DIPARTIMENTO DI MATEMATICA

Protocollo di scambio di chiavi Post-Quantistico in contesti IoT

Relatori:
Prof.ssa **Anna Bernasconi**
Prof. **Stefano Chessa**

Candidato:
Francesco Baldino

Anno Accademico 2021-2022

Indice

1 Introduzione

Lo scopo di questa tesi è di modificare un protocollo di scambio di chiavi post-quantistico per migliorarne l'efficienza in contesti di *Internet of Things* (IoT). Il protocollo originale considerato è quello proposto da J. Ding, X. Xie e X. Lin nel 2012 [1], che garantisce sicurezza da attacchi post-quantistici grazie ad una riduzione al problema Ring Learning with Errors ($RLWE$), che può essere a sua volta ridotto ad una approssimazione di un problema *NP-Hard*.

Per modificare il protocollo assumeremo che i due interlocutori conoscano, a meno di uno scarto piccolo, un certo valore difficilmente stimabile da avversari che cercano di attaccare il protocollo. Questa aggiunta è un'ipotesi sensata supponendo di utilizzare il protocollo in un contesto IoT: un esempio potrebbe essere dato da qualche quantità fisica che i due dispositivi stanno rilevando quali la qualità del canale radio di comunicazione tra i due, che dipende fortemente dalla posizione fisica di chi intende calcolarne il valore. Supporremo che questo valore abbia una distribuzione gaussiana centrata di varianza relativamente ampia per chiunque cerchi di attaccare il protocollo e non si trovi nella stessa posizione fisica di uno dei due interlocutori, e che abbia una distribuzione non centrata ma con varianza molto più stretta per i due interlocutori.

Utilizzeremo questa ipotesi per sostituire alcuni dei rumori generati necessari per il protocollo. Questa sostituzione avrà due risvolti computazionali positivi principali: come prima conseguenza, che citiamo in quanto risultato desiderabile ma non tratteremo estensivamente, risparmieremo un utilizzo di un generatore pseudorandomico crittograficamente sicuro, che può risultare eccessivamente pesante da un punto di vista computazionale per dispositivi a bassa potenza computazionale tipici dei contesti IoT; come seconda conseguenza, che comporrà l'oggetto di studio di questa tesi, otterremo una piccola riduzione sulle stime dei parametri per garantire la convergenza delle chiavi, che renderà il protocollo computazionalmente più leggero.

2 Premesse

2.1 Notazioni e definizioni

Nel corso della tesi faremo uso delle seguenti notazioni e definizioni

Definizione 2.1. Chiamiamo *transcript* l'insieme dei messaggi scambiati su un canale pubblico (non sicuro).

Definizione 2.2. Dato $n \in \mathbb{N}$ indichiamo con $[n]$ l'insieme $\{1, \ldots, n\}$.

Definizione 2.3. Diciamo che $f : \mathbb{N} \to \mathbb{R}$ è trascurabile se $\forall c \in \mathbb{N}\ \exists n_0 \in \mathbb{N}$ tale che $\forall n \geq n_0$ vale $f(n) < \frac{1}{n^c}$.

Definizione 2.4. Dato $q \in \mathbb{N}$ dispari indichiamo con $\mathbb{Z}_q$ l'anello $\mathbb{Z}/q\mathbb{Z}$ e lo identifichiamo con l'insieme dei rappresentanti bilanciati $\left\{-\frac{q-1}{2}, \ldots, -1, 0, 1, \ldots, \frac{q-1}{2}\right\}$.

Quando scriviamo $(x \mod q)$ con $x \in \mathbb{N}$ intendiamo il rappresentante in $\left\{-\frac{q-1}{2}, \ldots, \frac{q-1}{2}\right\}$ della classe di equivalenza di x modulo q.

Quando scriviamo $(x \mod 2)$ con $x \in \mathbb{N}$ intendiamo il rappresentante in $\{0, 1\}$ della classe di equivalenza di x modulo 2.

Definizione 2.5. Dato $\boldsymbol{a} \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ chiamiamo canonica la sua scrittura $\boldsymbol{a} = a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ a coefficienti in $\left\{-\frac{q-1}{2}, \ldots, \frac{q-1}{2}\right\}$.

Su $\mathbb{Z}_q[x]/\langle x^n + 1\rangle$ definiamo le norme

$$\|\boldsymbol{a}\|_\infty = \max_i |a_i|$$

$$\|\boldsymbol{a}\|_2 = \sqrt{a_{n-1}^2 + \cdots + a_0^2}$$

Osservazione 2.1. $\forall \boldsymbol{a} \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ *vale*

$$\|\boldsymbol{a}\|_\infty \leq \|\boldsymbol{a}\|_2$$

Lemma 2.1. $\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ *vale*

$$\|\boldsymbol{ab}\|_2 \leq \sqrt{n}\, \|\boldsymbol{a}\|_2\, \|\boldsymbol{b}\|_2$$

Dimostrazione. Scriviamo $\boldsymbol{a} = \sum_{i=0}^{n-1} a_i x^i$ e $\boldsymbol{b} = \sum_{i=0}^{n-1} b_i x^i$. Eseguendo semplicemente il prodotto tra polinomi in $\mathbb{Z}_q[x]$, senza preoccuparci del quoziente, otteniamo un polinomio di grado $2n-2$ avente coefficiente $k-$esimo

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} \mod q$$

ponendo a 0 i termini a_i e b_i per $i < 0 \vee i > n-1$. In particolare, per $k \geq n$ otteniamo

$$c_k = \sum_{i=k-(n-1)}^{n-1} a_i b_{k-i} \mod q$$

quindi per il coefficiente $n+k-$esimo otteniamo

$$c_{n+k} = \sum_{i=n+k-(n-1)}^{n-1} a_i b_{n+k-i} \mod q = \sum_{i=k+1}^{n-1} a_i b_{n+k-i} \mod q$$

Osserviamo che nel quoziente $\mathbb{Z}_q[x]/\langle x^n + 1\rangle$ valgono le congruenze

$$x^n \equiv -1 \mod (x^n + 1)$$
$$x^{n+k} \equiv -x^k \mod (x^n + 1)$$

quindi i termini di grado $n+k$-esimo in $\mathbb{Z}_q[x]$ diventano nel quoziente termini di grado k-esimo con coefficiente cambiato di segno.

Avendo chiamato c_k il coefficiente k del prodotto tra $\boldsymbol{a}$ e $\boldsymbol{b}$ in $\mathbb{Z}_q[x]$, e introducendo un coefficiente $c_{2n-1} = 0$, otteniamo che il coefficiente $k-$esimo del prodotto di $\boldsymbol{a}$ e $\boldsymbol{b}$ nel quoziente è

$$d_k = (c_k - c_{n+k}) \mod q = \left(\sum_{i=0}^{k} a_i b_{k-i} - \sum_{i=k+1}^{n-1} a_i b_{n+k-i} \right) \mod q$$

Osserviamo che possiamo vedere questa differenza di sommatorie come il prodotto scalare dei seguenti vettori

$$\boldsymbol{v}_k = \begin{pmatrix} a_0 \\ \vdots \\ a_k \\ -a_{k+1} \\ \vdots \\ -a_{n-1} \end{pmatrix}, \quad \boldsymbol{w}_k = \begin{pmatrix} b_k \\ \vdots \\ b_0 \\ b_{n-1} \\ \vdots \\ b_{k+1} \end{pmatrix}$$

Osserviamo che $\|\boldsymbol{v}_k\|_2 = \|\boldsymbol{a}\|_2$ e $\|\boldsymbol{w}_k\|_2 = \|\boldsymbol{b}\|_2$.

Per Cauchy-Schwarz abbiamo $(\boldsymbol{v}_k^T \boldsymbol{w}_k)^2 \leq \|\boldsymbol{v}_k\|_2^2 \, \|\boldsymbol{w}_k\|_2^2 = \|\boldsymbol{a}\|_2^2 \, \|\boldsymbol{b}\|_2^2$.

Inoltre, per come abbiamo definito l'operazione "mod q", abbiamo che $\forall x \in \mathbb{N}$ vale $|x \mod q| \leq |x|$ e quindi anche

$$(x \mod q)^2 \leq x^2$$

Allora otteniamo

$$\begin{aligned} \|\boldsymbol{ab}\|_2^2 = \sum_{i=0}^{n-1} d_k^2 = \sum_{i=0}^{n-1} (\boldsymbol{v}_k^T \boldsymbol{w}_k \mod q)^2 \leq \\ \leq \sum_{i=0}^{n-1} (\boldsymbol{v}_k^T \boldsymbol{w}_k)^2 \leq \sum_{i=0}^{n-1} \|\boldsymbol{a}\|_2^2 \, \|\boldsymbol{b}\|_2^2 = \\ = n \, \|\boldsymbol{a}\|_2^2 \, \|\boldsymbol{b}\|_2^2 \end{aligned}$$

da cui la tesi □

Definizione 2.6. Detto X un insieme e $\mathcal{D}$ una distribuzione di probabilità su X, indichiamo con $x \leftarrow_{\mathcal{D}} X$ l'estrazione di un elemento x dall'insieme X tramite la distribuzione $\mathcal{D}$.

Se X è un insieme su cui ha senso definire una distribuzione uniforme, indichiamo con $x \leftarrow_{\mathcal{U}} X$ l'estrazione di un x dall'insieme X in modo uniforme.

Definizione 2.7. Diciamo che una distribuzione $\mathcal{D}$ su uno spazio normato S è limitata da b se $v \leftarrow_{\mathcal{D}} S \Rightarrow \|v\| \leq b$.

Diciamo che una distribuzione $\mathcal{D}_n$ su uno spazio normato S_n è limitata da b_n eccetto che con probabilità trascurabile se vale

$$\mathbb{P}_{v \leftarrow_{\mathcal{D}_n} S_n} [\, \|v\| > b_n \,] \leq \mathrm{negl}(n)$$

con $\mathrm{negl}(n)$ una funzione trascurabile

Definizione 2.8. Indichiamo con χ_α la distribuzione Gaussiana discreta centrata di deviazione α su $\mathbb{Z}_q[x]/\langle x^n + 1 \rangle$, ovvero la distribuzione data da

$$\chi_\alpha(\boldsymbol{a}) = \frac{\mathtt{exp}(-\pi \|\boldsymbol{a}\|_2^2 / \alpha^2)}{c}$$

con c fattore di normalizzazione

Osservazione 2.2. *La distribuzione χ_α su $(\mathbb{Z}_q[x]/\langle x^n + 1 \rangle, \, \|\cdot\|_2)$ è limitata da $\alpha\sqrt{n}$ eccetto che con probabilità trascurabile*

2.2 Crittografia simmetrica e scambio di chiavi

I metodi crittografici di comunicazione attraverso canali non sicuri (ovvero in scenari in cui supponiamo che i messaggi inviati possano essere intercettati da avversari) si dividono in due macro categorie: i metodi di crittografia a chiave asimmetrica e i metodi di crittografia a chiave simmetrica.

I metodi di crittografia a chiave asimmetrica si basano sull'utilizzo di una coppia di chiavi, una pubblica e una privata, generate dalla persona che intende ricevere il messaggio. La chiave pubblica, utilizzata dal mittente, servirà esclusivamente per cifrare il messaggio in chiaro e la chiave privata, utilizzata dal destinatario, servirà esclusivamente per decifrare il messaggio cifrato. Il vantaggio offerto dalla crittografia a chiave asimmetrica consiste nel poter realizzare una comunicazione sicura tra due parti senza che queste si siano messe d'accordo precedentemente su qualche tipo di informazione segreta, poiché la chiave utilizzata per cifrare è pubblica e non contiene informazione sensibile. Lo svantaggio consiste nell'avere, a parità di livello di sicurezza offerto, messaggi cifrati e chiavi di dimensioni considerevolmente maggiori rispetto a metodi a chiave simmetrica. Per contesti dove è richiesto diminuire quanto possibile la dimensione dei pacchetti trasmessi e la complessità computazionale richiesta per cifrare e decifrare, conviene se possibile utilizzare un metodo a chiave simmetrica.

