ROBUST-6G

Smart, Automated, and Reliable Security Service Platform for 6G

Deliverable D5.3
Release of Physical Layer Security Challenges

Co-funded by the European Union

6GSNS

ROBUST-6G project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101139068.

Date of delivery:	30/04/2026	Version:	1.0
Project reference:	101139068	Call:	HORIZON-JU-SNS-2023
Start date of project:	01/01/2024	Duration:	30 months

Document properties:

Document Number:	D5.3
Document Title:	Release of Physical Layer Security Challenges
Editor(s):	Laura Luzzi
Authors:	Cem Ayyildiz, Sara Berri, Luan Chen, Arsenia Chorti, Mamady Delamou, Ramin Fouladi, Eunjeong Jeong, Laura Luzzi, Angelo Passah, Nikolaos Pappas, Mattia Piana, Mehdi Sattari, Linda Senigagliesi, Tommy Svensson, Azadeh Tabeshnezhad, Stefano Tomasin, Solomon Yese, Emre Yildiz
Contractual Date of Delivery:	30/4/2026
Dissemination level:	PU
Status:	Final
Version:	1.0
File Name:	ROBUST-6G D5.3_v1.0

Abstract

This deliverable presents nine public challenges on Physical Layer Security, which are meant to validate the security solutions proposed as part of Work Package 5.

The first chapter illustrates how the challenges fit into the proposed architecture and contribute to the validation of key performance indicators.

The main part of the deliverable describes the challenges, focusing on three physical layer security techniques:

- Physical Layer Authentication, including the use of the Angle of Arrival as an unforgeable feature in Massive Multiple Input Multiple Output systems, and the impact of Reconfigurable Intelligent Surfaces

- Secret Key Generation, including channel feature extraction from high-dimensional noisy data, feasibility in indoor environments when the passive attacker is close to a legitimate node, and robustness against attacks in the presence of reconfigurable surfaces

- Radio Frequency Fingerprinting for device identification, in the presence of device replacements, temporal changes and hardware impairments

Finally, the appendices to the deliverable describe the datasets for the challenges and summarize the contributions of Work Package 5 to the ROBUST-6G objectives.

Keywords

Physical Layer Security, Physical Layer Authentication, Secret Key Generation, Radio Frequency Fingerprinting, Angle of Arrival, Reconfigurable Intelligent Surfaces, Massive MIMO

Executive Summary

In this deliverable, we propose nine public challenges on Physical Layer Security (PLS), meant to validate the security solutions proposed as part of Work Package 5, to increase confidence for experts, industry partners, and end users. The challenges are available on the ROBUST-6G website at the following link: `https://robust-6g.eu/physical_layer_security_challenges/`.
The first chapter of the deliverable introduces the practice of security challenges, and illustrates how the challenges fit into the proposed architecture and contribute to the validation of key performance indicators for ROBUST-6G.
The following chapters introduce each challenge in detail, spanning different PLS techniques:

- Physical Layer Authentication (PLA) methods are addressed in Chapters 3 and 9; the first evaluates the use of Angle of Arrival (AoA) as an unforgeable feature in massive Multiple-Input Multiple-Output (mMIMO) systems, while the second aims at quantifying the impact of Reconfigurable Intelligent Surfaces (RISs) on authentication.

- Chapters 2, 4 and 8 focus on Secret-Key-Generation (SKG) at the physical layer. More precisely, Chapter 2 focuses on the problem of extracting channel features from high-dimensional and noisy Channel State Information (CSI); Chapter 4 aims to demonstrate the feasibility of SKG in indoor environments when the eavesdropper is located extremely close to a legitimate node; Chapter 8 evaluates the robustness of SKG against passive attacks in the presence of a RIS.

- Finally, Chapters 5, 6, 7, and 10 consider different challenges for radio frequency fingerprinting for device identification, such as the impact of device replacements, temporal changes, and the effect of hardware impairments (carrier frequency offsets and symbol timing offsets).

As part of the challenge release, we have created or collected new datasets that were not presented in the D5.1 library of datasets; the dataset description is provided in Appendix A.
Finally, Appendix B summarizes the contributions of the activities of Work Package 5 to the objectives, measurable results, and quantifiable targets stated in the ROBUST-6G proposal.

Table of Contents

Acronyms **5**

1 Deliverable Overview and Contribution to the Architecture **8**

2 Challenge 1: Contrastive Representation Learning for CSI-Based Secret Key Generation **10**

3 Challenge 2: Angle-of-Arrival-Based Physical Layer Authentication in Digital Massive MIMO Systems **13**

4 Challenge 3: Secret Key Generation in Massive MIMO OFDM Under One-Wavelength Eavesdropping **19**

5 Challenge 4: Receiver-Invariant Device Identification Under Single Receiver Replacement **24**

6 Challenge 5: Robust Device Identification Under Sequential Receiver Replacement **27**

7 Challenge 6: Device Identification Under Temporal Drift in RF Fingerprinting **30**

8 Challenge 7: SKG on BRISC Dataset **33**

9 Challenge 8: PLA Authentication With RIS **36**

10 Challenge 9: Device Classification Under Hardware Impairments **38**

A Additional datasets **41**
A.1 CDL dataset 41
A.2 RF Fingerprinting Migration Dataset 41
A.3 RFFI-Temporal Dataset 42
A.4 BRISC Dataset 42
A.5 RF Fingerprint Dataset for Device Classification under Hardware Impairments 44

B Validation of the Objectives of WP5 **46**
B.1 Measurable results 46
B.2 Quantifiable targets 48
B.3 WP5 Objectives 49

Acronyms

ADE Automatic Device Enrolment.

AES Advanced Encryption Standard.

AKA Authentication and Key Agreement.

AoA Angle of Arrival.

AUC Area-Under-the-Curve.

BPSK Binary Phase Shift Keying.

BRISC Broadband Reconfigurable Intelligent Surface Channel.

CC Channel Charting.

CDL Clustered Delay Line.

CFO Carrier Frequency Offset.

CNN Convolutional Neural Network.

CPS Cyber–Physical Systems.

CR Challenge-Response.

CRA Channel-Response Authentication.

CRB Cramer-Rao bound.

CRC Cyclic Redundancy Check.

CSI Channel State Information.

DET Detection Error Tradeoff.

dMIMO distributed MIMO.

DP Differential Privacy.

FA False Alarm.

GBM Gradient Boosting Machine.

IoT Internet of Things.

IQ In-phase and Quadrature.

ISAC Integrated Sensing And Communication.

KDR Key Disagreement Rate.

KGR Key Generation Rate.

LDPC Low-Density-Parity-Check.

LoS Line of Sight.

MD Missed Detection.

ML Machine Learning.

mMIMO massive Multiple-Input Multiple-Output.

NLoS Non Line-of-Sight.

NMSE Normalized Mean Squared Error.

NOMA Non-Orthogonal Multiple Access.

OFDM Orthogonal Frequency Division Multiplexing.

OTP One-Time Pad.

PLA Physical Layer Authentication.

PLS Physical Layer Security.

PLS-CL Physical Layer Security Closed Loop.

PSD Power Spectral Density.

RF Radio Frequency.

RFFI Radio Frequency Fingerprinting Identification.

RHI Residual Hardware Impairments.

RIS Reconfigurable Intelligent Surface.

RSA Rivest-Shamir-Adleman.

RTC Real-Time Clock.

SDR Software-Defined Radio.

SIC Successive Interference Cancelation.

SINR Signal-to-Interference-plus-Noise Ratio.

SKG Secret-Key-Generation.

SNR Signal-to-Noise Ratio.

STO Symbol Timing Offset.

TDD Time-Division Duplex.

ToF Time of Flight.

ULA Uniform Linear Array.

Chapter 1

Deliverable Overview and Contribution to the Architecture

In cryptography, it is common practice to issue public challenges as a way to involve the wider community in the testing of cryptographic schemes. Notable examples are the Rivest-Shamir-Adleman (RSA) Secret-Key Challenge and the Challenges for Advanced Encryption Standard (AES) [1].
The benefits of challenges are manifold: they attract the attention of the general public to new techniques, assess the strength and robustness of new schemes against different types of attacks, foster international collaborative efforts to solve difficult problems, increase the confidence of industrial partners and end users, and act as a catalyst to help standardization bodies in the transition to higher-security protocols.

In this deliverable, we aim to bring this practice to the physical layer security community. We provide adversarial observations for various security scenarios, and ask researchers to break the security of the proposed schemes, according to specific, measurable target outcomes. By including measured experimental data in some of these challenges, we aim to determine whether practical imperfections due to wireless propagation or hardware impairments can be exploited to break theoretical guarantees.

This is among the first examples of challenges for PLS and it is an important contribution of the ROBUST-6G project. Moreover, the challenges are aligned with the activities carried out in the project and are aimed to show and strengthen the security of the ROBUST-6G solutions.

Overview and contribution to the architecture

The challenges will contribute to validating the architecture components proposed in D6.1 [2], including the final calibration of parameters. The connections with architecture components and KPIs are illustrated in Table 1.1.

Chapter 2 focuses on learning compact and robust channel representations from CSI for secret key generation by maximizing mutual information between Alice's and Bob's observations using contrastive methods like InfoNCE. Participants design neural models and are evaluated based on Key Generation Rate (KGR), reliability (low Key Disagreement Rate (KDR)), and reconstruction accuracy. This challenge will contribute to KPI8 (schemes with less than 5 ms latency).

Chapter 3 proposes a security challenge to validate the unforgeability property of AoA as a feature for PLA in mMIMO systems, in connection with the AoA-based PLA architecture component CENS04 described in Section 4.8.4 of D6.1 [2], and with KPI7 (PLA accuracy). The challenge dataset uses real data collected in an indoor ultra-dense scenario, including Line of Sight (LoS) and Non Line-of-Sight (NLoS) components, spatial correlation, frequency selectivity, and multipath fading.
Chapter 4 focuses on SKG, evaluating its robustness in realistic propagation conditions when the eavesdropper is located extremely close to a legitimate node (up to one wavelength). This challenge will contribute to validating the Fast SKG component CENS05 (Section 4.8.5 of D6.1), in connection with KPI8 (reconciliation success).

Table 1.1: Detailed Contribution to the Architecture

Chapter Number	Architecture Component	KPIs
2	CCHA02	KPI8
3	CENS04	KPI7
4	CENS05	KPI8
5	CGHM01	KPI6, KPI7
6	CGHM01	KPI6, KPI7
7	CGHM02	KPI6, KPI7
8	CUPD04	KPI8
9	CUPD04	KPI8
10	CEBY04	KPI6, KPI7

Chapters 5, 6, and 7 focus on enhancing Radio Frequency Fingerprinting Identification (RFFI) for robust PLA in Internet of Things (IoT) systems. RFFI serves dual purposes: it enables high-accuracy device authentication (KPI7) and provides a powerful defense mechanism against spoofing and Sybil attacks (KPI6). Chapters 5 and 6 address receiver variability through single and sequential receiver replacement scenarios, while Chapter 7 tackles temporal drift using the RFFI-Temporal dataset. Together, they aim to improve the accuracy, long-term robustness, and resilience of RFFI-based authentication against impersonation and Sybil attacks, thereby supporting both KPI6 and KPI7.

Chapters 8 and 9 employ the Broadband Reconfigurable Intelligent Surface Channel (BRISC) dataset to validate the PLA and SKG techniques. By leveraging RIS, these approaches ensure rapid detection capabilities—meeting KPI8 targets with latencies below 5 ms for static nodes—while also achieving strong authentication accuracy, enabled by the high position discrimination provided by the RIS.

Chapter 10 uses a dataset that enables RF fingerprinting-based device classification under hardware impairments using raw signals and statistical features. It directly supports KPI6 (Sybil attack detection) and can be extended to jamming/interference detection, while indirectly contributing to KPI7 (resilience).

Chapter 2

Challenge 1: Contrastive Representation Learning for CSI-Based Secret Key Generation

Abstract

SKG based on channel reciprocity is a promising technique for enabling physical-layer security in Time-Division Duplex (TDD) wireless systems. However, practical SKG faces significant challenges in extracting reliable and high-entropy keys from high-dimensional, noisy, and time-varying CSI, particularly in frequency-selective and multi-antenna channels.
In this challenge, we focus on the feature extraction stage of the SKG pipeline and investigate mutual information–driven representation learning for CSI-based key generation. Specifically, we consider neural encoders that learn compact latent representations by maximizing the shared information between Alice's and Bob's reciprocal CSI observations.
To address the intractability of mutual information in high-dimensional settings, we employ neural estimators based on contrastive learning objectives, particularly the InfoNCE criterion [3]. The goal is to learn representations that are robust to channel impairments and highly informative for secret key generation, enabling high key generation rate while maintaining low key disagreement.

