SECRET KEY GENERATION SCHEME FROM WIFI AND LTE REFERENCE SIGNALS

Christiane L. Kameni Ngassa (Thales Communications and Security (TCS), Gennevilliers, France; Christiane.Kameni@thalesgroup.com); Renaud Molière (TCS; Renaud.Moliere@thalesgroup.com); François Delaveau (TCS; Francois.Delaveau@thalesgroup.com); Alain Sibille (TPT; Alain.Sibille@telecom-paristech.fr) ; Nir Shapira (Celeno Communications, Ra'anana, Israel; Nir.Shapira@celeno.com)

ABSTRACT

Physical layer security has emerged as a promising approach to strengthen security of wireless communications. Particularly, extracting secret keys from channel randomness has attracted an increasing interest from both academic and industrial research groups. In this paper, we present a complete implantation of a Secret Key Generation (SKG) protocol which is compliant with existing widespread Radio Access Technologies. This protocol performs the Quantization of the Channel State Information (CSI), then Information Reconciliation and Privacy Amplification. We also propose an innovative algorithm to reduce the correlation between quantized channel coefficients that significantly improves the reliability and the resilience of the complete SKG scheme. Finally we assess the performance of our protocol by evaluating the quality of secret keys generated in various propagation environments from real single sense LTE signals, and real single and dual sense WiFi signals.

1. INTRODUCTION

Recent news highlighting security failures of public wireless communication systems have recalled the limits of the cryptographic key distribution approach and the urge to improve security of the information exchanged over the air interface [1, 2, 3, 4]. The emergence of Physical layer Security (Physec) has provided an alternative approach for designing robust secret keys by leveraging the intrinsic randomness of wireless channels. This technique is referred to as Secret Key Generation (SKG) [5].

In § 2, we detail the typical scenario where two legitimate users (Alice and Bob) can communicate securely in presence of an eavesdropper (Eve), and how this principle works. The vast majority of existing works on SKG use the Received Signal Strength Indication (RSSI) since it is easily accessible. However, RSSI does not capture the entire richness of the channel as it ignores the phase of channel coefficients, which usually provide more randomness than the power of the signal. In this paper we present a full SKG scheme based on full Channel State Information (CSI) or its Fourier transform (Channel Frequency Response – CFR, §3). Our SKG protocol is composed of the following steps: Channel estimation (§3), Channel Coefficient de-correlation (§4), Quantization of the CSI (§5), Information Reconciliation (§6) and Privacy Amplification (§7).

In order to evaluate the performance of our scheme (quality of the generated keys, complexity of the processing), we apply our secret key generation protocol on single sense real field WiFi and LTE networks (§8) and on dual sense real WiFi signals (§9). Signals are captured in several indoor and outdoor locations, keys and estimates of channel entropy are computed from Channel Frequency Responses extracted from these real field records.

2. SECRET KEY GENERATION PRINCIPLE

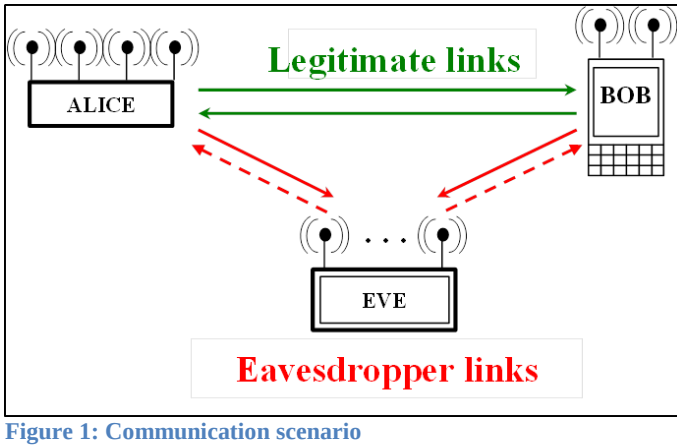


Figure 1: Communication scenario

2.1. Communication scenario

The legitimate users Alice and Bob attempt to communicate securely in presence of an eavesdropper Eve. For this, Alice and Bob observe and estimate the channel then they extract a common secret key from their channel estimates. When Eve is located at a distance of a few wavelengths from Bob, her channel measurements will be de-correlated of the legitimate channel and therefore any measure of Eve will be de-correlated to the secret key.

2.2. On channel randomness

The main reasons a secret key can be extracted from the random radio propagation are the following.

In indoor and outdoor environments, waveforms transmitted from Alice to Bob and Eve follow multiple paths and come across various obstacles with distinct angles of incidence. As a result, they are altered very differently when they are received by Bob and Eve. A few wavelengths are enough to ensure a complete de-correlation between Bob and Eve's channels, especially when the scatterers' Angular Spread (AS) is large [6].

Besides, due to complex wave propagation and unpredictable scatterers in the communication channel, Eve cannot predict or recover the legitimate channel.

Finally, in TDD mode and for each carrier, waveforms from Alice to Bob hit obstacles in the forward and return direction with the same angle of incidence. Therefore the legitimate users see the same randomness and thus have similar channel measurements. This phenomenon is referred to as "channel reciprocity".

Consequently, the channel coefficients measured by Alice and Bob characterize the legitimate link and cannot be reconstructed by Eve. Thus, Alice and Bob can use this shared pool of randomness to generate secret keys.

2.3. Secret key generation steps

The proposed SKG protocol is composed of the following steps:

Channel Estimation: the first step of the SKG scheme estimates the radio channel and computes CSI or CFR

Channel Coefficient de-correlation: in this second step, we apply a new algorithm to select channel coefficients with low cross correlation. This optimizes the randomness selection in stationary environments.

Quantization: this step uses the Channel Quantization Alternate (CQA) algorithm introduced by Wallace to quantize selected channel coefficients [5], that minimizes key mismatch between the legitimate users Alice and Bob.

Information Reconciliation: this step corrects the remaining mismatch between Alice and Bob keys. We employ secure sketch and error correcting codes to correct Bob's errors on Alice's key. To do so, Alice has to send the secure sketch over the public channel, possibly leaking a controlled amount of information to the eavesdropper Eve.

Privacy Amplification: this step improves the randomness of the secret key and removes the redundant information that could be used by Eve. To do so, we use hash functions and, when necessary, reduce key length. This final step guarantees that the generated secret key is fully de-correlated from the key computed by the eavesdropper.

Note: searching for practical implementation inside communication devices, we focused in each step on most robust and simple algorithms. For example, we choose a simple algebraic forward error correcting (FEC) code to reconcile Alice and Bob keys and a classical family of 2-universal hash function in the privacy amplification step [7].

3. CHANNEL ESTIMATION

When considering an Orthogonal Frequency Division Multiplexing signal (OFDM, such as encountered in WiFi and LTE networks) in the frequency domain, the component of the Channel Frequency Response (CFR) H_f quantifies the fading applying on each subcarrier. In a sampled system, considering a finite response and band, the k^{th} frequency component f_k of the CFR can be calculated as follows:

$$\widehat{H}_f\ (k)\ = Y(f_k)/X(f_k)$$

where Y is the received signal, and X is the emitted signal (or reference signal).

In the time domain, an equivalent Channel Input Response (CIR) estimation can be deduced from the CFR by IFFT, as follows:

$$\widehat{H}_{IFFT}\ =\ IFFT(\ \widehat{H}_f)$$

When considering now TDMA of CDMA wave forms encountered in 2G and 3G radio Access technologies (RAT), CIR can be computed directly in the time domain by applying filter estimations techniques to reference signal X.

4. CHANNEL DECORRELATION

Secret key bits should be completely random to keep them unpredictable by Eve, therefore any deterministic component in the radio propagation channel should be removed. Same apply to any time or frequency correlation between quantized bits: the quantization algorithm should not only generate bits with equal probability but also the channel coefficients that are quantized to generate these bits should be as random and de-correlated as possible.

The goal of this step is to decrease the negative effect of channel correlation by a careful selection of the channel coefficient to be quantized.

First, time correlation is decreased between channel coefficients. To do so:

- Channel coefficients computed at a given time acquisition constitute a frame.
- Cross-correlation coefficients are computed between the two first frames
- Only frames with low cross-correlation coefficient (under a given threshold T_t) are selected.
- Cross-correlation coefficients are computed between the previous selected frame and the next frame.

Then, same procedure applies to frequency correlation:

- Cross-correlation coefficients are computed between two consecutive frequency carriers
- Only frequency carriers for which the cross-correlation coefficient is below a given threshold T_f are selected. In addition, lowest and highest frequency carriers are dropped.

Finally, Alice sends to Bob the position of the channel coefficients over the public channel. Hence, Eve also knows which coefficients were dropped and which ones were selected but she does not have any information on their value. Therefore there is no information leakage during the channel de-correlation step.

