UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

"TECNICHE DI ADVANTAGE DISTILLATION PER LA GENERAZIONE DI CHIAVI SICURE"

Relatore: Prof. Stefano Tomasin

Laureando: Stefano Lattenero

ANNO ACCADEMICO 2022 – 2023

Data di laurea 16/11/2023

Sommario

La protezione dei dati e la sicurezza delle comunicazioni giocano un ruolo fondamentale nell'era digitale. La crittografia si dimostra di primaria importanza in questo contesto poiché permette la trasmissione sicura di informazioni sensibili attraverso reti wireless. Tali reti presentano una grande debolezza, consentire di inviare e ricevere trasmissioni da qualsiasi utente nel raggio d'azione del dispositivo, rendendolo vulnerabile a qualsiasi tipo di attacco. Per ovviare a questo problema vengono utilizzate delle chiavi segrete ad eventuali utenti indesiderati. In questa tesi si vogliono definire alcuni metodi elementari per generare tali chiavi ricorrendo ad algoritmi primitivi, una volta poste le fondamenta per la comunicazione che prendiamo in considerazione.

Indice

Capitolo 1

Chiavi sicure nelle reti wireless

1.1 Introduzione

Il problema della generazione di chiavi segrete è stato studiato per la prima volta da Maurer [1], e da Ahlswede e Csiszár [2]. In un problema di generazione di chiavi segrete di base, chiamato *basic source model*, due terminali (*Alice* e *Bob*)[1] osservano una sorgente casuale comune inaccessibile ad un terminale esterno definito *eavesdropper* o *Eve*.

Per definire il sistema di riferimento che prendiamo in considerazione necessaria risulta la definizione del termine *memoryless* attribuito ad un canale. Utilizzando l'aggettivo senza memoria si indica una situazione nella quale i simboli vengono comunicati attraverso il canale uno ad uno e ogni simbolo ricevuto dipende solo dal simbolo corrispondente inviato.
Modellando le osservazioni in relazione a quanto appena descritto, possiamo definire il modello come segue: *Alice* e *Bob* osservano rispettivamente ripetizioni indipendenti e equamente distribuite (i.i.d.) delle variabili casuali dipendenti X e Y, definite da $X^n = (X_1, ..., X_n)$ e $Y^n = (Y_1, ..., Y_n)$. In qualsiasi istanza temporale, la coppia osservata (X_i, Y_i) è dipendente. Sulla base delle loro osservazioni, *Alice* e *Bob* generano una chiave segreta comune, comunicando su un canale pubblico privo di errori.
Una variabile casuale κ con range finito K rappresenta una chiave segreta ϵ per *Alice* e *Bob*, realizzabile tramite la comunicazione $\mathbf{V}$, se esistono due funzioni f_A, f_B tali che $\kappa_A = f_A(X^n, \mathbf{V}), \kappa_B = f_B(Y^n, \mathbf{V})$ e per qualsiasi ϵ vale:

$$Pr(\kappa = \kappa_A = \kappa_B) \geq 1 - \epsilon \tag{1.1}$$

$$I(\kappa; \mathbf{V}) \leq \epsilon \tag{1.2}$$

$$H(\kappa) \leq log|K| - \epsilon \tag{1.3}$$

Qui, la condizione (1.1) assicura che *Alice* e *Bob* generino la stessa chiave segreta con un'alta probabilità; la condizione (1.2) assicura che tale chiave segreta sia effettivamente nascosta all'*eavesdropper* che osserva la comunicazione pubblica; e la condizione (1.3) assicura che tale chiave segreta sia distribuita quasi uniformemente.

Alice e *Bob* cercano di stabilire una chiave univoca sfruttando la *randomness* del canale tramite cui comunicano. Il terzo utente *Eve* intercetta il canale e si mette in ascolto senza agire. Definito il canale di comunicazione tra le due parti nei capitoli successivi spiegheremo nel dettaglio le caratteristiche

necessarie affinchè si possa creare la situazione ottimale per la generazione di chiavi segrete, in secondo luogo andremo a definire e spiegare nel dettaglio l'intero processo di generazione delle suddette chiavi.

1.2 Reti wireless

Le comunicazioni senza fili hanno conosciuto un notevole sviluppo e sono utilizzate per diverse applicazioni. Nonostante la comodità che deriva direttamente dall'utilizzo di un tale tipo di comunicazione, uno dei principali punti a sfavore rimane la sicurezza, la quale desta preoccupazioni in diverse applicazioni. I protocolli di sicurezza tradizionali si basano principalmente su funzioni di crittografia e hashing e altre proprietà matematiche per raggiungere i loro obiettivi eppure, oggi, questi protocolli non rappresentano la soluzione adeguata a causa della natura stessa di una comunicazione *wireless*.

1.2.1 Sicurezza delle reti wireless

La natura intrinseca di trasmissione delle comunicazioni senza fili consente di ricevere le trasmissioni da qualsiasi utente all'interno del raggio d'azione, consentendo agli aggressori di avviare diversi attacchi passivi, quali intercettazioni, analisi e monitoraggio del traffico, ecc., o di eseguire attacchi attivi, quali jamming, spoofing, modifica, replay e denial-of-service (DoS), ecc.

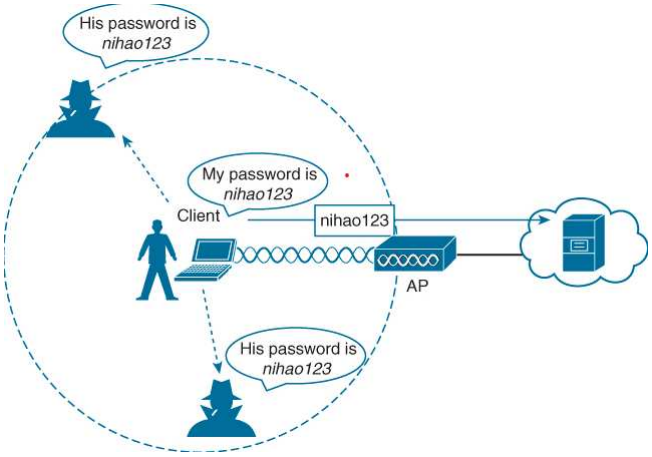

Figura 1.2.1: Due utenti nel raggio d'azione del client recepiscono la password.

Tradizionalmente, i dati sono protetti da sistemi di cifratura di classe, che partono dal presupposto che l'algoritmo sia sufficientemente complesso da far sì che il tempo impiegato dagli intercettatori per decifrare il sistema crittografico sia molto più lungo della validità dell'informazione stessa. Tali schemi di cifratura sono suddivisi in simmetrici e asimmetrici, a seconda delle chiavi utilizzate dalle due parti crittografiche. I sistemi di cifratura simmetrica utilizzano la stessa chiave e sono generalmente impiegati per la protezione dei dati grazie alla loro efficienza nella cifratura delle informazioni stesse. I sistemi di cifratura asimmetrica invece, noti anche come crittografia a chiave pubblica, utilizzano la stessa chiave pubblica ma diverse chiavi private e sono generalmente applicati per la distribuzione delle chiavi.
I sistemi di criptazione classici devono far fronte a diverse vulnerabilità. Prendiamo come esempio la crittografia a chiave pubblica. In primo luogo, dipende direttamente dalla durezza computazionale del problema matematico su cui si appoggia, per esempio, il logaritmo discreto. La sicurezza computazionale potrebbe non durare in futuro a causa del rapido sviluppo della tecnologia hardware. In secondo luogo, richiede un'infrastruttura di gestione fondamentale che dovrebbe essere anch'essa messa in sicurezza. Questo approccio risulta pertanto meno interessante per molte reti di sensori wireless (WSN) e applicazioni di rete ad hoc, poiché i nodi dei sensori hanno una capacità di calcolo limitata mentre le reti ad hoc sono decentralizzate.

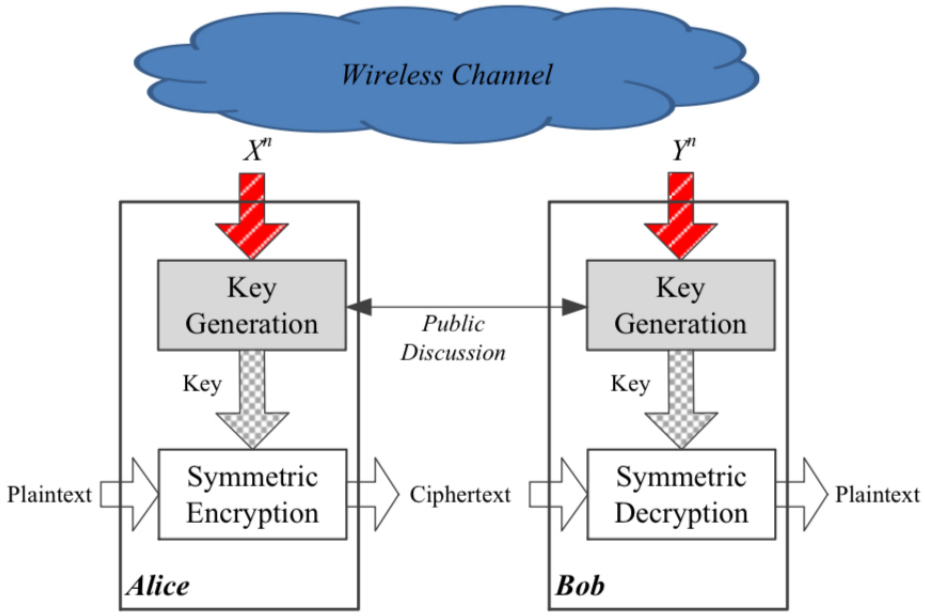

Figura 1.2.2: Schema cifratura simmetrica ed asmmetrica in un canale Wireless.

