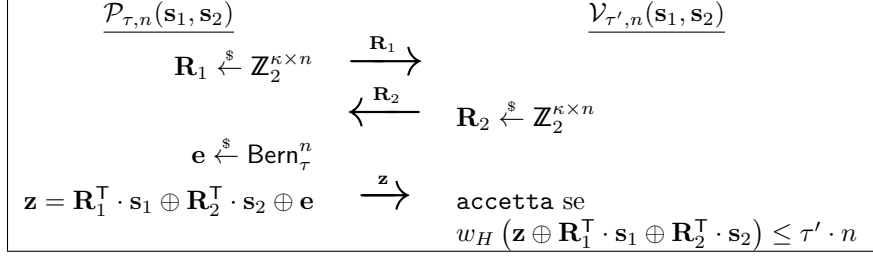


Figura 2: Il protocollo HB^+ , sicuro contro attaccanti attivi.

Dimostrazione. Mostriamo che se esiste un attaccante PPT \mathcal{A} in grado di violare la sicurezza del protocollo HB attraverso un attacco attivo con probabilità ϵ' , allora possiamo usare \mathcal{A} per costruire un attaccante PPT \mathcal{D} in grado di risolvere la versione decisionale del problema $\text{LPN}_{\tau,\kappa}$ con probabilità ϵ come nell'enunciato del teorema. L'attaccante \mathcal{D} ha accesso ad un oracolo che restituisce stringhe in $\mathbb{Z}_2^{\kappa+1}$ e deve distinguere se tale oracolo è $\Lambda_{\tau,\kappa}(\mathbf{s}_1)$ (per un qualche $\mathbf{s}_1 \xleftarrow{\$} \mathbb{Z}_2^{\kappa}$) oppure la distribuzione uniforme $U_{\kappa+1}$. Per fare ciò \mathcal{D} usa \mathcal{A} come segue:

1. Estrai $\mathbf{s}_2 \xleftarrow{\$} \mathbb{Z}_2^{\kappa}$.
2. Nella prima fase dell'attacco, quando \mathcal{A} (al posto di \mathcal{V}) vuole interagire con \mathcal{P} , interroga l'oracolo n volte, ottenendo $(\mathbf{R}_1[i], z_i)$ con $i = 1, \dots, n$. Invia quindi ad \mathcal{A} la matrice $\mathbf{R}_1 = (\mathbf{R}_1[1], \dots, \mathbf{R}_1[n])$. Quando \mathcal{A} risponde con la matrice \mathbf{R}_2 calcola la risposta $\bar{\mathbf{z}} = \mathbf{R}_2^{\top} \cdot \mathbf{s}_2 \oplus \mathbf{z}$ avendo posto $\mathbf{z} = (z_1, \dots, z_n)$.
3. Nella seconda fase dell'attacco, quando \mathcal{A} (al posto di \mathcal{P}) tenta di autenticarsi a \mathcal{V} , ricevi innanzitutto la sfida di \mathcal{A} , della forma $\mathbf{R}_1 = (\mathbf{R}_1[1], \dots, \mathbf{R}_1[n])$. Estrai quindi $\mathbf{R}'_2 \xleftarrow{\$} \mathbb{Z}_2^{\kappa \times n}$ e sfida \mathcal{A} con $\mathbf{R}'_2 = (\mathbf{R}'_2[1], \dots, \mathbf{R}'_2[n])$. Sia $\mathbf{z}' = (z'_1, \dots, z'_n)$ la risposta di \mathcal{A} .
4. Riavvolgi \mathcal{A} (al punto in cui ha già inviato \mathbf{R}_1) e sfidalo con $\mathbf{R}''_2 \xleftarrow{\$} \mathbb{Z}_2^{\kappa \times n}$ della forma $\mathbf{R}''_2 = (\mathbf{R}''_2[1], \dots, \mathbf{R}''_2[n])$. Sia $\mathbf{z}'' = (z''_1, \dots, z''_n)$ la risposta di \mathcal{A} .
5. Poni $\mathbf{z}^{\oplus} = \mathbf{z}' \oplus \mathbf{z}''$. Siano poi $\mathbf{R}^*[i] = \mathbf{R}'_2[i] \oplus \mathbf{R}''_2[i]$, $\mathbf{R}^* = (\mathbf{R}^*[1], \dots, \mathbf{R}^*[n])$ e $\mathbf{z}^* = (\mathbf{R}^*)^{\top} \cdot \mathbf{s}_2$. Restituisci 1 se e solo se \mathbf{z}^{\oplus} e \mathbf{z}^* differiscono in al più $2\tau' \cdot n$ valori.