5. QUANTIZATION

After measuring the radio channel, Alice and Bob jointly employ an algorithm to quantize the channel taps that they have estimated in order to generate a common sequence of key bits from their instantiation of the shared channel, under reciprocity assumption.

However, due to noise and channel estimation errors, Alice and Bob may disagree on some key bits. Several quantization algorithms employing censoring schemes have been developed to limit this mismatch between Alice and Bob keys.

A typical censoring algorithm defines guard band intervals and discards any channel measurement falling into it [5]; leading to an inefficient exploitation of channel measurements and to a lower number of generated key bits.

Thus, other schemes employ different quantization maps where each one is adapted to the channel observations, e.g. channel quantization alternating (CQA) algorithm [5]. The principle consists in choosing the adaptive quantization map where the current observation is less sensitive to mismatch. Consequently, we apply the CQA algorithm to complex channel coefficients to generate secret key bits.

6. INFORMATION RECONCILIATION

This step suppress remaining mismatches between Alice and Bob keys by using secure sketch based on error-correcting codes [8]. The key computed by Alice is considered as the secret key and Bob wants to retrieve Alice's key using the key K_b he extracts from his channel measurements.

The processing can be described as follows:

Alice:

- selects a random codeword c from an error-correcting code C
- computes the secure sketch $s = K_a \oplus c$
- sends s to Bob over the public channel

Bob:

- subtracts s from its computed key K_b :
$$c_b = K_b \oplus s \ (= K_b \oplus K_a \oplus c)$$
- decodes c_b to recover c and gets $\hat{c}$
- computes K_a by shifting back and gets:
$$\widehat{K_a} = \hat{c} \oplus s$$

Perfect reconciliation is achieved when Bob perfectly retrieves the random codeword chosen by Alice, meaning that $\hat{c} = c$. As a result, no mismatch occurs between Alice and Bob keys ($K_a = K_b$).

Therefore the secure sketch s, sent over the public channel, allows the exact recovery of the secret key without revealing the exact value of the key.

However, s might leak some information on the secret key over the public channel as Eve can also use the secure sketch to retrieve the secret key K_a.

Thus, a final step is then necessary to suppress the leaked information and to improve the quality of the secret key.

7. PRIVACY AMPLIFICATION

The objective of the privacy amplification step is to erase the information leaked to Eve on the secret key during the information reconciliation step and to improve the randomness of the key.

For our SKG scheme we interpret the secret key K as an element of the Galois Field $GF(2^n)$ and we choose the following two-universal family of hash functions [9] where n is the number of bits of the key K.

For $1 \leq r \leq n$ and for $a \in GF(2^n)$, the functions $\{0,1\}^n \rightarrow \{0,1\}^r$ assigning to the key K the first r bits of key $a.K \in GF(2^n)$ define a two-universal family of hash functions. r is the final length of the secret key.

In practice, at each new key computation, the parameter a is randomly chosen by Alice who sends it to Bob over the public channel. Alice and Bob then compute the product $a.K \in GF(2^n)$.

The hash mechanism spreads any bit error all over the final key $(a.K)_{r\ bits}$ (first r bits of $a.K$), thus:

- When Eve tries to recover the initial key K (at the reconciliation step), any error on K will make the final key $(a.K)_{r\ bits}$ unusable for her.
- Bob has to perfectly recover the initial key K (i.e. reconciliation should be perfectly achieved) in order to get the usage of the final key $(a.K)_{r\ bits}$.

8. EXPERIMENTAL RESULTS FROM SINGLE SENSE SIGNALS

In this section we generate keys from real LTE and WiFi signals acquired using the test bed of figure 2 and we analyze their quality and the processing complexity.

8.1. Impact of channel de-correlation

Figure 3 is relevant to a very stationary the propagation environment (empty indoor tennis court, static Alice and Bob, static scatterers) and shows the direct output of the CQA algorithm (§5) with 4 Quantization Regions (QR). CFR computed from LTE signals over 5 seconds

(frequency: 2627.5MHz, bandwidth: 1.4 MHz) produced 1000 frames detections and 122 secret bits per frames. However, we can notice a repetitive pattern on the generated keys meaning that CFR coefficients are highly correlated in time and in frequency. This high correlation represents a major vulnerability as the generated secret key bits will not be random enough.

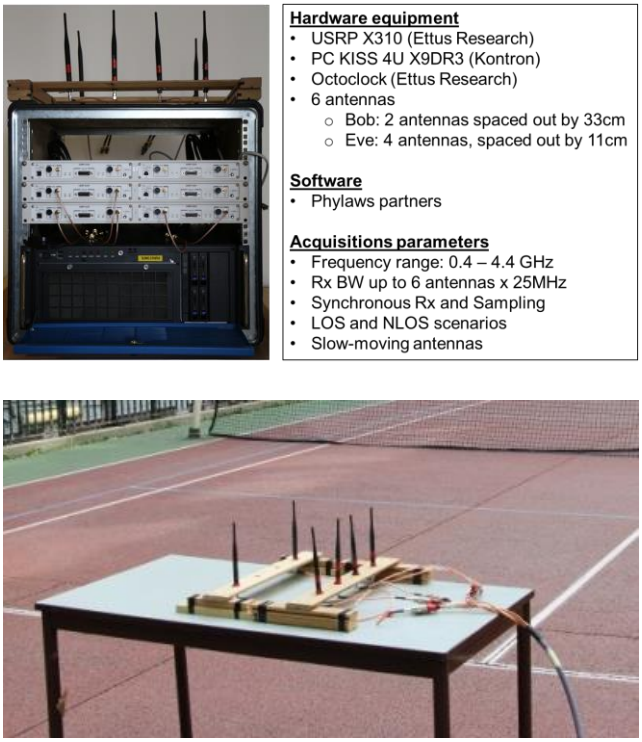


Figure 2: PHYLAWS test bed ([10]) and configuration of antennas during recordings

Figure 4 shows key bits obtained on the same record with the same processing after applying our channel coefficient selection (§4) on the original CFRs: the correlation between bits has significantly decreased both in time and frequency (our algorithm managed to extract the repeating pattern of the key bit). However the price to pay is fewer secret key bits.

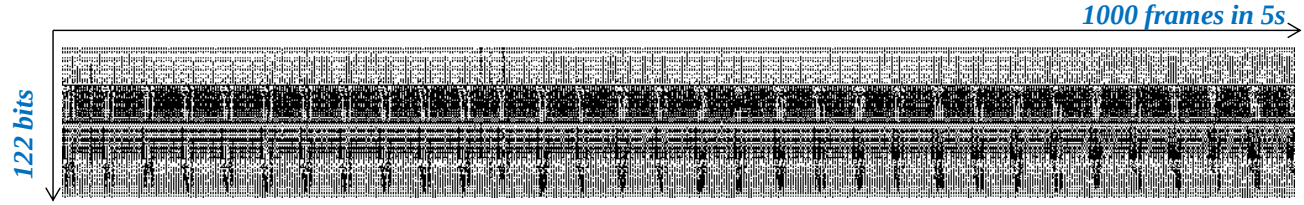


Figure 3: Resulting key bits after quantization of all available channel coefficients