I metodi di crittografia a chiave simmetrica, invece, presuppongono che i due interlocutori siano riusciti precedentemente a mettersi d'accordo su un'unica chiave segreta condivisa, e utilizzano la stessa chiave per cifrare i messaggi in chiaro e decifrare i messaggi cifrati. Formalmente, fissati uno spazio delle chiavi $\mathcal{K}$, uno spazio dei messaggi in chiaro $\mathcal{M}$ e uno spazio dei messaggi cifrati $\mathcal{C}$, un metodo di crittografia a chiave simmetrica consiste in una tripla di algoritmi (`Gen`, `Enc`, `Dec`) con le seguenti richieste:

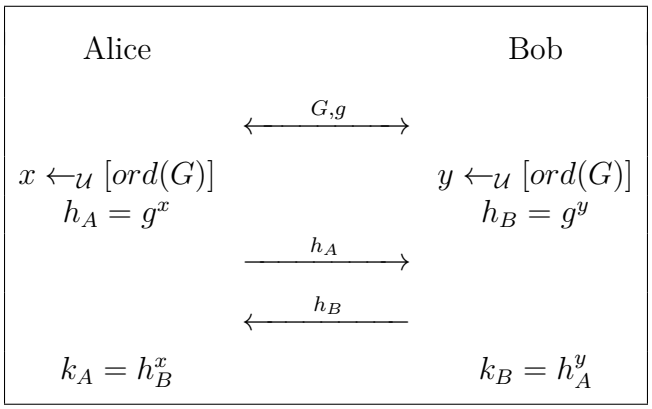
- `Gen`: $\texttt{()} \to \mathcal{K}$ è un algoritmo non deterministico che genera una chiave per il protocollo. Si suppone che i due interlocutori riescano in qualche modo a mettersi d'accordo sulla stessa chiave k.
- `Enc`: $\mathcal{K} \times \mathcal{M} \to \mathcal{C}$
- `Dec`: $\mathcal{K} \times \mathcal{C} \to \mathcal{M}$
- $\forall k \in \mathcal{K}, \forall m \in \mathcal{M} \quad \texttt{Dec}(k, \texttt{Enc}(k, m)) = m$

Se supponiamo che due interlocutori riescano a mettersi d'accordo su una chiave segreta in uno spazio $\mathcal{K}$ sufficientemente grande (per rendere gli attacchi a forza

bruta computazionalmente improponibili), risulta sufficientemente facile costruire metodi di crittografia simmetrica computazionalmente efficienti e difficilmente attaccabili (basti pensare al *One Time Pad*, che utilizzando solo uno `XOR` per cifrare e decifrare riesce ad essere *perfectly secure*). L'ovvio problema dietro questa supposizione è che è necessario che le due parti siano riuscite a mettersi precedentemente d'accordo (in modo sicuro) su una chiave. Risulta quindi utile l'avere a disposizione dei protocolli di scambio di chiavi (indicati anche con *KE* per *key exchange* o *key establishment*) che permettano a due parti di mettersi d'accordo su una chiave segreta comune attraverso un canale di comunicazione non sicuro.

Con protocollo di scambio di chiavi indichiamo una successione di computazioni e trasmissioni di messaggi attraverso i quali le due parti comunicanti riescano a mettersi d'accordo su un qualche valore (o qualche informazione) in modo che un avversario che intercetta il *transcript* non riesca a calcolare il valore scambiato a meno che con probabilità trascurabile (in funzione del livello di sicurezza garantito).

Un esempio classico di protocollo di scambio di chiavi è il *Diffie-Hellman Key Exchange* [2]. Il protocollo, che si basa sull'*assumption* del logaritmo discreto (che afferma che in alcuni gruppi sia computazionalmente difficile calcolare il logaritmo discreto), è così composto:



1. Alice e Bob stabiliscono come parametri pubblici un gruppo G ciclico (sul quale il logaritmo discreto sia difficile) e un generatore del gruppo g.

2. Alice sceglie in modo uniforme un esponente x, calcola l'elemento del gruppo $h_A = g^x$ e manda questo elemento a Bob. Simmetricamente, Bob sceglie in modo uniforme un esponente y, calcola l'elemento del gruppo $h_B = g^y$ e manda questo elemento ad Alice.

3. Alice calcola $k_A = h_B^x = g^{xy}$. Bob calcola $k_B = h_A^y = g^{xy}$. Alice e Bob riescono a mettersi d'accordo sull'elemento $k_A = k_B$.

Come vedremo nella prossima sezione, l'arrivo dei computer quantistici rende necessaria la ricerca di nuovi protocolli crittografici che si basino su *assumption* diverse.

2.3 Crittografia post-quantistica e RLWE

Molti dei protocolli crittografici utilizzati ad oggi si basano su *assumption* quali la fattorizzazione (la quale afferma che dato il prodotto di due numeri primi "grandi" è computazionalmente difficile risalire ai due primi) o il logaritmo discreto.

Tuttavia, grazie all'algoritmo di Shor [3] per la fattorizzazione, che necessita di un passaggio da eseguire su un computer quantistico, è possibile risolvere sia il problema della fattorizzazione sia il problema del logaritmo discreto in tempo polinomiale. Risulta quindi necessario sviluppare nuovi protocolli crittografici che basino la loro sicurezza su nuove *assumption* resistenti ad attacchi classici, quantistici e ibridi.

A questo proposito risulta di particolare interesse la crittografia basata sui reticoli. Introduciamo qualche concetto di base della crittografia sui reticoli che ci servirà per poter parlare della riduzione a caso peggiore dell'*assumption* che useremo.

Definizione 2.9. Fissato un insieme di vettori linearmente indipendenti $\mathcal{B} = \{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\}$ in $\mathbb{R}^n$ definiamo il reticolo generato da $\mathcal{B}$ come

$$L(\mathcal{B}) = \left\{ \sum_{i=1}^{n} a_i \boldsymbol{b}_i \;\mid\; a_i \in \mathbb{Z} \right\}$$

È importante osservare che basi diverse possono definire lo stesso reticolo, e la risoluzione di molti problemi dipende da quanto sia ben condizionata la base con cui decidiamo di caratterizzare il reticolo, dove il condizionamento di una base aumenta al crescere dalle norme dei vettori che la compongono. Dei problemi definibili sui reticoli ci interessa principalmente α-*approximate shortest vector problem* (α-*SVP*), che definiamo:

Definizione 2.10 (SVP)**.** Dato un reticolo L, lo *shortest vector problem* consiste nel trovare un vettore $\bar{v}$ del reticolo di norma minima, ovvero

$$\bar{v} \in L \quad \text{tc} \quad \|\bar{v}\| = \lambda_1 = \min_{v \in L} \|v\|$$

Definizione 2.11 (α-SVP)**.** Dato un reticolo L e $\alpha > 1$, l' α-*approximate shortest vector problem* consiste nel trovare un vettore $\bar{v}$ del reticolo di norma minore di α volte la norma minima nel reticolo, ovvero

$$\bar{v} \in L \quad \text{tc} \quad \|\bar{v}\| = \alpha\lambda_1$$

Grazie a M. Ajtai[7] sappiamo che *SVP* in norma 2 è *NP-hard* per basi randomizzate.

Considereremo ora il problema *Ring-Learning with errors* (*RLWE*), una restrizione agli anelli dei polinomi del problema generale *Learning with errors* (*LWE*) introdotto da O. Regev nel 2005 [4]. Diamo le seguenti definizioni:

Definizione 2.12 (LWE)**.** Fissati $q \in \mathbb{N}$ primo, $n \in \mathbb{N}$, ϕ una distribuzione di probabilità su $\mathbb{R}/\mathbb{Z}$ e $\boldsymbol{s} \in \mathbb{Z}_q^n$, definiamo una distribuzione su $\mathbb{Z}_q^n \times (\mathbb{R}/\mathbb{Z})$ data dal seguente algoritmo:

1. $\boldsymbol{a} \leftarrow_{\mathcal{U}} \mathbb{Z}_q^n$
2. $e \leftarrow_{\phi} \mathbb{R}/\mathbb{Z}$
3. $t = \langle \boldsymbol{a}, \boldsymbol{s} \rangle / q + e \in \mathbb{R}/\mathbb{Z}$
4. Output $(\boldsymbol{a}, t)$

Il problema *Learning with errors* consiste nel ricavare $\boldsymbol{s}$ dati una quantità polinomiale di campioni $(\boldsymbol{a}, t)$ così generati.

Definizione 2.13 (RLWE)**.** Fissati $q \in \mathbb{N}$ primo, $n \in \mathbb{N}$, ϕ una distribuzione limitata da b su $\mathbb{Z}_q[x]/\langle x^n + 1\rangle$ e $\boldsymbol{s} \leftarrow_{\phi} \mathbb{Z}_q[x]/\langle x^n + 1\rangle$, definiamo una distribuzione su $(\mathbb{Z}_q[x]/\langle x^n + 1\rangle)^2$ data dal seguente algoritmo:

1. $\boldsymbol{a} \leftarrow_{\mathcal{U}} \mathbb{Z}_q[x]/\langle x^n + 1\rangle$
2. $\boldsymbol{e} \leftarrow_{\phi} \mathbb{Z}_q[x]/\langle x^n + 1\rangle$
3. $\boldsymbol{t} = \boldsymbol{a} \cdot \boldsymbol{s} + \boldsymbol{e} \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$
4. Output $(\boldsymbol{a}, \boldsymbol{t})$

Il problema *Ring-Learning with errors* consiste nel ricavare $\boldsymbol{s}$ dati una quantità polinomiale di campioni $(\boldsymbol{a}, \boldsymbol{t})$ così generati.

La *assumption RLWE* afferma che la distribuzione degli $(\boldsymbol{a}, \boldsymbol{t})$ generati da *RLWE* è indistinguibile dalla distribuzione uniforme, ovvero che risolvere il problema *RLWE* è computazionalmente difficile. È stato recentemente dimostrato [5] che esiste una riduzione da *RLWE* nel caso medio a un'istanza di α-SVP nel caso peggiore.

Nel nostro caso utilizzeremo una versione leggermente modificata di *RLWE* in cui invece di considerare distribuzioni limitate utilizzeremo delle distribuzioni limitate eccetto che con probabilità trascurabile. Si dimostra [6] che il caso di distribuzioni limitate eccetto che con probabilità trascurabile si riduce al caso di distribuzioni limitate.

2.4 *Internet of Things*

Con contesto *Internet of Things* (*IoT*) intendiamo qualsiasi agglomerato di dispositivi che interagiscono e comunicano tra loro attraverso canali di comunicazione a breve distanza (ad esempio WiFi o Bluetooth) o canali di comunicazione a lunga distanza (l'Internet).

Questo tipo di contesto rappresenta l'avanguardia dell'automatizzazione moderna e può essere trovato in svariati ambiti e scale: da piccoli agglomerati di dispositivi *smart* nell'automatizzazione e la domotica delle case, a vaste quantità di sensori per la rilevazione e studio di dati fisici e ambientali. Uno dei concetti caratterizzanti per i dispositivi pensati per contesti *IoT* è la possibilità di essere sempre attivi o attivabili automaticamente, per poter creare sistemi che non necessitino di interventi umani. Per questa necessità e per efficienza dei consumi di questi, si tende a preferire dispositivi a richiesta energetica limitata (e quindi necessariamente a potenza computazionale limitata, crescente con gli sviluppi sulle tecnologie per processori a bassa potenza). I dispositivi impiegati in contesti *IoT* sono principalmente microcontrollori a bassa potenza.

Oltre ai limiti computazionali imposti dalla tipologia di dispositivi impiegati, un'altra problematica è data dalla tipologia di canali di comunicazioni utilizzati. Raramente i dispositivi utilizzati sono fisicamente collegati all'interlocutore (o ad un intermediario che gestisca le interazioni di più dispositivi): più comunemente (sia per le comunicazioni a breve distanza, sia per le comunicazioni a lunga distanza che premettono una comunicazione a breve distanza tra il dispositivo e l'intermediario) abbiamo comunicazioni che includono almeno un passaggio tra un canale di comunicazione meno affidabile, tipicamente radio. Se si utilizzassero dispositivi per cui l'utilizzo di potenza computazionale non fosse limitato, potremmo ovviare

al problema dell'affidabilità nelle comunicazioni tramite metodi di ridondanza e di correzione di errori. Poiché per i contesti IoT la regola è il risparmio energetico, una scelta tipica è quella di evitare di utilizzare metodi di rilevazione e correzione di errori nella trasmissione dei dati, riducendo così il consumo energetico dovuto alle relative computazioni e trasmissioni.