Dataset

The dataset is available at this link: `https://doi.org/10.5281/zenodo.19493713`
The proposed framework is evaluated using a standardized 3GPP Clustered Delay Line (CDL) channel model based on TR 38.901 at mmWave frequency (28 GHz). The dataset captures realistic multipath propagation, spatial correlation, and temporal channel evolution under mobility.
The channel configuration is as follows: the base station is equipped with $N_t = 16$ antennas arranged in a uniform linear array, and the user equipment has $N_r = 1$ antenna. Each CSI sample consists of $T = 100$ OFDM symbols (time steps) and $N_{\text{sub}} = 16$ subcarriers. The CSI is provided in the frequency domain across all subcarriers, and time evolution is modeled through temporally correlated fading. The system operates in TDD mode, enabling reciprocal channel observations at Alice and Bob.
The CSI is provided as complex-valued tensors

$$\mathbf{H} \in \mathbb{C}^{T \times N_{\text{sub}} \times N_t},$$

which can be equivalently represented as real-valued tensors of size $T \times N_{\text{sub}} \times N_t \times 2$ corresponding to real and imaginary components.
The dataset includes multiple propagation scenarios (CDL-A to CDL-E), different mobility regimes (e.g., 3 km/h, 30 km/h, 120 km/h).

Challenge Description

The objective of this challenge is to develop contrastive, mutual information–driven representation learning methods for CSI-based SKG in TDD systems. Participants are expected to design neural models that exploit channel reciprocity by maximizing the shared information between Alice's and Bob's CSI using contrastive objectives such as InfoNCE.
The learned representations should be robust to channel noise, mobility, and hardware impairments, while preserving sufficient shared information to enable reliable secret key generation.
Contrastive learning methods based on the InfoNCE objective [3] provide a tractable way to approximate mutual information in high-dimensional settings. The use of standardized 3GPP CDL channel models [4] ensures realistic and reproducible evaluation conditions.

Input and Output

The input to the participant's model consists of CSI observations from Alice and Bob, denoted as $\mathbf{H}^{(A)}$ and $\mathbf{H}^{(B)}$, where each tensor has dimensions $\mathbf{H} \in \mathbb{C}^{T \times N_{\text{sub}} \times N_t}$ (or equivalently $T \times N_{\text{sub}} \times N_t \times 2$ in real-valued form).
The observed CSI at Alice and Bob is modeled as

$$\mathbf{H}^{(A)} = \mathbf{H} + \mathbf{N}^{(A)}, \quad \mathbf{H}^{(B)} = \mathbf{H} + \mathbf{N}^{(B)}, \tag{2.1}$$

where $\mathbf{H}$ denotes the underlying reciprocal channel, and $\mathbf{N}^{(A)}$ and $\mathbf{N}^{(B)}$ are independent additive white Gaussian noise tensors.
This models the practical scenario where Alice and Bob observe the same underlying channel but with independent noise realizations due to receiver noise and hardware impairments.
The output of the model must be binary key sequences for both Alice and Bob,

$$\mathbf{k}^{(A)}, \mathbf{k}^{(B)} \in \{0, 1\}^L,$$

obtained from their respective CSI observations.
Participants must submit a trained neural model together with inference code that maps CSI samples to quantized bit sequences. All submissions must be compatible with the provided evaluation pipeline.

Evaluation Metric

Submissions will be evaluated using the following metrics.
The KGR is defined as

$$\text{KGR} = \frac{L}{T}, \tag{2.2}$$

where L is the total number of generated key bits and T is the number of channel uses.
The key disagreement rate (KDR) is defined as

$$\text{KDR} = \frac{1}{L} \sum_{i=1}^{L} \mathbf{1}\left(k_i^{(A)} \neq k_i^{(B)}\right), \tag{2.3}$$

where $k_i^{(A)}$ and $k_i^{(B)}$ denote the i-th key bit generated at Alice and Bob, respectively, and $\mathbf{1}(\cdot)$ is the indicator function.

The Normalized Mean Squared Error (NMSE) is defined as

$$\text{NMSE} = \frac{\mathbb{E}\left[\|\mathbf{H} - \hat{\mathbf{H}}\|^2\right]}{\mathbb{E}\left[\|\mathbf{H}\|^2\right]}, \tag{2.4}$$

where $\hat{\mathbf{H}}$ denotes the reconstructed channel obtained from the learned representation.
The final ranking is based on performance at a reference operating point corresponding to SNR = 15 dB with 5-bit quantization. Submissions must satisfy KDR $< 5 \times 10^{-2}$ and NMSE < -25 dB. Among all valid submissions, methods are ranked according to their achieved KGR, where higher values indicate better performance.
In case of similar KGR, preference will be given to methods with lower computational complexity (inference time and model size) and better generalization to unseen environments.

Chapter 3

Challenge 2: Angle-of-Arrival-Based Physical Layer Authentication in Digital Massive MIMO Systems

Abstract

Angle-of-arrival (AoA) has been identified as a geometry-constrained feature for physical layer authentication (PLA) in digital antenna array systems. Recent theoretical results demonstrate that AoA impersonation is only feasible when an adversary shares the same spatial direction as the legitimate transmitter, even under multi-antenna precoding [5]. This chapter introduces a reproducible security challenge designed to evaluate this unforgeability property in a practical massive MIMO orthogonal frequency division multiplexing (OFDM) environment. Using a 64-antenna uniform linear array (ULA) and 100-subcarrier channel state information (CSI) measurements from the IEEE DataPort Ultra-Dense Indoor MaMIMO dataset [6], the challenge invites the research community to attempt spatial impersonation attacks under controlled geometric constraints. The objective is to empirically validate AoA robustness in realistic propagation environments and establish a benchmark for spatial-feature-based authentication in 6G systems.

Physical layer authentication (PLA) has emerged as a complementary mechanism to upper-layer cryptographic authentication in beyond-5G and 6G networks. Most PLA schemes rely on amplitude-dependent channel features such as channel impulse response (CIR), channel state information (CSI), or received signal strength (RSS). These features, however, can be manipulated through transmit-side precoding, particularly when adversaries are equipped with multiple antennas.

Angle-of-arrival (AoA) constitutes a fundamentally different class of feature. When estimated using a digital uniform linear array (ULA), AoA is intrinsically tied to transmitter geometry. Under far-field narrowband assumptions, the received baseband signal at antenna index m is

$$x_m = s\, e^{-j\frac{2\pi}{\lambda} m d \sin(\theta)} + n_m, \tag{3.1}$$

where x_m denotes the received complex baseband signal at the m-th antenna element, s is the transmitted complex baseband signal, m is the antenna index, d is the inter-element spacing, λ is the signal wavelength, and θ denotes the angle of arrival (AoA) measured relative to the array broadside. The noise term n_m represents additive complex Gaussian noise at antenna m. The steering vector structure constrains spatial signatures observable at the receiver. Recent analysis in [5] proves that even with arbitrary multi-antenna precoding, an adversary cannot impersonate the AoA of a legitimate transmitter unless they are aligned in the same spatial direction. This result motivates a community-level security challenge to test the robustness of AoA-based authentication in practical massive MIMO OFDM environments.

The PLA protocol operates as follows: let $\mathbf{h}(f_k) \in \mathbb{C}^{M\times 1}$ denote the CSI vector at subcarrier f_k. AoA estimation is performed using subspace-based methods such as MUSIC across antennas. Authentication follows a two-phase protocol:

- *Enrollment Phase*: The legitimate transmitter's AoA is estimated as $\hat{\theta}_{\text{enr}}$ and stored.
- *Verification Phase*: For a new transmission, AoA is estimated as $\hat{\theta}_{\text{ver}}$ and compared:

$$\xi = |\hat{\theta}_{\text{enr}} - \hat{\theta}_{\text{ver}}|. \tag{3.2}$$

A threshold τ determines authentication decisions and is typically determined based on the target false alarm probability.

Dataset

The original dataset is available at this link: `https://ieee-dataport.org/open-access/ultra-dense-indoor-mamimo-csi-dataset`. The security challenges proposed are based on the Ultra-Dense Indoor Massive MIMO (MaMIMO) Channel State Information (CSI) dataset publicly available through IEEE DataPort [6]. The dataset was collected in an indoor ultra-dense deployment scenario designed to emulate realistic beyond-5G and 6G environments characterized by high spatial reuse, strong multipath propagation, and fine-grained spatial resolution. The measurement campaign employs a 64-element uniform linear array (ULA) as the receiver, operating in a time-division duplex (TDD) configuration. The antenna elements are arranged with $0.87 \times$ half-wavelength inter-element spacing, ensuring spatial sampling consistent with classical array processing assumptions. The system operates using orthogonal frequency division multiplexing (OFDM) with 100 active subcarriers, enabling high-resolution frequency-domain channel estimation. The moving single antenna user transmits a pilot signal every 5 mm, and for each spatial position, the dataset provides complex baseband CSI measurements across all antenna elements and subcarriers. Specifically, each CSI snapshot can be represented as a matrix

$$\mathbf{H}(f_k) \in \mathbb{C}^{64\times 1}, \quad k = 1, \ldots, 100,$$

where each element captures the complex channel coefficient between a transmit antenna and a receive antenna at subcarrier f_k. The measurements reflect realistic propagation conditions with line-of-sight (LoS), spatial correlation, frequency selectivity, and multipath fading. The indoor environment in which the dataset was recorded exhibits rich scattering due to walls, furniture, and reflective surfaces. As a result, the dataset captures realistic angular spreads and spatial covariance structures that are critical for evaluating both angle-of-arrival estimation robustness and spatial decorrelation properties relevant to secret key generation. The dataset structure enables spatial analysis across antenna elements, frequency-domain analysis across subcarriers, and joint spatial-frequency processing. This makes it particularly suitable for evaluating geometry-based authentication mechanisms as well as high-dimensional secret key extraction schemes. Table 3.1 summarizes the key parameters of the MaMIMO CSI original dataset.
In addition to raw CSI values, the dataset provides the ground-truth coordinates of each spatial position, and the midpoint of the ULA is the reference. This makes it possible to compute the ground-truth AoA and user positions, enabling controlled evaluation of angular separation constraints and proximity-based eavesdropping scenarios.
The combination of large-scale antenna arrays, frequency-selective channel observations, and spatially dense measurement positions makes this dataset an appropriate benchmark for evaluating geometry-constrained security mechanisms in massive MIMO systems.

Challenge Description

The software needed for data processing is available at this link: `https://6jwvqfw2j7lzakeepstj74.streamlit.app/`. In this challenge, we consider a subset of 26 datapoints extracted from a chunk of the

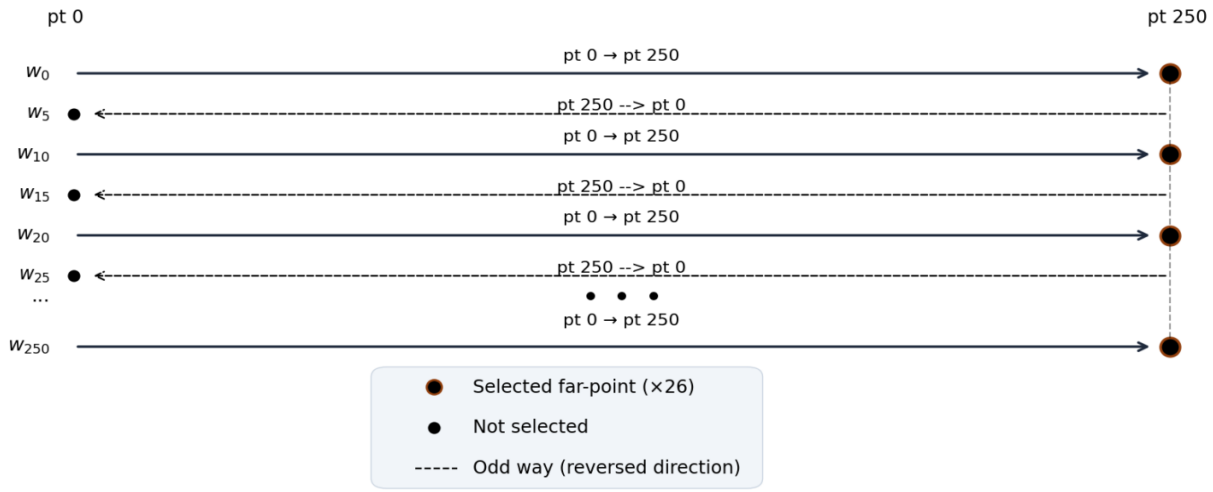


Figure 3.1: Subsampling of the original dataset. The user moves back and forth following a zigzag. Instead of using all the 251 ways, we keep only a single datapoint on every 10th way.