Figure 4: Resulting key bits after channel de-correlation

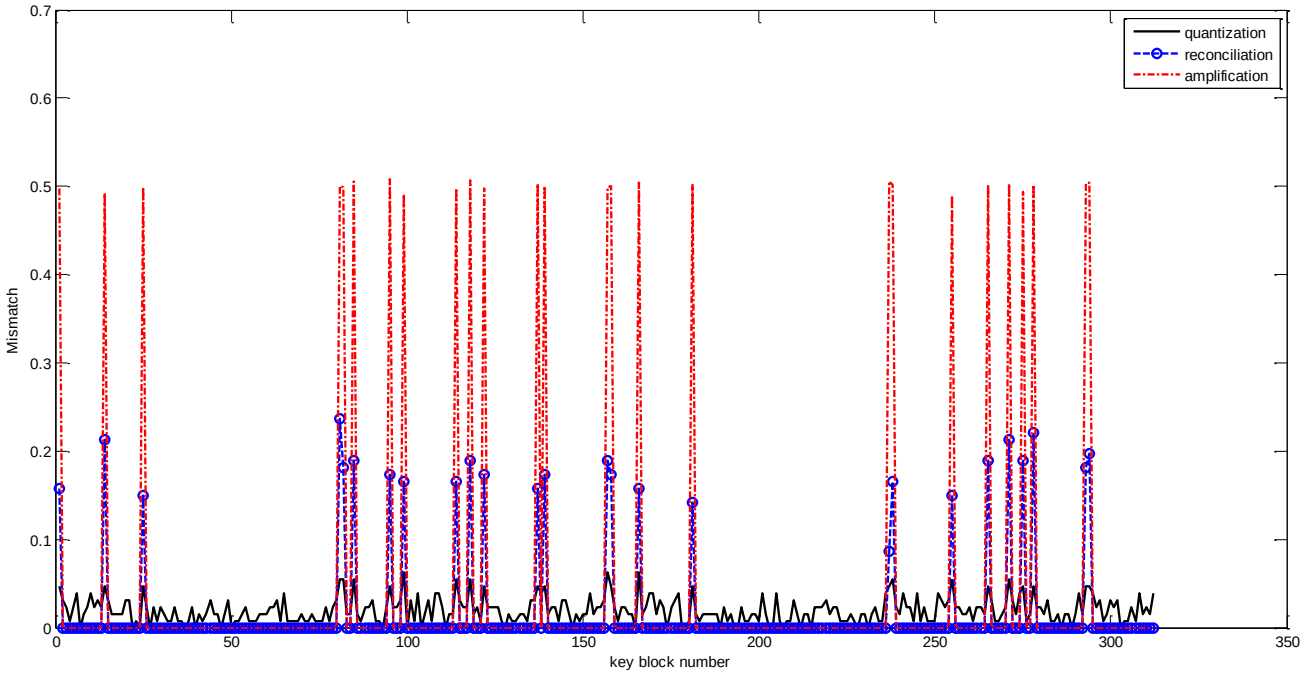


Figure 5: Mismatch between Alice and Bob for a low SNR value of legitimate link

8.2. Analysis of the Mismatch between Alice and Bob

In this section we analyze the key bit error rate (or "key mismatch") between Alice and Bob's generated keys. 320 keys of 127 bits were generated under a Wifi carrier (IEEE 802.11a, frequency 2462 MHz, Bandwidth: 20 MHz) in an open space environment (office) and three different SNR (20 dB, 25 dB and 28 dB) were considered.

For each SNR we plot the mismatch of Alice and Bob keys after each step of our SKG scheme.

- Black curves represent the mismatch after quantization
- Blue curves represent the mismatch after information reconciliation
- Red curves represent the mismatch after privacy amplification

Our quantization step (§5) uses the CQA Algorithm with 4 regions. Our reconciliation step (§6) uses secure sketch based on a (127, 92, 11) BCH code. Our privacy amplification step uses the 2-universal family of hash functions of §7.

Figure 5 plots the key mismatch between Alice and Bob for a low value of the SNR: here the number of errors is much higher than the error-correction capability of the BCH code, and key mismatched remain. A more powerful FEC code would optimize the information reconciliation.

We note that when the information reconciliation step fails, it increases the key mismatch compare to its value after the quantization step. Moreover the privacy amplification induces two extreme behaviors.

- When there is no error between the Alice and Bob's key, the mismatch remains null
- However, for any non-zero value, the mismatch is driven to 0.5. Thus privacy amplification increases the confusion on key mismatches when Bob's does not success to extract the same key than Alice. Same applies to Eve.

Figure 6 shows the same results when considering SNR=28dB. Here all errors were corrected by the information reconciliation step thus Bob and Alice generate the same keys after privacy amplification.