Mentre i classici schemi di crittografia vengono applicati ai livelli superiori dei protocolli di comunicazione, il livello fisico può anche essere sfruttato per migliorare la sicurezza wireless. Gli schemi di Physical Layer Security (PLS) sfruttano le caratteristiche imprevedibili e casuali dei canali wireless al fine di ottenere la sicurezza teorica dell'informazione. I sistemi PLS sono composti da sicurezza senza chiave e segretezza basata su chiave segreta.
La sicurezza senza chiave non richiede chiavi per la cifratura, ma utilizza la progettazione del codice e le proprietà di canale degli utenti legittimi e degli intercettatori per ottenere la segretezza. Tuttavia, gli utenti legittimi di solito richiedono informazioni complete o parziali sullo stato del canale, Channel State Information (CSI), istantanee/statistiche dagli intercettatori, che nella pratica non sono sempre disponibili e comportano un'attuazione molto complessa.
La segretezza basata sulla chiave segreta risale al 1919 quando il concetto di one-time pad, noto anche come cifrario Vernam, veniva utilizzato per crittografare ogni bit di messaggio con un bit di chiave segreta casuale. In seguito, Shannon pose le basi teoriche per una perfetta segretezza[3]. Il messaggio M è codificato nella parola in codice C che non rivela alcuna informazione sul messaggio, vale a dire:

$$H(M|C) = H(M), \tag{1.4}$$

dove H(·) indica l'entropia. Ciò richiede che le informazioni della sequenza di chiavi siano maggiori, o almeno uguali, alle informazioni contenute nel messaggio.
Un modo possibile per stabilire la chiave è quello di generare chiavi dai canali wireless. Tuttavia, nella pratica è molto difficile, se non impossibile, stabilire in modo efficiente chiavi casuali tra utenti legittimi che non possono essere riutilizzate. In alternativa, si può costruire un criptosistema ibrido combinando la generazione di chiavi e la cifratura simmetrica. Il livello di sicurezza del sistema è migliorato sostituendo la crittografia a chiave pubblica con la generazione di chiavi[4].

Capitolo 2

Principi

Al fine di studiare al meglio la generazione di chiavi segrete e rappresentare con maggiore chairezza gli esperimenti che visioneremo risulta necessario portare alla luce l'importanza di *randomness*, *key generation rate*, *key disagreement rate* e *coefficienti di correlazione* definiti in questo paragrafo. Quanto presentato in questo capitolo è preso da [5].

2.1 Randomness

La *randomness* delle chiavi utilizzate è il carattere fondamentale per l'utilizzo e l'efficacia delle stesse in ambito crittografico. L'utilizzo di una chiave non randomica implica un tempo di ricerca decisamente inferiore per attacchi *brute force*, rendendo il sistema in discussione vulnerabile. Il *National Institute of Standards and Technology* (NIST) fornisce un strumento per testare la randomness di un *random number generator*(RNG)[6], per generare chiavi il più sicure possibile.

2.2 Key generation rate

Il KGR definisce il numero di bit chiave generati in 1 secondo/misura. Per applicazioni in ambiente crittografico risultano necessarie sequenze di bit di una certa lunghezza per la generazione di chiavi.

2.3 Key disagreement rate

Il KDR quantifica il numero di bit non corrispondenti tra i messaggi degli utenti, successivamente alla quantizzazione, definito da:

$$KDR_{uv,u'v'} = \frac{\sum_{i=1}^{N} |K_{uv}(i) - K_{u'v'}(i)|}{N}, \tag{2.1}$$

dove K_{uv} e $K_{u'v'}$ sono le chiavi quantizzate e N la lunghezza delle chiavi.

2.4 Coefficienti di correlazione

La correlazione definisce il grado di somiglianza tra due sequenze prese in considerazione. Se gli indicatori sono simili, il coefficiente di correlazione sarà 1 e 0 nel caso opposto. Se confrontate due sequenze indipendenti introdurremo il termine cross-correlation.

Definiti i criteri di valutazione possiamo passare direttamente ai capisaldi del processo di key generation. La generazione di chiavi poggia le sue fondamenta su tre principi: variazione temporale, reciprocità nel canale e de-correlazione spaziale. Questi tre principi si dimostrano indispensabili per descrivere al meglio le metriche da seguire per generare delle chiavi sicure in una rete wireless. A titolo esemplificativo utilizzeremo una serie determinata di esperimenti[5], dei quali sarebbe possibile studiare tre scenari differenti: camera anecoica, camera di riverbero, e ambiente d'ufficio. Prenderemo in cosiderazione solo l'ultimo poichè sufficiente a coprire integralmente il nostro argomento di studio.

2.5 Variazione temporale

In un canale dinamico la *randomness* viene introdotta da utenti e oggetti e come interagiscono fra loro. Da questo ne deriva che maggiore è la velocità a cui il canale cambia, maggiore è la *randomness* del canale stesso. La variazione temporale presa in considerazione può essere quantificata utilizzando la *funzione di auto-correlazione*(ACF) del canale scritta come:

$$R_{X_{uv}}(t, \Delta t) = \frac{E\{(X_{uv}(t) - \mu_{X_{uv}})(X_{uv}(t + \Delta t) - \mu_{X_{uv}})\}}{E\{\mid X_{uv}(t) - \mu_{X_{uv}} \mid^2\}}, \tag{2.2}$$

dove $E\{\cdot\}$ rappresenta il calcolo atteso, $X_{uv}(t)$ la misura del canale, $\mu_{X_{uv}}$ è il valore medio di $X_{uv}(t)$, u e v rispettivamente trasmettitore e ricevitore.

Nei grafici sono rappresentati i dati risultanti dall'esperimento condotto, mostrano nell'asse delle ordinate l'intervallo di tempo nel quale vengono prese le misurazioni e nell'asse delle ascisse i valori dell'ACF normalizzato. Le curve sono generate dai valori ottenuti rispettivamente dal CSI in blu e dal Received Signal Strength (RSS) in rosso. Prendendo in considerazione l'intera gamma di risultati ottenuti dall'esperimento possono essere studiati altri scenari, ma è sufficiente descriverne tre: scenario statico, scenario statico con oggetti in movimento e scenario dinamico (utente in movimento).

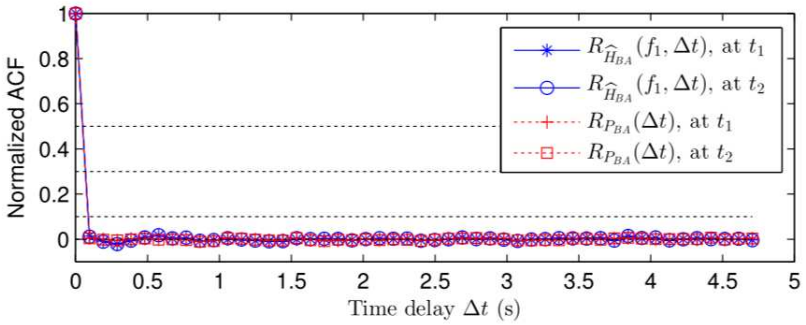

Figura 2.5.1: ACF normalizzati, $R_{\hat{H}_{BA}}(f_1, \Delta t)$ e $R_{P_{BA}}(\Delta t)$, al tempo t nell'ambiente Ufficio con scenario statico[5].

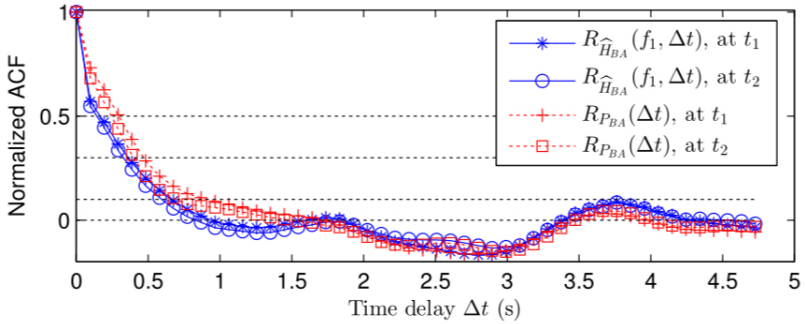

Figura 2.5.2: ACF normalizzati, $R_{\hat{H}_{BA}}(f_1, \Delta t)$ e $R_{P_{BA}}(\Delta t)$, al tempo t nell'ambiente Ufficio con scenario con oggetti in movimento[5].