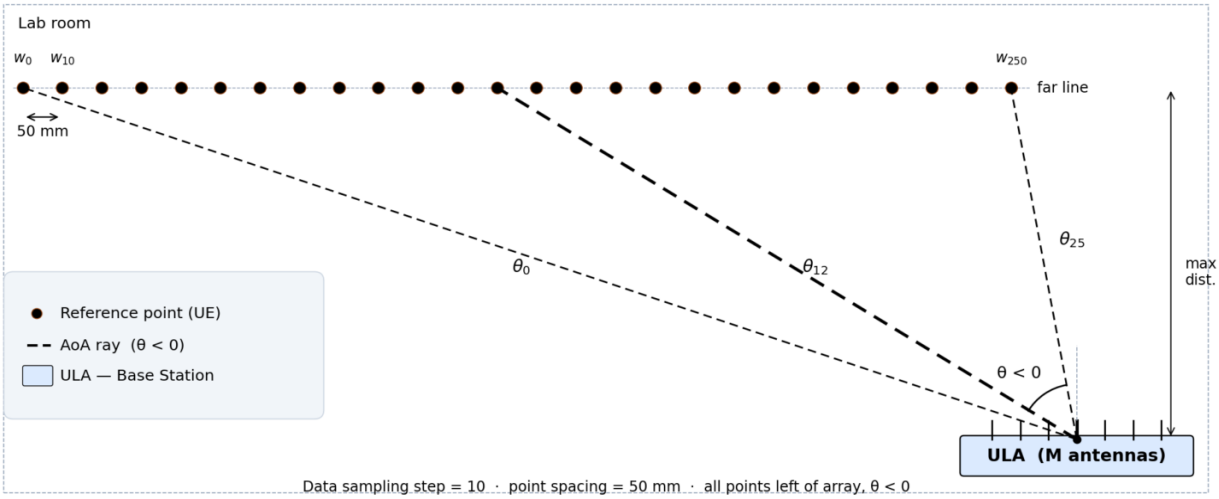


Figure 3.2: System model of the challenge. The AoA-based authentication is performed on 26 reference datapoints selected from the original dataset and spaced with 50 mm.

Table 3.1: Ultra-Dense Indoor MaMIMO CSI Dataset Parameters

Parameter	**Specification**
Deployment Scenario	Indoor ultra-dense environment
Array Configuration	Uniform Linear Array (ULA)
Number of Antenna Elements	64
Antenna Spacing	$0.87 \times \lambda/2$
Duplexing Mode	Time-Division Duplex (TDD)
Modulation Scheme	OFDM
Number of Subcarriers	100
CSI Representation	Complex baseband coefficients
Channel Dimension per Snapshot	64×100 (spatial $\times$ frequency)
Propagation Conditions	LoS and NLoS
Spatial User Positions	Multiple indoor locations
Data Availability	IEEE DataPort

original dataset that comprises 251 ways and 251 datapoints per way. The subsampling is performed by retaining a single datapoint per way, selecting every 10th way ($w_0, w_{10}, w_{20}, \ldots, w_{250}$), and always picking the datapoint located at the farthest distance from the antenna array as depicted in Fig. 3.1. This last criterion follows from the snake structure of the measurement grid: the farthest point alternates between index 250 for even-numbered ways and index 0 for odd-numbered ways. The resulting 26 points are physically separated by 50 mm from one another, all lying along the same line. This selection is motivated by two properties: first, the uniform spatial separation ensures that the 26 reference AoA values are well-spread and geometrically distinct, eliminating any angular ambiguity in the authentication process; second, operating at maximum distance from the array places all points in a consistent far-field regime, where AoA estimation via MUSIC is most accurate and reliable.
The system model for the challenge is illustrated in Fig. 3.2. The reference coordinate system is centered at the midpoint of the ULA. The 26 selected reference points are positioned along the far line. All the reference positions lie to the left of the array center, resulting in strictly negative AoA values. For each reference point, the AoA θ is defined as the angle between the incoming wavefront and the broadside direction of the array, and is estimated at the ULA using the MUSIC algorithm.
For each datapoint, the AoA is estimated as follows:

- We first compute the ground-truth (GT) AoA θ_{GT} of each datapoint using the given ground-truth coordinates.
- Since the original collected data are affected by hardware impairments, we calibrated the datapoints to mitigate these impairments.
- After calibration, AoA estimation is performed using the MUSIC algorithm with spatial smoothing applied over a subarray of 16 antennas. The smoothed spatial covariance matrix is constructed from the calibrated CSI, decomposed into signal and noise subspaces via eigen decomposition, and the AoA is extracted as the peak of the MUSIC pseudo-spectrum evaluated over a fine angular grid of

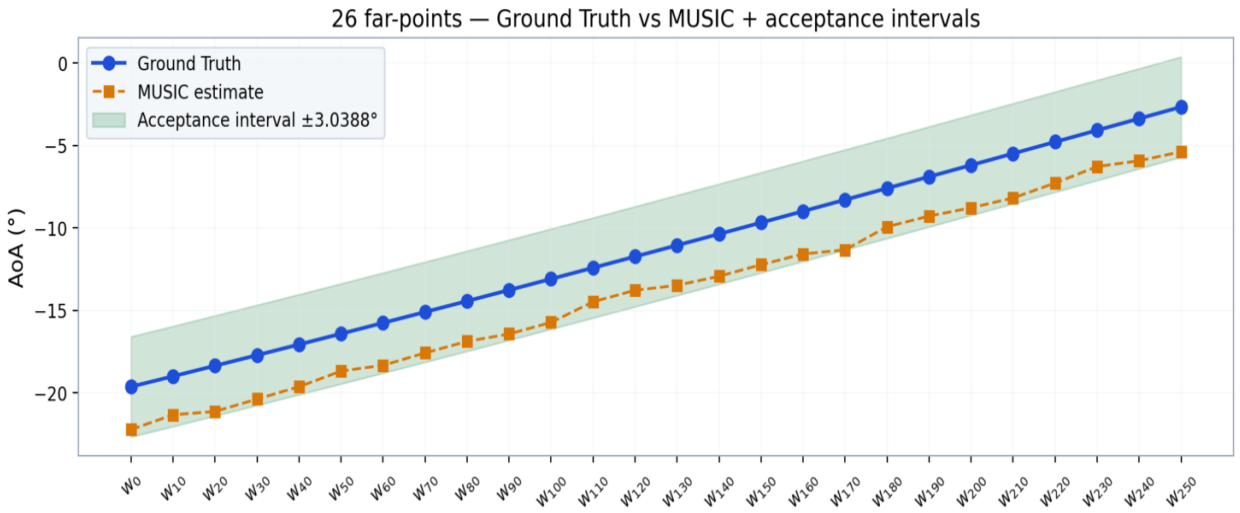


Figure 3.3: GT AoA vs. estimated AoA for the 26 datapoints and the derived acceptance interval.

3600 points spanning −90° to +90°. Using θ_{GT} as reference, we compute the estimation error for each datapoint. Finally, the maximum error is used as the acceptance interval to ensure that any AoA estimate stays within the interval. The maximum angle error determines the threshold τ.

Finally, the impersonation challenge can be stated as follows:

An adversary at angle $\hat{\theta}$ cannot impersonate a legitimate transmitter at angle θ, regardless of any complex precoding, as long as $\hat{\theta} \notin \mathcal{I} = [\theta_{GT} - \tau, \theta_{GT} + \tau]$.

Fig. 3.3 illustrates the ground-truth, the estimated AoAs (MUSIC) of each datapoint, along with the acceptance interval.

Input and Output

Select any datapoint. You know its ground-truth AoA and the acceptance interval $\mathcal{I}$. Choose your physical angle $\hat{\theta} \notin \mathcal{I}$ and any complex precoding scalar q (a complex scalar will be represented by its modulus $|q|$ and its phase $\angle q$). The MUSIC estimator will output your true angle. Can you make it output an angle in the acceptance interval $\mathcal{I}$? If so, you've passed the challenge.
Based on Proposition 1 of [5], in an ideal setting MUSIC always outputs $\hat{\theta}$ regardless of q, since scaling the channel by a complex scalar does not alter the spatial covariance structure nor the noise subspace.
Participants should submit a datapoint, a valid physical angle and a precoding scalar q, together with a description of their methodology, including well-commented code and an explanation of the proposed solution.

Evaluation Metric

The evaluation metric is the absolute angular error between the MUSIC-estimated AoA and the ground-truth AoA of the target point, defined as:

$$\xi = \left|\hat{\theta}_{\text{MUSIC}} - \theta_{\text{GT}}\right| \tag{3.3}$$

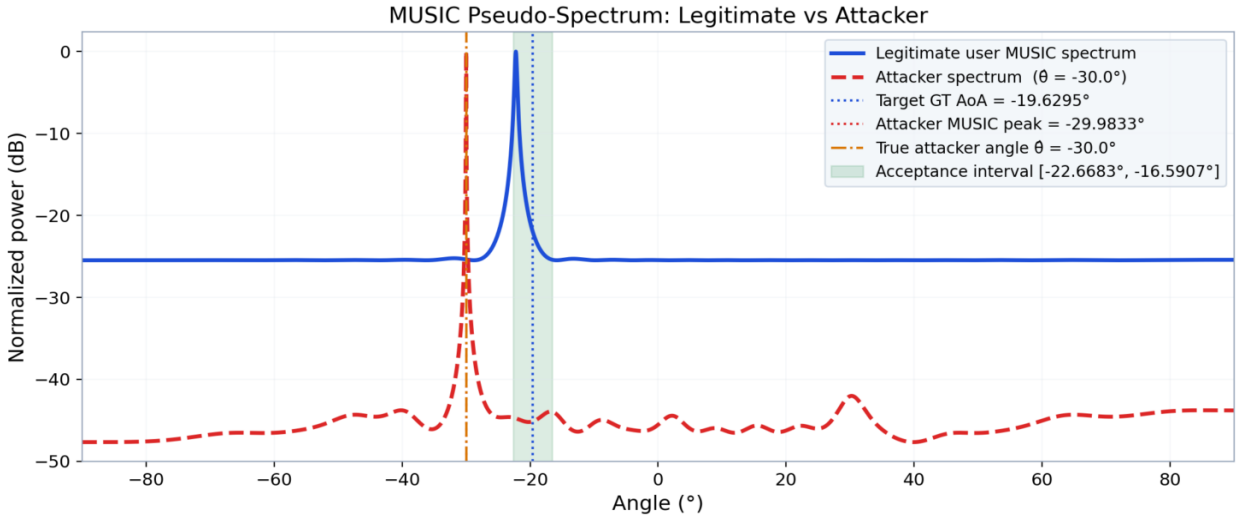


Figure 3.4: Challenge output for GT $= -19.6295°$, $\mathcal{I} = [-22.6683°, -16.5907°]$. The submitted precoder is such that $\hat{\theta} = -30°$ $|q| = 10$ and $\angle q = 40°$.

A submission is declared successful if $\xi < \tau$, where τ is the acceptance margin derived from the maximum MUSIC estimation error observed across the 26 legitimate reference points:

$$\tau = \max_{i \in \{1,\ldots,26\}} \left| \hat{\theta}_i^{\text{MUSIC}} - \theta_i^{\text{GT}} \right| \tag{3.4}$$

The challenge is considered broken if the challenger achieves $\xi < \tau$ while transmitting from an angle $\hat{\theta} \notin \mathcal{I}$, i.e., from outside the acceptance interval $\mathcal{I}$ of the target point.

Fig. 3.4 depicts a sample of the challenge performed on GT $= -19.6295°$, and $\mathcal{I} = [-22.6683°, -16.5907°]$. When $\hat{\theta} = -30°$, $|q| = 10$ and $\angle q = 40°$, as expected, the algorithm returns $\hat{\theta} = -30°$.

Chapter 4

Challenge 3: Secret Key Generation in Massive MIMO OFDM Under One-Wavelength Eavesdropping

Abstract

Secret key generation (SKG) exploits channel reciprocity and randomness to derive shared cryptographic keys from wireless fading coefficients. While classical models assume decorrelation beyond half-wavelength separation, practical indoor propagation may violate this assumption. This chapter introduces a security challenge evaluating SKG robustness in a massive MIMO orthogonal frequency division multiplexing (OFDM) system when a passive eavesdropper is located approximately one wavelength from a legitimate node. Using 64-antenna ULA and 100-subcarrier channel state information from the Ultra-Dense Indoor MaMIMO dataset [6], and building upon the full-chain SKG framework in [7], the challenge provides reconciliation artifacts and invites the research community to attempt key recovery or entropy reduction attacks.