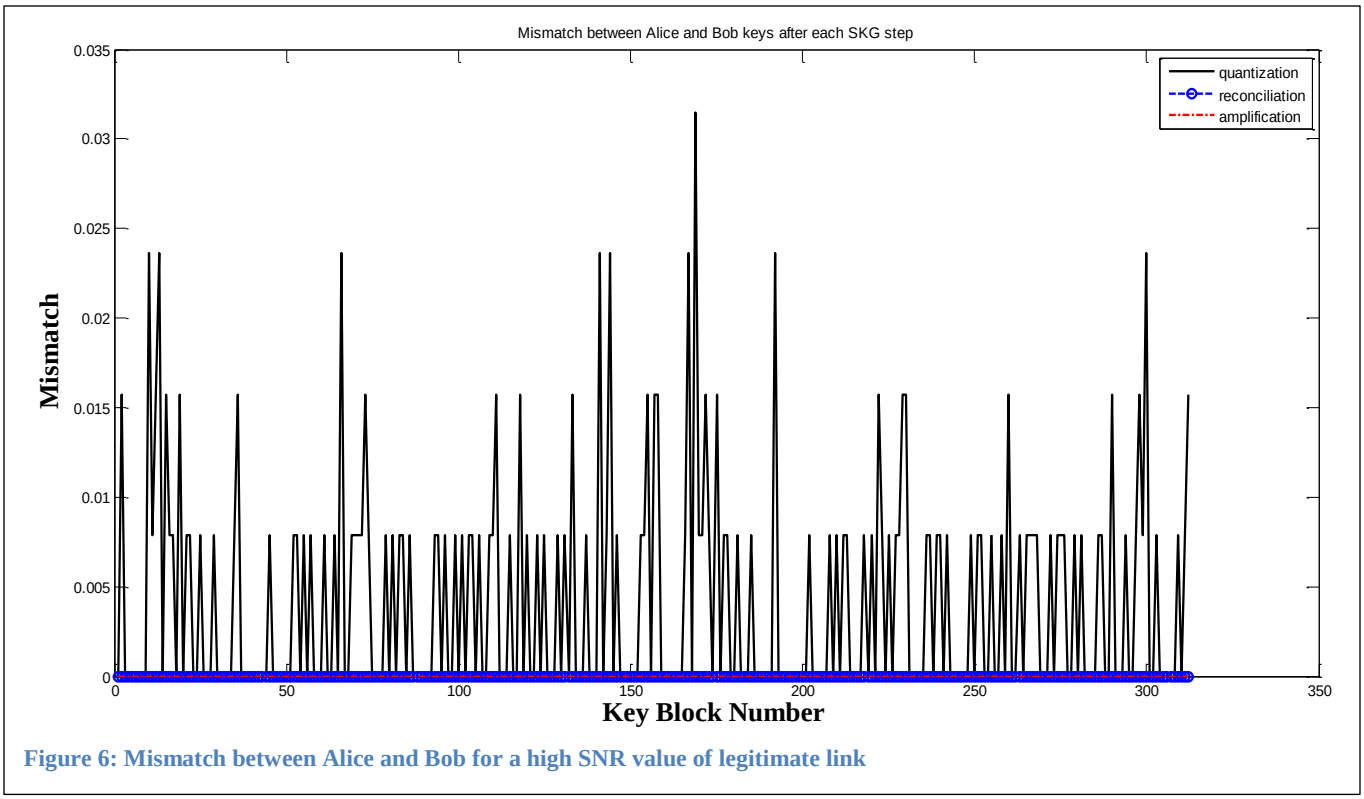


Figure 6: Mismatch between Alice and Bob for a high SNR value of legitimate link

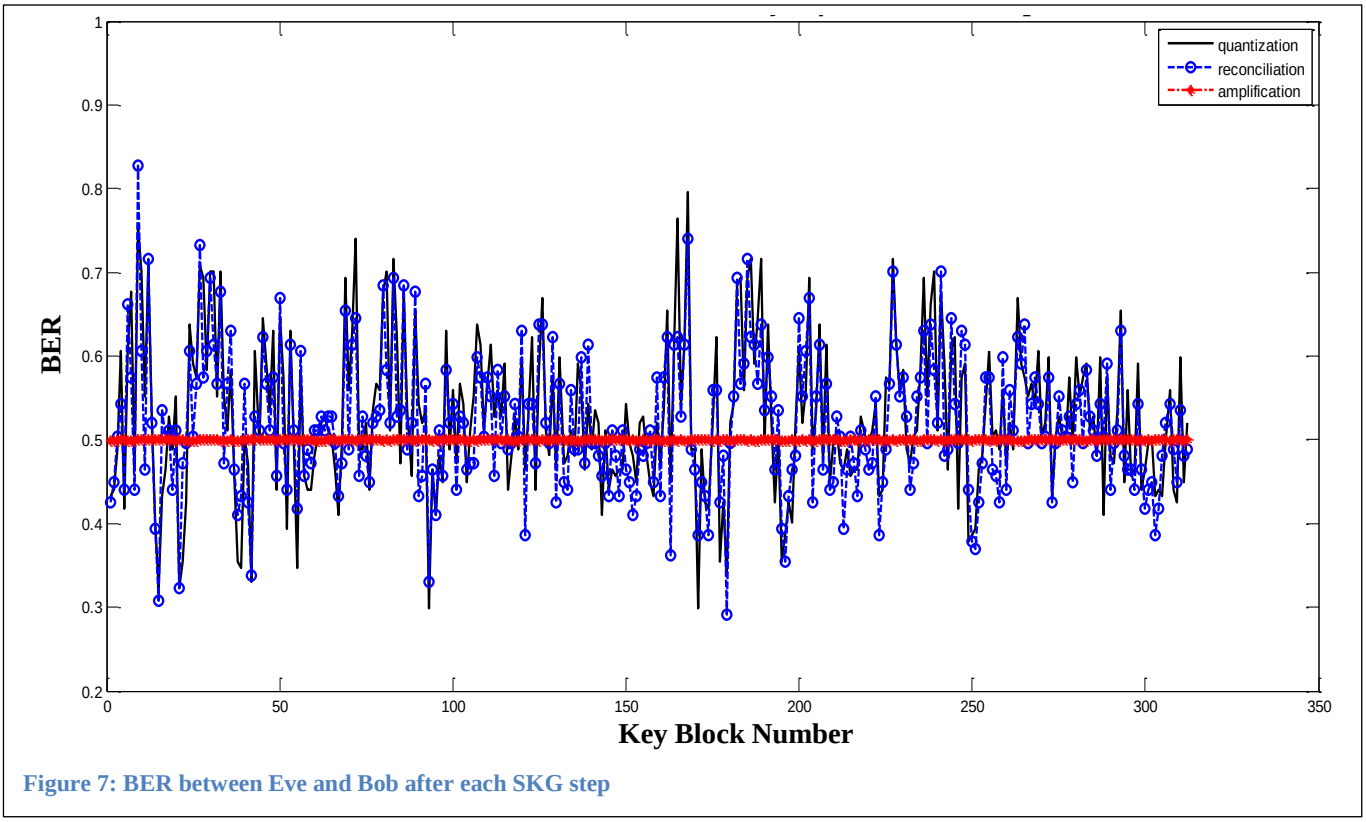


Figure 7: BER between Eve and Bob after each SKG step

8.3. Analysis of the BER between Eve and Bob

In this section, under the same WiFi carrier as above, we evaluate the number of errors that Eve makes on Bob's key.

We analyze the Bit Error Rate between Eve and Bob when Eves perform the same process than Bob to Alice's signals with some antenna advantage (Eve has four antennas for her CFR estimations while Bob has only two antennas).

Figure 7 plots the BER between Eve and Bob for each of the 320 generated keys.

We note that BER does not change much after information reconciliation. However, after privacy amplification, the BER is driven to 0.5 after.
In other words, reconciliation has low impact on Eve but privacy amplification highly increases the confusion of Eve on Bob's key and ensures that Eve's key is de-correlated from Bob's Key.

Figure 7 shows that Eve's BER after privacy amplification is 0.5. Thus, Eve has no information on the value for each bit of Bob's key. Hence, further investigation showed that no vulnerability occurred to particular bits.

Nevertheless, theoretically, information was leaked during the information reconciliation step (exchange of the secure sketch s). Therefore a corresponding number of bits should be removed from the key.

Denote N the length of the FEC code used for information reconciliation and R the rate of the code. The secure sketch s sent over the public channel leaks information on N(1-R) bits of Bob's key. Therefore the secret key length should be decreased to N*R.

8.4. Analysis of the randomness of the keys

In this section we study secret keys computed from WiFi Carrier (2462 MHz, Bandwidth: 20 MHz) and LTE signals (Frequency: 2627.5 MHz, Bandwidth: 1.4 MHz).

Wifi Carrier - Indoor environment (open space office) with slow mobile antennas and LOS configuration).

Figure 8 and figure 9 show that a significant number of keys are generated thanks to the mobility of the antennas and that the keys after quantization appear relatively random.

Wifi Carrier - Indoor environment (open space office) with slow mobile antennas and NLOS configuration).

Figure 10 and figure 11 show that a larger number of keys are generated. From observation, the randomness of the keys seems convenient.

LTE Carrier - Indoor environment (classroom) with static antennas and limited mobility of scatterers).

- Figure 12 shows quantization outputs: some patterns can be detected and the keys do not look really random.

- Figure 13 shows the keys after privacy amplification: the keys seem more random. It confirms that this step provides an extra level of security and improves the randomness of the keys.

LTE Carrier - Urban outdoor environment with static antennas and mobile people and cars:

- Figure 14 shows that more keys are generated outdoor compared to indoor. However some patterns exist within key bits after quantization.

- Figure 15 shows privacy amplification outputs: it confirms that the key are numerous and that key randomness is significantly improved.

Finally, all these figures show the following trend

- more mobility and richness in the channel provide more keys of better randomness

- secret keys can be rapidly generated: 49 keys in 5 seconds in a static environment to 152 keys in 2 seconds when antennas are mobile.

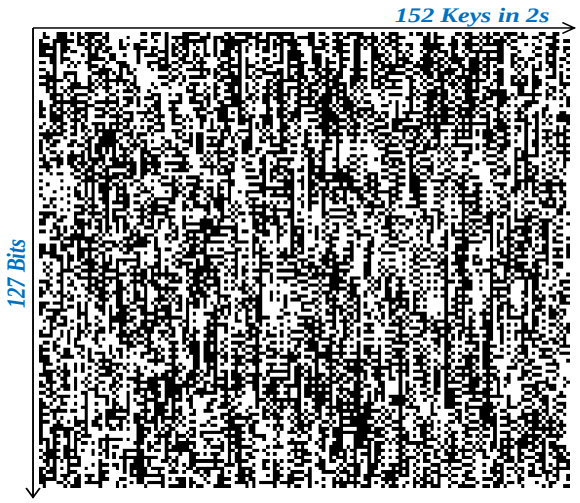


Figure 8: Key bits after quantization (WiFi, 2462 MHz , indoor LOS)

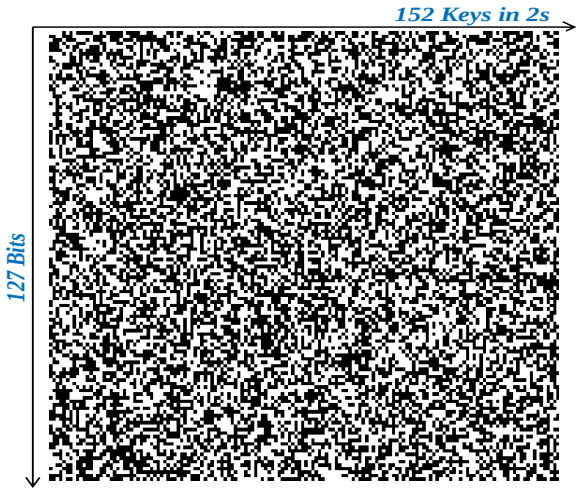


Figure 9: Key bits after privacy amplification (WiFi, 2462 MHz, indoor LOS)

Figure 10: Key bits after quantization (WiFi, 2462 MHz , indoor NLOS)

Figure 11: Key bits after privacy amplification (WiFi, 2462 MHz, indoor NLOS)

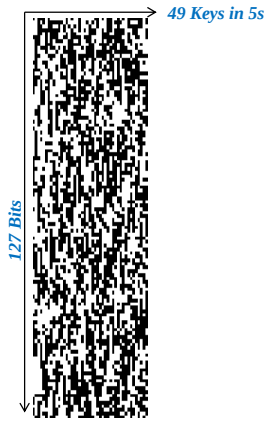


Figure 12: Key bits after quantization (LTE, indoor classroom, 2627.5 MHz)

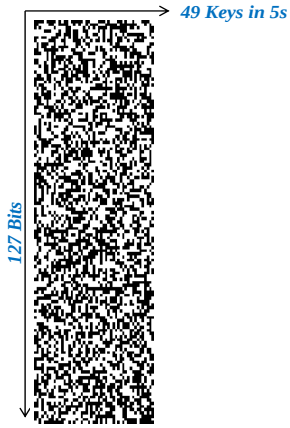


Figure 13: Key bits after privacy amplification (LTE, indoor classroom, 2627.5 MHz)

Figure 14: Key bits after quantization (LTE, 2627.5 MHz, outdoor urban street)

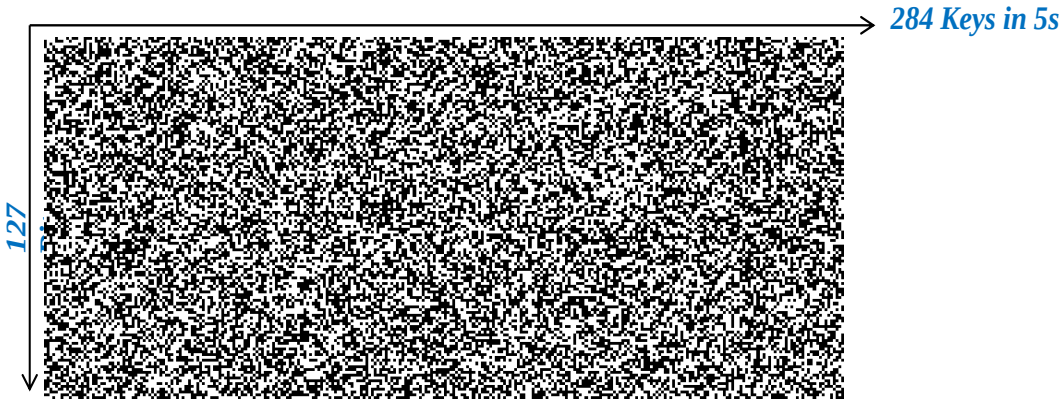


Figure 15: Key bits after privacy amplification (LTE, 2627.5 MHz, outdoor urban street)

8.5. NIST Statistical tests of computed keys

In this section we evaluate the quality of the keys by performing two randomness tests defined in the NIST Statistical Test Suite [11].