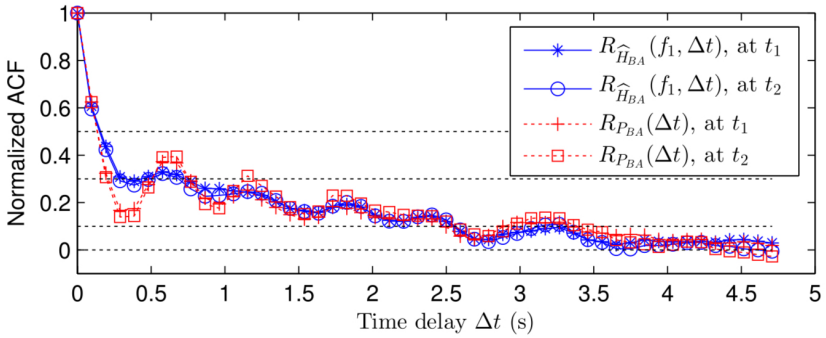

Figura 2.5.3: ACF normalizzati, $R_{\hat{H}_{BA}}(f_1, \Delta t)$ e $R_{P_{BA}}(\Delta t)$, al tempo t nell'ambiente Ufficio con scenario mobile[5].

Dalle figure si può evincere che vi sia una divisione degli scenari in due casi principali.
Poiché non vi è alcuna interferenza da altre reti wireless, il canale rimane lo stesso nello scenario statico (Fig 2.2.1) a meno della presenza di rumore hardware.
In ambiente dinamico, cioè nel caso di movimento di oggetti e scenario mobile nell'ambiente d'ufficio, possiamo notare come cambino i valori degli ACF tracciando delle curve pressochè simili, rispettivamente per CSI e RSS, nonostante differiscano leggermente in relazione a come varia l'ambiente e ai movimenti di utenti/oggetti.

La randomizzazione della sequenza di chiavi, valutata dalla National Institute of Standards and Technology(NIST) random test suite, è stata ampiamente utilizzata nelle applicazioni di generazione chiave. Ci sono 15 test in totale, ognuno dei quali valuta una caratteristica specifica di casualità, esempio, il test di frequenza si concentra sulla proporzione di uno e zero, il test DFT rileva la caratteristica periodica della sequenza, ecc. Ogni test restituisce un valore P, che viene confrontato con un valore di significatività , che assume tipicamente valori nell'intervallo di [0,001, 0,01]. Quando il valore P è maggiore di α, la sequenza è accettata come casuale.

In un secondo esperimento condotto, ad α è stato assegnato il valore di 0,01. Vengono eseguiti 8 test, oltre la metà della suite di test, che soddisfano ancora i requisiti del NIST. Alcuni dei test richiedono sequenze estremamente lunghe. Per esempio, il test della variante delle escursioni casuali raccomanda una sequenza di input più lunga di 10^6.
L'esperimento preso in considerazione, impostato affinchè assomigliasse il più possibile allo scenario mobile dell'ambiente d'ufficio, viene sondato per 300 secondi per raccogliere più dati possibli al fine di valutare la randomness del canale. Il canale è stato originariamente campionato a periodi di 0,96 ms ma poichè si è evinto che esistesse ridondanza tra campioni di dati adiacenti, in seguito, è stato ricampionato a periodi di T_p e poi quantizzato a valori binari.

Dai risultati ottenuti dai test di casualità delle chiavi quantizzate da $|\hat{H}_{AB}(f_1, t)|$ e $P_{AB}(t)$ si possono dedurre due importanti valutazioni. In primo luogo, quando la correlazione tra due misure adiacenti è elevata, la sequenza chiave fallisce, in secondo luogo, l'ACF temporale descrive la velocità con cui il segnale varia nel tempo e può quindi essere utilizzato per determinare l'intervallo di rilevamento ottimale(T_p). Un intervallo di rilevamento troppo breve comporterà una ridondanza del campione e inciderà sulla casualità della sequenza di chiavi, mentre un intervallo troppo lungo comporterà un basso tasso di generazione di chiavi(KGR) e ne limiterà l'applicazione pratica. In questo esempio, il sistema non può generare una chiave casuale sequenza da $|\hat{H}_{AB}(f_1, t)|$ fino a quando il coefficiente di correlazione tra campioni adiacenti è inferiore al $20,2\%$ e la velocità di campionamento T_p raggiunge valori maggiori di 1,5 s, che rappresenta la velocità di campionamento ottimale.

2.6 Reciprocità nel canale

Per reciprocità nel canale si intende una situazione nella quale diverse caratteristiche di un link, e.g. fase, ritardo e attenuazione risultano identiche al termine del link stesso. La maggior parte dei sistemi di generazione di chiavi moderni utilizzano un hardware *half-duplex* dove si utilizza un singolo canale di trasmissione che implica un passaggio di dati unidirezionale. Nel caso in cui il canale cambi troppo velocemente, una misurazione non simultanea va a influire negativamente sullo scambio di messaggi tra *Alice* e *Bob*. In secondo luogo la reciprocità di canale viene influenzata anche dalla presenza di rumore hardware. Essa può essere quantificata misurando la relazione tra le misure di canale con i Pearson correlation coefficients in:

$$\rho_{uv,u'v'} = \frac{E\{X_{uv}X_{u'v'}\} - E\{X_{uv}\}E\{X_{u'v'}\}}{\sigma_{X_{uv}}\sigma_{X_{u'v'}}}, \tag{2.3}$$

dove $\sigma_{X_{uv}}$ rappresenta la deviazione standard di $X_{uv}(t)$.

La dissolvenza del canale a ciascuna estremità del link è reciproca. Tuttavia, i segnali misurati da ciascun utente sono asimmetrici a causa delle misure non simultanee e del rumore hardware non correlato. La somiglianza tra i segnali ricevuti di Alice e Bob può essere quantificata dalla relazione di cross-correlazione e KDR, sostituendo $X_{AB}(t_A)$ e $X_{BA}(t_B)$.

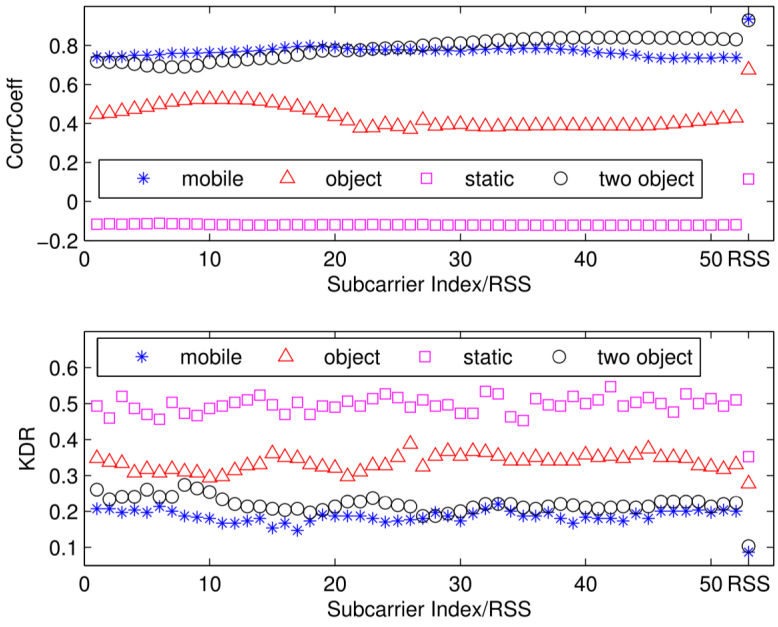

Figura 2.6.1: Coefficienti di Cross-correlazione ${\rho^X}_{AB,BA}$, KDR, ${KDR^X}_{AB,BA}$, di CSI e RSS nell'ambiente Ufficio con scenario statico, di oggetti in movimento e mobile[5].