SKG Protocol: We consider a time-division duplex (TDD) massive MIMO OFDM system where Alice and Bob aim to generate a shared secret key from reciprocal channel observations $\mathbf{H}_{AB}(f_k) \approx \mathbf{H}_{BA}(f_k) \in \mathbb{C}^{M\times 1}$, where M is the number of ULA antenna elements, N the number of active OFDM subcarriers, and f_k the k-th subcarrier frequency with $k = 1, \ldots, N$. A passive Eve at approximately one wavelength λ from Bob observes a correlated channel $\mathbf{H}_{AE}(f_k) \in \mathbb{C}^{M\times 1}$. The SKG protocol extracts shared randomness through quantization, reconciliation via public syndrome s_A, and privacy amplification. Let r_A, r_B, and r_E denote the bit sequences derived by Alice, Bob, and Eve, respectively, after quantization. Security is evaluated via the conditional min-entropy

$$H_\infty(r_A \mid r_E, s_A) = -\log_2\left(\max_{r_A} p(r_A \mid r_E, s_A)\right), \tag{4.1}$$

which captures the residual uncertainty in Alice's bit sequence given all information available to Eve. Unlike classical fading assumptions, indoor multipath may induce structured spatial correlation even at one-wavelength separation, making robust key extraction non-trivial.

Dataset

The original Ultra-Dense Indoor MaMIMO dataset [6] is available at this link: `https://dx.doi.org/10.21227/nr6k-8r78`. The dataset is collected using the KU Leuven ESAT-TELEMIC massive MIMO testbed, which consists of a base station equipped with 64 patch antennas and four user equipments (UEs). Measurements are performed in an indoor 3 $m \times 3$ m area using a time-division duplex (TDD) system with OFDM modulation. During data acquisition, the base station simultaneously receives orthogonal pilot

signals from the four UEs and performs channel estimation. Each channel state information (CSI) sample is represented as a complex matrix of size 64×100 corresponding to 64 antennas and 100 subcarriers. The system operates at a center frequency of 2.61 GHz with a 20 MHz bandwidth. The dataset is designed for high-precision indoor localization and sensing applications. It includes measurements under three antenna array configurations: a uniform linear array (ULA), a uniform rectangular array (URA), and distributed linear arrays (DIS). The UEs are moved along a controlled zigzag trajectory. The ground truth positions are collected with less than 1 mm error, and the movement step size is 5 mm. In total, 252004 CSI samples are collected across the four UEs. The dataset supports research in fine-grained positioning, multi-device localization, and applications such as smart environments and robotic positioning.

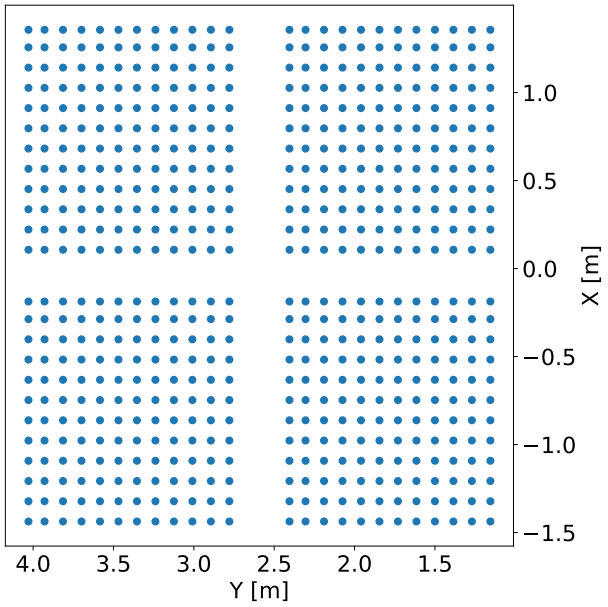


Figure 4.1: UE positions in the wavelength-scaled sampling grid.

Wavelength-scaled Spatial Sampling: We constructed a wavelength-based spatial sampling grid from the original CSI data. The grid is shown in Fig. 4.1. This approach enables the structured collection of CSI snapshots at precise spatial locations, allowing for coherent analysis of how channels evolve across distinct propagation paths.

The spatial grid is constructed using wavelength-scaled sampling. Given the wavelength $\lambda \approx 11.5\,cm$ and the UEs' moving stride $0.5\,cm$, we define a sampling resolution parameter that represents the number of measurement samples per wavelength. For each UE, the grid construction process identifies measurement indices at discrete wavelength intervals, creating a regular pattern of CSI observations spaced approximately one wavelength apart. The resulting grid (Fig. 4.1) spans multiple wavelengths, yielding a 12×12 grid structure for each of the four UEs. The obtained dataset includes 576 total measurement points for all the 4 UEs. All extracted grid indices and corresponding CSI measurements and position coordinates undergo validation procedures to ensure the grid dataset provides reliable spatial wavelength-scaled sampling.

To support both uplink and downlink channel analysis for SKG purposes, the grid indices identify paired measurement points representing channel reciprocity conditions. For each grid point designated as an uplink measurement, we designate a corresponding downlink measurement taken at a consecutive snapshot from the original dataset ($5\,mm$ away). This enables direct comparison of reciprocal channel conditions. This uplink-downlink pairing facilitates the study of channel reciprocity properties for SKG.

Challenge description

No additional software is required beyond standard Matlab libraries.
Despite the principle of channel reciprocity, practical channel measurements are corrupted by noise, hardware imperfections, and time-variant phenomena. This can lead to random variations and inconsistencies in reciprocal CSI observations. When quantizing these CSI estimations, the granularity of quantization levels and the magnitude of channel mismatch can result in different binary representations between Alice and Bob. These mismatches directly limit the achievable secret key rate if no corrective measures are taken. To overcome this fundamental challenge, distributed source coding techniques are adopted, specifically Slepian-Wolf coding [8]. This enables efficient reconciliation [8, 9] of the mismatched sequences through structured error correction, allowing Alice and Bob to agree on a common binary key despite their imperfect observations.
The key generation process includes three main steps: quantization, information reconciliation, and privacy amplification. To extract binary key material from the continuous-valued channel observations, uniform quantization is applied to both uplink and downlink CSI. We use a 2-bit uniform quantizer to quantize the CSI values.
Reconciliation: To reconcile the mismatched binary sequences between Alice and Bob, Slepian-Wolf decoding is implemented using Polar codes with cyclic redundancy check (CRC) as the error correction code. Alice then generates a syndrome using the Slepian-Wolf coding approach and transmits it over a public channel to Bob. This enables Bob to decode and recover Alice's key with high reliability despite the channel mismatch using Polar codes. These reconciled sequences form the input to privacy amplification.
Concerning the considered Polar codes for the reconciliation block, we use Gaussian approximation based polar code construction [10,11], which assumes that log-likelihood ratios (LLRs) remain Gaussian distributed through the polar transformation. The idea of Gaussian approximation is to evolve the densities and estimate the precise reliability of each channel. For AWGN with variance σ^2, the channel LLR is modeled as LLR $\sim \mathcal{N}(\mu, 2\mu)$ and $\mu_0 = \frac{2}{\sigma^2}$, where μ_0 is the initial LLR mean. Gaussian approximation propagates only the mean LLR μ through the polarization process, yielding μ_i for each synthesized bit-channel $W^{(i)}$. Let $\phi(\cdot)$ denote the Gaussian approximation check-node function [10], with an accurate empirical fit

$$\phi(\mu) \approx \begin{cases} \exp(-0.4527\,\mu^{0.86} + 0.0218), & 0 \le \mu \le 10, \\ \sqrt{\frac{\pi}{\mu}}\left(1 - \frac{10}{\mu}\right)\exp(-\mu/4), & \mu > 10. \end{cases} \tag{4.2}$$

The polarized outputs are given by

$$\mu^- = \phi^{-1}\big(1 - (1 - \phi(\mu))^2\big), \qquad \mu^+ = 2\mu, \tag{4.3}$$

where μ^- is the Gaussian approximation estimated reliability of the bad channel and μ^+ is the Gaussian approximation estimated reliability of the good channel. Then the reliabilities of the polar code are found by sorting the μ_i values. The larger μ_i correspond to the more reliable bit-channels.
The reconciliation process is comprehensively evaluated across a range of code rates at an SNR of 20 dB for a codelength of 256. The reconciliation scheme operates independently at each of the 576 spatial grid points. Hence, for a chosen point on the grid, the reconciliation can output the reconciled vectors for both Alice and Bob. For example, for the point at the coordinates $(-1.202, 2.405, 0.4)$, we show in Fig. 4.2 the error probability after reconciliation for different code rates. As expected, the error probability increases with the code rate.

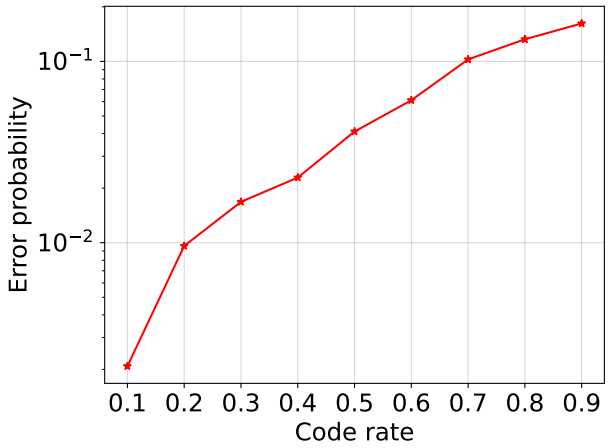


Figure 4.2: Error probability after reconciliation vs Code rate: SNR = 20 dB, Codelength = 256.

Privacy amplification: Following quantization and information reconciliation, privacy amplification is applied to compress the reconciled bit sequence into a shorter secret key while removing any residual information potentially available to the eavesdropper. In this challenge, we do not explicitly run a statistical test to estimate the optimal hashing rate for the considered dataset. Instead, we rely on previously obtained results under a LoS static scenario, which is compatible with the propagation conditions of the current dataset [7, 12]. Based on these results, we adopt a conservative hashing rate of $R = 0.1$, chosen deliberately lower than the estimated requirement to ensure a sufficient security margin. This value may be refined in future updates of the challenge as more precise entropy estimates become available. Once the hashing rate is fixed, privacy amplification is performed by applying a cryptographic hash function to the reconciled bit sequence. Specifically, the reconciled vectors are processed using the AES-128 hash function, producing a secret key of length consistent with the selected compression rate. A Davies-Meyer compression function is used to build the hash function (Fig. 4.3). Suppose $E : K \times \{0,1\}^n \to \{0,1\}^n$ a block cipher. Davies-Meyer compression function is given by

$$h(H, m) = E(m, H) \oplus H, \tag{4.4}$$

where m is the message block and H the current hash value. Let r be an output vector after reconciliation. At any point, we take 2 blocks of 128 bits from r as inputs and the current hash value is also XORed with the output of that iteration. Therefore, at each iteration, the size of r is reduced by half, then we move on to the next iteration, and so on.

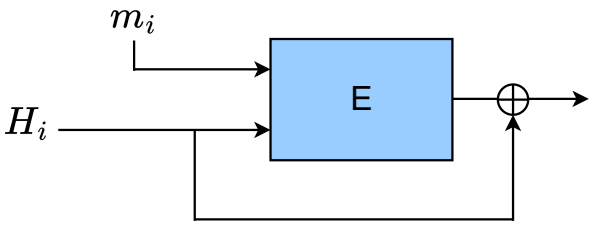


Figure 4.3: Davies-Meyer compression

This obtained key is then used for subsequent cryptographic operations, including one-time pad encryption in the challenge setting.

Input and Output

Inputs provided to the participant:

- Eve's CSI $\mathbf{H}_{AE} \in \mathbb{C}^{64\times 1}$, where the 64 rows correspond to the Uniform Linear Array (ULA) antenna elements and the 100 columns correspond to the Orthogonal Frequency Division Multiplexing (OFDM) subcarriers, measured at approximately one wavelength from Bob over a 12×12 spatial grid (576 total measurement points across 4 UEs).
- Public reconciliation information s_A: the syndrome generated by Alice using Polar codes with Cyclic Redundancy Check (CRC), transmitted over the public channel and therefore fully observable by Eve.
- Ciphertext $c = m \oplus k$: the One-Time Pad (OTP)-encrypted message, publicly disclosed.
- Full SKG protocol specification:
  - 2-bit uniform quantization applied to CSI observations.
  - Slepian-Wolf reconciliation via Polar codes with CRC (codelength = 256, evaluation Signal-to-Noise Ratio (SNR) = 20 dB).
  - Privacy amplification with hashing rate $R = 0.1$ using AES-128 based hashing function, implemented using Davies-Meyer compression function.
- No access to Alice's and Bob's channel observations $\mathbf{H}_{AB}$ and $\mathbf{H}_{BA}$.

Output expected from the participant:

- Recovered plaintext m, obtained by reconstructing the secret key k and computing $m = c \oplus k$.