The NIST tests are well suited for off-line evaluation of the randomness of generated keys. However since several NIST tests are required to guarantee the randomness of a sequence of bits, the NIST tests cannot be used for online testing.

- **NIST frequency mono-bit test**

The goal of this test is to determine whether the numbers of 0s and 1s in the key are approximately the same as would be expected for a truly random sequence.

Table 1 provides the percentage of keys that successfully passed the frequency mono-bit test for the previous LTE and WiFi signals.

According to the results, almost all the keys pass the test after quantization and the privacy amplification increase the percentage of successful keys to 99% and 100%.

- **NIST runs tests**

The goal of this test is to determine whether the oscillation between 0s and 1s is too fast or too slow compared to what it is expected for a truly random sequence.

Table 2 provides the percentage of keys that successfully passed the runs test for the previous LTE and WiFi signals.

When considering quantization only, and according to the previous results,

- only a small percentage of keys generated in the indoor environment with limited mobility passed the tests
- a high percentage of keys generated with dispersive channels passed the test.

Note about the LTE Indoor case after quantization:

- Most of the keys that did not pass the runs test passed the frequency mono-bit test which is less stringent (since the CQA algorithm divides the CFR in equi-probable regions, it is expected that the number of 0s and 1s in each key should be approximately equal, which matches the frequency mono-bit test).
- The runs test better captures the randomness of a sequence. (Since CFRs captured on 1.4 MHz bandwidth only in indoor environment were a little correlated, the keys steam after quantization provides time and frequency correlation which are rejected).

Note about the benefit of privacy amplification:

After privacy amplification step, the success to NIST test is always improved, even in the static indoor environment. This final step of our SKG scheme appears really necessary for processing low dispersive radio environments and narrow band signals.

Table 1: Frequency monobit test results

LTE	**Indoor (2.6GHz)**	**Outdoor (2.6GHz)**
Quantization	**98%** (48/49)	**99%** (281/284)
Amplification	**100%** (49/49)	**100%** (284/284)

WIFI	**LOS (2.4 GHz)**	**NLOS (2.4 GHz)**
Quantization	**87%** (132/152)	**100%** (171/171)
Amplification	**99%** (151/152)	**100%** (171/171)

Table 2: Run test results

LTE	**Indoor (2.6GHz)**	**Outdoor (2.6GHz)**
Quantization	**27%** (13/49)	**80%** (228/284)
Amplification	**100%** (49/49)	**100%** (284/284)

WIFI	**LOS (2.4 GHz)**	**NLOS (2.4 GHz)**
Quantization	**84%** (128/152)	**99%** (169/171)
Amplification	**98%** (149/152)	**99%** (170/171)

Table 3: Entropy estimates for LTE, static environment (tennis court), LOS, 2.6 GHz

	Min-entropy estimates					
	Antenna 1	Antenna 2	Antenna 3	Antenna 4	Antenna 5	Antenna 6
Min-entropy	0.195466	0.323639	0.226037	0.499656	0.289194	0.323639

	Joint entropy estimates					
	Ant. 1 - 2	Ant. 1 - 4	Ant. 2 - 4	Ant. 5 - 6	Ant. 1 - 6	Ant. 2 - 3
Min-entropy	0.315329	0.663064	0.693907	0.296657	0.315329	0.264215
Max min-entropy	1.8317	1.91236	1.78921	1.37881	1.8317	1.37596

	Mutual information					
	Ant. 1 - 2	Ant. 1 - 4	Ant. 2 - 4	Ant. 5 - 6	Ant. 1 - 6	Ant. 2 - 3
Max mutual information	0.165279	0.196586	0.385967	0.739011	0.248614	0.839857

Table 4: Entropy estimates for WiFi, NLOS, 2.4 GHz

	Min-entropy estimates					
	Antenna 1	Antenna 2	Antenna 3	Antenna 4	Antenna 5	Antenna 6
Min-entropy	0.630948	0.740257	0.697361	0.652379	0.761595	0.740257

	Joint entropy estimates					
	Ant. 1 - 2	Ant. 1 - 4	Ant. 2 - 4	Ant. 5 - 6	Ant. 1 - 6	Ant. 2 - 3
Min-entropy	1.20996	1.18903	1.18662	0.652329	1.20996	0.640132
Max min-entropy	1.89685	1.95025	1.97559	1.35019	1.89685	1.34821

	Mutual information					
	Ant. 1 - 2	Ant. 1 - 4	Ant. 2 - 4	Ant. 5 - 6	Ant. 1 - 6	Ant. 2 - 3
Max mutual information	0.180416	0.198405	0.206016	0.849523	0.19369	0.797486

8.6. Entropy estimation and analysis versus the channel stationarity

The aim of this section is to evaluate the percentage of entropy bits extractable from the radio channel in realistic radio environment. To do so, we estimate the min-entropy of channels, first between Alice and Bob, then between Alice and Eve, at the output of the quantization step of the SKG scheme (without applying the channel de-correlation). Our computation uses NIST's tests for Estimating the Min-Entropy of non-IID Sources described in [12].

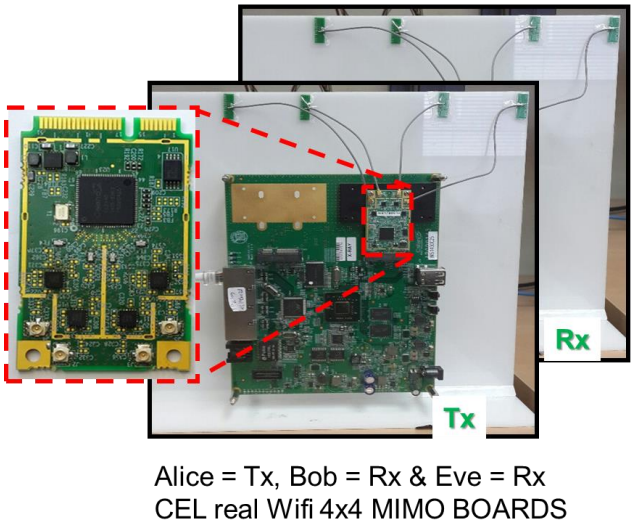
We also estimate the joint entropy and the mutual information between pairs of antennas in order to evaluate the dependence between two distinct antennas.

Therefore, the mutual information can be used as an indicator of the common information shared by two receivers. Hence, for a given pair of antennas, the entropy and the mutual information can provide us an experimental insight on the percentage of secure entropy bits.

Table 3 and 4 provide the results for the six antennas of the test bed (figure 2). Recall that Antennas 1 and 4 where dedicated to Bob while antennas 2, 3, 5 and 6 where dedicated to Eve, the six antennas being closed to each other. The results are provided for two extreme propagation environments. The first one, very stationary, is an empty tennis indoor court surrounded by building and a LTE e-node, the geometry is fixed and LOS. The second one, much less stationary is an indoor office where antennas where slightly mobile and WiFi signals come from by NLOS access points.

The results show that there are at least 20% of entropy bits in the first (worst) case and around 70% of entropy bits in the second (better) case.

In addition, the computed maximum value of the mutual information between pairs of antennas reveals that one antenna on Eve's array only shares around 20% of information with one antenna on Bob's array.



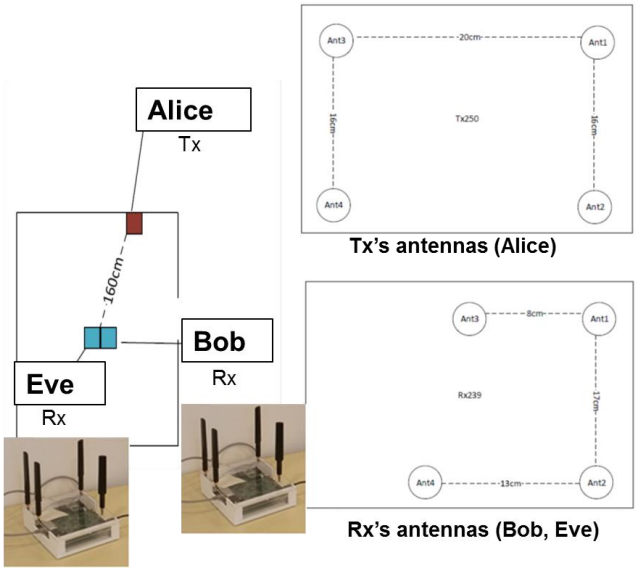


Figure 16: 4x4 WiFi chipsets and measurement environment

9. EXPERIMENTAL RESULTS FROM DUAL SENSE SIGNALS

In this section we generate keys from dual sense real signals emitted and received by WiFi chipsets designed by Celeno Communication Ltd. We then evaluate the randomness and secrecy of generated keys.