Dobbiamo distinguere due casi.

L'oracolo di \mathcal{D} è $U_{\kappa+1}$. Nella prima fase dell'attacco in questo caso i valori z_i sono bit random. Ne segue che la risposta simulata $\bar{\mathbf{z}}$ è anch'essa random ed in particolare indipendente da \mathbf{s}_2 . Notare che anche il vettore \mathbf{z}^{\oplus} è completamente random.

Siccome poi i vettori $\mathbf{R}^*[i]$ (con $i = 1, \dots, n$) sono indipendenti ed uniformemente distribuiti, essi sono anche linearmente indipendenti con probabilità $\frac{2^n}{2^{\kappa}}$.⁸ Quando ciò accade, il vettore \mathbf{z}^* è anch'esso a distribuzione uniforme e la probabilità che i due vettori \mathbf{z}^* e \mathbf{z}^{\oplus} differiscano in al più

⁸Ciò può essere mostrato come segue. Siano $\{\mathbf{r}_i\}_{i=1}^n$ vettori random in \mathbb{Z}_2^{κ} . Indichiamo con E_i l'evento che il vettore \mathbf{r}_i sia linearmente *dipendente* da uno dei vettori $\mathbf{r}_1, \dots, \mathbf{r}_{i-1}$ (per $i = 0$ tale evento è l'evento in cui \mathbf{r}_1 sia il vettore tutto nullo). Siccome il sottospazio definito da $i - 1$ vettori ha dimensione al più 2^{i-1} , la probabilità di E_i è al più $\frac{2^{i-1}}{2^{\kappa}}$. Applicando il limite dell'unione otteniamo

$$\mathbb{P} \left[\bigvee_{i=1}^n E_i \right] \leq 2^{-\kappa} \sum_{i=0}^{n-1} 2^i < \frac{2^n}{2^{\kappa}},$$

come desiderato.

$2\tau' \cdot n$ valori è esattamente $2^{-n} \cdot \sum_{i=0}^{2\lceil \tau' \cdot n \rceil} \binom{n}{i} \leq 2^{-\Theta(n)}$ (usando il limite di Chernoff). Quindi in questo caso \mathcal{D} restituisce 1 con probabilità al più $\frac{2^n}{2^\kappa} + 2^{-\Theta(n)}$.

L'oracolo di \mathcal{D} è $\Lambda_{\tau, \kappa}(\mathbf{s}_1)$. In questo caso la simulazione della prima fase è perfetta. Sia ω la randomicità usata per simulare la prima fase dell'attacco (i.e., i vettori $\mathbf{s}_1, \mathbf{s}_2$, la randomicità di \mathcal{A} e la randomicità usata per rispondere alle richieste di \mathcal{A}). Per un ω fissato sia ϵ'_ω la probabilità, presa sulle scelte random dei vettori $\mathbf{R}_2[1], \dots, \mathbf{R}_2[n]$ in \mathbf{R}_2 , che \mathcal{A} venga autenticato da \mathcal{V} . La probabilità che \mathcal{A} risponda correttamente ad entrambe le sfide \mathbf{R}'_2 ed \mathbf{R}''_2 è quindi $(\epsilon'_\omega)^2$. Mediando ed applicando la disuguaglianza di Jensen otteniamo

$$\mathbb{E}_\omega [(\epsilon'_\omega)^2] \geq (\mathbb{E}_\omega [\epsilon'_\omega])^2 = (\epsilon')^2.$$