I coefficienti di correlazione incrociata e i KDR degli esperimenti nell'ambiente ufficio sono illustrati rispettivamente nella Fig 2.3.1.
Quando il canale è statico, situazione in cui il rumore indipendente dell'hardware è l'unico fattore che contribuisce alla variazione del segnale, i coefficienti di cross-correlazione sono quasi nulli. I KDR corrispondenti nel canale statico sono circa 0.5, che non sono tanto meglio di ipotizzare un valore a caso. Ciò rende la generazione di chiavi non operativa in quanto gli utenti non sono in grado di raggiungere un accordo.
Negli scenari mobili i coefficienti di correlazione sono elevati e tutti i KDR sono accettabili e potrebbero essere successivamente corretti mediante tecniche di riconciliazione delle informazioni (spiegata nel capitolo successivo). Ad esempio, il codice BCH può correggere fino al 25% di disaccordo sulle chiavi[7]. Nello scenario degli oggetti in movimento nell'ambiente dell'ufficio, la correlazione non è così elevata come nello scenario mobile. Ciò è dovuto al fatto che quando un utente si muove, il canale cambia in modo più significativo rispetto all movimento di oggetti, in cui sono interessati solo alcune variabili. Tuttavia, quando due oggetti si muovono nell'ambiente d'ufficio come indicato dai simboli ∘ nella Fig 2.3.1, la correlazione è elevata come quella dello scenario mobile, il che significa che l'aumento del movimento contribuisce a migliorare la correlazione.

2.7 De-correlazione spaziale

La decorrelazione spaziale è essenziale per la sicurezza dei sistemi di generazione chiave. Il valore del KDR viene solitamente usato per quantificare il grado di disaccordo (de-correlazione) tra Alice e Bob. Tuttavia, può anche essere esteso per quantificare il disaccordo tra utenti legittimi e intercettatori(eavesdropper).
Una stima del valore medio del coefficiente di correlazione può essere quantificata tramite la seguente formula:

$$\rho^{-\hat{H}}_{uv,u'v'} = \frac{1}{M}\sum_{i=0}^{M-1} \rho^{\hat{H}_m}_{uv,u'v'}. \tag{2.4}$$

Una stima del KDR medio puo essere espressa mediante:

$$\overline{KDR}^{\hat{H}}_{uv,u'v'} = \frac{1}{M}\sum_{i=0}^{M-1} KDR^{\hat{H}_m}_{uv,u'v'}. \tag{2.5}$$

2.7.1 Eavesdropper in posizionamento lineare

Numerosi esperimenti sono stati effettuati con diverse configurazioni di distanze. Le forme delle curve negli stessi ambienti ottenute da CSI e RSS sono abbastanza simili, mentre i valori assoluti sono leggermente diversi.

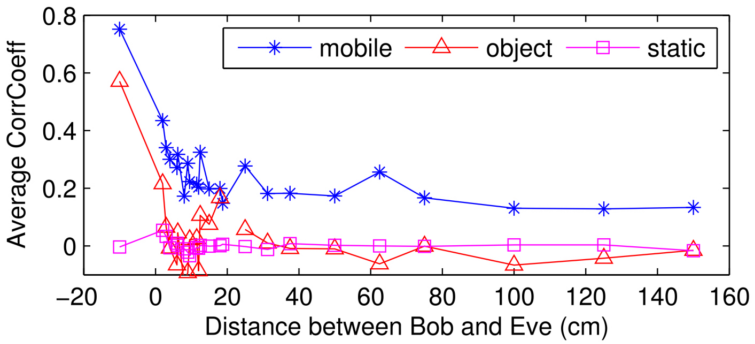

Figura 2.7.1: Coefficiente medio di correlazione $\rho^{-\hat{H}}_{AB,AE_j}$, con scenario statico, di oggetti in movimento e mobile[5].

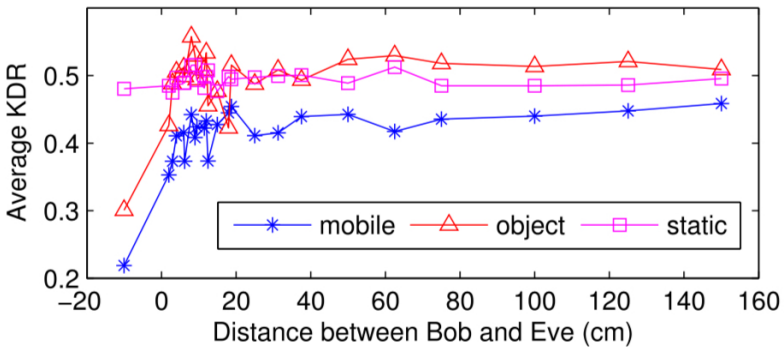

Figura 2.7.2: KDR medio $\overline{KDR}^{\hat{H}}_{AB,AE_j}$, nell'ambiente Ufficio con scenario statico, di oggetti in movimento e mobile[5].

Relativamente all'ambiente d'ufficio si può notare come il multipath aiuti la decrescita della correlazione spaziale tra gli utenti. Questo, influenza positivamente la sicurezza nella generazione di chiavi e indica che l'eavesdropper non può recepire nessuna informazione utile riguardante le chiavi. Si può inoltre evidenziare che, quando gli intercettatori sono molto vicini agli utenti, i loro segnali ricevuti risultano molto differenti.
Tuttavia, in un ambiente con Line-of-Sight(LoS) forte come una camera anecoica, anche quando gli eavesdropper si trovano a diverse lunghezze d'onda (3λ come da esempio nell'esperimento), posso ancora osservare un segnale altamente correlato a quello proveniente direttamente dagli Users autorizzati.
Per LoS si intende la linea immaginaria tra l'osservatore e l'obiettivo. Nella comunicazione, la linea-of-sight è il percorso diretto da un trasmettitore al ricevitore e gli ostacoli che possono cadere in quel percorso. Visto quanto detto in precedenza, risulterà necessario prestare particolare attenzione a contrastare correttamente gli eavesdroppers in ambienti con forte LoS.
Il multipath inoltre è generalmente considerato dannoso per i sistemi wireless in quanto aumenta la complessità dell'equalizzatore, tuttavia, risulta vantaggioso nell'applicazione della generazione chiave a causa dell'incertezza introdotta.

2.7.2 Eavesdropper in posizionamento circolare

Ulteriori esperimenti sono stati effettuati mettendo sei eavesdropper intorno a Bob, come mostrato in Fig 2.4.3.

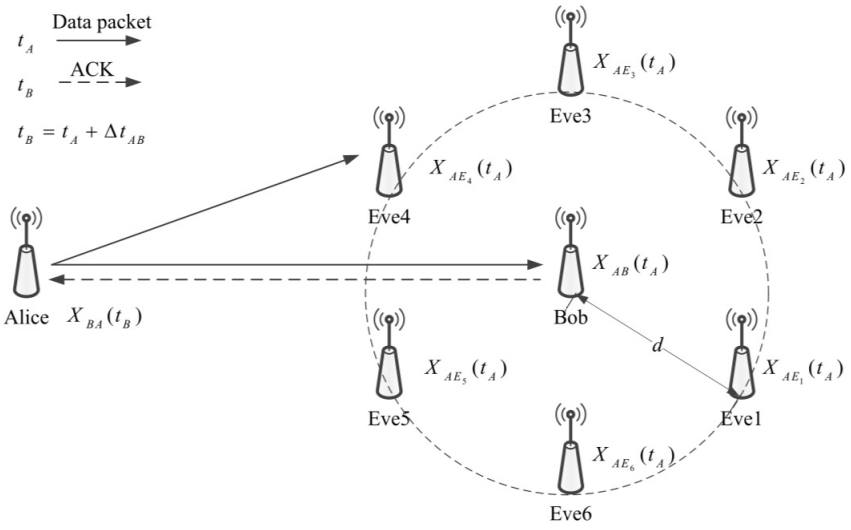

Figura 2.7.3: Disposizione circolare degli eavesdropper intorno a Bob[5].

Tali esperimenti hanno fornito come risultati i valori rappresentati nelle tabelle sottostanti.

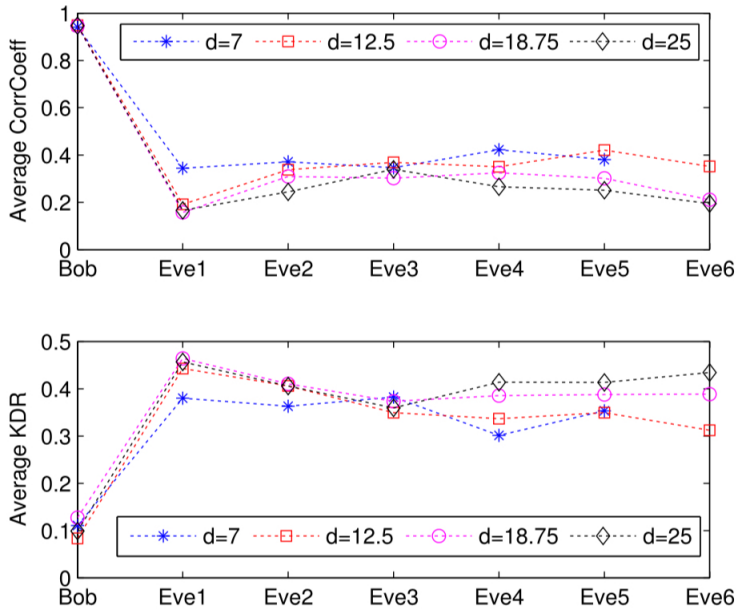

Figura 2.7.4: Coefficiente medio di Cross-correlazione $\rho^{-\hat{H}}_{AB,AE_j}$,KDR medio $\overline{KDR}^{\hat{H}}_{AB,AE_j}$, con scenario mobile[5].