Participants must also submit a description of their methodology, including well-commented code and an explanation of the proposed solution.

Evaluation metric

The objective of challenge participants is to **recover the OTP-encrypted plaintext** m, obtained from the publicly provided ciphertext $c = m \oplus k$. The resulting secret key is subsequently used as an OTP encryption key to assess its cryptographic strength under adversarial observation. Successful recovery is defined as exact reconstruction of the plaintext, while partial success is measured by the fraction of correctly recovered bits.
A submission is considered successful if it achieves either exact key recovery, exact plaintext recovery, or a statistically significant improvement over random guessing.

Chapter 5

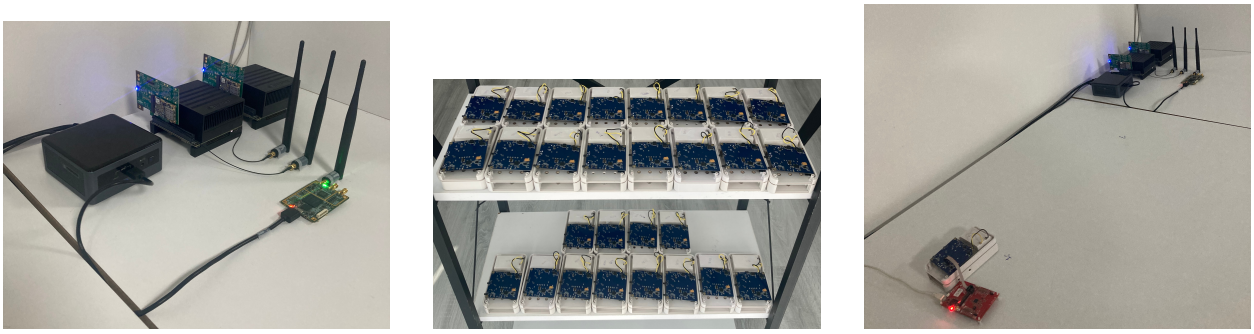
Challenge 4: Receiver-Invariant Device Identification Under Single Receiver Replacement

Abstract

RFFI is a physical-layer authentication technique that identifies wireless transmitters by learning unique hardware-induced signal impairments [13]. While RFFI has shown strong performance under controlled conditions, practical long-term deployments introduce a critical challenge: receiver hardware may need to be replaced due to failures, maintenance, or upgrades [14]. When this happens, the new receiver introduces its own hardware-specific distortions, causing a domain shift that degrades the performance of a previously trained RFFI model. This problem is particularly relevant for IoT deployments in 6G networks, where continuous and autonomous device authentication is required.
This challenge addresses the problem of maintaining reliable Radio Frequency (RF) fingerprinting performance when the monitoring receiver is replaced by another device in a deployed IoT system. When a new receiver is introduced, its hardware characteristics differ from those of the original receiver, leading to a domain shift and degraded identification performance. The challenge is based on an experimental dataset collected from 30 identical IoT transmitters, captured simultaneously by multiple software-defined radio receivers, enabling a controlled and reproducible study of this effect. The goal is to develop methods that adapt a trained RFFI model to a new, unseen receiver using only unlabeled signals collected after the replacement. This reflects a realistic post-deployment condition where relabeling data is not feasible. The scenario aligns with zero-touch operational requirements in 6G IoT systems, where security mechanisms must remain functional under hardware changes without manual intervention.

Dataset

The RF Fingerprinting Migration Dataset [15], introduced in Deliverable 5.1 [16] (Section 3.2.1), is publicly available at: `https://zenodo.org/records/14801935`
The dataset consists of raw In-phase and Quadrature (IQ) samples collected from 30 identical TI CC13XX IoT transmitters operating at 866 MHz using 2-GFSK modulation, as illustrated in Figure 5.1.
Each transmitted packet is captured by all receivers under synchronized conditions, and only successfully decoded packets are retained. This ensures that the same transmission instance can be directly compared across different receiver hardware. The dataset contains 815,367 packets collected in a controlled indoor environment, with transmitter–receiver distances of 0.5, 1, and 1.5 m. The consistent packet structure and sequence numbering allow reliable alignment of transmissions across receivers. This dataset enables the study of receiver-induced domain shift in RF fingerprinting and supports the development of methods for robust device identification under varying receiver conditions.



(a) Receiver infrastructure (b) 30 IoT transmitters (c) Transmission scenario

Figure 5.1: Experimental setup for receiver-invariant RF fingerprinting, including (a) Software-Defined Radio (SDR)-based receiver infrastructure, (b) 30 identical IoT transmitters used to collect data, and (c) an example transmission scenario illustrating data collection.

Challenge Description

No additional software is required beyond standard deep learning libraries.
In this challenge, an RFFI model is trained on labeled data collected from an initial receiver. After deployment, that receiver is replaced by a new one for which no labeled data is available. The goal is to adapt the model so that transmitter identification remains reliable on the new receiver. This challenge follows an open benchmark format. Participants train their own models, run their own experiments, and self-report results. They are encouraged to publish their findings citing the dataset. No base model is provided; participants must train their source model from scratch.
Data Split: The following fixed chronological split must be applied independently per receiver, per transmitter, and per transmission distance. Data from all three transmission distances (0.5, 1, and 1.5 m) must be used in training, adaptation, and evaluation:

- **Source Training:** Packets 0–599 (600 packets per transmitter per distance), labeled data for training the source model.
- **Source Validation:** Packets 600–699 (100 packets per transmitter per distance), labeled data for source model selection.
- **Source Test:** Packets 700–799 (100 packets per transmitter per distance), held-out labeled data for evaluating the source model before deployment.
- **Adaptation:** Packets 800–1399 (600 packets per transmitter per distance), unlabeled packets from the target receiver for unsupervised adaptation.
- **Oracle:** Packets 1400–1499 (100 packets per transmitter per distance), labeled packets from the target receiver. These may optionally be used strictly for adapted model selection. They must not be used for training or fine-tuning the model. Any use of the Oracle split must be explicitly declared.
- **Challenge Test:** Packets 1500–1599 (100 packets per transmitter per distance), held-out packets from the target receiver used to compute the final reported score.
- **Constraint:** Data used for training the source model must not be reused during the adaptation phase. Any use of labeled target data must be explicitly declared and must not exceed the scope permitted by the chosen track category.

Example: A participant trains the source model using packets 0–599 from receiver R_2. When adapting to R_1 or R_3, only packets 800–1399 from the target receiver may be used for adaptation. Packets 0–599 from R_1 or R_3 must not be used in any adaptation or evaluation step, as that interval is reserved exclusively for source model training.

Input and Output

Input: Raw IQ packets from the Challenge Test split of the target receiver, collected at all three transmission distances (0.5, 1, and 1.5 m), with receiver identity and sequence number information provided per packet. The challenge must be evaluated across all six source-target receiver pairs.
Output: The average Delta-F1 score across all six source-target transfer directions, as defined in the Evaluation Metric section. Participants must also submit a description of their methodology, including data preprocessing steps, model architecture, training procedure, adaptation strategy, and whether the Oracle split was used for model selection.

Evaluation Metric

The primary ranking metric is the Average Delta-F1, defined as:

$$\overline{\Delta F1} = \frac{1}{6}\sum_{k=1}^{6}\left(F1^{(k)}_{\text{adapted}} - F1^{(k)}_{\text{baseline}}\right) \tag{5.1}$$

where k indexes the six source-target receiver pairs, $F1^{(k)}_{\text{adapted}}$ is the macro-averaged F1-score of the adapted model on the Challenge Test split of pair k, and $F1^{(k)}_{\text{baseline}}$ is the macro-averaged F1-score of the unadapted source model on the same split. A higher $\overline{\Delta F1}$ indicates greater and more consistent performance recovery across all transfer directions. The macro-averaged F1-score is defined as:

$$F1_{\text{macro}} = \frac{1}{C}\sum_{c=1}^{C}\frac{2 \cdot P_c \cdot R_c}{P_c + R_c} \tag{5.2}$$

where $C = 30$ is the number of transmitter classes, and P_c and R_c are the precision and recall for class c, respectively. Macro-averaging is used to account for class imbalance, as the number of captured packets varies across transmitters. Participants must also report the per-direction Delta-F1 and absolute $F1_{\text{adapted}}$ scores as secondary metrics.

Chapter 6

Challenge 5: Robust Device Identification Under Sequential Receiver Replacement

Abstract

In long-term IoT deployments, receiver hardware may be replaced not once but multiple times over the system lifetime. Each successive replacement introduces new receiver-dependent distortions, which may compound over time and cause progressive degradation of the RFFI model. The question of how to maintain reliable identification across multiple sequential hardware changes has direct implications for system robustness, resource efficiency, and long-term authentication reliability.
This challenge extends the receiver replacement problem of Chapter 5 to deployments where multiple hardware changes occur over time. Each new receiver introduces additional distortions that can accumulate and cause the identification model to drift progressively. The challenge uses the same experimental dataset as in Chapter 5, described in the Dataset section of that challenge, exploiting its multi-receiver structure to simulate a sequence of receiver replacements in a controlled setting. The goal is to develop adaptation strategies that preserve transmitter identification capability across successive receiver changes, without any labeled data from replacement receivers, by building incrementally on each previous adaptation step.

Dataset

The dataset used in this challenge is the same as in Chapter 5, described in the Dataset section of that challenge and introduced in Deliverable D5.1 [16] (Section 3.2.1). It is publicly available at: `https://zenodo.org/records/14801935`
The multi-receiver structure of the dataset naturally supports the simulation of sequential receiver replacements, as three receivers are available to form all possible replacement chains.

Challenge Description

No additional software is required beyond standard deep learning libraries.
This challenge considers a scenario where the receiver is replaced more than once over the system lifetime. At each replacement stage, no labeled data is available for the new receiver, and the model must be adapted using only unlabeled signals from the current receiver. The adaptation is performed sequentially: at each replacement stage, the model is adapted from the most recently adapted model, building incrementally on previous adaptation steps. The objective is to maintain reliable transmitter identification performance across all replacement stages, avoiding progressive degradation over time. This challenge follows an open benchmark format. Participants train their own models, run their own experiments, and self-report results. They are encouraged to publish their findings citing the dataset. No base model is provided; participants must train their source model from scratch.

The dataset contains three receivers, giving rise to six possible sequential replacement chains across all permutations (e.g., $R_1 \rightarrow R_2 \rightarrow R_3$, $R_1 \rightarrow R_3 \rightarrow R_2$, and so on). Participants are expected to evaluate all six chains and report results for each chain independently.
Data Split: The source model split defined in Chapter 5 applies here. The following additional fixed chronological splits are used for the two sequential replacement phases, applied independently per receiver, per transmitter, and per transmission distance. Data from all three transmission distances (0.5, 1, and 1.5 m) must be used in adaptation and evaluation:

- **Phase 1 Adaptation:** Packets 800–1399 (600 packets per transmitter per distance), unlabeled packets from the first replacement receiver.
- **Phase 1 Oracle:** Packets 1400–1499 (100 packets per transmitter per distance), labeled packets from the first replacement receiver. These may optionally be used strictly for adapted model selection after Phase 1. They must not be used for training or fine-tuning the model. Any use must be explicitly declared.
- **Phase 1 Test:** Packets 1500–1599 (100 packets per transmitter per distance), held-out packets for evaluating performance after Phase 1 adaptation. Results must be reported as secondary metrics but are not used for ranking.
- **Phase 2 Adaptation:** Packets 1600–2199 (600 packets per transmitter per distance), unlabeled packets from the second replacement receiver.
- **Phase 2 Oracle:** Packets 2200–2299 (100 packets per transmitter per distance), labeled packets from the second replacement receiver. These may optionally be used strictly for adapted model selection after Phase 2. They must not be used for training or fine-tuning the model. Any use must be explicitly declared.
- **Phase 2 Test:** Packets 2300–2399 (100 packets per transmitter per distance), held-out packets used to compute the primary ranking metric.
- **Constraint:** Data from any previous receiver stage must not be reused during adaptation to a new receiver. Any use of labeled target data must be explicitly declared.

  Example: A participant trains the source model using packets 0–599 from receiver R_2. In Phase 1, the model is adapted to R_1 or R_3 using packets 800–1399 from that receiver. In Phase 2, the model is adapted to the remaining target receiver using packets 1600–2199 from that receiver. Packets 0–599 from any receiver must not be reused at any adaptation stage. Similarly, packets 800–1399 used in Phase 1 must not be reused during Phase 2 adaptation.