Supponiamo dunque che \mathcal{A} risponda correttamente ad entrambe le sfide \mathbf{R}'_2 ed \mathbf{R}''_2 . Ciò significa che il vettore \mathbf{z}' differisce in al più $\tau' \cdot n$ valori dal vettore delle risposte corrette

$$\begin{aligned} \boldsymbol{\xi}' &\stackrel{\text{def}}{=} \mathbf{R}_1^\top \cdot \mathbf{s}_1 \oplus (\mathbf{R}'_2)^\top \cdot \mathbf{s}_2 \\ &= (\mathbf{R}_1[1]^\top \cdot \mathbf{s}_1 \oplus \mathbf{R}'_2[1]^\top \cdot \mathbf{s}_2, \dots, \mathbf{R}_1[n]^\top \cdot \mathbf{s}_1 \oplus \mathbf{R}'_2[n]^\top \cdot \mathbf{s}_2), \end{aligned}$$

ed allo stesso tempo il vettore \mathbf{z}'' differisce in al più $\tau' \cdot n$ valori dal vettore delle risposte corrette

$$\begin{aligned} \boldsymbol{\xi}'' &\stackrel{\text{def}}{=} \mathbf{R}_1^\top \cdot \mathbf{s}_1 \oplus (\mathbf{R}''_2)^\top \cdot \mathbf{s}_2 \\ &= (\mathbf{R}_1[1]^\top \cdot \mathbf{s}_1 \oplus \mathbf{R}''_2[1]^\top \cdot \mathbf{s}_2, \dots, \mathbf{R}_1[n]^\top \cdot \mathbf{s}_1 \oplus \mathbf{R}''_2[n]^\top \cdot \mathbf{s}_2), \end{aligned}$$

Ma allora il vettore $\mathbf{z}' \oplus \mathbf{z}'' = \mathbf{z}^\oplus$ differisce in al più $2\tau' \cdot n$ valori dal vettore

$$\begin{aligned} \boldsymbol{\xi}' \oplus \boldsymbol{\xi}'' &= \mathbf{R}_1^\top \cdot \mathbf{s}_1 \oplus (\mathbf{R}'_2)^\top \cdot \mathbf{s}_2 \oplus \mathbf{R}_1^\top \cdot \mathbf{s}_1 \oplus (\mathbf{R}''_2)^\top \cdot \mathbf{s}_2 \\ &= (\mathbf{R}'_2[1]^\top \cdot \mathbf{s}_2 \oplus \mathbf{R}''_2[1]^\top \cdot \mathbf{s}_2, \dots, \mathbf{R}'_2[n]^\top \cdot \mathbf{s}_2 \oplus \mathbf{R}''_2[n]^\top \cdot \mathbf{s}_2) \\ &= ((\mathbf{R}'_2[1] \oplus \mathbf{R}''_2[1])^\top \cdot \mathbf{s}_2, \dots, (\mathbf{R}'_2[n] \oplus \mathbf{R}''_2[n])^\top \cdot \mathbf{s}_2) \\ &= (\mathbf{R}^*)^\top \cdot \mathbf{s}_2 = \mathbf{z}^*. \end{aligned}$$

Dunque in questo caso \mathcal{D} restituisce 1 con probabilità $(\epsilon')^2$.