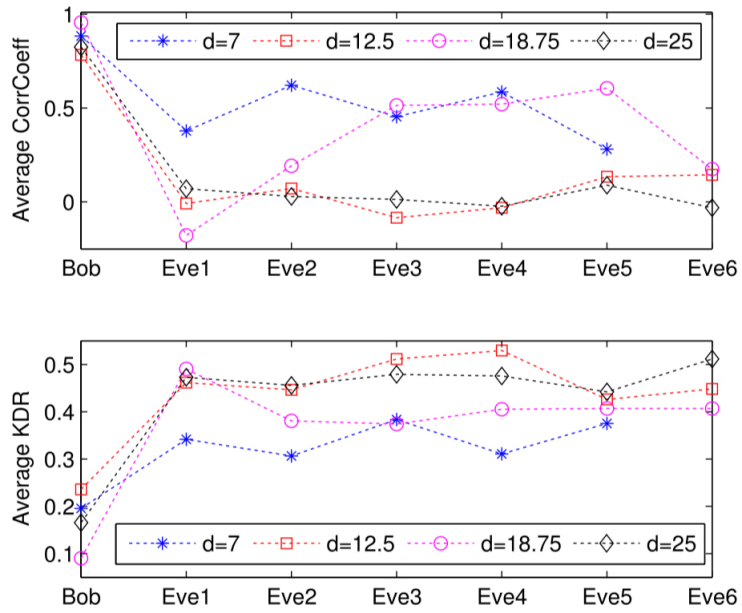

Figura 2.7.5: Coefficiente medio di Cross-correlazione $\rho^{-\hat{H}}_{AB,AE_j}$,KDR medio $\overline{KDR}^{\hat{H}}_{AB,AE_j}$, con oggetti in movimento[5].

Tuttavia, non sembra esserci alcuna relazione tra i coefficienti di correlazione risultanti e la posizione degli eavesdropper; questo perchè, in un ambiente multipath, il segnale proviene da tutte le direzioni a causa degli effetti di riflessione, dispersione, rifrazione, ecc. Questa proprietà è molto vantaggiosa per la generazione di chiavi, in quanto anche se gli intercettatori si trovano tra gli utenti, non riescono comunque ad ottenere un grado di correlazione migliore. [5].

Capitolo 3

Procedura di generazione di chiavi

Poggiandosi su quanto detto in precedenza andremo a definire in modo preciso il vero e proprio processo di generazione di chiavi segrete. Questo processo viene studiato prendendo in considerazione un ambiente di comunicazione costituito da reti wireless. In questo scenario i due utenti *Alice* e *Bob* devono comunicare tra loro tramite messaggi che devono rimanere nascosti ad un terzo interlocutore, *Eve*. La generazione di chiavi al fine di criptare i messaggi fra loro, consta fondamentalmente di quattro fasi.

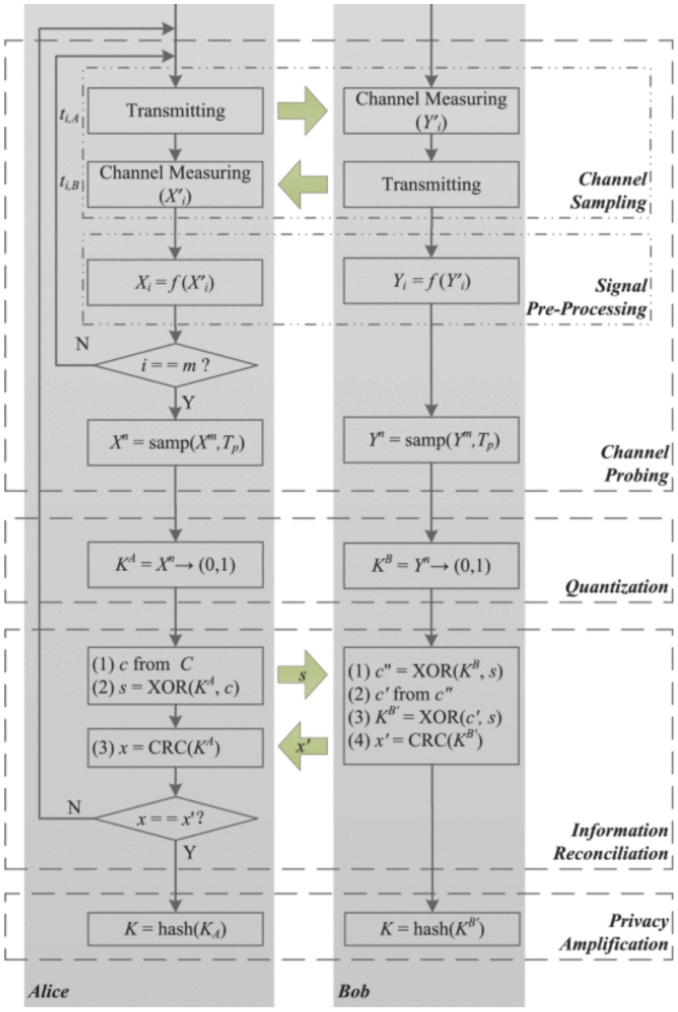

Figura 3.0.1: Procedura di generazione[8].

3.1 Channel probing

La procedura di channel probing definisce la casualità del canale di comunicazione appoggiandosi direttamente alla randomness del canale.

In primo luogo Alice e Bob cercano di calcolare la randomness del canale tramite il quale comunicano, inviandosi a vicenda dei segnali pilota pubblici. Sondare il canale è il punto chiave per raccoglierne informazioni. Come mostrato in Fig. 3.0.1, al tempo $t_{i,A}$, Alice trasmette il segnale di sonda i^{th} a Bob il quale misurerà alcuni parametri del canale attraverso il segnale ricevuto e lo memorizzerà in Y_i . Al momento $t_{i,B}$, Bob trasmette il suo segnale di sonda ad Alice che misurerà rispettivamente lo stesso parametro del canale e lo memorizzerà in X_i . La differenza di tempo di campionamento $\Delta t_i = |t_{i,A} - t_{i,B}|$ viene volutamente mantenuta minore del tempo di coerenza del canale T_c(misura della durata temporale in cui la risposta impulsiva del canale rimane invariata). Questo di modo che il canale, durante le due sonde, possa essere considerato costante. Alice e Bob continueranno a ripetere il processo sopra citato fino ad ottenere risultati sufficienti ai fini del sondaggio.

La ricerca nel sondare i canali considera tre fattori: il channel parameter, il signal pre-processing, e il channel probing rate.
Sebbene le caratteristiche dei canali a ciascuna estremità del collegamento siano reciproche, i segnali ricevuti misurati sono asimmetrici, principalmente a causa di misurazioni non simultanee (ad esempio, $\Delta t_i \neq 0$) e del rumore hardware delle due separate piattaforme. Pertanto, il signal pre-processing viene utilizzato per aumentare il grado di cross-correlazione tra i segnali ricevuti, vale a dire, $X_i = f(X_i)$ in Fig. 3.0.1. Gli effetti delle misurazioni non simultanee e del rumore possono essere mitigati rispettivamente mediante interpolazione e filtraggio.
Può esistere una ridondanza all'interno delle misure campionate X_m e Y_m, che vengono quindi ricampionate con una frequenza di campionamento T_p descritta al capitolo precedente, scelta per essere maggiore del T_c descritto prima.
Tuttavia, la randomness del canale è causata da un movimento imprevedibile, che porta a un diverso tasso di variazione della condizione del canale. Di conseguenza, è stato progettato un sistema di rilevamento adattivo basato su un proportional-integral-derivative (PID) controller-based per regolare la frequenza di rilevamento in base alle condizioni del canale. Si potrebbe quindi generare sequenze chiave in modo sicuro ed efficace anche se in ambiente che cambia dinamicamente. Dopo una sufficiente serie di segnali i due utenti raccolgono una determinata quantità di parametri, definendo che tipologia di chiave sia meglio utilizzare.

3.2 Quantization

Il quantizzatore viene utilizzato per mappare in valori binari le misure di canale ottenute. Un quantizzatore è caratterizzato da due fattori, il livello di quantizzazione (QL) e i threshold. Con QL si intende il numero di bit in cui ogni misura viene convertita e generalmente viene determinato utilizzando il *signal-to-noise ratio* (SNR) del canale.

Calcolato tramite le seguente formule:

$$SNR_{q_{dB}} = 10\ log_{10}\ SNR_q = 4.77\ +\ 6.02b\ -\ 20\ log_{10}\left(\frac{v_{sat}}{\sigma_A}\right) \tag{3.1}$$

dove b rappresenta il numero di bit, v_{sat} il valore di saturazione(il quale indica il valore massimo e preso negativo il valore minimo da considerare) e σ_A la deviazione standard del rumore.