Sia la richiesta di efficienza computazionale che la richiesta di minimizzazione dell'informazione inviata sono in netto contrasto con le idee alla base dei metodi crittografici, caratterizzati dall'offuscamento dei messaggi tramite ridondanza e aggiunta di piccoli errori, e da algoritmi computazionalmente impegnativi, specialmente per la crittografia post-quantistica. L'idea di questa tesi è quella di sfruttare proprietà tipiche dei contesti IoT per provare a migliorare protocolli crittografici post-quantistici in suddetti contesti.

Come mostrato in [8] è possibile sfruttare il rilevamento di dati fisici fortemente dipendenti dalla posizione, come ad esempio la qualità del canale di comunicazione tra due dispositivi, per stabilire un piccolo segreto tra dispositivi comunicanti. Formalizzeremo e sfrutteremo questo concetto per diminuire i parametri necessari per la correttezza del protocollo di scambio di chiavi considerato, che contemporaneamente ridurrà la dimensione dei messaggi (e quindi dei pacchetti e dell'informazione trasmessa) e ridurrà la complessità delle operazioni da svolgere nel protocollo.

3 Protocollo di scambio di chiavi

3.1 Definizione

Introduciamo ora il protocollo di scambio di chiavi come riportato originariamente da [1]. Definiamo prima le seguenti funzioni:

Definizione 3.1 (Sig)**.** Definiamo le funzioni segnale $\texttt{Sig}_0 : \mathbb{Z}_q \to \mathbb{Z}_2$ e $\texttt{Sig}_1 : \mathbb{Z}_q \to \mathbb{Z}_2$ nel seguente modo:

$$\texttt{Sig}_0(a) = \begin{cases} 0 & \text{se } a \in \left\{-\lfloor \frac{q}{4} \rfloor, \ldots, \lfloor \frac{q}{4} \rfloor\right\} \\ 1 & \text{altrimenti} \end{cases}$$

$$\texttt{Sig}_1(a) = \begin{cases} 0 & \text{se } a \in \left\{-\lfloor \frac{q}{4} \rfloor + 1, \ldots, \lfloor \frac{q}{4} \rfloor + 1\right\} \\ 1 & \text{altrimenti} \end{cases}$$

Definiamo la versione randomizzata $\texttt{Sig}_* : \mathbb{Z}_q \to \mathbb{Z}_2$ come:

$$\texttt{Sig}_*(a) = \texttt{Sig}_{b \leftarrow_{\mathcal{U}} \{0,1\}}(a)$$

Definiamo infine l'estensione all'anello dei polinomi $\texttt{Sig} : \mathbb{Z}_q[x]/\langle x^n + 1\rangle \to \mathbb{Z}_2^n$ come:

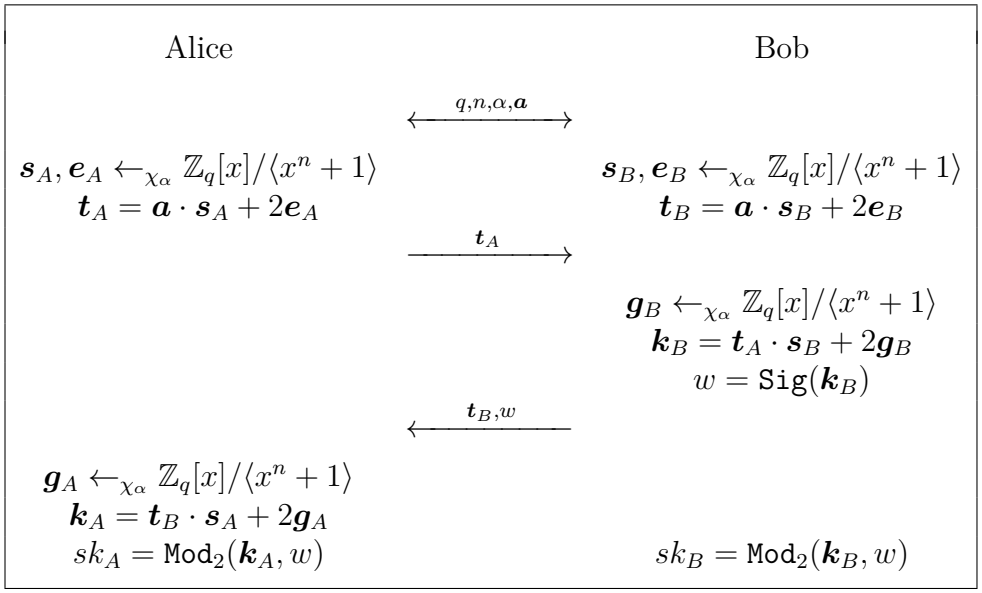
$$\texttt{Sig}(\boldsymbol{a}) = \texttt{Sig}(a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) = (\texttt{Sig}_*(a_{n-1}), \ldots, \texttt{Sig}_*(a_0))$$

Definizione 3.2 (Mod_2)**.** Definiamo l'estrattore $\texttt{Mod}_2 : (\mathbb{Z}_q[x]/\langle x^n+1\rangle) \times \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ definito componente per componente da:

$$\texttt{Mod}_{2,i}(\boldsymbol{a}, w) = \left\lfloor \left(a_i + w_i \frac{q-1}{2} \right) \mod q \right\rceil \mod 2$$

Utilizzeremo queste due funzioni per estrarre una chiave segreta condivisa da due valori segreti generati vicini, che per costruzione avranno scarto "pari" (più precisamente, lo scarto sarà un polinomio avente coefficienti pari). Il problema di base è che anche se $x, y \in \mathbb{Z}_q$ differiscono modulo q di una quantità pari piccola, non è detto che abbiano la stessa classe di equivalenza modulo 2 [1]. Come vedremo nel Lemma 3.1, lo scopo di queste due funzioni è di modificare i valori considerati garantendo che sommando un errore pari di norma piccola la classe di equivalenza modulo 2 non cambi.

Possiamo ora definire il protocollo di scambio di chiavi



1. Vengono concordati i parametri pubblici del protocollo $q \in \mathbb{N}$ primo, $n \in \mathbb{N}$, $\alpha > 0$ e $\boldsymbol{a} \leftarrow_{\mathcal{U}} \mathbb{Z}_q[x]/\langle x^n+1\rangle$.

2. Sia Alice che Bob generano una coppia di chiavi effimere $(\boldsymbol{s}_i, \boldsymbol{e}_i)$ da una distribuzione gaussiana centrata di deviazione standard α su $\mathbb{Z}_q[x]/\langle x^n+1\rangle$ e calcolano i rispettivi valori $\boldsymbol{t}_i = \boldsymbol{a}\cdot\boldsymbol{s}_i + 2\boldsymbol{e}_i$ con $i = A, B$.

3. Alice invia a Bob $\boldsymbol{t}_A$.

4. Bob usa $\boldsymbol{t}_A$ e un piccolo errore $\boldsymbol{g}_B \leftarrow_{\chi_\alpha} \mathbb{Z}_q[x]/\langle x^n+1\rangle$ generato per ottenere $\boldsymbol{k}_B = \boldsymbol{t}_A\cdot\boldsymbol{s}_B + 2\boldsymbol{g}_B$, di cui calcola il segnale $w = \mathtt{Sig}(\boldsymbol{k}_B)$.

5. Bob invia ad Alice $\boldsymbol{t}_B$ e w.

[1] Ad esempio, per $q = 7$ e $\mathbb{Z}_q = \{-3, \ldots, 3\}$, consideriamo $y = 3$ e $2\varepsilon = 2$. Allora $x := (y + 2\varepsilon) \bmod q = 5 \bmod q = -2$ e y differiscono in modulo q di una quantità pari piccola (la più piccola possibile non nulla), ma x è pari e y è dispari

6. Simmetricamente, Alice calcola $\boldsymbol{k}_A = \boldsymbol{t}_B \cdot \boldsymbol{s}_A + 2\boldsymbol{g}_A$ con $\boldsymbol{t}_B$ appena ricevuto e un piccolo errore $\boldsymbol{g}_A \leftarrow_{\chi_\alpha} \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ generato.

7. Alice e Bob calcolano le chiavi segrete sk_A e sk_B tramite l'estrattore $\mathtt{Mod}_2$ con i rispettivi $\boldsymbol{k}_i$ e il segnale di $\boldsymbol{k}_B$.

3.2 Correttezza

La convergenza delle chiavi sk_A e sk_B che garantisce la correttezza del protocollo dipende dai seguenti lemmi.

Lemma 3.1. *Sia $q > 8$ un intero dispari. Dati $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}_q[x]/\langle x^n+1\rangle$ e $w = \mathtt{Sig}(y)$, se esiste $\boldsymbol{\varepsilon} \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ tale che $\boldsymbol{x} = \boldsymbol{y} + 2\boldsymbol{\varepsilon} \mod q$, con $\|2\boldsymbol{\varepsilon}\|_\infty < \frac{q}{4} - 2$ allora*

$$\mathtt{Mod}_2(\boldsymbol{x}, w) = \mathtt{Mod}_2(\boldsymbol{y}, w)$$

Dimostrazione. Poiché $\|\cdot\|_\infty$, $\mathtt{Sig}$ e $\mathtt{Mod}_2$ agiscono componente per componente (identificando un polinomio con il vettore dei suoi coefficienti), basta dimostrare la tesi per le singoli componenti. Siano quindi $x_i, y_i \in \mathbb{Z}_q$ tali che $x_i = y_i + 2\varepsilon_i \mod q$ con $|2\varepsilon_i| < \frac{q}{4} - 2$.

Verifichiamo che $\left|\left(y_i + w_i \frac{q-1}{2}\right) \mod q\right| \leq \frac{q}{4} + 1$. Ricordiamoci che poiché $\mathtt{Sig}$ sulla singola componente usa $\mathtt{Sig}_*$ che è casuale, bisogna considerare i casi in cui w_i sia stato calcolato con $\mathtt{Sig}_0$ o con $\mathtt{Sig}_1$. Facciamo questa distinzione utilizzando un parametro $b \in \{0, 1\}$.

- Se $w_i = \mathtt{Sig}_b(y_i) = 0$, allora vale $y_i \in \left\{-\lfloor \frac{q}{4} \rfloor + b, \ldots, \lfloor \frac{q}{4} \rfloor + b\right\}$, da cui

$$\left|\left(y_i + w_i \frac{q-1}{2}\right) \mod q\right| = |y_i \mod q| = |y_i| \leq \frac{q}{4} + 1$$

- Se $w_i = \mathtt{Sig}_b(y_i) = 1$, allora vale $y_i \notin \left\{-\lfloor \frac{q}{4} \rfloor + b, \ldots, \lfloor \frac{q}{4} \rfloor + b\right\}$, da cui
  - Se $y_i \in \left\{-\frac{q-1}{2}, \ldots, -\lfloor \frac{q}{4} \rfloor - 1 + b\right\}$ allora

$$\left(y_i + \frac{q-1}{2}\right) \mod q = y_i + \frac{q-1}{2}$$

    e vale

$$\begin{aligned} 0 \leq y_i + \frac{q-1}{2} &\leq -\left\lfloor \frac{q}{4} \right\rfloor - 1 + b + \frac{q-1}{2} \leq \\ &\leq \frac{q-1}{2} - \frac{q}{4} + b = \frac{q}{4} - \frac{1}{2} + b \leq \\ &\leq \frac{q}{4} + 1 \end{aligned}$$

da cui

$$\left| y_i + \frac{q-1}{2} \right| \leq \frac{q}{4} + 1$$

– Se $y_i \in \left\{ \lfloor \frac{q}{4} \rfloor + 1 + b, \ldots, \frac{q-1}{2} \right\}$ allora

$$\left(y_i + \frac{q-1}{2} \right) \mod q = y_i + \frac{q-1}{2} - q$$

e vale

$$\begin{aligned} -1 \geq y_i + \frac{q-1}{2} - q \geq \left\lfloor \frac{q}{4} \right\rfloor + 1 + b + \frac{q-1}{2} - q \geq \\ \geq \frac{q}{4} + b + \frac{q-1}{2} - q = -\frac{q}{4} - \frac{1}{2} + b \geq \\ \geq -\frac{q}{4} - 1 \end{aligned}$$

da cui

$$\left| y_i + \frac{q-1}{2} \right| \leq \frac{q}{4} + 1$$

Quindi in ogni caso abbiamo la maggiorazione $\left| y_i + w_i \frac{q-1}{2} \mod q \right| \leq \frac{q}{4} + 1$. Osserviamo che poiché vale

$$\left| \left[\left(y_i + w_i \frac{q-1}{2} \right) \mod q \right] + 2\varepsilon_i \right| \leq \left| \left(y_i + w_i \frac{q-1}{2} \right) \mod q \right| + |2\varepsilon_i| \leq \frac{q-1}{2}$$

vale

$$\begin{aligned} \left(x_i + w_i \frac{q-1}{2} \right) \mod q = \left(y_i + w_i \frac{q-1}{2} + 2\varepsilon_i \right) \mod q = \\ = \left[\left(y_i + w_i \frac{q-1}{2} \right) \mod q \right] + 2\varepsilon_i \end{aligned}$$

e quindi abbiamo

$$\begin{aligned} \mathtt{Mod}_{2,i}(\boldsymbol{x}, w) &= \left[\left(x_i + w_i \frac{q-1}{2} \right) \mod q \right] \mod 2 = \\ &= \left[\left(y_i + w_i \frac{q-1}{2} \right) \mod q \right] \mod 2 = \\ &= \mathtt{Mod}_{2,i}(\boldsymbol{y}, w) \end{aligned}$$

e quindi la tesi. □

Lemma 3.2. *Se $q > 16\alpha^2 n\sqrt{n}+16\alpha\sqrt{n}+8$ allora le chiavi generate dal protocollo convergono eccetto che con probabilità trascurabile*