Input and Output

Input: Raw IQ packets from the Phase 2 Test split of each target receiver in the chain, collected at all three transmission distances (0.5, 1, and 1.5 m), with receiver identity and sequence number information provided per packet. This must be evaluated for all six receiver chains.
Output: The Average Delta-F1 score across all six sequential receiver chains at Phase 2, as defined in the Evaluation Metric section. Participants must also submit a description of their methodology, including data preprocessing steps, model architecture, training procedure, and whether Oracle splits were used for model selection.

Evaluation Metric

The primary ranking metric is the Average Delta-F1, defined as:

$$\overline{\Delta F1} = \frac{1}{6} \sum_{k=1}^{6} \left(F1_{\text{adapted}}^{(k)} - F1_{\text{baseline}}^{(k)} \right) \tag{6.1}$$

where k indexes the six sequential receiver chains, $F1_{\text{adapted}}^{(k)}$ is the macro-averaged F1-score of the adapted model on the Phase 2 Test split of chain k, and $F1_{\text{baseline}}^{(k)}$ is the macro-averaged F1-score of the unadapted source model on the same split. A higher $\overline{\Delta F1}$ indicates greater and more consistent performance recovery across all transfer directions. The macro-averaged F1-score is defined in Equation 5.2. Participants must also report the per-chain Delta-F1 and absolute $F1_{\text{adapted}}$ scores as secondary metrics.

Chapter 7

Challenge 6: Device Identification Under Temporal Drift in RF Fingerprinting

Abstract

While many RFFI systems achieve high identification accuracy when training and test data are collected at the same time, their performance degrades significantly when evaluated on data collected weeks later. This degradation is primarily caused by temporal variations in the transmitter hardware characteristics due to thermal effects during the hardware warm-up phase and the impact of power cycling the devices between data collections [17, 18].
This challenge addresses the long-term stability of RF fingerprinting in real IoT deployments. Participants must train their model exclusively on the short controlled baseline phase (Phase 1) and then evaluate its performance on the long-term phase (Phase 2) spanning several weeks. The goal is to explore effective strategies for temporally robust physical-layer authentication under realistic drift conditions. The challenge uses the RFFI-Temporal dataset [19].

Dataset

The RFFI-Temporal dataset is available via Zenodo (`https://zenodo.org/records/18952487`). It is designed to study the long-term stability of Radio Frequency Fingerprint Identification. It provides packet-aligned complex baseband IQ recordings from 30 TI CC13xx IoT devices captured over approximately nine weeks using three software-defined radio receivers. Only data from receiver R02 is used in this challenge. The dataset includes two collection phases: a controlled baseline phase (Phase 1) with uniform transmission intervals over approximately 45 hours, and a long-term evaluation phase (Phase 2) with device-specific intervals ranging from 15 seconds to 24 hours over approximately nine weeks. To support transition-based analysis, each packet is stored with pre- and post-packet margins that preserve transmitter startup and shutdown transients. Rich per-packet metadata, including internal temperature, battery level, and Real-Time Clock (RTC) timestamps, is available for additional analysis. The experimental testbed used for the longitudinal data collection is shown in Figure 7.1.

Challenge Description

No additional software is required beyond standard deep learning libraries.
In this challenge, participants must train their RF fingerprinting model exclusively on Phase 1 data from receiver R02. The trained model is then evaluated on the entirety of Phase 2 data from receiver R02. No data from Phase 2 may be used at any stage of training or model selection.
Data Split: The following fixed chronological split must be applied to Phase 1 data from receiver R02, independently per transmitter:

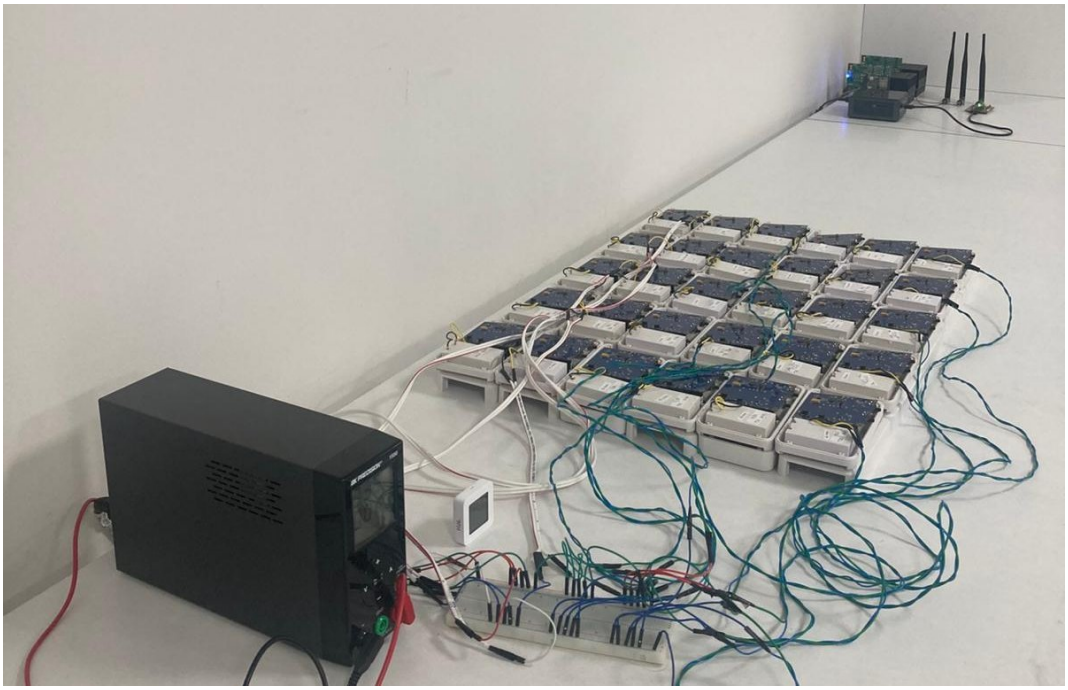


Figure 7.1: Experimental testbed showing the array of 30 TI CC13xx IoT transmitters arranged for the longitudinal data collection of the RFFI-Temporal dataset.

- **Training:** First 70% of Phase 1 packets per transmitter, labeled data for training the RF fingerprinting model.
- **Validation:** Next 15% of Phase 1 packets per transmitter, labeled data for model selection.
- **Test:** Last 15% of Phase 1 packets per transmitter, held-out labeled data for evaluating in-phase performance before long-term evaluation.
- **Final Evaluation:** All Phase 2 packets from receiver R02, used to compute the primary ranking metric. No Phase 2 data may be used during training or model selection.

Feature Extraction Rules:

- Only the preamble, sync word, and transient portions (pre-packet margins) may be used for feature extraction and model training.
- The payload part of the packet must not be used to avoid data leakage.
- Per-packet metadata (temperature, battery level, RTC timestamp) may be used for additional analysis but must not be included as input features for the main classifier.
- **Constraint:** Chronological ordering must be strictly preserved across all splits. No data from Phase 2 may be used at any stage before final evaluation.

Input and Output

Input: Raw IQ packets from Phase 2 of receiver R02, with transmitter identity and sequence number information provided per packet.

Output: The macro-averaged F1-score across all 30 transmitters on Phase 2, as defined in the Evaluation Metric section. Participants must also submit a description of their methodology, including data preprocessing steps, model architecture, training procedure, and any drift compensation strategy used.

Evaluation Metric

The primary ranking metric is the macro-averaged F1-score across all 30 transmitters on Phase 2, as defined in Equation 5.2. A higher $F1_{\text{macro}}$ indicates better long-term identification performance under temporal drift. Participants must also report the Phase 1 test split $F1_{\text{macro}}$ as a secondary metric to allow comparison between in-phase and cross-phase performance.

Chapter 8

Challenge 7: SKG on BRISC Dataset

Abstract

This challenge consists of evaluating the robustness of SKG techniques at the physical layer when an RIS is present. Specifically, two devices named Bob and Alice are located in two positions and transmit known pilot signals, allowing them to estimate the channel where an RIS is present. The channel is assumed to be symmetric for uplink and downlink transmissions. Upon the channel estimation, the two devices quantize their estimate with a uniform quantizer, decode their bits using a Low-Density-Parity-Check (LDPC) code, and then apply a hash function on the resulting bitstream. Trudy, on the other hand, is located in a different position and estimates the channel between her and Bob. From this estimate, she aims to guess the maximum number of secret bits extracted by Bob-Alice.

Dataset

The dataset is available at this link: `https://doi.org/10.5281/zenodo.18714621`.
The BRISC dataset is a large-scale, indoor measurement dataset designed to characterize wireless propagation in the presence of RIS. It provides experimentally collected CSI using standard compliant IEEE 802.11ac VHT frames over an 80 MHz bandwidth under a wide range of RIS configurations, enabling realistic evaluation of RIS-assisted communication systems.

Hardware and Experimental Setup The BRISC dataset is collected using synchronized SDRs operating at a carrier frequency of 5.53 GHz. The main components of the system are:

- **RIS:** A programmable metasurface composed of 256 reflecting elements. Each element can be configured in a binary manner (ON/OFF), controlling the reflection of incident electromagnetic waves.
- **SDRs:** Used at both the transmitter and receiver to generate, transmit, and receive signals. In the experiments two identical Ettus USRP-X310 are used.
- **Synchronization System:** An OctoClock distribution system ensures phase coherence and time synchronization between all SDR units.
- **Antennas:** the transmitter uses a single Log-periodic antenna (HyperLOG 60100), while the receiver uses two antennas: a Log-periodic (HyperLOG 30100) and a dipole.

The RIS elements are grouped into *blocks* to reduce configuration complexity. All elements within a block share the same state, and multiple block sizes are considered.

Measurement Procedure The dataset is generated by systematically varying:

- RIS configurations: either blocks of 64 elements are considered, resulting in a partition of the RIS into 4 blocks, or configurations using varying block sizes that divide the RIS into 9 blocks. For these two partitions, a total number of $2^4 + 2^9 = 528$ configurations is explored. Finally, random configurations are set, reaching a total number of 10k configurations.
- Transmitter positions: a total number of 9 transmitting positions are explored, as depicted in Fig. 8.1.
- Pilot transmission: for each transmitting position and RIS configuration, approximately 100 packets are transmitted, thus approximately 100 CSI estimates are provided.

Challenge Description

The software needed for data processing is available at this link: `https://github.com/MattiaP1999/BRISC_RIS_dataset_scripts`.
Bob is located in a fixed position, and Alice can be located in multiple positions, namely $[\boldsymbol{p}_1, \ldots, \boldsymbol{p}_9]$, as shown in Fig. 8.1. Bob selects configurations according to a distribution to derive secret bits: this distribution can be uniform, or optimized to maximize the mutual information or key capacity.

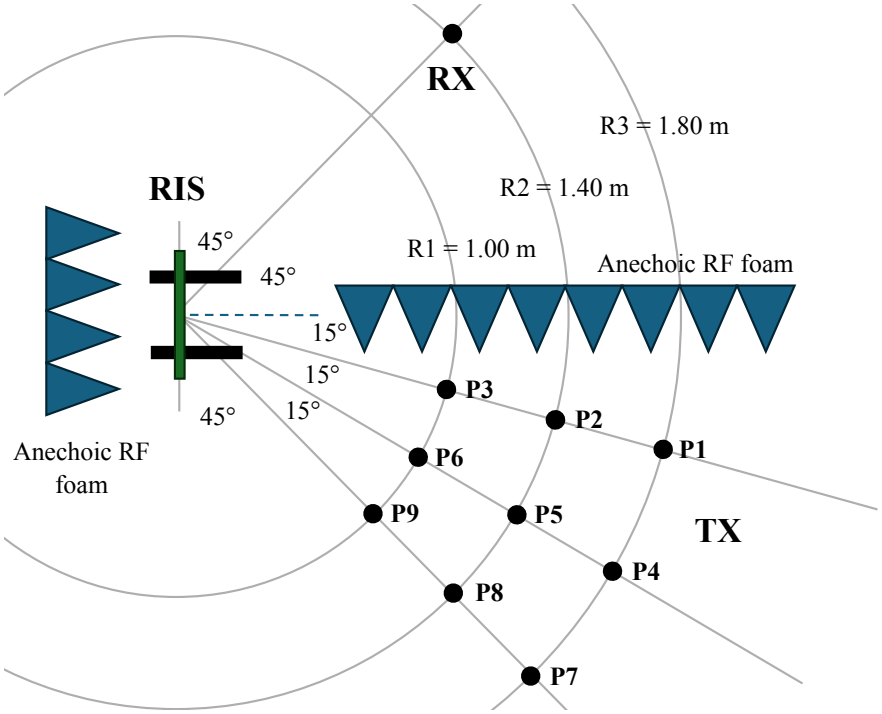


Figure 8.1: Measurement setup and geometry of the experimental scenario, showing the positions of the transmitter TX, receiver RX, and the RIS.