Riferendosi ad una quantizzazione multi-bit viene utilizzata inoltre la codifica di Gray che permette una riduzione del numero di bit differenti all'interno della sequenza.

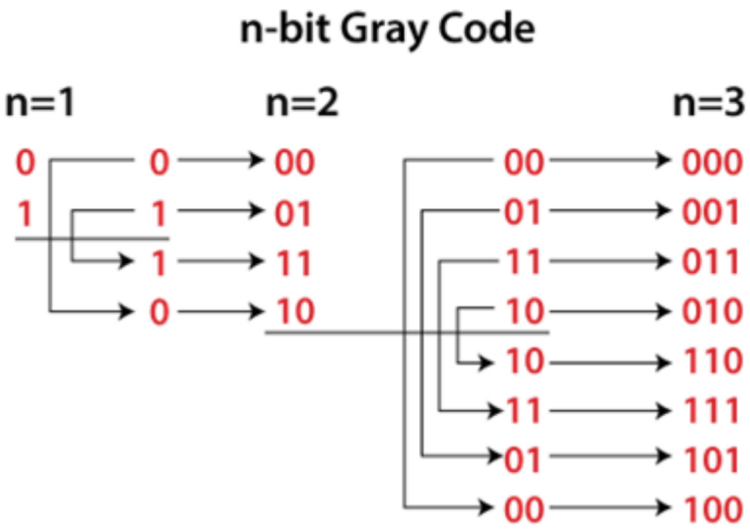

Figura 3.2.1: Codifica di Gray.

I threshold infine, sono i livelli di riferimento presi in considerazione e vengono definiti in relazione alla media e alla deviazizone standard, o utilizzando la funzione di distribuzione cumulativa(CDF):

$$F(x) = Pr(X^n \leq x) \tag{3.2}$$

Tali livelli di riferimento vengono utilizzati per suddividere le misurazioni in diversi gruppi.

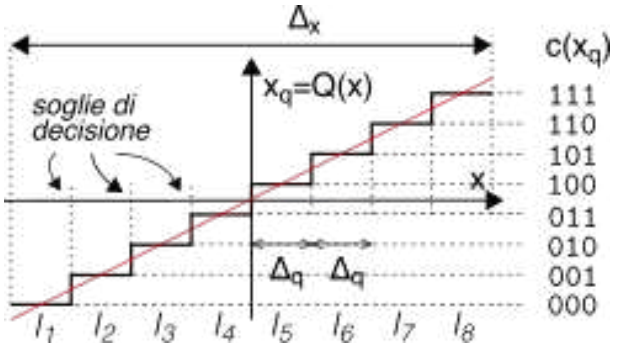

Figura 3.2.2: Esempio di quantizzatore (uniforme).

Per determinare le soglie si utilizzano comunemente il valore medio μ (o insieme alla deviazione standard σ) e la funzione di distribuzione cumulativa (CDF). Lo schema di quantizzazione basato sul valore medio e sulla deviazione standard ha una semplice implementazione.

Le soglie sono determinate come segue:

$$\begin{aligned} \eta+ &= \mu + \alpha \times \sigma; \\ \eta- &= \mu - \alpha \times \sigma; \end{aligned} \tag{3.3}$$

Quando $\alpha \neq 0$, le misurazioni tra $\eta+$ e $\eta-$ vengono omesse. I campioni sopra $\eta+$ /sotto $\eta-$ saranno convertiti in 1/0. Il quantizzatore basato su CDF è spiegato nel dettaglio nell'Algoritmo 1. Inoltre, le sue soglie possono essere regolate per garantire la stessa proporzione di 0 e 1, caratteristica importante per la casualità.

Algoritmo 1: Algoritmo di quantizzazione basato su CDF

1) $F(x) = Pr(X^n \leq x)$

2) $\eta_i = F^{-1}(\frac{i}{2^{QL}})$, $i = 1, 2, ..., 2^{QL} - 1$

3) $\eta_0 = -\infty$

4) $\eta_{2^{QL}} = \infty$

5) Codifica di Gray b_i e lo assegno a un intervallo differente $[\eta_{i-1}, \eta_i]$

6) $K(j, QL) = b_i$, if $\eta_{i-1} \leq X_j \leq \eta_i$

In sostanza, il progetto del quantizzatore è la regolazione del livello di quantizzazione e della soglia al fine di avvicinarsi ad una prestazione ottimale della randomness, KGR e KDR. Ciò può essere esguito in diverse varianti, ad es. tramite regolazione adattiva della soglia per seguire la lenta variazione del segnale e infine evitare lunghi *uni* o *zeri* e migliorare la caratteristica di randomicità; quantizzazione multi-bit per un KGR più elevato ; abbassamento dei bit che non sono sullo stesso lato della soglia per un accordo migliore; ecc...

3.3 Information reconciliation

Anche se gli algoritmi di pre-elaborazione del segnale possono essere adottati per migliorare la correlazione incrociata delle misure dei canali, ci può essere ancora un disaccordo fondamentale tra Alice e Bob dopo la quantizzazione. Il disallineamento può essere corretto ricorrendo a tecniche di riconciliazione delle informazioni. Tali protocolli mirano a generare una sequenza casuale identica tra i due terminali sfruttando il canale pubblico.

Per un migliore tasso di generazione di chiave segreta, l'entropia di questa sequenza casuale dovrebbe essere massimizzata, mentre la quantità di informazioni trasmesse sul canale pubblico dovrebbe essere minimizzata. Molto spesso capita che le sequenze di bit quantizzate non coincidano per qualche bit. La riconciliazione di informazioni è designata alla risoluzione di questo problema. Permette ai due utenti di accordarsi sull'utilizzo di una stessa chiave, ad esempio utilizzando protocolli come Cascade o Error Correcting Code (ECC) come Low-density Parity-Check (LDPC), BCH Code, Reed-Solomon Code, Golay Code, Turbo Code, Secure sketch ecc.

I sistemi di riconciliazione basati su ECC sono più efficienti di Cascade, ma forniscono anche maggiori informazioni e maggiore complessità. La scelta dell'ECC dipende dalla complessità e dalla capacità di correzione. Ad esempio, il tasso massimo di capacità di correzione del codice $[n, k, t]$ BCH è dato come

$$\zeta = \frac{t_{max}}{n} = \frac{2^{m-2}/1}{2^m - 1} \tag{3.4}$$

che si avvicina a 0,25 quando m diventa grande.

A titolo esemplificativo viene introdotto lo sketch di sicurezza, illustrato anche nella figura 3.0.1. Viene adottato un ECC C per correggere il disaccordo.
Alice prima seleziona casualmente una parola in codice c da C e poi calcola s attuando la funzione XOR tra la sua sequenza di chiavi K^A e c, cioè $s = XOR(K^A, c)$, che viene poi inviata a Bob dal canale pubblico. Bob calcolerà c'' facendo lo XOR della sua sequenza di chiavi K^B con la sequenza di chiavi s ricevuta correttamente, cioè $c'' = XOR(K^B, s)$, e decodificherà c' da c''. Calcola K^B facendo l'XOR di c' con s, cioè $K^{B'} = XOR(c', s)$. Quando la distanza di Hamming tra c e c'' è minore della capacità di correzione t del codice di correzione, cioè $dis(c - c'') \leq t$, Bob può concordare la stessa chiave di Alice, cioè $K^{B'} = K^A$.

L'accordo fondamentale può essere confermato mediante l'applicazione del controllo ciclico della ridondanza (CRC) o di altri protocolli e strumenti, ad esempio il software AVISPA (Automatic Validation of Internet Security Protocols and Applications). Sussiste il rischio che il KDR superi il tasso di correzione della capacità di riconciliazione delle informazioni, con conseguente fallimento e riavvio dell'intero processo di generazione di chiavi a partire dal rilevamento dei canali.

3.4 Privacy amplification

Ultima l'amplificazione della privacy atta al fine di rimuovere eventuali perdite di informazioni.
Durante lo scambio di informazioni tra i due utenti attraverso il canale pubblico nello stadio di information reconciliation alcune informazioni vengono trasmesse pubblicamente. Tali informazioni, recepite dall'eavesdropper, possono rivelarsi cruciali e potenzialmente compromettere la sicurezza della sequenza di chiavi. L'amplificazione della privacy viene utilizzata per eliminare queste perdite di informazioni.
Essa può essere implementata tramite estrattori o funzioni universali di hash, come leftover hash lemma, cryptographic hash functions e Merkle-Damgarn hash function.

L'amplificazione della privacy e la riconciliazione delle informazioni appaiono sempre insieme, il che richiede una progettazione incrociata tra queste due fasi. Tuttavia, in pratica, è difficile quantificare la quantità di informazioni trapelate, o identificare dove si verifica la perdita nei dati[8].