Dimostrazione. Osserviamo che

$$\begin{aligned}\boldsymbol{k}_A &= \boldsymbol{a}\boldsymbol{s}_A\boldsymbol{s}_B + 2\boldsymbol{e}_B\boldsymbol{s}_A + 2\boldsymbol{g}_A\\ \boldsymbol{k}_B &= \boldsymbol{a}\boldsymbol{s}_A\boldsymbol{s}_B + 2\boldsymbol{e}_A\boldsymbol{s}_B + 2\boldsymbol{g}_B\end{aligned}$$

ovvero

$$\boldsymbol{k}_A - \boldsymbol{k}_B = 2(\boldsymbol{e}_B\boldsymbol{s}_A + \boldsymbol{g}_A - \boldsymbol{e}_A\boldsymbol{s}_B - \boldsymbol{g}_B) = 2\boldsymbol{\varepsilon}$$

Per quanto riguarda la norma di $2\boldsymbol{\varepsilon}$ vale

$$\begin{aligned}\|2\boldsymbol{\varepsilon}\|_\infty &= \|2(\boldsymbol{e}_B\boldsymbol{s}_A + \boldsymbol{g}_A - \boldsymbol{e}_A\boldsymbol{s}_B - \boldsymbol{g}_B)\|_\infty \leq\\ &\leq \|2(\boldsymbol{e}_B\boldsymbol{s}_A + \boldsymbol{g}_A - \boldsymbol{e}_A\boldsymbol{s}_B - \boldsymbol{g}_B)\|_2 \leq\\ &\leq 2\left[\|\boldsymbol{e}_B\boldsymbol{s}_A\|_2 + \|\boldsymbol{e}_A\boldsymbol{s}_B\|_2 + \|\boldsymbol{g}_A\|_2 + \|\boldsymbol{g}_B\|_2\right] \leq\\ &\leq 2\left[\sqrt{n}\|\boldsymbol{e}_B\|_2\|\boldsymbol{s}_A\|_2 + \sqrt{n}\|\boldsymbol{e}_A\|_2\|\boldsymbol{s}_B\|_2 + \|\boldsymbol{g}_A\|_2 + \|\boldsymbol{g}_B\|_2\right]\end{aligned}$$

dove abbiamo utilizzato la disuguaglianza del Lemma 2.1.
Poiché la somma finita di funzioni trascurabili è trascurabile, possiamo stimare contemporaneamente tutte le norme con la stima dell'osservazione 2.2 eccetto che con probabilità trascurabile, ottenendo

$$\begin{aligned}\|2\boldsymbol{\varepsilon}\|_\infty &\leq 2\left[\sqrt{n}\|\boldsymbol{e}_B\|_2\|\boldsymbol{s}_A\|_2 + \sqrt{n}\|\boldsymbol{e}_A\|_2\|\boldsymbol{s}_B\|_2 + \|\boldsymbol{g}_A\|_2 + \|\boldsymbol{g}_B\|_2\right] \leq\\ &\leq 2\left[\sqrt{n}\alpha\sqrt{n}\alpha\sqrt{n} + \sqrt{n}\alpha\sqrt{n}\alpha\sqrt{n} + \alpha\sqrt{n} + \alpha\sqrt{n}\right] =\\ &= 4\alpha^2 n\sqrt{n} + 4\alpha\sqrt{n} <\\ &< \frac{q}{4} - 2\end{aligned}$$

Utilizzando il lemma precedente, abbiamo che $sk_A = sk_B$ eccetto che con probabilità trascurabile. □

3.3 Sicurezza

Il transcript del protocollo è dato dalla tripla $(\boldsymbol{t}_A, \boldsymbol{t}_B, w)$. Per stabilire se il protocollo è sicuro, definiamo un gioco in cui immaginiamo di eseguire un'istanza del protocollo fornendo il transcript ad un avversario, estraiamo un bit $b \leftarrow_{\mathcal{U}} \{0,1\}$ e se $b = 0$ forniamo all'avversario la chiave segreta, se $b = 1$ forniamo un valore casuale uniforme. Diciamo che l'avversario vince se riesce ad indovinare il bit b, ovvero se riesce a capire dal transcript se gli è stata fornita l'effettiva chiave segreta o un valore casuale.

L'idea che lega questa definizione all'effettiva sicurezza del protocollo è che se un avversario riuscisse a calcolare l'effettiva chiave dal transcript riuscirebbe a distinguere la chiave vera dal valore casuale e riuscirebbe quindi ad indovinare il bit generato uniformemente. Otteniamo in realtà una definizione di sicurezza più forte, poiché soddisfacendo questa definizione riusciamo a difenderci anche da avversari che riescono ad ottenere solo qualche informazione parziale sulla chiave e non la chiave completa.

Chiaramente un avversario riesce a garantire come probabilità di vittoria almeno $\frac{1}{2}$, restituendo come possibile valore di b un altro bit casuale. Diciamo che il protocollo è sicuro se ogni avversario a tempo polinomiale riesce a vincere al più con probabilità $\frac{1}{2}$ più uno scarto trascurabile.

Chiamiamo $\mathrm{KE}_{q,n,\alpha,\boldsymbol{a}}$ (o più semplicemente KE) la distribuzione dei valori generati da un protocollo di scambio di chiavi con parametri pubblici q, n, α e $\boldsymbol{a}$. Quando scriviamo $x \leftarrow \mathrm{KE}$ con x una tupla di valori del protocollo intendiamo la generazione di quei parametri di una istanza del protocollo (ad esempio con $(\boldsymbol{t}_A, \boldsymbol{t}_B, w) \leftarrow KE$ intendiamo il generare una istanza del protocollo e restituirne il transcript; la scelta della tupla di valori è una scelta di notazione per poter scrivere solo i valori che effettivamente intendiamo utilizzare). Chiamiamo avversario a tempo polinomiale probabilistico (PPT) un algoritmo $\mathcal{A}$ eseguibile in tempo polinomiale tale che

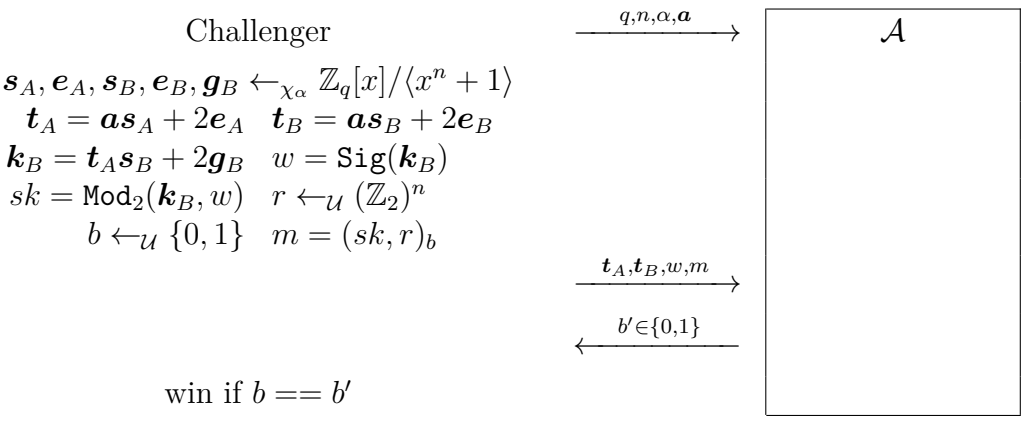
$$\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) \in \{0,1\} \quad \text{con } (\boldsymbol{t}_A, \boldsymbol{t}_B, w) \leftarrow \mathrm{KE},\ sk \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$$

Dati due valori x e y, indichiamo con $(x,y)_0 = x$ e $(x,y)_1 = y$, ovvero indichiamo con $(x,y)_b$ la funzione che dato x, y e b restituisce x se $b = 0$ e y se $b = 1$. Definiamo come vantaggio di un avversario $\mathcal{A}$ la quantità

$$\mathrm{Adv}_{\mathcal{A}} = \left| \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, (sk, r)_b) = b\right] - \frac{1}{2} \right|, \quad \text{con} \quad \begin{array}{l} (\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) \leftarrow \mathrm{KE} \\ r \leftarrow_{\mathcal{U}} \mathbb{Z}_q[x]/\langle x^n + 1\rangle \\ b \leftarrow_{\mathcal{U}} \{0,1\} \end{array}$$

Diciamo che il protocollo è sicuro se per ogni avversario PPT $\mathcal{A}$ il vantaggio di $\mathcal{A}$ è trascurabile in n.

Schematicamente, definiamo come gioco di sicurezza il seguente gioco



dove la probabilità di vittoria di $\mathcal{A}$ è $\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, (sk, r)_b) = b\right]$, e chiedere che il protocollo sia sicuro equivale a chiedere che per ogni avversario PPT $\mathcal{A}$, la probabilità di vittoria di A disti da $\frac{1}{2}$ di una quantità trascurabile.

Osserviamo che

$$
\begin{aligned}
&\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, (sk, r)_b) = b\right] = \\
&= \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right]\mathbb{P}\left[b = 0\right] + \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 1\right]\mathbb{P}\left[b = 1\right] = \\
&= \frac{1}{2}\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right] + \frac{1}{2}\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 1\right] = \\
&= \frac{1}{2}\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right] + \frac{1}{2} - \frac{1}{2}\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 0\right] = \\
&= \frac{1}{2} + \frac{1}{2}\left(\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right] - \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 0\right]\right) \\
&= \frac{1}{2} + \frac{1}{2}\left(\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 1\right] - \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 1\right]\right)
\end{aligned}
$$

quindi chiedere che il protocollo sia sicuro equivale a chiedere che la quantità

$$\left|\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right] - \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 0\right]\right|$$

(o equivalentemente $\left|\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 1\right] - \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 1\right]\right|$) sia trascurabile. Mostriamo quindi che sotto *assumption RLWE*, per ogni avversario PPT $\mathcal{A}$ questa quantità è trascurabile.

Il metodo di dimostrazione si basa sul creare una successione di giochi $G_1, \ldots, G_8$, uno indistinguibile dal successivo sotto *RLWE assumption*, dove in G_1 forniamo all'avversario $(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk)$, e in G_8 forniamo all'avversario $(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r)$. Allora

per assurdo se esistesse un avversario $\mathcal{A}$ per il quale la quantità che stiamo considerando fosse non trascurabile, questo avversario riuscirebbe a distinguere G_1 e G_8 con probabilità non trascurabile. Esisterebbe allora $i \in \{1, \dots, 7\}$ per cui sia possibile costruire un avversario che utilizzando internamente $\mathcal{A}$ riesca a distinguere i giochi G_i e G_{i+1}, cosa impossibile sotto *RLWE assumption* per costruzione.

Prima di esibire la successione di giochi, dimostriamo il seguente lemma necessari per la dimostrazione di indistinguibilità di quelli che saranno i giochi G_4 e G_5.

Lemma 3.3. *La distribuzione di* $\texttt{Mod}_2(\boldsymbol{a}, \texttt{Sig}(\boldsymbol{a}))$ *noto* $\texttt{Sig}(\boldsymbol{a})$ *con* $\boldsymbol{a} \leftarrow_{\mathcal{U}} \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ *è indistinguibile dalla distribuzione uniforme su* $(\mathbb{Z}_2)^n$.