Mettendo tutto insieme abbiamo trovato

$$\left| \mathbb{P} \left[\mathcal{D}^{\Lambda_{\tau, \kappa}(\mathbf{s}_1)}(1^\kappa) = 1 \right] - \mathbb{P} \left[\mathcal{D}^{U_{\kappa+1}}(1^\kappa) = 1 \right] \right| \geq (\epsilon')^2 - \frac{2^n}{2^\kappa} - 2^{-\Theta(n)},$$

e quindi \mathcal{D} può risolvere il problema $\text{LPN}_{\tau, \kappa}$ decisionale almeno con probabilità $\epsilon = (\epsilon')^2 - 2^{-\Theta(n)}$, come richiesto. \square

Notare che la riduzione perde un fattore $\sqrt{\epsilon}$. Inoltre il protocollo richiede 3 messaggi esclusivamente perchè in questo modo \mathcal{P} è “costretto” ad impegnarsi all'inizio del protocollo sulla matrice \mathbf{R}_1 il che rende possibile applicare il trucco del riavvolgimento. La tecnica usata nella dimostrazione fallisce se modifichiamo il protocollo usando solo 2 round (all'inizio \mathcal{V} sfida \mathcal{P} con \mathbf{R}_2 , quindi \mathcal{P} estrae \mathbf{R}_1 e calcola la risposta \mathbf{z} come di consueto). D'altra parte non è noto alcun attacco contro questa versione più compatta del protocollo HB^+ . Sarebbe interessante vedere un protocollo di autenticazione con sicurezza attiva in solamente 2 round e con una riduzione “stretta”.

Attacchi MiM. Un attacco attivo non è tanto potente quanto un attacco MiM. In effetti il protocollo HB^+ non è resistente a tale tipo di attacchi [GRS05]. L'avversario si pone in mezzo tra Alice e Bob, sceglie un vettore $\boldsymbol{\Delta} \in \mathbb{Z}_2^\kappa$ e, ogni qualvolta \mathcal{V} invia una sfida \mathbf{r}_2 la sostituisce con $\mathbf{r}_2 \oplus \boldsymbol{\Delta}$. Alla fine del round ottiene la risposta $z = \mathbf{r}_1^\top \cdot \mathbf{s}_2 \oplus (\mathbf{r}_2 \oplus \boldsymbol{\Delta})^\top \cdot \mathbf{s}_2 \oplus e$. La decisione di \mathcal{V}

può essere utilizzata per recuperare un singolo bit di \mathbf{s}_2 . Infatti, se l'autenticazione (al termine degli n round del protocollo) ha successo (risp. fallisce) con alta probabilità abbiamo $\Delta^T \cdot \mathbf{s}_2 = 0$ (risp. $\Delta^T \cdot \mathbf{s}_2 = 1$). In questo modo è possibile recuperare \mathbf{s}_2 bit per bit cambiando Δ progressivamente.

Varie soluzioni, per lo più euristiche, sono state proposte per risolvere il problema, si vedano ad esempio [BCD06, MP07, BC08, GRS08b, GRS08a, GRS08c, GMZZ08, LMM08, OOV08]. Purtroppo al momento la maggior parte degli sforzi si è rivelata vana, in quanto le soluzioni proposte non hanno una prova formale di sicurezza o addirittura si sono rivelate essere insicure (ad esempio la soluzione proposta in [BC08], violata immediatamente da Frumkin e Shamir [FS09] subito dopo essere stata pubblicata).