Capitolo 4

Case study

Nel nuovo scenario che prendiamo in considerazione ci ritroviamo in una situazione in cui:

$$y_B = x + w_B, \qquad y_A = f(x) + w_A, \tag{4.1}$$

dove y rappresenta il segnale risultante dato da x (o una sua funzione) sommato al rumore w.
Studieremo tre diversi metodi d'approccio per ottimizzare, utilizzando dei metodi non ottimali ma agevolmente implementabili, il seganle x che arriva ad *Alice*.

4.1 Primo scenario

Come prima opzione prendiamo in considerazione l'applicazione di una semplice trasformazione prima $h_A = f^{-1}$ e quantizzazzione su y senza ulteriori arrangiamenti. Facendo ciò tutto quello che potrebbe generare un errore a livello di valore quantizzato deriva direttamente dalla presenza di rumore w_A nel segnale da quantizzare.

$$\begin{aligned} &Q_B(y_B), \qquad && Q_A(h_A(y_A)), \\ & && h_A = f^{-1}(), \\ & && Q_A(x + h_A(w_A)). \end{aligned} \tag{4.2}$$

Tale scelta risulta, ovviamente, ottimale a livello di ricerca del valore poichè applichiamo una semplice funzione inversa e ottenere in maniera precisa x. Non è invece ottimo per quanto riguarda la presenza di w_A che purtroppo è situato al di fuori di ciò che è nostro controllo.
Di conseguenza proseguendo nel processo di generazione chiavi e applicando ad esempio una funzione di hashing basta anche un singolo bit scorretto per portare *Alice* a recepire un messaggio completamente errato.

4.2 Secondo scenario

Per il secondo caso studieremo una procedura molto simile al criterio Mean squared error (MSE). Noto anche come deviazione quadratica media (MSD), ricerca il valore atteso dell'errore quadratico medio che misura la quantità di errore nei modelli statistici. Tale valore rappresenta la differenza media al quadrato tra i valori osservati e quelli previsti.
Quando un modello non presenta errori, MSE è uguale a zero. Con l'aumento dell'errore del modello, il suo valore aumenta. Un modello con meno errori produce previsioni più precise.

L'errore quadratico medio è quantificabile tramite la seguente formula:

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n}, \tag{4.3}$$

dove y_i è l'i^{th} valore osservato, $\hat{y}_i$ è il corrispondente valore previsto e n rappresenta numero di osservazioni.

I calcoli per l'errore quadratico medio sono simili alla varianza.

Definito con precisione tale principio è possibile utilizzarlo apportando qualche modifica. È da ricercare l'errore quadratico medio tra y_A e x. Definito da formula come sommatoria ne calcoleremo il valore atteso E. Al fine di raggingere il nostro scopo, in secondo luogo, ne ricercheremo il valore minimo tramite la funzione min generando una formulazione come quella trascritta di seguito.

$$\begin{aligned}
\min_{h_A(y_A)} E(|h_A(y_A) - x|^2) &= \int_{-\infty}^{+\infty} P_x(a) \left(\int_{-\infty}^{+\infty} P_{y_A|x}(b|a)\ |h_A(b) - a|^2\ db \right) da \\
&= \min_{h_A(b)} \int_{-\infty}^{+\infty} P_x(a) \left(\int_{-\infty}^{+\infty} P_{y_A|x}(b|a)\ |h_A(b) - a|^2\ da \right) db \\
&= \min_{h_A(b)} \int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ (h_A(b) - a)^2\ da \qquad \forall b \\
&= \min_{h_A(b)} \int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ \{[h_A(b)]^2 - 2a[h_A(b)] + a^2\}\ da \qquad \forall b.
\end{aligned} \tag{4.4}$$

Nell'equazione di qui sopra per prima cosa si invertono i domini di integrazione per poi semplificare l'integrale in b, poichè ricerco il minimo che assume la funzione per ogni valore di b. Successivamente risolvo il quadrato di binomio.

Necessaria risulta una sostituzione per semplificare l'equazione:

$$\begin{aligned} A(b) &= \int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ da \\ B(b) &= \int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ 2a\ da \\ C(b) &= \int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ a^2\ da. \end{aligned} \tag{4.5}$$

Sostituendo m iritrovo in una situazione di questo tipo:

$$= \min_{h_A(b)} \Big\{ [h(b)]^2\ A(b)\ -\ [h(b)]\ B(b)\ +\ C(b) \Big\}. \tag{4.6}$$

Per terminare infine sicalcola il minimo derivando rispetto ad $h(b)$ e si risolve ricavando $h(b)$ in funzione di b :

$$\begin{aligned} \tfrac{\partial}{\partial[h(b)]} \Big\{ [h(b)]^2\ A(b)\ -\ [h(b)]\ B(b)\ +\ C(b) \Big\} &= 0 \\ 2[h(b)]\ A(b)\ -\ B(b) &= 0 \\ h(b) &= \tfrac{B(b)}{2A(b)}, \end{aligned} \tag{4.7}$$

ed infine risostituisco:

$$\begin{aligned} h(b) &= \frac{\displaystyle\int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ 2a\ da}{\displaystyle 2\int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ da} \\ &= \frac{\displaystyle\int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ a\ da}{\displaystyle\int_{-\infty}^{+\infty} P_x(a) P_{y_A|x}(b|a)\ da}. \end{aligned} \tag{4.8}$$

Il senso di tale ricerca è quello di, dopo aver applicato h_A, minimizzare l'errore quadratico medio. Ciò significa che siamo piu vicini possibile ad un calcolo esatto stimando nella maniera migliore possibile il valore di x.
La principale carenza di correttezza che deriva dall'utilizzo di questa metodologia, risulta dal fatto che venga stimato al meglio il valore di x. Questo non significa che anche il valore quantizzato che ne deriva sia il migliore possibile. Necessario sarebbe eseguire in seguito un MAP su tale versione quantizzata scomponendo il problema in più casistiche. Per quanto concerne il nostro studio dello scenario quindi ci fermeremo qui.

4.3 Terzo scenario

Nel terzo scenario prendiamo in considerazione il criterio Maximum a posteriori (MAP) di x dato y_A. La MAP può essere utilizzata per ottenere una stima di una quantità non osservata sulla base di dati empirici. È strettamente correlato al metodo Maximum likelihood (ML), ma utilizza delle informazioni aggiuntive disponibili attraverso la conoscenza preliminare di un evento correlato, sulla quantità che si vuole stimare. La stima MAP può quindi essere vista come una regolarizzazione della Maximum likelihood estimation. Per *regolarizzazione* si intende l'introduzione di un'ulteriore informazione allo scopo di risolvere un problema con poche informazioni o per prevenire l'overfitting.
Di seguito definiamo il criterio MAP.

Si suppone di voler stimare un parametro di popolazione non osservato θ, sulla base delle osservazioni x. Presa f come distribuzione dei campioni di x, in modo che $f(x|\theta)$ sia la probabilità di x quando il parametro della popolazione è θ. Si assume infine che esista un funzione di distribuzione a priori g data da θ. Il metodo MAP stima quindi θ come parametro della funzione di distribuzione a posteriori:

$$\hat{\theta}_{MAX}(x) = \underset{\theta}{argmax}\ f(x|\theta)\ g(\theta) \tag{4.9}$$

Definito il criterio MAP bisogna apportare qualche modifica.
Tale metodologia massimizza la probabilità di decisione corretta, dalla quale ne deriverà la probabilità che x sia proprio uguale a $\hat{x}$, $P(x = \hat{x})$.
Come prossimo passaggio si mappa h_A (che prende in ingresso y_A) in x e in uscita restituisce $\hat{x}$ (una stima di x), $\hat{x} = h_A(y_A)$.
Di seguito verrà applicato il quantizzatore ottenendo una stima di y_A, $Q_A(\hat{x})$. Questo rappresenta l'ultimo passaggio al fine di ottenere il risultato cercato.

$$\hat{x} = \underset{a}{argmax}\ P_{y_A|x}(y_A|a)\ P_x(a), \tag{4.10}$$

Applicato il criterio MAP alla situazione esaminata risulta tale espressione, da suddividere in due parti.
La prima parte dell'espressione, $P_{y_A|x}(y_A|a)$, per la presenza di x e necessariamente dell'errore che ne consegue, risulterà descrivibile al meglio da una probabilità gaussiana.
Tale curva rappresenta come risulta distribuita la probabilità di y_A dato un certo valore di x uguale ad a.
Ritrovandoci nella situazione in cui, a varianza definita e media $f(a)$(come quella nel caso da noi esaminato), sommando alla funzione un numero, ne consegue un cambiamento della media. Tale media assumerà proprio il valore corrispondente a quello della costante aggiunta.