Dimostrazione. Poiché $\texttt{Mod}_2$ e $\texttt{Sig}$ agiscono componente per componente e le singole componenti (ovvero i coefficienti) sono indipendenti, basta verificare che per una singola componente vale

$$\forall b, c \in \{0,1\} \quad \mathbb{P}_{a_i \leftarrow_{\mathcal{U}} \mathbb{Z}_q}\left[\texttt{Mod}_{2,i}(a_i, b) = c \mid \texttt{Sig}_*(a_i) = b\right] = \frac{1}{2}$$

Per definizione di $\texttt{Sig}_*$, utilizzando $\mathbb{P}_{a_i \leftarrow_{\mathcal{U}} \mathbb{Z}_q}\left[\texttt{Sig}_0(a_i) = b\right] = \mathbb{P}_{a_i \leftarrow_{\mathcal{U}} \mathbb{Z}_q}\left[\texttt{Sig}_1(a_i) = b\right]$, abbiamo (omettendo il pedice $a_i \leftarrow_{\mathcal{U}} \mathbb{Z}_q$ per semplicità di notazione)

$$\begin{aligned}
&\mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c \mid \texttt{Sig}_*(a_i) = b\right] = \mathbb{P}_{d \leftarrow_{\mathcal{U}} \{0,1\}}\left[\texttt{Mod}_{2,i}(a_i, b) = c \mid \texttt{Sig}_d(a_i) = b\right] = \\
&= \frac{1}{2}\mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c \mid \texttt{Sig}_0(a_i) = b\right] + \frac{1}{2}\mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c \mid \texttt{Sig}_1(a_i) = b\right] = \\
&= \frac{1}{2}\frac{\mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c,\ \texttt{Sig}_0(a_i) = b\right]}{\mathbb{P}\left[\texttt{Sig}_0(a_i) = b\right]} + \frac{1}{2}\frac{\mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c,\ \texttt{Sig}_1(a_i) = b\right]}{\mathbb{P}\left[\texttt{Sig}_1(a_i) = b\right]} = \\
&= \frac{1}{2}\frac{\mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c,\ \texttt{Sig}_0(a_i) = b\right] + \mathbb{P}\left[\texttt{Mod}_{2,i}(a_i, b) = c,\ \texttt{Sig}_1(a_i) = b\right]}{\mathbb{P}\left[\texttt{Sig}_0(a_i) = b\right]}
\end{aligned}$$

Per il caso $b = 0$, indichiamo $I = \{-\frac{q-1}{2}, \dots, \frac{q-1}{2}\}$, $I_0 = \{x \in I \mid x \mod 2 = 0\}$, $I_1 = \{x \in I \mid x \mod 2 = 1\}$ e in modo analogo $(I+1)_0$ e $(I+1)_1$.

Osserviamo che, valendo $I = I_c \cup I_{1-c}$ ed essendo I_{1-c} in bigezione con $(I+1)_c$,

vale $|I_c| + |(I+1)_c| = |I|$ e quindi

$$\begin{aligned}
&\mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 0) = c \mid \mathtt{Sig}_*(a_i) = 0\right] = \\
&= \frac{1}{2}\frac{\mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 0) = c,\ \mathtt{Sig}_0(a_i) = 0\right] + \mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 0) = c,\ \mathtt{Sig}_1(a_i) = 0\right]}{\mathbb{P}\left[\mathtt{Sig}_0(a_i) = 0\right]} = \\
&= \frac{1}{2}\frac{\mathbb{P}\left[a_i \in I_c\right] + \mathbb{P}\left[a_i \in (I+1)_c\right]}{\mathbb{P}\left[a_i \in I\right]} = \\
&= \frac{1}{2}\frac{\frac{|I_c|}{q} + \frac{|(I+1)_c|}{q}}{\frac{|I|}{q}} = \frac{1}{2}\frac{|I_c| + |(I+1)_c|}{|I|} = \\
&= \frac{1}{2}
\end{aligned}$$

Per il caso $b = 1$ osserviamo che poiché $\mathbb{Z}_q \setminus I$ ha cardinalità pari, abbiamo

$$\begin{aligned}
&\mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 1) = c,\ \mathtt{Sig}_0(a_i) = 1\right] = \mathbb{P}\left[a_i \in (\mathbb{Z}_q \setminus I)_c\right] = \frac{|(\mathbb{Z}_q \setminus I)_c|}{q} = \\
&= \frac{|(\mathbb{Z}_q \setminus I)_{1-c}|}{q} = \mathbb{P}\left[a_i \in (\mathbb{Z}_q \setminus I)_{1-c}\right] = \mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 1) = 1 - c,\ \mathtt{Sig}_0(a_i) = 1\right]
\end{aligned}$$

da cui

$$\mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 1) = c,\ \mathtt{Sig}_0(a_i) = 1\right] = \frac{1}{2}\mathbb{P}\left[\mathtt{Sig}_0(a_i) = 1\right]$$

e analogamente

$$\mathbb{P}\left[\mathtt{Mod}_{2,i}(a_i, 1) = c,\ \mathtt{Sig}_1(a_i) = 1\right] = \frac{1}{2}\mathbb{P}\left[\mathtt{Sig}_1(a_i) = 1\right] = \frac{1}{2}\mathbb{P}\left[\mathtt{Sig}_0(a_i) = 1\right]$$

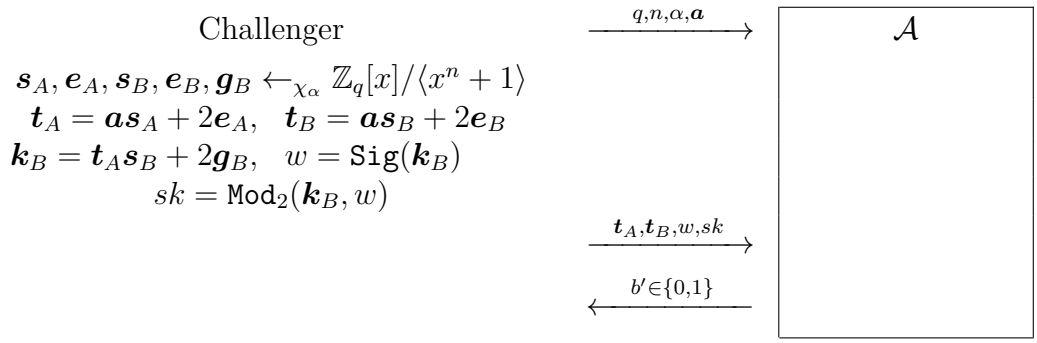
Sostituendo nel risultato precedente al caso $b = 0$ otteniamo la tesi. □

Possiamo ora dare la dimostrazione di sicurezza del protocollo

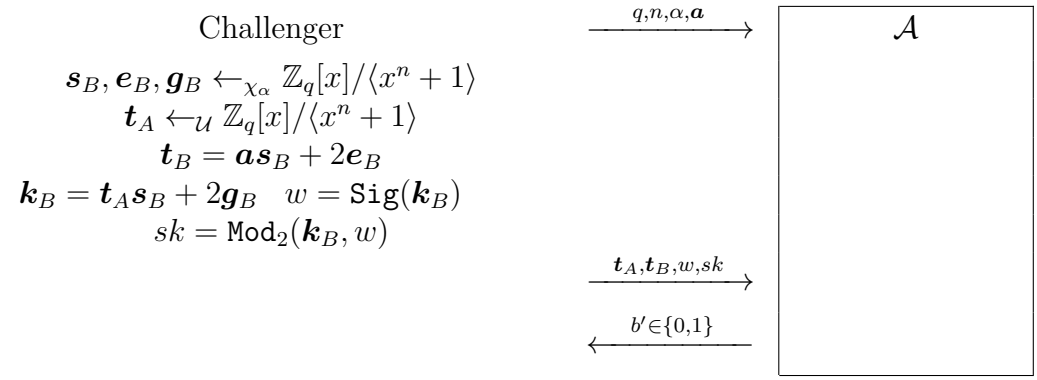
Lemma 3.4. *Se vale la RLWE assumption, ogni avversario PPT può avere vantaggio al più trascurabile al gioco di sicurezza del protocollo.*

Dimostrazione. Cominciamo esibendo la seguente successione di giochi, ognuno indistinguibile dal precedente:

G_1

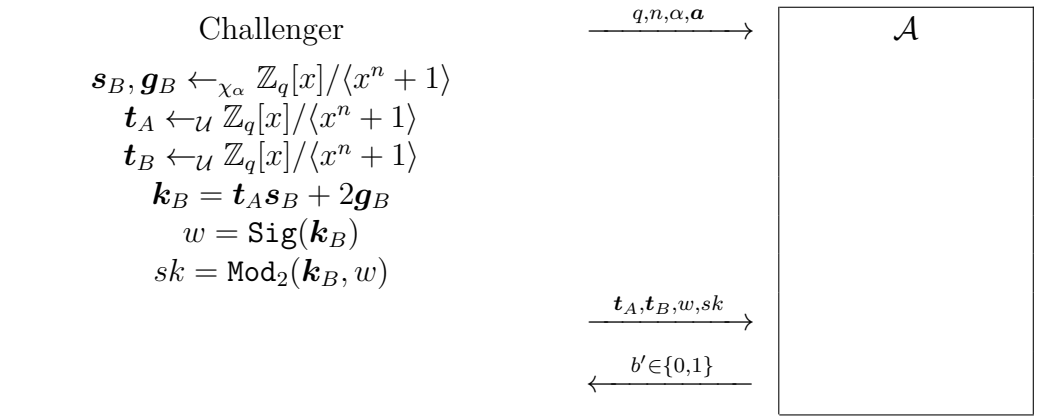


G_2



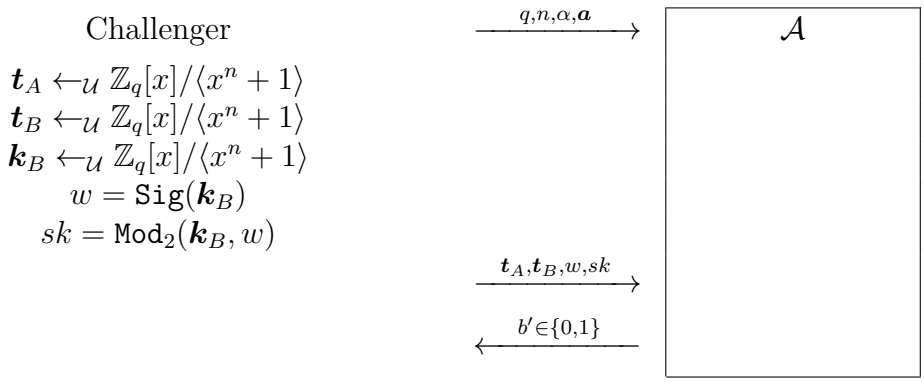
G_2 è indistinguibile da G_1 poiché per *RLWE assumption* $\boldsymbol{t}_A = \boldsymbol{a}\boldsymbol{s}_A + 2\boldsymbol{e}_A$ è indistinguibile da $\boldsymbol{t}_A$ uniforme.

G_3



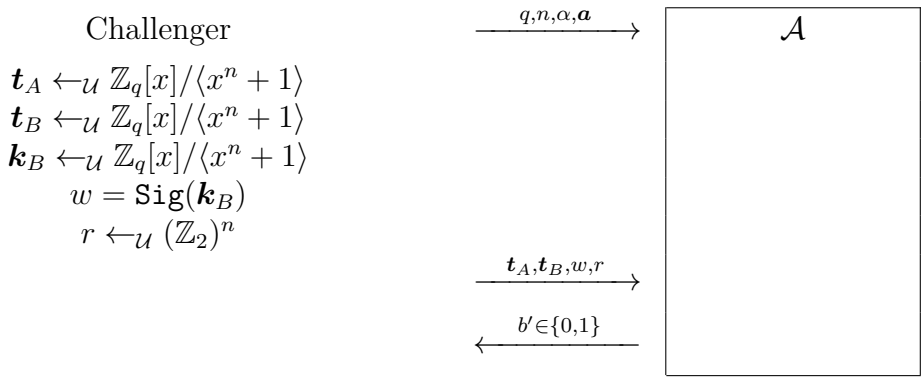
G_3 è indistinguibile da G_2 poiché per *RLWE assumption* $\boldsymbol{t}_B = \boldsymbol{a}\boldsymbol{s}_B + 2\boldsymbol{e}_B$ è indistinguibile da $\boldsymbol{t}_B$ uniforme.