9.1. Test bed and measurement environment

The test bed depicted in Figure 16 is based on a state of the art 4x4 MIMO chipsets made by Celeno. Each Chipset is based on a Software Defined Radio architecture, using a Digital Signal Processing core that enables to implement algorithms in the physical layer on top a real WiFi system.

The test bed supports operation in both 5GHz and 2.4GHz bands by using two different chips: the CL2440 is a 4x4 AP chip supporting 5GHz operation (for up to 80 MHz bandwidth), while the CL2442 is a 4x4 AP chip supporting 2.4GHz operation (for up to 40 MHz bandwidth). The test bed is also hooked to the local network via Ethernet for control and for data extraction.

A typical placement of the antennas for transmitter (Tx) and receiver (Rx) boards is shown in Figure 16. The antenna spacing on the test bed is always more than half of a wave length (2.7cm in 5.5 GHz and 6.25 cm in 2.4 GHz) to provide adequate diversity.

Experiments are carried out in Celeno's testing apartment. The apartment provides a clean testing environment that is relatively interference free. Various indoor NLOS and LOS scenarios can be emulated.

9.2. Description of a bi-directional sounding exchange

Alice and Bob exchange WiFi sounding frames (2462 MHz, Bandwidth: 20 MHz).

- Alice first sends a sounding frame which is captured by Bob (and Eve).
- Bob sends back to Alice a sounding frame.
- Alice, Bob and Eve extract 4x4 channel estimates.

CSI estimates are then processed in Matlab offline.

- In the first phase Alice Bob and Eve compensate their channel estimation for timing errors and normalizes each channel coefficient.
- In the second phase, a Matlab processing script involves secret key extraction from channel estimates and evaluation of the generated keys.

The main steps of the SKG scheme are recalled below.

- CSI coefficient selection in a pre-processing step.
- Dual sense CSI quantization using CQA algorithm.
- Information reconciliation with BCH codes.
- Privacy amplification using two-universal family of hash functions and, when necessary, key length reduction avoiding any capability for Eve to exploit the FEC reconciliation code redundancy.

Generated keys are evaluated as follows:

- Test of key randomness by using the Intel Health Check applied on keys after quantization and privacy amplifications steps.
- Computation of the mismatch between Alice and Bob's keys.
- Computation of the BER between Bob and Eve.

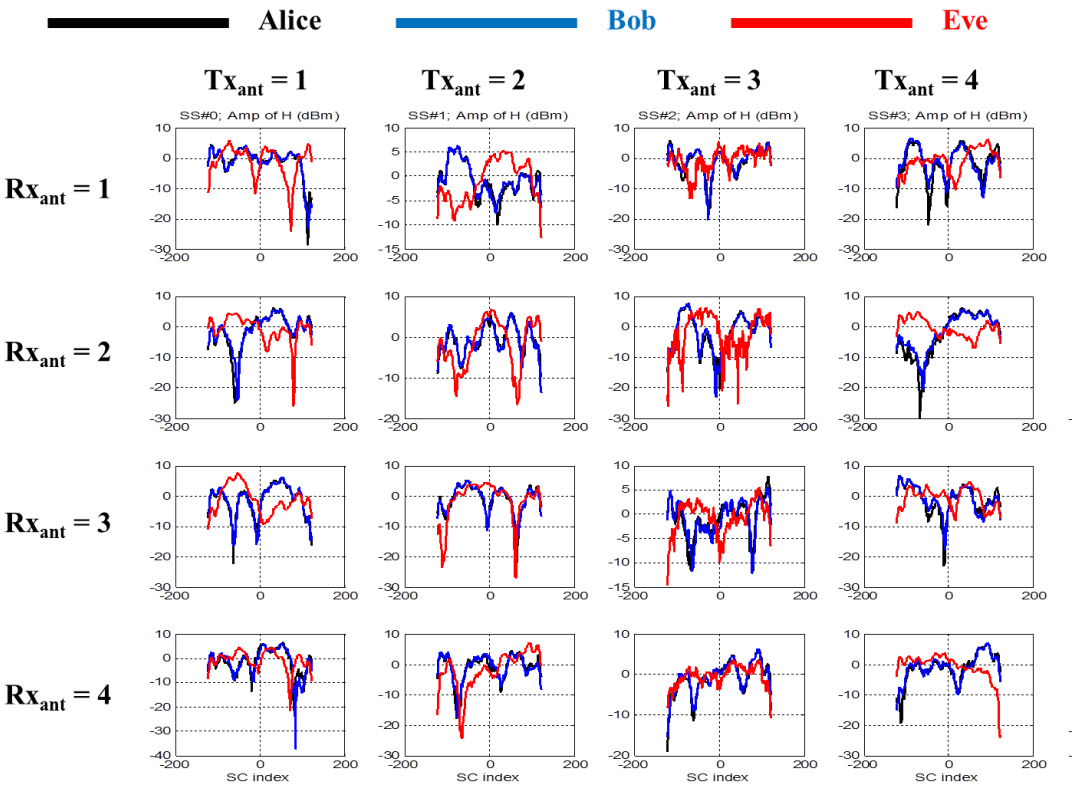


Figure 17: Amplitude channel measurements for Alice, Bob and Eve

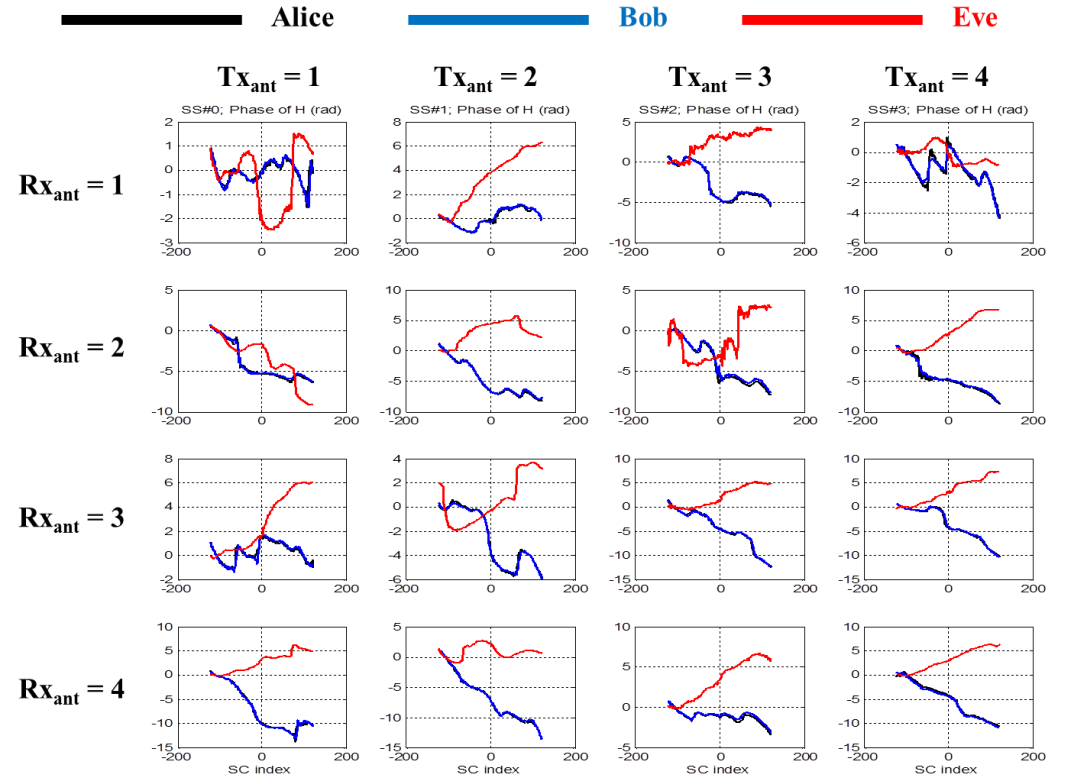


Figure 18: Phase of channel measurements for Alice, Bob and Eve

9.3. Measured CSI

Figures 17 and 18 plot the amplitude and phase of CSI computed by Alice, Bob and Eve. These figures show that Alice and Bob's channel measurements are quite similar (channel reciprocity) while they differ significantly from Eve's measurements (channel spatial diversity).

9.4. Key extraction from bi-directional CSI

After channel measurements, a Matlab script runs the SKG scheme on three consecutive channel sounding exchanges between Alice and Bob. Eve also captures the signal sent by Alice in order to compute her keys.

The SKG protocol at Alice's side can be described as follows.