4.3.1 Caso con rumore gaussiano

Tale curva è rappresentabile mediante la funzione:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{4.11}$$

dove y_A si sostituisce ad y traccaindo una curva gaussiana con media $f(a)$ e varianza σ^2. La se-

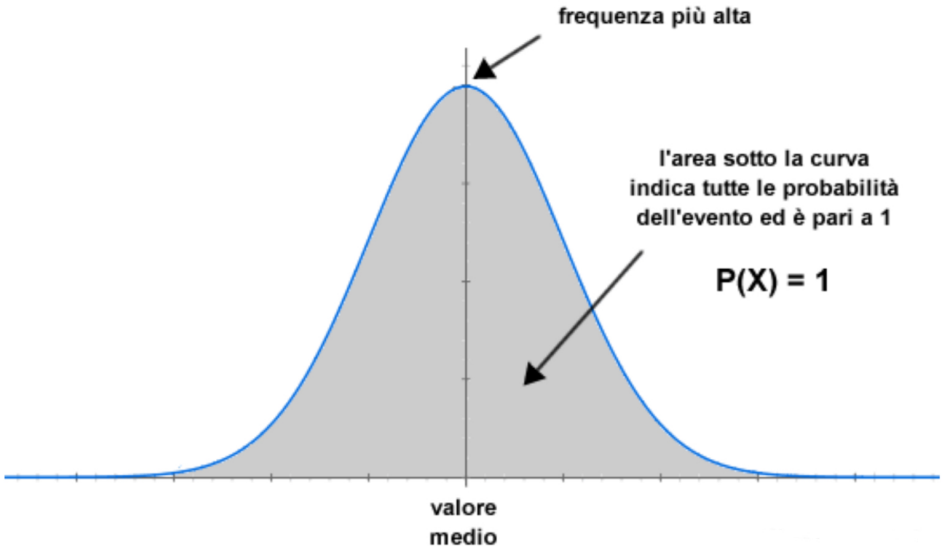

Figura 4.3.1: Rappresentazione curva Gaussiana.

conda, $P_x(a)$, presa x uniforme risulta uguale 1 diviso l'intervallo esaminato. Seguendo i canoni delle simulazioni rappresentate a fine capitolo prendiamo come limite un v_{sat} uguale ad 1. Ne risulta quindi che l'intervallo tra $-v_{sat}$ e v_{sat}, limiti inferiore e superiore risulti come il range compreso tra -1 e 1. Ultima condizione, risulta che l'intervallo sia uguale a 2 e di conseguenza $P_x(a) = \frac{1}{2}$.

Il malore massimo assunto dalla funzione gaussiana risulta nel così detto "picco della campana" (rappresentato dalla **frequenza più alta nella Fig 4.3.1**. Risultato del MAP quindi sarà che la stima migliore che può risultare in una tale funzione di x è proprio y_A.
Scegliendo quindi il valore di a per cui $f(a) = y_A$, per rimanere in linea con i canoni delle simulazioni, risulterà uguale a $a = f^{-1}(y_A)$. Ciò comporta che $y_A = f(a)$ e di conseguenza nel caso uniforme ci troveremo in una condizione analoga a quella descritta del Primo Scenario.

4.4 Risultati simulazione

Nella tabella sottostante sono riportati i risultati delle simulazioni eseguite al fine di valutare quale dei tre scenari precedenti risulti il più efficace nella trasmissione di un segnale tra *Bob* e *Alice*.

	PRIMO CASO	SECONDO CASO	TERZO CASO	QUARTO CASO
2 L.	9.2e-02	9.1e-02	8.0e-05	1.0e-04
4 L.	2.5e-01	2.6e-01	7.1e-05	7.1e-05
8 L.	2.9e-01	3.1e-01	6.5e-05	5.8e-05
16 L.	3.3e-01	3.4e-01	5.0e-05	5.0e-05

Nella tabella vengono rappresentati: nel "PRIMO CASO" l'errore generato dalla trasmissione di un segnale nel Primo Scenario e con funzione $f = e^x$, nel "SECONDO CASO" il Primo Scenario con funzione $f = x^3$, nel "TERZO CASO" il Secondo Scenario con funzione $f = e^x$ e nel "QUARTO CASO" il Secondo Scenario con funzione $f = x^3$.

Dai risultati delle simulazioni si possono notare alcune peculiarità.
Nel Primo Scenario, all'aumentare del numero di livelli di quantizzazione, aumenta anche l'errore quadratico medio, a differenza del Secondo Scenario che, salvo errori di approssimazione, si comporta in maniera completamente speculare.
In secondo luogo, sebbene si utilizzino funzioni diverse, in generale, a pari numero di livelli, i due risultati nello stesso scenario assumono valori molto simili.
Altro fattore, che ai fini della nostra ricerca risulta il più significativo, è che, a pari funzione, in qualsiasi risultato della simulazione utilizzare il Secondo Scenario porta in generale a ottimizzare la trasmissione. Questo deriva direttamente dal fatto che ci si sia concentrati nel minimizzare il più possibile il valore che assume l'errore quadratico medio. Da ciò si conclude quindi che, nonostante la randomness del canale, in ogni caso, si presenta come via migliore agire come possibile anche sul rumore hardware, evitando una casistica come quella descritta nel Primo Scenario.

Risolvendo il terzo scenario in funzione del nostro caso si ottiene una casistica identica a quella del Primo Scenario, rendendo superflue ulteriori simulazioni.

Conclusioni

Questa tesi ha esaminato le tecniche di advantage distillation per la generazione di chiavi sicure, mettendo in evidenza l'importanza di queste metodologie nell'ambito della crittografia e della sicurezza delle comunicazioni. Durante l'analisi dei vari approcci e algoritmi, abbiamo constatato la loro efficacia nel migliorare la sicurezza dei sistemi di comunicazione crittografica.

Le simulazioni condotte hanno evidenziato i progressi raggiunti attraverso l'impiego di tecniche di advantage distillation definite nei tre scenari presi in considerazione, evidenziando il ruolo cruciale che svolgono nella creazione di chiavi robuste e nella protezione delle informazioni sensibili. L'approfondimento delle peculiarità di questi metodi ha consentito di apprezzarne la semplicità e la versatilità, sottolineando la loro adattabilità ad alcuni contesti e scenari di utilizzo. Sono state affrontate casistiche elementari, ma esemplificative, per la comunicazione tra due utenti in un sistema wireless. Tali approcci sono stati definiti esclusivamente nel contesto in cui *Alice* e *Bob* comunicano, escludendo lo studio dell'intromissione di un eventuale *eavesdropper*. Questo poichè, in generale, un intercettatore viene presentato come un utente che ha a propria disposizione risorse infinite e capacità di calcolo ottimali. Nel nostro caso però, abbiamo affrontato casistiche basilari, barriere facilmente sormontabili da un calcolatore come quello descritto.

Le limitazioni associate agli approcci considerati riguardano la semplicità computazionale degli algoritmi impiegati. Nonostante vengano successivamente implementate funzioni di hash per cifrare le trasmissioni, l'utilizzo di funzioni elementari rende il sistema vulnerabile ad eventuali attacchi.

La sicurezza delle comunicazioni digitali rimane una priorità cruciale, e le conoscenze acquisite attraverso questa ricerca possono costituire una base per ulteriori innovazioni e miglioramenti nel campo della crittografia moderna. Fondamentale risulta, continuare ad esplorare nuovi sviluppi e affinamenti per garantire una sicurezza informatica sempre più avanzata, considerando l'evoluzione costante delle minacce digitali.

Bibliografia

[1] U. M. Maurer, «Secret key agreement by public discussion from common information», *IEEE transactions on information theory*, vol. 39, n. 3, pp. 733–742, 1993.

[2] R. Ahlswede, «Common Randomness in Information Theory and Cryptography CR Capacity», in *Identification and Other Probabilistic Models: Rudolf Ahlswedes Lectures on Information Theory 6*, Springer, 2021, pp. 231–269.

[3] D. ALIFFI e S. CORNACCHIA, «VERNAM E SHANNON LA CRITTOGRAFIA DA ARTE A SCIENZA»,

[4] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe e N. B. Mandayam, «Information-theoretically secret key generation for fading wireless channels», *IEEE Transactions on Information Forensics and Security*, vol. 5, n. 2, pp. 240–254, 2010.

[5] J. Zhang, R. Woods, T. Q. Duong et al., «Experimental study on key generation for physical layer security in wireless communications», *IEEE Access*, vol. 4, pp. 4464–4477, 2016.

[6] A. Rukhin, J. Soto, J. Nechvatal et al., *A statistical test suite for random and pseudorandom number generators for cryptographic applications*. US Department of Commerce, Technology Administration, National Institute of, 2001, vol. 22.

[7] J. Zhang, A. Marshall, R. Woods e T. Q. Duong, «Efficient key generation by exploiting randomness from channel responses of individual OFDM subcarriers», *IEEE Transactions on Communications*, vol. 64, n. 6, pp. 2578–2588, 2016.

[8] J. Zhang, T. Q. Duong, A. Marshall e R. Woods, «Key generation from wireless channels: A review», *Ieee access*, vol. 4, pp. 614–626, 2016.