G_4



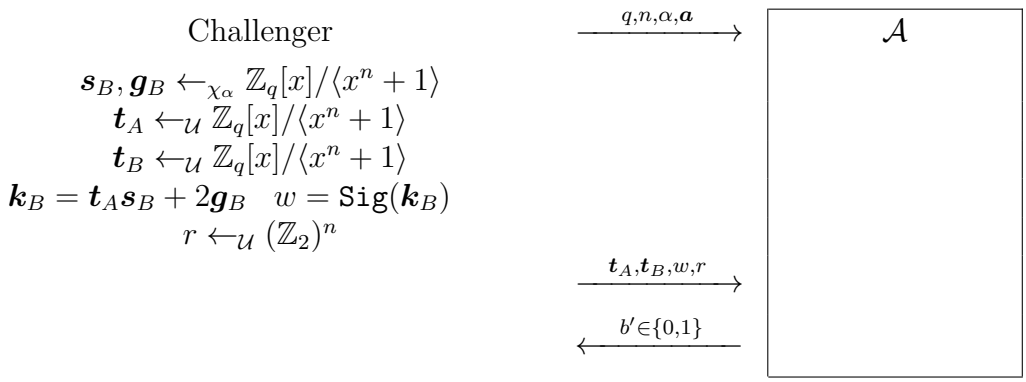
G_4 è indistinguibile da G_3 poiché per *RLWE assumption* $\boldsymbol{k}_B = \boldsymbol{t}_A\boldsymbol{s}_B + 2\boldsymbol{g}_B$ è indistinguibile da $\boldsymbol{k}_B$ uniforme.

G_5



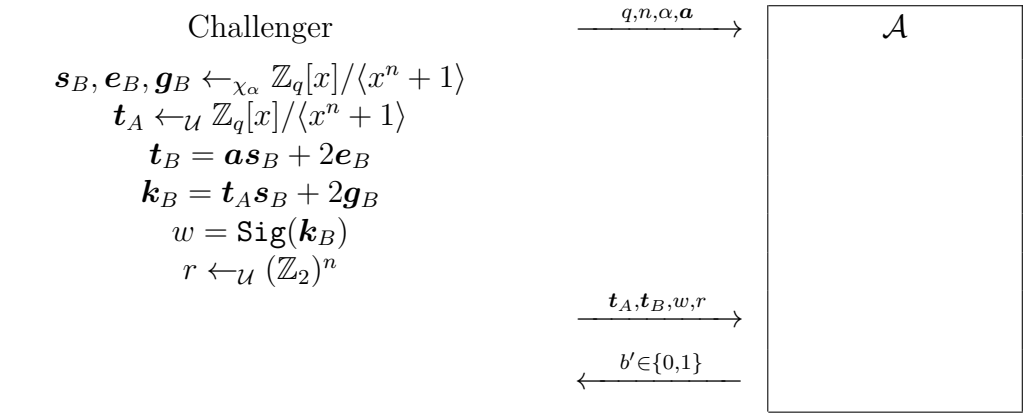
G_5 è indistinguibile da G_4 poiché per Lemma 3.3 $sk = \mathtt{Mod}_2(\boldsymbol{k}_B, \mathtt{Sig}(\boldsymbol{k}_B))$ con $\boldsymbol{k}_B$ uniforme e $\mathtt{Sig}(\boldsymbol{k}_B)$ noto è indistinguibile da r uniforme.

G_6



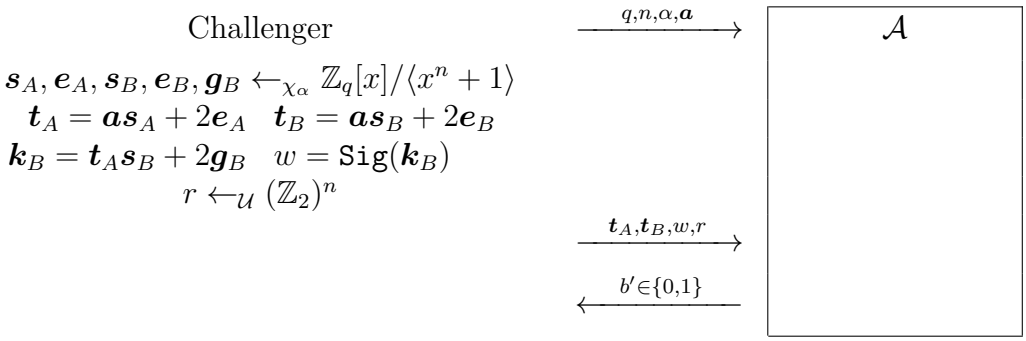
G_6 è indistinguibile da G_5 poiché per *RLWE assumption* $\boldsymbol{k}_B$ uniforme è indistinguibile da $\boldsymbol{k}_B = \boldsymbol{t}_A\boldsymbol{s}_B + 2\boldsymbol{g}_B$.

G_7



G_7 è indistinguibile da G_6 poiché per *RLWE assumption* $\boldsymbol{t}_B$ uniforme è indistinguibile da $\boldsymbol{t}_B = \boldsymbol{a}\boldsymbol{s}_B + 2\boldsymbol{e}_B$.

G_8



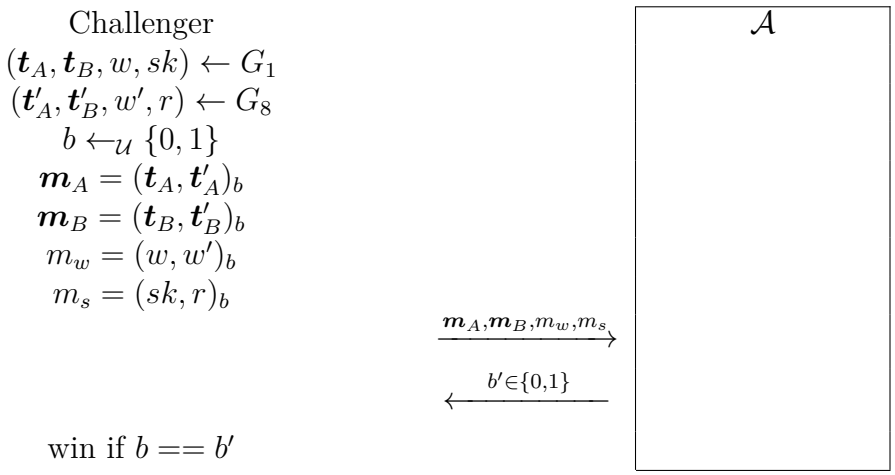
G_8 è indistinguibile da G_7 poiché per *RLWE assumption* $\boldsymbol{t}_B$ uniforme è indistinguibile da $\boldsymbol{t}_A = \boldsymbol{a}\boldsymbol{s}_A + 2\boldsymbol{e}_A$.

A questo punto $\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right]$ corrisponde alla probabilità che $\mathcal{A}$ restituisca 0 al gioco G_1, e $\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 0\right]$ corrisponde alla probabilità che $\mathcal{A}$ restituisca 0 al gioco G_8. Ricordiamo che avendo dimostrato che G_i è indistinguibile da G_{i+1} per $i = 1, \ldots, 7$, per transitività abbiamo che G_1 è indistinguibile da G_8.

Supponiamo per assurdo che queste probabilità differiscano di una quantità non trascurabile. Sia quindi senza perdita di generalità

$$\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right] = \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 0\right] + v(n)$$

Mostriamo che $\mathcal{A}$ riesce a distinguere i giochi G_1 e G_8. Per farlo, consideriamo il seguente gioco:



Allora il vantaggio con cui $\mathcal{A}$ riesce a restituire il bit corretto (e quindi distinguere i due giochi) è proprio $v(n)$, che per ipotesi assurda era non trascurabile, quindi $\mathcal{A}$ riesce a distinguere G_1 e G_8 che avevamo dimostrato essere indistinguibili, che è assurdo.

Allora la quantità

$$\left|\mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, sk) = 0\right] - \mathbb{P}\left[\mathcal{A}(\boldsymbol{t}_A, \boldsymbol{t}_B, w, r) = 0\right]\right|$$

è trascurabile e per quanto osservato precedentemente lo è anche il vantaggio dell'avversario $\mathcal{A}$ nel gioco di sicurezza, quindi il protocollo è sicuro sotto *RLWE assumption*. □

4 Manipolazione degli errori

4.1 Definizione

Modifichiamo adesso il protocollo per abbassare la stima del Lemma 3.2 per il parametro q, aggiungendo un'ipotesi che possiamo supporre vera in contesti IoT. Vogliamo sfruttare qualche valore su cui i due interlocutori possano mettersi d'accordo per sostituire parte degli errori di norma limitata introdotti nel protocollo per ridursi al problema *RLWE*. Possiamo immaginare che questo valore provenga da qualche fenomeno fisico che dipenda da certe proprietà non ripetibili dei dispositivi, come la qualità di un canale radio di comunicazione può dipendere dalla posizione fisica dei dispositivi. A meno di esplicitare una mappa di codifica dallo spazio dei valori rilevati a $\mathbb{Z}_q[x]/\langle x^n + 1\rangle$, possiamo supporre che i valori rilevati siano già polinomi nel quoziente. La sicurezza di questo nuovo protocollo dipende necessariamente dalla segretezza di questo valore. Per dimostrarla considereremo un modello in cui supponiamo che questo valore sia stimabile solo grossolanamente da avversari passivi.

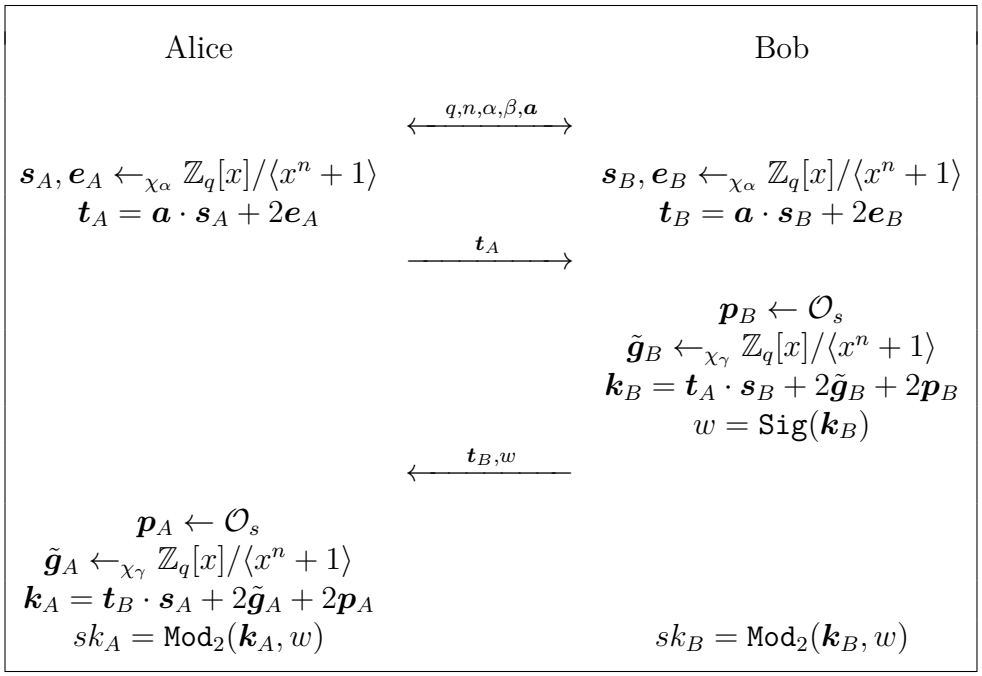
Formalmente, supponiamo che esista una famiglia di oracoli $(\mathcal{O}_s)_{s\in[0,1]}$ con $\mathcal{O}_s : () \to \mathbb{Z}_q[x]/\langle x^n+1\rangle$ tali per cui la distribuzione di $x \leftarrow_{\mathcal{O}_{s\leftarrow\mathcal{U}[0,1]}} \mathbb{Z}_q[x]/\langle x^n+1\rangle$ sia gaussiana centrata di deviazione standard $\beta > 0$, ovvero χ_β, e che per s fissato la distribuzione $x \leftarrow_{\mathcal{O}_s} \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ sia uniforme su una palla con centro in un certo $\boldsymbol{p}_s$ di raggio $\delta/2$ con $0 < \delta < \beta$.
Supponiamo che Alice e Bob abbiano entrambi a disposizione l'oracolo $\mathcal{O}_s$ con $s \leftarrow_{\mathcal{U}} [0, 1]$ fissato che cambia ad ogni istanza del protocollo.