Riferimenti bibliografici

- [BC08] Julien Bringer and Hervé Chabanne. Trusted-HB: A low-cost version of HB^+ secure against man-in-the-middle attacks. *IEEE Transactions on Information Theory*, 54(9):4339–4342, 2008.
- [BCD06] Julien Bringer, Hervé Chabanne, and Emmanuelle Dottax. HB^{++} : a lightweight authentication protocol secure against some attacks. In *SecPerU*, pages 28–33. IEEE Computer Society, 2006.
- [BCK98] Mihir Bellare, Ran Canetti, and Hugo Krawczyk. A modular approach to the design and analysis of authentication and key exchange protocols (extended abstract). In *STOC*, pages 419–428, 1998.
- [BFGM01] Mihir Bellare, Marc Fischlin, Shafi Goldwasser, and Silvio Micali. Identification protocols secure against reset attacks. In Birgit Pfitzmann, editor, *EUROCRYPT*, volume 2045 of *Lecture Notes in Computer Science*, pages 495–511. Springer, 2001.
- [BKW03] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, 2003.
- [BMvT78] Elwin R. Berlekamp, Robert J. McEliece, and Henk C. A. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24:384–386, 1978.
- [FS09] Dmitry Frumkin and Adi Shamir. Un-trusted-HB: Security vulnerabilities of trusted-hb. Cryptology ePrint Archive, Report 2009/044, 2009. <http://eprint.iacr.org/>.
- [GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, April 1984.
- [GMZZ08] Zbigniew Golebiewski, Krzysztof Majcher, Filip Zagorski, and Marcin Zawada. Practical attacks on HB and HB^+ protocols. Cryptology ePrint Archive, Report 2008/241, 2008. <http://eprint.iacr.org/>.
- [GRS05] Henri Gilbert, Matt Robshaw, and Herve Sibert. An active attack against HB^+ - a provably secure lightweight authentication protocol. Cryptology ePrint Archive, Report 2005/237, 2005. <http://eprint.iacr.org/>.
- [GRS08a] Henri Gilbert, Matthew J. B. Robshaw, and Yannick Seurin. Good variants of HB^+ are hard to find. In Gene Tsudik, editor, *Financial Cryptography*, volume 5143 of *Lecture Notes in Computer Science*, pages 156–170. Springer, 2008.

- [GRS08b] Henri Gilbert, Matthew J. B. Robshaw, and Yannick Seurin. $HB^\#$: Increasing the security and efficiency of HB^+ . In Nigel P. Smart, editor, *EUROCRYPT*, volume 4965 of *Lecture Notes in Computer Science*, pages 361–378. Springer, 2008.
- [GRS08c] Henri Gilbert, Matthew J. B. Robshaw, and Yannick Seurin. How to encrypt with the LPN problem. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfssdóttir, and Igor Walukiewicz, editors, *ICALP (2)*, volume 5126 of *Lecture Notes in Computer Science*, pages 679–690. Springer, 2008.
- [HB01] Nicholas J. Hopper and Manuel Blum. Secure human identification protocols. In Colin Boyd, editor, *ASIACRYPT*, volume 2248 of *Lecture Notes in Computer Science*, pages 52–66. Springer, 2001.
- [JW05] Ari Juels and Stephen A. Weis. Authenticating pervasive devices with human protocols. In Victor Shoup, editor, *CRYPTO*, volume 3621 of *Lecture Notes in Computer Science*, pages 293–308. Springer, 2005.
- [KS06] Jonathan Katz and Ji Sun Shin. Parallel and concurrent security of the HB and HB^+ protocols. In Serge Vaudenay, editor, *EUROCRYPT*, volume 4004 of *Lecture Notes in Computer Science*, pages 73–87. Springer, 2006.
- [KSS10] Jonathan Katz, Ji Sun Shin, and Adam Smith. Parallel and concurrent security of the HB and HB^+ protocols. *J. Cryptology*, 23(3):402–421, 2010.
- [LMM08] Xuefei Leng, Keith Mayes, and Konstantinos Markantonakis. HB-MP+ Protocol: An Improvement on the HB-MP Protocol. *IEEE International Conference on RFID – IEEE RFID 2008*, pages 118–124, April 2008.
- [MP07] Jorge Munilla and Alberto Peinado. HB-MP: A further step in the HB-family of lightweight authentication protocols. *Computer Networks*, 51(9):2262–2267, 2007.
- [MU07] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2007.
- [OOV08] Khaled Ouafi, Raphael Overbeck, and Serge Vaudenay. On the security of $HB^\#$ against a man-in-the-middle attack. In Josef Pieprzyk, editor, *ASIACRYPT*, volume 5350 of *Lecture Notes in Computer Science*, pages 108–124. Springer, 2008.
- [Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *STOC*, pages 84–93. ACM, 2005.