Methodology The procedure for converting the physical channel estimate into bitstreams is the following. Let $\boldsymbol{h}_{\mathrm{B}}(\boldsymbol{\Phi}, \boldsymbol{p}, n) \in \mathbb{C}^{N_s}$, where $N_s = 242$ is the number of subcarriers, be the measured channel vector by Bob with the RIS configuration $\boldsymbol{\Phi}$ when Alice transmits from position $\boldsymbol{p}$ at frame n. Assuming reciprocity of the wireless channel, we model the channel measured by Alice when Bob is transmitting as $\boldsymbol{h}_{\mathrm{A}} = \boldsymbol{h}_{\mathrm{B}}(\boldsymbol{\Phi}, \boldsymbol{p}, n')$,

with $n' \neq n$, which is taken from the dataset. This is equivalent to having a reciprocal channel with a different noise realization.
The bit extraction then is as follows:

- Bob measures the channel $\boldsymbol{h}_{\mathrm{B}}$ and, for each subcarrier, quantizes it with a uniform quantizer with K levels. Doing so, he attains a sequence of quantized channels $\{\boldsymbol{q}_s = [Q_K(\Re(h_{\mathrm{B}}^{(s)})), Q_K(\Im(h_{\mathrm{B}}^{(s)})]^T\}$, $s = 1, \ldots, N_s$, where $\Re(\cdot)$ and $\Im(\cdot)$ are the real-part and imaginary-part operators, respectively.
- The quantization levels are mapped into bit sequences using a Gray code, and concatenated together (horizontally). In formulae, Bob attains the sequence $\boldsymbol{b} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{N_s}]^T \in [0,1]^{2KN_s}$ where $\boldsymbol{b}_s = [\mathcal{G}(q_s^{(0)}), \mathcal{G}(q_s^{(1)})]$ and then the corresponding bit sequences are decoded with an LDPC code.
- The LDPC syndrome is then forwarded back to Alice, in clear, and she will correct her bits to end up in the same codeword as Bob.
- A hash function is finally applied to those decoded bits, and the output is the key.

Eve, positioned elsewhere in another position of Figure 8.1, has access to the channels corresponding to Bob's chosen configurations and attempts to reconstruct the key.

Input and Output

Each participant will be given:

- Eve's channel measurements and the corresponding bits extracted by Alice-Bob, which follow the pipeline previously explained, for a set of RIS configurations (training set).
- The hash function and LDPC code to use for extracting bits from channel measurements.

Each participant should provide the bit sequence associated with Eve's channel, such that it is as close as possible to Alice-Bob, for given RIS configurations, as explained below. Moreover, they should provide well-commented code and an explanation of the proposed solution.

Evaluation Metric

Using Eve's channels, the participant must be able to recover the bits that Alice-Bob extracted. In particular, the participants will have a new dataset (testing dataset) containing Eve's channels for unseen RIS configurations, and the corresponding key bits Alice-Bob extracted. The winner is the participant who attains the highest accuracy, defined as the number of correctly recovered secret bits on the test set. The current best score will be made available in a leaderboard.

Chapter 9

Challenge 8: PLA Authentication With RIS

Abstract

The goal of the challenge is to evaluate the effectiveness of PLA techniques and to quantify the impact of an RIS on authentication performance. Specifically, two legitimate devices, Alice and Bob, are located at fixed positions and exchange known pilot signals to estimate the wireless channel in the presence of an RIS. The scenario considers a game between the legitimate pair (Alice/Bob) and an adversary, Trudy. Each agent selects an action defined as a pair consisting of a spatial position and an RIS configuration. Based on a prior training phase, Alice and Bob aim to learn and optimize the joint probability distribution $p(\mathbf{p}_A, \mathbf{\Phi})$ over Alice's position and RIS configurations in order to minimize the False Alarm (FA) and Missed Detection (MD) probabilities. Conversely, Trudy optimizes moves uniformly at random across positions (except the current Alice position) to maximize the FA and MD probabilities.

Dataset

The dataset is available at this link: `https://doi.org/10.5281/zenodo.18714621`. The dataset used in this challenge is the BRISC dataset, whose details are provided in Chapter 8.

Challenge Description

The software needed for data processing is available at this link `https://github.com/MattiaP1999/BRISC_RIS_dataset_scripts`.
The considered scenario involves a wireless communication system assisted by an RIS, where a legitimate transmitter (Alice) communicates with a receiver (Bob) in the presence of an adversary (Trudy) attempting to impersonate Alice. The position of Alice is assumed to be known by Bob.
The PLA protocol consists of two main phases:

1. *Association Phase*: During this phase, Bob performs channel estimation using pilot signals transmitted by Alice. This allows Bob to learn the statistical characteristics of the legitimate channel under different RIS configurations. The BRISC dataset is used to obtain RIS-assisted channel realizations.

2. *Optimization Phase*: After the training phase, Alice and Bob jointly optimize the probability distribution $p(\mathbf{p}_A, \mathbf{\Phi})$, where $\mathbf{p}_A$ denotes the position of Alice and $\mathbf{\Phi}$ represents the RIS configuration. The objective is to enhance the separability between legitimate and adversarial channel distributions.

In contrast, Trudy moves uniformly at random across positions, avoiding the current Alice position.
Authentication at Bob is performed via hypothesis testing, where Bob checks whether the estimated channel belongs to the null hypothesis $\mathcal{H}_0$, i.e., Alice is transmitting, and the under-attack hypothesis $\mathcal{H}_1$, i.e., the received message comes from Trudy. This gives rise to two types of errors, namely FA, i.e., the message

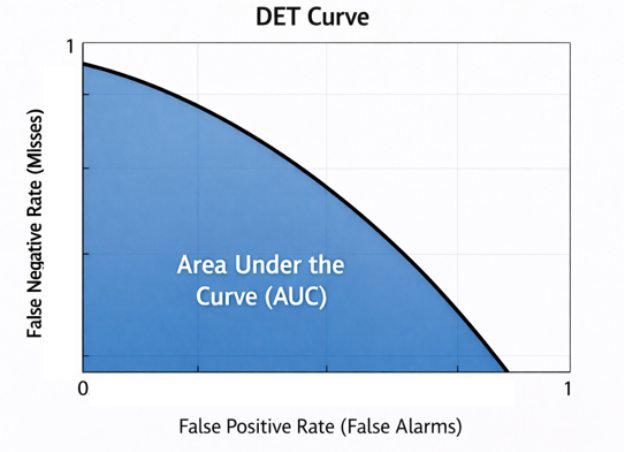


Figure 9.1: AUC of the DET curve

was authentic but it was labeled as fake, and the MD, i.e., the message was from Trudy but it was labeled as legitimate.

Input and Output

Each participant will be given:

- The measured channels by Bob when the transmitter is in multiple positions for given RIS configurations (training set). Note that artificial noise will be added to the measured channels in the processing scripts to make the challenge more challenging.
- Other testing channels similar to point 1), that cannot be used for training/fitting the algorithm, but are used to evaluate it (testing set). Also here, artificial noise is added.

Each participant should design an algorithm to find the joint probability of positions and configurations $p(\mathbf{p}_A, \mathbf{\Phi})$ that minimizes both the MD and FA error rates. The authors shall provide well-commented code and an explanation of the proposed solution.

Evaluation Metric

We use the Area-Under-the-Curve (AUC) of the Detection Error Tradeoff (DET) depicted in Fig. 9.1 as a performance metric. We recall that DET shows the False Negative probability (i.e., the MD) against the False Alarm probability (i.e., the FA). A lower AUC reflects stronger authentication performance and system robustness. The challenger who obtains the lowest AUC wins.

Chapter 10

Challenge 9: Device Classification Under Hardware Impairments

Abstract

This challenge introduces a controlled RF fingerprint dataset designed to evaluate device classification under hardware-induced signal distortions. The dataset is generated from OFDM transmissions affected by residual hardware impairments, Carrier Frequency Offset (CFO), and Symbol Timing Offset (STO).
Multiple feature families are extracted in the frequency domain to capture device-specific signal characteristics. The framework supports supervised classification and enables systematic evaluation of robustness under varying impairment conditions.
The objective is to assess how well machine learning-based classification methods can distinguish devices when the observed signals are distorted by realistic hardware imperfections, reflecting practical deployment conditions in emerging 6G systems.

Dataset

The dataset is available at this link: `https://doi.org/10.5281/zenodo.19481618`
The dataset consists of OFDM signal realizations generated from two transmitting devices operating under distinct hardware impairment conditions. Each device is characterized by a unique combination of residual hardware impairments (Residual Hardware Impairments (RHI)), CFO, and STO.
Signals are generated using Binary Phase Shift Keying (BPSK)-modulated OFDM waveforms over Rayleigh fading channels with additive white Gaussian noise. The receiver captures the impaired signals, which are transformed into the frequency domain prior to feature extraction.
Residual hardware impairments refer to non-idealities such as amplifier nonlinearities, I/Q imbalance, and phase noise that remain after calibration and act as device-specific fingerprints.
The impairment parameters are configured as follows:

- Residual hardware impairment coefficients: $\kappa_1 = 1$, $\kappa_2 = 1.3$
- Carrier frequency offsets: 1.25 kHz and 2.5 kHz
- Symbol timing offsets: 5 and 10 samples
- Noise variance: $\sigma^2 = 10^{-5}$

Each transmitting device has 1000 OFDM realizations, resulting in a balanced dataset.
For each received signal, multiple feature families are extracted, including:

- Power Spectral Density (PSD) statistics

- Spectral flatness and bandwidth
- Statistical descriptors of I/Q components
- L-moments
- Autocorrelation-based metrics
- Amplitude and phase statistics

Challenge Description

The objective of this challenge is to design a machine learning-based classification model capable of identifying the transmitting device based on RF fingerprints extracted from impaired signals.
The classification task is performed under realistic hardware impairment conditions, where distortions caused by RHI, CFO, and STO affect signal characteristics. These impairments introduce variability in the feature space, making the classification problem more challenging and representative of practical 6G deployment scenarios.
Participants are expected to:

- Develop robust classification models
- Analyze feature effectiveness under impairment conditions
- Ensure generalization across different distortion scenarios

This challenge establishes a controlled and reproducible benchmark for evaluating RF fingerprinting techniques in physical layer security and lightweight device authentication for 6G systems.

Input and Output

Input:

- Feature vectors extracted from received OFDM signals (frequency-domain features such as PSD, statistical descriptors, etc.)

Output:

- Predicted device label (e.g., Device 1 or Device 2) for each input sample

The task is a supervised classification problem where each input feature vector must be mapped to the correct transmitting device. The authors shall provide well-commented code and an explanation of the proposed solution.

Evaluation Metric

Performance is evaluated using the following metrics:

- **Accuracy:**

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}}$$

- **Macro-averaged F1-score:**

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Accuracy measures the overall proportion of correctly classified samples, while the macro-averaged F1-score evaluates classification performance across all classes equally, regardless of class imbalance.
These metrics jointly assess classification effectiveness and robustness under hardware impairment conditions.

Appendix A

Additional datasets

A.1 CDL dataset

This dataset is generated using time-varying 3GPP-compliant clustered delay line (CDL) channel models at mmWave frequency. It is sampled under diverse mobility and propagation conditions by varying the CDL profile, user velocity, and delay spread.

Dataset ID | Name	DAT85 | 3GPP CDL CSI Dataset
Key Features	Time-varying wideband CSI generated from 3GPP-compliant CDL models at 28 GHz; multiple channel profiles; varying user mobility and delay spread.
Quick Overview	Simulated CSI dataset for a 16×1 MIMO-OFDM system. CDL profiles are randomly selected from {`CDL-A`, `CDL-B`, `CDL-C`, `CDL-D`, `CDL-E`}. Each sample contains 100 OFDM symbols and CSI from 16 evenly spaced subcarriers.
Threat Coverage	Covers channel variability due to mobility, multipath richness, and delay spread variations; useful for studying channel aging and distribution shifts in wireless environments.
6G-TE — Tech Domain	Wireless communications | CSI feature extraction | 3GPP channel modeling
Data Type	Simulated complex-valued CSI sequences from MIMO-OFDM channels
Data size	Each CSI sample consists of 100 time steps and 16 selected subcarriers for a 16×1 MIMO system
Transmitter | Receiver	Base station with 16 transmit antennas | user terminal with 1 receive antenna
Owner | Access | Licence	CHA | Public | CC BY 4.0
Link	`https://zenodo.org/records/19493713`

A.2 RF Fingerprinting Migration Dataset

The RF Fingerprinting Migration Dataset consists of synchronized wireless signal captures from 30 identical IoT transmitters and 3 software-defined radio receivers, collected simultaneously in a controlled indoor environment. It is specifically designed to support the study of receiver-induced domain shift in RF fingerprinting systems, enabling controlled evaluation of cross-receiver adaptation methods for physical-layer authentication in IoT deployments.