In questa formalizzazione, χ_β è la distribuzione generica del valore della quantità fisica che stiamo misurando senza informazioni aggiuntive, $\boldsymbol{p}_s$ è l'effettivo valore della quantità fisica nella specifica istanza del protocollo e $\delta/2$ è la precisione con cui Alice e Bob riescono a determinare questo valore. Poiché solo Alice e Bob hanno a disposizione lo specifico oracolo $\mathcal{O}_s$, l'unico modo che ha un avversario passivo

per stimare i valori generati tramite questo oracolo è utilizzando la distribuzione nota χ_β.

La modifica che intendiamo fare è andare a sostituire gli errori $\boldsymbol{g}_A$ e $\boldsymbol{g}_B$ del protocollo con i valori ottenuti invocando l'oracolo. Se avessimo $\beta \geq \alpha$ potremmo sostituire direttamente questi errori con i valori dell'oracolo mantenendo la stessa distribuzione degli errori. Tuttavia, questa supposizione non è sempre lecita poiché potrebbe essere che il fenomeno fisico che stiamo osservando non sia soggetto ad assumere valori con una variazione sufficientemente ampia. Nel caso di $\beta < \alpha$ vogliamo ampliare l'incertezza sui valori ottenuti dall'oracolo sommando un altro errore da una distribuzione gaussiana di deviazione standard $\gamma := \sqrt{\alpha^2 - \beta^2}$ per raggiungere la varianza desiderata.

Definiamo quindi il protocollo modificato (nel caso di $\alpha > \beta$) come



1. Vengono concordati i parametri pubblici del protocollo $q \in \mathbb{N}$ primo, $n \in \mathbb{N}$, $\alpha > 0, \beta > 0$ e $\boldsymbol{a} \leftarrow_{\mathcal{U}} \mathbb{Z}_q[x]/\langle x^n+1\rangle$.

2. Sia Alice che Bob generano una coppia di chiavi effimere $(\boldsymbol{s}_i, \boldsymbol{e}_i) \leftarrow_{\chi_\alpha} \mathbb{Z}_q[x]/\langle x^n+1\rangle$ e calcolano i rispettivi valori $\boldsymbol{t}_i = \boldsymbol{a} \cdot \boldsymbol{s}_i + 2\boldsymbol{e}_i$.

3. Alice invia a Bob $\boldsymbol{t}_A$.

4. Bob invoca l'oracolo ottenendo $\boldsymbol{p}_B$ e genera un piccolo errore $\tilde{\boldsymbol{g}}_B \leftarrow_{\chi_\gamma} \mathbb{Z}_q[x]/\langle x^n+1\rangle$. Usa questi due valori per ottenere $\boldsymbol{k}_B = \boldsymbol{t}_A \cdot \boldsymbol{s}_A + 2\tilde{\boldsymbol{g}}_B + 2\boldsymbol{p}_B$, di cui calcola il segnale $w = \mathtt{Sig}(\boldsymbol{k}_B)$.

5. Bob invia ad Alice $\boldsymbol{t}_B$ e w.

6. Simmetricamente, Alice calcola $\boldsymbol{k}_A = \boldsymbol{t}_B \cdot \boldsymbol{s}_A + 2\tilde{\boldsymbol{g}}_A + 2\boldsymbol{p}_A$ con $\boldsymbol{t}_B$ appena ricevuto, un piccolo errore $\tilde{\boldsymbol{g}}_A \leftarrow_{\chi_\gamma} \mathbb{Z}_q[x]/\langle x^n+1\rangle$ generato e il valore $\boldsymbol{p}_A$ ottenuto dall'oracolo.

7. Alice e Bob calcolano le chiavi segrete sk_A e sk_B tramite l'estrattore $\mathtt{Mod}_2$ con i rispettivi $\boldsymbol{k}_i$ e il segnale di $\boldsymbol{k}_B$.

4.2 Correttezza

La convergenza delle chiavi sk_A e sk_B che garantisce la correttezza del protocollo modificato segue da una modificazione del Lemma 3.2, ovvero

Lemma 4.1. *Se $q > 16\alpha^2 n\sqrt{n} + 16\gamma\sqrt{n} + 8\delta + 8$ allora le chiavi generate dal protocollo convergono eccetto che con probabilità trascurabile.*

Dimostrazione. La dimostrazione è simile a quella del Lemma 3.2. In questo caso osserviamo che poiché

$$\begin{aligned}\boldsymbol{k}_A &= \boldsymbol{a}\boldsymbol{s}_A\boldsymbol{s}_B + 2\boldsymbol{e}_B\boldsymbol{s}_A + 2\tilde{\boldsymbol{g}}_A + 2\boldsymbol{p}_A\\ \boldsymbol{k}_B &= \boldsymbol{a}\boldsymbol{s}_A\boldsymbol{s}_B + 2\boldsymbol{e}_A\boldsymbol{s}_B + 2\tilde{\boldsymbol{g}}_B + 2\boldsymbol{p}_B\end{aligned}$$

vale

$$\boldsymbol{k}_A - \boldsymbol{k}_B = 2(\boldsymbol{e}_B\boldsymbol{s}_A + \tilde{\boldsymbol{g}}_A - \boldsymbol{e}_A\boldsymbol{s}_B - \tilde{\boldsymbol{g}}_B + \boldsymbol{p}_A - \boldsymbol{p}_B) = 2\boldsymbol{\varepsilon}$$

Osserviamo che poiché $\boldsymbol{p}_A$ e $\boldsymbol{p}_B$ sono estratti da una distribuzione uniforme su una palla di raggio $\delta/2$, vale $\|\boldsymbol{p}_A - \boldsymbol{p}_B\|_2 < \delta$. Allora analogamente al Lemma 3.2 vale

$$\begin{aligned}\|2\boldsymbol{\varepsilon}\|_\infty &= \|2(\boldsymbol{e}_B\boldsymbol{s}_A + \tilde{\boldsymbol{g}}_A - \boldsymbol{e}_A\boldsymbol{s}_B - \tilde{\boldsymbol{g}}_B + \boldsymbol{p}_A - \boldsymbol{p}_B)\|_\infty \le\\ &\le \|2(\boldsymbol{e}_B\boldsymbol{s}_A + \tilde{\boldsymbol{g}}_A - \boldsymbol{e}_A\boldsymbol{s}_B - \tilde{\boldsymbol{g}}_B + \boldsymbol{p}_A - \boldsymbol{p}_B)\|_2 \le\\ &\le 2\left[\|\boldsymbol{e}_B\boldsymbol{s}_A\|_2 + \|\boldsymbol{e}_A\boldsymbol{s}_B\|_2 + \|\tilde{\boldsymbol{g}}_A\|_2 + \|\tilde{\boldsymbol{g}}_B\|_2 + \|\boldsymbol{p}_A - \boldsymbol{p}_B\|_2\right] \le\\ &\le 2\left[\sqrt{n}\|\boldsymbol{e}_B\|_2\|\boldsymbol{s}_A\|_2 + \sqrt{n}\|\boldsymbol{e}_A\|_2\|\boldsymbol{s}_B\|_2 + \|\tilde{\boldsymbol{g}}_A\|_2 + \|\tilde{\boldsymbol{g}}_B\|_2 + \|\boldsymbol{p}_A - \boldsymbol{p}_B\|_2\right] \le\\ &\le 2\left[\sqrt{n}\alpha\sqrt{n}\alpha\sqrt{n} + \sqrt{n}\alpha\sqrt{n}\alpha\sqrt{n} + \gamma\sqrt{n} + \gamma\sqrt{n} + \delta\right] =\\ &= 4\alpha^2 n\sqrt{n} + 4\gamma\sqrt{n} + 2\delta <\\ &< \frac{q}{4} - 2\end{aligned}$$

Utilizzando il Lemma 3.1 precedente, abbiamo che $sk_A = sk_B$ eccetto che con probabilità trascurabile. □

Confrontando con il risultato del Lemma 3.2 del protocollo originale, vediamo che la stima inferiore per i valori di q che garantiscano la correttezza del protocollo è stata abbassata di un fattore $16\alpha\sqrt{n} - 16\gamma\sqrt{n} - 8\delta \approx 8(2\beta\sqrt{n} - \delta)$

4.3 Sicurezza

Vogliamo mostrare che il protocollo modificato (MKE) è sicuro quanto il protocollo originale. Per farlo, mostriamo che i giochi di sicurezza per MKE e per KE sono indistinguibili.

Lemma 4.2. *Se vale la RLWE assumption, il protocollo modificato è sicuro quanto il protocollo originale.*

Dimostrazione. Cominciamo osservando che se $\boldsymbol{k}_B = \boldsymbol{t}_A\boldsymbol{s}_B + 2\boldsymbol{g}_B$ è indistinguibile da $\boldsymbol{k}_B$ uniforme per *RLWE assumption*, allora anche $\boldsymbol{k}_B = \boldsymbol{t}_A\boldsymbol{s}_B + 2(\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B)$ è indistinguibile da $\boldsymbol{k}_B$ uniforme.

Ricordiamo che per un avversario che non conosce s uniforme e dipendente dall'istanza del protocollo, il valore $\boldsymbol{p}_B$ segue la distribuzione χ_β.

Se χ_α, χ_β e χ_γ fossero gaussiane "normali" (ovvero non discrete ma nel caso continuo), allora $\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B$ avrebbe esattamente distribuzione gaussiana χ_α (poiché per costruzione $\alpha^2 = \beta^2 + \gamma^2$). Varrebbe quindi l'indistinguibilità di $\boldsymbol{k}_B$ del protocollo modificato da $\boldsymbol{k}_B$ uniforme sotto *RLWE assumption*.

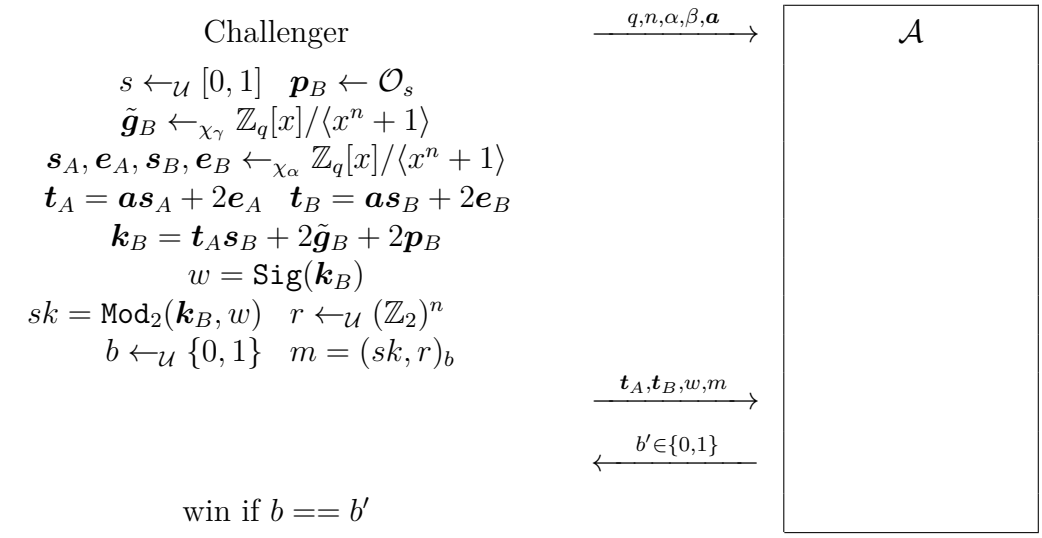
Poiché stiamo considerando gaussiane discrete, non è vero che $\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B$ ha distribuzione gaussiana discreta χ_α nonostante sia somma di gaussiane discrete. Tuttavia, vale comunque che la varianza di $\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B$ è la stessa di $\boldsymbol{g}_B$. Per quanto riguarda la norma, abbiamo che, eccetto che con probabilità trascurabile, vale

$$\|\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B\|_2 \leq \|\tilde{\boldsymbol{g}}_B\|_2 + \|\boldsymbol{p}_B\|_2 \leq \gamma\sqrt{n} + \beta\sqrt{n}$$