- Pre-processing : selection of low-decorrelated CSI frames
- Quantization of CSI to get secret keys of 127 bits length
- Computation of secure sketches used by Bob for information reconciliation using BCH (127,15,27)
- Privacy amplification of the secret keys
- Key concatenation (final 256-bits)
- Test of the key randomness after quantization and amplification with the Intel Heath Check [13]
- Selection of amplified version of successful 256-bits secret keys both after quantization and amplification. Note that all keys should pass the test after privacy amplification since a hash function is used during this step.

Alice also sends over the public channel a message containing indexes of the selected CSI frames and quantization map, secure sketches, hashing parameters and indexes of successful 256-bit secret keys.

Although this message helps Bob's to compute same secret keys than Alice, secure sketches sent for reconciliation might leak some information to Eve as it allows her to correct errors she made on Alice's keys. This leaked information is mitigated by reducing the length of extracted keys during the privacy amplification step.

The SKG protocol at Bob's side can be summarized as follows.

- Pre-processing : selection of CSI frames according to the indexes sent by Alice
- Quantization of CSI using the quantization map indexes sent by Alice but his quantization maps are computed using his own channel measurements.
- Information reconciliation step using secure sketches sent by Alice and using BCH (127,15, 27).
- Privacy amplification of the keys using the hashing parameters sent by Alice.
- Key concatenation to 256-bits.
- Selection of successful 256-bit secret keys according to the indexes sent by Alice.

In our simulation, Eve performs exactly the same SKG steps as Bob.

9.5. Results when no channel de-correlation is performed

Figure 19 shows the keys extracted after quantization by Alice from channel measurements when no pre-processing step is performed.
78 keys of length 127-bits were generated but none of them passed the NIST runs test. 38 keys of length 256-bits were obtained by concatenating previous keys and none of them passed the Intel health Check.

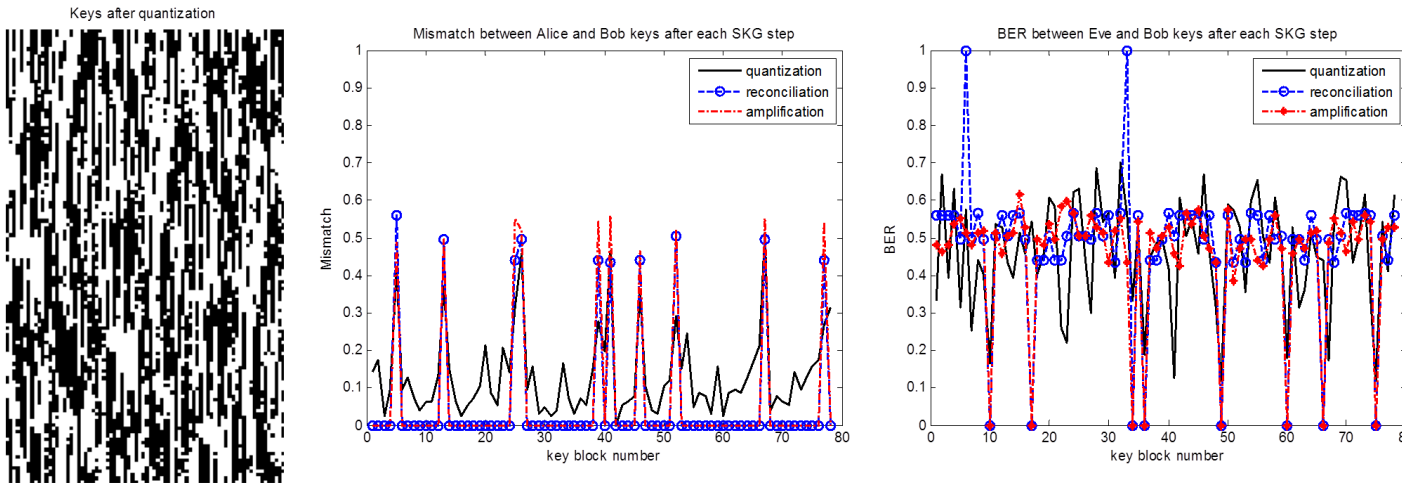


Figure 19: SKG results with no pre-processing step

Figure 19 also shows the mismatch between Alice and Bob, and the BER between Bob and Eve's keys at the end of the SKG processing when amplitude and phase of CSI are quantized using the CQA algorithm with 4 regions, information reconciliation and amplification being achieved as described in § 6 and 7.

According to the results, Bob often computes different keys than Alice while Eve manages to recover some of the secret keys: SKG performances are poor in this case

9.6. Note on the evaluation of the randomness of generated keys

We recall that NIST tests are well suited for off-line evaluation of the randomness of generated keys. However since several NIST tests are required to guarantee the randomness of a sequence of bits, the NIST tests cannot be used for online testing. Hence, we should not perform several testing tests during the SKG process in order to reduce the latency of the whole processing. We need one test that can allow the selection of generated keys with good randomness properties.

The Intel health test is very appropriate for online testing. Indeed it is composed of only one test that manages to detect non-random sequence of bits. In addition, the Intel health check is more stringent than both NIST frequency mono-bit test and runs test. Finally, the Intel health check has been evaluated and we can be confident on its performance [13].

9.7. Results when channel de-correlation is performed

Figure 20 shows the keys extracted after quantization by Alice from channel measurements when the pre-processing described in § 4 is performed with thresholds values $T_t = 1$ (no selection in time domain in this particular test case, because only 3 time instances were available in the records) and $T_f = 0.4$. Here, 5 keys of length 127-bits were generated and 4 of them passed the NIST runs test. 2 keys of length 256-bits were obtained by concatenating previous keys and both of them passed the Intel health Check.

After privacy amplification, all keys passed both NIST runs test and Intel Health Check.

Figure 20 also shows the mismatch between Alice and Bob, and the BER between Bob and Eve's keys.

As previously, amplitude and phase of CSI are quantized using the CQA algorithm with 4 regions. Information reconciliation is achieved using the BCH (127, 15, 27) code.

Here, Bob successfully computes the same keys than Alice while Eve's BER is always close to 0.5. Thus, Eve has no information on the secret keys computed by Alice and Bob. Finally the SKG perfectly works.

These results show that although the channel de-correlation pre-processing step reduces the number of generated keys, it not only improves the agreement between Alice and Bob, but also reduces the number of vulnerable key bits. By selecting only frames with low cross-correlation, the pre-processing step increases the available entropy and decreases the mutual information between Alice and Eve's channel measurements, leading finally to more secure random keys.

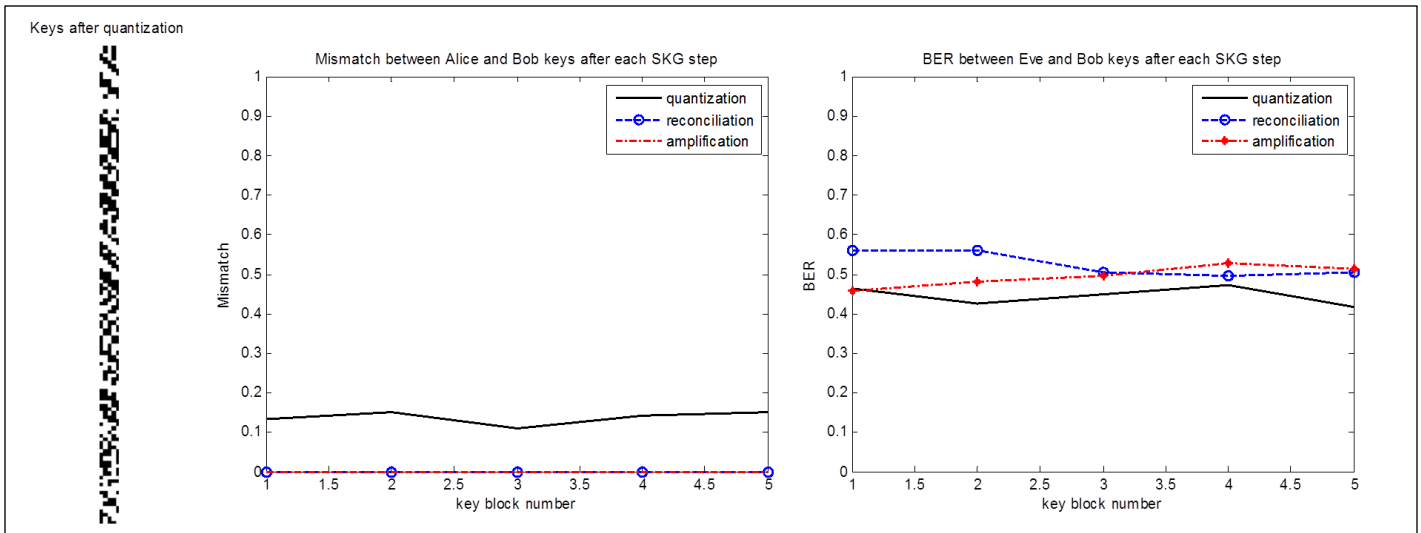


Figure 20: SKG results with pre-processing step

10. CONCLUSION

After recalling the basic schemes and principle of Secret Key Generation, and after describing particular implementation case to WiFi and LTE carrier, this paper outlined practical results performed in various radio-environments.