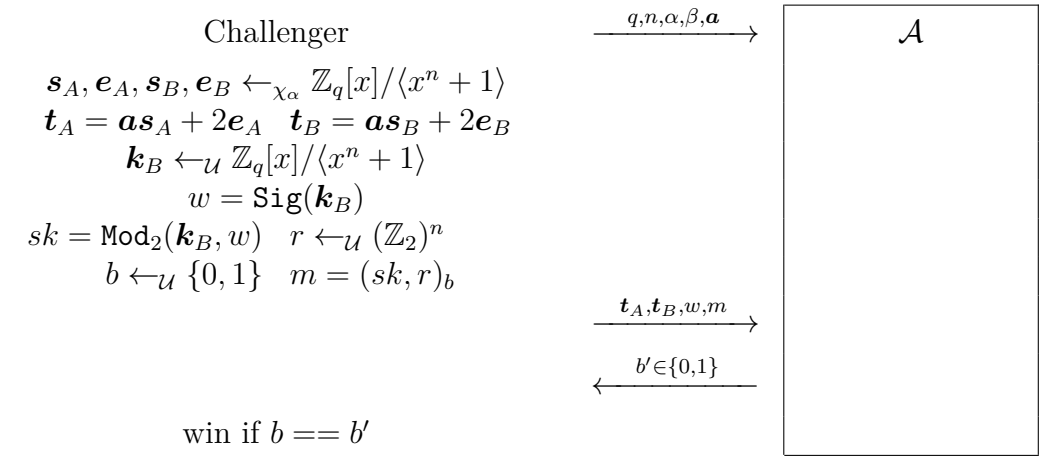
Quindi $\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B$ ha distribuzione di varianza identica a χ_α e limitata eccetto che con probabilità trascurabile. Se allora non valesse la *RLWE assumption* per $\boldsymbol{k}_B$ del secondo protocollo, per la stessa scelta di parametri non potrebbe valere per $\boldsymbol{k}_B$ del primo protocollo (poiché $\boldsymbol{g}_B$ è estratto da una distribuzione potenzialmente più concentrata di $\tilde{\boldsymbol{g}}_B + \boldsymbol{p}_B$). Quindi *RLWE* per $\boldsymbol{k}_B$ di MKE è sicuro almeno quanto *RLWE* per $\boldsymbol{k}_B$ di KE

Consideriamo quindi la seguente successione di giochi di sicurezza:

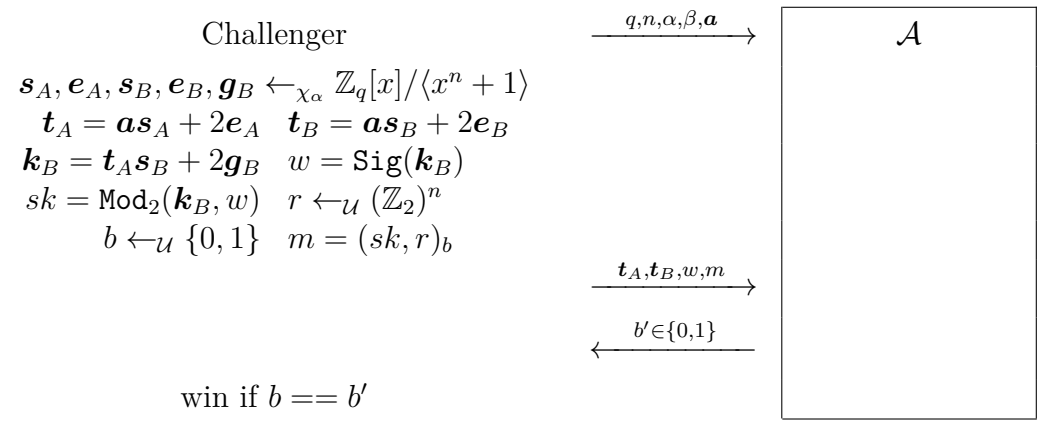
G_1


Challenger
$s \leftarrow_{\mathcal{U}} [0,1] \quad \boldsymbol{p}_B \leftarrow \mathcal{O}_s$
$\tilde{\boldsymbol{g}}_B \leftarrow_{\chi_\gamma} \mathbb{Z}_q[x]/\langle x^n+1\rangle$
$\boldsymbol{s}_A, \boldsymbol{e}_A, \boldsymbol{s}_B, \boldsymbol{e}_B \leftarrow_{\chi_\alpha} \mathbb{Z}_q[x]/\langle x^n+1\rangle$
$\boldsymbol{t}_A = \boldsymbol{a}\boldsymbol{s}_A + 2\boldsymbol{e}_A \quad \boldsymbol{t}_B = \boldsymbol{a}\boldsymbol{s}_B + 2\boldsymbol{e}_B$
$\boldsymbol{k}_B = \boldsymbol{t}_A\boldsymbol{s}_B + 2\tilde{\boldsymbol{g}}_B + 2\boldsymbol{p}_B$
$w = \mathtt{Sig}(\boldsymbol{k}_B)$
$sk = \mathtt{Mod}_2(\boldsymbol{k}_B, w) \quad r \leftarrow_{\mathcal{U}} (\mathbb{Z}_2)^n$
$b \leftarrow_{\mathcal{U}} \{0,1\} \quad m = (sk, r)_b$
win if $b == b'$
$q,n,\alpha,\beta,\boldsymbol{a}$
$\mathcal{A}$
$\boldsymbol{t}_A,\boldsymbol{t}_B,w,m$
$b' \in \{0,1\}$

G_2


Challenger
$\boldsymbol{s}_A, \boldsymbol{e}_A, \boldsymbol{s}_B, \boldsymbol{e}_B \leftarrow_{\chi_\alpha} \mathbb{Z}_q[x]/\langle x^n+1\rangle$
$\boldsymbol{t}_A = \boldsymbol{a}\boldsymbol{s}_A + 2\boldsymbol{e}_A \quad \boldsymbol{t}_B = \boldsymbol{a}\boldsymbol{s}_B + 2\boldsymbol{e}_B$
$\boldsymbol{k}_B \leftarrow_{\mathcal{U}} \mathbb{Z}_q[x]/\langle x^n+1\rangle$
$w = \mathtt{Sig}(\boldsymbol{k}_B)$
$sk = \mathtt{Mod}_2(\boldsymbol{k}_B, w) \quad r \leftarrow_{\mathcal{U}} (\mathbb{Z}_2)^n$
$b \leftarrow_{\mathcal{U}} \{0,1\} \quad m = (sk, r)_b$
win if $b == b'$
$q,n,\alpha,\beta,\boldsymbol{a}$
$\mathcal{A}$
$\boldsymbol{t}_A,\boldsymbol{t}_B,w,m$
$b' \in \{0,1\}$

G_3



Per ipotesi sulla famiglia di oracoli e per *RLWE assumption* G_2 è indistinguibile da G_1. Per *RLWE assumption* G_3 è indistinguibile da G_2 e quindi per transitività G_3 è indistinguibile da G_1. Poiché G_1 è esattamente il gioco di sicurezza per MKE e G_3 è esattamente il gioco di sicurezza di KE, per indistinguibilità il vantaggio di un avversario $\mathcal{A}$ in G_1 può differire al più di una quantità trascurabile dal vantaggio di $\mathcal{A}$ in G_3, che per quanto dimostrato nella sezione 2 del capitolo precedente è al più trascurabile.

Allora per ogni avversario *PPT* $\mathcal{A}$, il vantaggio di $\mathcal{A}$ contro MKE è al più somma di due quantità trascurabili e quindi trascurabile, quindi il protocollo è sicuro (sotto RLWE assumption). □

5 Considerazioni conclusive

Specializzando il protocollo di scambio di chiavi proposto in [1] ad un contesto IoT, siamo riusciti ad abbassare i parametri del protocollo (a parità di livello di sicurezza) facendo utilizzo di un'ipotesi ottenuta dal contesto. Il miglioramento sui parametri ottenuto è sicuramente relativamente piccolo rispetto alla dimensione assoluta dei parametri originali, ma lavorando in un contesto in cui ogni bit e ogni ciclo risparmiato conta, è comunque rilevante.

Si potrebbe osservare che il guadagno ottenuto è minore di quello che ci si potrebbe aspettare partendo dall'ipotesi in cui i due interlocutori già dispongono di un'informazione segreta e condivisa. Si potrebbe decidere ad esempio di utilizzare questa informazione come parte della chiave finale, e generare il resto della chiave finale con lo stesso protocollo ma con parametri per una chiave più piccola (e quindi essi stessi più piccoli). In verità questo approccio è meno semplice di quanto possa sembrare a primo impatto. Supponiamo per semplicità che l'informazione condivisa a meno di un piccolo errore sia una stringa finita di bit. Per via dell'incertezza sul dato fisico rilevato, i valori dei due interlocutori differiranno sugli ultimi bit. Per via della distribuzione del dato fisico rilevato, un avversario può facilmente stimare l'ordine di grandezza e quindi i primi bit dei valori dei due interlocutori. Possiamo raggruppare i bit in tre insiemi:

- Un insieme di bit di coda, sui quali probabilmente i valori rilevati dai due interlocutori non coincidono, e che quindi devono essere scartati per ottenere chiavi finali coincidenti
- Un insieme di bit di testa, facilmente stimabili da un avversario, e che quindi devono essere scartati per ottenere una chiave finale sicura
- Un insieme di bit centrali, che possono essere effettivamente utilizzati per la chiave finale

Per utilizzare quindi questo approccio "semplice" bisogna fare un processo di estrazione dall'informazione rilevata, che vada a determinare e scartare i bit di testa e i

bit di coda, che ne diminuisce notevolmente le dimensioni. Il protocollo modificato proposto, invece, risolve automaticamente questo problema utilizzando la *RLWE assumption*, che richiede un errore della stessa forma e compie implicitamente l'estrazione, garantendo una sicurezza dimostrabile.

Infine, osserviamo che il tipo di sicurezza garantito dal protocollo (sia modificato che originale) è relativa esclusivamente ad avversari passivi, ovvero avversari che possono attaccare il protocollo solo attraverso informazioni pubbliche (i parametri e il transcript). Non c'è garanzia (ed anzi si dimostra che il protocollo è attaccabile) in merito ad avversari attivi, ovvero attaccanti che possono controllare alcune delle parti attive del protocollo, ovvero contesti in cui l'interlocutore con cui stiamo svolgendo il protocollo sia corrotto.

La scelta di restringersi ad un protocollo che garantisca sicurezza solo da avversari passivi è una conseguenza diretta dell'ipotesi aggiuntiva fatta per il miglioramento del protocollo: l'idea è di sfruttare un piccolo segreto condiviso dagli interlocutori a meno di un piccolo errore e difficilmente stimabile dagli avversari per abbassare i parametri del protocollo mantenendo lo stesso livello di sicurezza. Se in questa ipotesi cercassimo di difenderci da avversari attivi, dovremmo difenderci anche da avversari che conoscono il piccolo segreto, che diventa quindi equivalente ad un valore pubblicamente noto, e se fosse possibile (in qualche modo non ancora noto) migliorare il protocollo con l'aggiunta di un parametro pubblico, questa sarebbe una modifica che potremmo fare anche al di fuori di contesti IoT.

Bibliografia

[1] Jintai Ding, Xiang Xie, and Xiaodong Lin. "A simple provably secure key exchange scheme based on the learning with errors problem".

[2] Whitfield Diffie, Martin E. Hellman, (1976). "New Directions in Cryptography".

[3] P.W. Shor, (1994). "Algorithms for quantum computation: discrete logarithms and factoring".

[4] O. Regev, (2005). "On lattices, learning with errors, random linear codes, and cryptography".

[5] Lyubashevsky, Vadim; Peikert, Chris; Regev, Oded (2012). "On Ideal Lattices and Learning with Errors Over Rings".

[6] B. Applebaum, D. Cash, C. Peikert, and A. Sahai. (2009) Fast cryptographic primitives and circular-secure encryption based on hard learning problems.

[7] Ajtai, Miklós (1998). "The shortest vector problem in L2 is NP-hard for randomized reductions".

[8] P. Barsocchi, S. Chessa, I. Martinovic, G. Oliger (2011). "A cyber-physical approach to secret key generation in smart environments". Journal of Ambient Intelligence and Humanized Computing.