In dispersive radio-environments (with some scatterers and some mobility), a significant number of keys (of hundreds of bits each) can be extracted in a very short time under Wifi carriers and under LTE carriers. At the output of the processing, these keys have basically high entropy, low cross correlation and they are quite robust to correlation attacks since the quantification step.

In stationary environments (with very few scatterers and no mobility, such as encountered in some indoor cases, in IoT applications, etc.) and when no channel coefficient de-correlation algorithm is applied, the channel entropy is reduced, the extracted keys may be highly correlated and this vulnerability can be exploited by Eve to recover Bob's key.

Still in stationary environments, the quantization processing takes a large benefit of our channel coefficient de-correlation algorithm: the key rate is quite decreased but the extracted keys present lower cross correlation, higher entropy, and better robustness to correlation attack.

In any case, the proposed simplified reconciliation step with classical FEC codes provides a significant resilience of the key agreement between Alice and Bob. Only the FEC capability has to be adapted to quantization error at receiving, which is linked to the Signal to Noise ratio (used as practical criteria).

In any case, the proposed simplified amplification step with classical 2-Universal hash functions provides significant resilience of the key randomness against Eve's attacks, with a limited reduction of the key lengths.

NIST statistical tests and Intel Health Check were used to assess the randomness of generated keys.

The agreement between Alice and Bob's keys was evaluated by computing the Bit Error Rate between keys of length 127-bits extracted from their respective channel measurements.

Similarly, the secrecy of generated keys was assessed by computing the Bit Error Rate between keys generated by Bob and Eve.

To the best of our knowledge, this is the first work on a full secret key generation scheme with experimental CSI results using real field WiFi and LTE signals. Our promising results are evidence that the studied Secret Key Generation scheme can provide significant secrecy capabilities to users of public Radio Access technologies and that it can be practically implemented in existing wireless communication systems with minor modifications of the software architecture of nodes and terminals.

11. REFERENCES

[1] ZEIT, "Wie Merkels Handy abgehört werden konnte," 18 12 2014. [Online]. Available: http://www.zeit.de/digital/datenschutz/2014-12/umts-verschluesselung-umgehen-angela-merkel-handy.

[2] Ccc-Tv, «SS7map : mapping vulnerability of the international mobile roaming infrastructure,» [En ligne]. Available: https://media.ccc.de/v/31c3_-_6531_-_en_-_saal_6_-_201412272300_-_ss7map_mapping_vulnerability_of_the_international_mobile_roaming_infrastructure_-_laurent_ghigonis_-_alexandre_de_oliveira#video.

[3] I. Surveillance, «Rayzone-piranha-lte-imsi-catcher,» [En ligne]. Available: https://insidersurveillance.com/rayzone-piranha-lte-imsi-catcher/.

[4] T. intercept, «The Great SIM Heist, Hows Spies kept the key of the Encrypton Castle,» [En ligne]. Available: https://firstlook.org/theintercept/2015/02/19/great-sim-heist/.

[5] J. Wallace and R. Sharma, "Automatic secret keys from reciprocal MIMO Wireless channels: measurement and analysis," IEEE Trans. on Info. Foren. and Sec., vol. 5, no. 3, pp. 381-392, Sep 2010.

[6] H. D. X. He, «Is link signature dependable for Wireless Security?,» proceeding IEEE INFOCOM, pp. 200-204, 2013.

[7] U. Maurer et S. Wolf, «Secret-key agreement over unauthenticated public channels. II. Privacy amplification,» Information Theory, IEEE Transactions on, vol. 49, n° %14, pp. 839-851, 2003.

[8] Y. Dodis, L. Reyzin et A. Smith, «Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,» Advances in cryptology-Eurocrypt, pp. 523-540, 2004.

[9] C. Bennett, G. Brassard, C. Crepeau and U. Maurer, "Generalized privacy amplification," Information Theory, IEEE Transactions on , vol. 41, no. 6, pp. 1915-1923, 1995.

[10] PHYLAWS, «www.Phylaws-ict.org,» [En ligne].

[11] NIST (National Institute of standards and technology), «A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications,» Special Publication 800-22 Rev. 1a 2010.

[12] NIST (National Institute of Standards and Technology), «Recommendation for the Entropy Sources Used for Random Bit Generation,», (Second Draft) Special Publication 800-90B 2016.

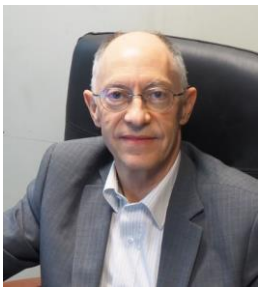
[13] M. Hamburg, P. Kocher and M. E. Marson, "Analysis of Intel's Ivy Bridge Digital Random Number Generator," Technical Report Cryptographic Research INC., March 2012.

9. BIODATAS

Christiane L. Kameni Ngassa received the "Diplôme d'ingénieur" (M.S. degree in electrical engineering) from Ecole Supérieure d'Electricité (Supelec), France, in 2011, the M.S. degree in electrical and computer engineering from Georgia Institute of Technology, USA, in 2011. From 2011 to 2014, she was a Ph.D. student at CEA - LETI, France, and received, in 2014, the Ph.D. degree in Information and Communication Sciences and Technologies from Cergy-Pontoise University, France. Since 2014, she has been a Research Engineer at Thales Communications and Security, France. Her research interests include error-correction coding, physical-layer security, and cryptography.
Dr. Kameni Ngassa has served as a Thales delegate for 3GPP and ITU standardization bodies. She has also been a European Commission expert for ICT topics since 2016.

Renaud Molière is graduated from Supelec (Ecole Supérieure d'Electricité, France) in 2014 and received an MSc in Space Science and Engineering from UCL (University College London, United Kingdom) in 2013. For his studies, he was rewarded by the prizes of best overall achievement and best individual project. He joined Thales in 2014 and works on signal processing for radio communication systems both for military and civilian applications. He was involved in the Phylaws FP7 project and published several patents and publications about physical layer security.
In parallel, he is the Thales delegate for the working group SA3 in 3GPP.

Alain Sibille graduated from Ecole Polytechnique (1977) and from Telecom-ParisTech (1979) in France and obtained the PhD/habilitation degree from University Paris 7 in 1985. He first conducted basic research in physics of semiconductor quantum and ultra high frequency devices, first within France-Telecom R&D then as a part time independent consultant. In 1992 he moved to ENSTA-ParisTech, where he led the Electronics and Computer Engineering department until 2010. He is currently professor and Director of Doctoral Education at Telecom ParisTech. His scientific interests lie in wireless communications systems, especially in the statistical modeling of antennas, radio channel and their combination. He has been actively participating in a number of FP4 to FP7 projects and is acting as EC expert in this area. Prof. Alain Sibille also chaired the European Wireless Conference 2007 in Paris, co-chaired the TPC of IEEE-PIMRC 2008 and chaired EuCAP 2017 in Paris. He has been a National delegate in a series of European COST Actions (COST 273, COST 2100, COST IC1004). He is also Secretary General of URSI-France.



François Delaveau received the M.Sc degrees from Ecole Nationale Supérieure de Techniques Avancées (ENSTA), Paris, France, in 1987, and from Mathematics University (maîtrise Paris VII - 1988; agrégation - 1990). His various activities (development, marketing support, project manager and head of laboratory...) covered several domains: radio communications, RADARS, SONARS, infra-red and Acoustic sensors and systems. Since 1997 within Thales Communications, he led researches and developments of new instruments for spectrum monitoring; of smart antennas for anti-jammed modems, direction finders and COMINT sensors (terrestrial, maritime, airborne and satellite); and of passive radars (Thales award 2007). More recently, he focused on advanced security schemes for Radio Access Technologies of wireless networks (coordinator of the EC-FP7 PHYLAWS project).
As a Thales' expert for radio communications, signal processing and electronic warfare, François Delaveau is author or co-author of numerous papers, tutorials and patents, and of several ITU-R recommendations.

Nir Shapira received a B.Sc. in Electrical Engineering and Physics (summa cum laude) from the Technion, Israel's Institute of Technology, and an MBA from the Tel-Aviv University. He is the CTO of Celeno Communications and has brought to Celeno over two decades of experience in the fields of communication theory, signal processing and wireless communications. He joined Celeno from the day of its inception and was responsible for forming the technological infrastructure and intellectual property of the company. He also represents Celeno in IEEE 802.11 standardization activities. Prior to joining Celeno, He held various managerial roles in Conexant, developing state-of-the-art wireless, DSL and voice band communication technologies. He served in an elite R&D unit of the IDF.