Dataset ID | Name	DAT86 | RF Fingerprinting Migration
Key Features	• 30 identical IoT transmitters (same manufacturer and batch) • 3 simultaneous SDR receivers (2 same-model, 1 heterogeneous) • Synchronized multi-receiver captures with packet-level correspondence • Three transmitter–receiver distances: 0.5 m, 1.0 m, 1.5 m • 2-GFSK modulation at 866 MHz, 400 kHz sample rate
Quick Overview	• Enables controlled evaluation of receiver-induced domain shift in RFFI • Supports development of unsupervised domain adaptation methods • Facilitates study of single and sequential receiver replacement scenarios • Suitable for benchmarking receiver-invariant physical-layer authentication
Threat Coverage	Spoofing, Tampering, Information Disclosure
6G-TE | Tech Domain	IoT | Short Range Communications
Data Type	Complex baseband I/Q samples (complex64, `.cfile`)
Data Size	815,367 packets, ≈6.7 GB
Transmitter | Receiver	30 custom IoT sensors, 2-GFSK, 866 MHz | 2× Fairwaves XTRX, 1× Ettus B200 Mini
Owner | Access | Licence	GOHM Electronics | Public | CC BY 4.0
Link	`https://zenodo.org/records/14801935`

A.3 RFFI-Temporal Dataset

RFFI-Temporal is a longitudinal dataset designed to study the long-term stability of Radio Frequency Fingerprint Identification (RFFI). RFFI systems are known to degrade over time due to temporal changes. RFFI-Temporal addresses this gap by providing packet-aligned complex baseband IQ recordings from 30 TI CC13xx IoT devices captured over approximately nine weeks. The dataset includes two collection phases: a controlled baseline phase with uniform transmission intervals, and a long-term evaluation phase with device-specific intervals ranging from 15 seconds to 24 hours. To support transition-based analysis, each packet is stored with pre- and post-packet margins that preserve transmitter startup and shutdown transients. Rich per-packet metadata, including internal temperature, battery level, and RTC timestamps, enables fine-grained study of hardware aging and battery decay effects on RF fingerprints. The dataset is intended to support the development and benchmarking of temporally robust physical-layer authentication methods for IoT deployments.

A.4 BRISC Dataset

UNIPD prepared the BRISC dataset. The dataset comprises CSI measurements collected at 5.53 GHz using a 256-element RIS with binary states. The measurement campaign utilized two identical single-antenna SDRs, phase-synchronized via an OctoClock. To manage complexity, the RIS elements are grouped into blocks, where all elements within a block share the same state. We captured CSI under multiple transmitter positions (and fixed receiver location), across various block sizes and state configurations, enabling an exhaustive search of all possible combinations when degrees of freedom are restricted. Furthermore, we calibrated the parameters of state-of-the-art RIS channel models to best fit the measured CSI. With approximately 10,000

Dataset ID | Name	DAT87 | RFFI-Temporal
Key Features	• 30 TI CC13xx IoT transmitters (15 battery-powered, 15 DC-powered) • 3 simultaneous SDR receivers (2 same-model, 1 heterogeneous) • Per-packet metadata: transmitter ID, sequence number, temperature, battery level, RTC timestamp • Pre- and post-packet margins preserving transmitter startup and shutdown transients • Two collection phases: ≈45 h baseline (Phase 1) and ≈9-week long-term evaluation (Phase 2) • 2-GFSK modulation at 866 MHz, 400 kHz sample rate
Quick Overview	• Enables longitudinal study of temporal drift and hardware aging in RFFI • Supports investigation of battery decay effects on RF fingerprints • Facilitates transition-based analysis using startup and shutdown transients • Suitable for benchmarking long-term stability of physical-layer authentication
Threat Coverage	Spoofing, Tampering, Information Disclosure
6G-TE | Tech Domain	IoT | Short Range Communications
Data Type	Complex baseband I/Q samples (complex64, HDF5)
Data Size	6,360,482 valid packets, ≈140.4 GB (6 HDF5 files)
Transmitter | Receiver	30 TI CC13xx IoT sensors, 2-GFSK, 866 MHz | 2× Fairwaves XTRX, 1× Ettus B200 Mini
Owner | Access | Licence	GOHM Electronics | Public | CC BY-NC 4.0
Link	`https://zenodo.org/records/18952487`

configurations explored per transmitting position, BRISC serves as a robust benchmark in communication applications. We also show here an example of its use for physical layer authentication. The detailed description of the dataset is available in [20].

Dataset ID | Name	DAT88 | BRISC
Key Features	• CSI collected at 5.53GHz using a 256-element RIS with binary states. • Transmitter and receiver are two software-defined radios (SDRs), phase-synchronized via an OctoClock • Multiple transmitting positions and a single receiving position
Quick Overview	The dataset contains CSI estimates attained from standard-compliant signals in an indoor scenario where an RIS is present. The RIS spans 10k configurations per transmitting position, while the receiver is in a fixed position
Threat Coverage	Spoofing and Confidentiality
6G-TE — Tech Domain	IEEE 802.11ac
Data Type	Complex baseband CSI samples
Data size	40 GB
Transmitter | Receiver	Ettus USRP X310 | Ettus USRP X310
Owner | Access | Licence	UNIPD | Public | CC BY 4.0
Link	`https://doi.org/10.5281/zenodo.18714621`

A.5 RF Fingerprint Dataset for Device Classification under Hardware Impairments

This dataset contains synthetically generated RF signals designed for device classification using radio frequency (RF) fingerprinting techniques under realistic hardware impairments. It includes both raw complex I/Q signal samples and extracted statistical features such as mean, variance, skewness, and kurtosis. The signals are generated using an OFDM-based model and incorporate key physical-layer impairments, including residual hardware impairments (RHI), carrier frequency offset (CFO), symbol timing offset (STO), Rayleigh fading, and additive white Gaussian noise (AWGN). The dataset consists of two classes representing different devices and is suitable for evaluating machine learning models in physical layer security, device authentication, and RF-based identification scenarios.

Dataset ID | Name	DAT89 | Title
Key Features	• RF fingerprint features extracted from OFDM signals • Includes statistical features (mean, variance, skewness, kurtosis) • Raw I/Q signal samples (128-length complex signals) • Simulated hardware impairments (RHI, CFO, STO) • Labeled dataset for binary device classification
Quick Overview	This dataset contains synthetically generated RF signals from two devices with distinct hardware impairments. It supports machine learning-based classification and analysis of physical layer security mechanisms.
Threat Coverage	• Device impersonation detection • RF spoofing resilience • Hardware impairment-based fingerprinting • Physical layer authentication scenarios
6G-TE — Tech Domain	Physical Layer Security (PLS), RF Fingerprinting, AI/ML-based Device Identification
Data Type	• Structured tabular data (CSV) • Complex RF signals (I/Q format) • Statistical feature vectors
Data size	• ~ 2000 samples • 2 classes (1000 samples per device) • Raw signal: 256 features (I/Q) per sample • Statistical dataset: 8+ features per sample
Transmitter | Receiver	• Transmitter: 2 simulated user devices with distinct impairments • Receiver: Single receiver under Rayleigh fading channel
Owner | Access | Licence	EBY | Public | CC BY 4.
Link	`https://doi.org/10.5281/zenodo.19481618`

Appendix B

Validation of the Objectives of WP5

In this Appendix, we summarize the main contributions of the Work Package 5 in relation to the measurable results, quantifiable targets, and objectives stated in the ROBUST-6G proposal.

B.1 Measurable results

R5.1 - Develop a library of RF fingerprints of attacks to build corresponding classifiers for real-time identification of the attack and, whenever possible of the attacker.

We have collected a library of RF fingerprints and attacks online datasets related to various threat scenarios in Chapter 5 of D5.1 [16], covering spoofing, tampering, information disclosure, and denial of service. One of these datasets, the IEEE DataPort Ultra-Dense Indoor MaMIMO data set [6] was used for Security Challenges 2 and 3 in the context of AoA-PLA spoofing and SKG eavesdropping. It has been further used in the Physical Layer Security Closed Loop (PLS-CL) ROBUST-6G Demonstrator 4 for which additionally two approaches were introduced for the identification of a jamming attack: i) a distributed jamming attack identification approach that employs a double change point detector in order to identify and localize a jammer; ii) a local (on the user location) change point detector that uses historical data and can be used to identify the attack (but is less accurate in identifying the location of the jammer). The achieved accuracy for the jamming detection on the given dataset approaches 100% in both the distributed and localized scenarios and is being incorporated in the PLS-CL to feedback decisions regarding the power adaptation necessary to counter the jamming attack. The details of the proposed real-time jamming detection approaches and the implementation of the PLS-CL can be found in [21] and are being integrated in the ROBUST-6G Demonstrator 4.
Furthermore, we have collected a dataset of RF fingerprints from 30 identical TI CC13XX IoT transmitters, captured by three receivers consisting of two Fairwaves XTRX SDRs and one Ettus USRP B200 Mini, operating at 866 MHz with 2-GFSK modulation. The dataset does not contain pre-defined attack scenarios. However, because all transmitters are identical and produce highly similar packet structures, any subset of the devices can later be designated as rogue transmitters attempting to impersonate legitimate ones. This enables the realistic study of impersonation and spoofing threats. A deep learning classifier was developed and evaluated on this real hardware testbed to authenticate devices based on their physical RF characteristics. The dataset is publicly available on Zenodo [15].
Further datasets created for the PLS challenges are described in Appendix A in this deliverable, complementing the datasets in D5.1 [16].

R5.2 - Propose mitigation strategies for the identified attacks

We focused on mitigation of spoofing for AoA-PLA [5, 22–25], jamming attack (through power adaptation) [21] and eavesdropping in SKG schemes [12, 26, 27], all of which are incorporated in the PLS-CL proposed in deliverables D6.2 [28] and D2.3 [29] and will be exemplified in the ROBUST-6G Demonstrator 4. Another set of contributions relates to either authentication [30–33] and Chapters 23 and 24 of D5.2 [34], or anti-jamming [35–37], and are detailed below in specific requirements.

R5.3 - Investigate the use of physical layer wiretap coding in mmWave mMIMO systems, including code design, information leakage analysis (secrecy maps), and enhancements using RIS and ML.

Chapter 14 and Appendix E of D5.2 [34] present bounds on the information leakage of polar wiretap codes in finite blocklength, measured in terms of total variation distance (TVD). More general results are presented in [38], where we show that for all polarizing codes, the leakage can be bounded by the sum of the TVDs of the bit-channels corresponding to the confidential and frozen bits. This suggests a simple code design criterion and allows for computing lower bounds for the secrecy rate of different families of polarizing wiretap codes over a binary erasure wiretap channel (BEC). We consider the case where Bob's channel is a BEC(0.05) for Bob, and Eve's channel is a BEC(0.4), and impose an error probability constraint $\epsilon = 0.001$ and a TVD secrecy constraint $\delta = 0.01$. We evaluate our constructions by numerical simulations for Polar, Adjacent Bit-Swapped polar (ABS), Multi-kernel polar (MK), and Reed-Muller (RM) codes, which perform respectively at 49%, 52.8%, 60.4%, and 88.7% of the second order lower bound for the secrecy capacity at blocklength 512. These results suggest that ABS codes and MK codes strike a balance between achievable secrecy rate and decoding complexity, outperforming polar codes while remaining more practical than RM codes.

R5.4 - Propose novel authentication and key agreement (AKA) solutions for low-latency and low-complexity scenarios, use physical layer authentication to identify false base stations.

Chapter 3 in D5.2 [34] presents novel techniques for PLA based on AoA, [22–25], and Chapter 13 proposes enhanced reconciliation and preprocessing methods for PLA [9, 39–41], with low time and memory complexity. Fast SKG is proposed in Chapter 12 of D5.2, presenting the works in [12, 26, 27]. Importantly, a demonstrator with sub-msec SKG run-time, leveraging physical context for parameter fine-tuning, is presented (awarded the Best Demo award at IEEE CNCS 2025). Furthermore, the use of RISs and drones for PLA is investigated in Chapters 15 and 16 of D5.2. In particular, in [30], a novel Challenge-Response (CR)-PLA mechanism is proposed for a cellular system that leverages the reconfigurability property of a RIS (under the control of the verifier) in an authentication mechanism, and in [31] novel attacks against CR-PLA are presented. In Chapter 23 of D5.2, we study a scenario where multiple BSs collaborate to authenticate the transmitting device at the physical layer using personalized federated learning, while Chapter 24 considers a cross-layer authentication mechanism using information from the network and physical layer. Such solutions can be deployed to identify false base stations. Moreover, in [32], the authors propose a novel CR-PLA protocol using drones to modify the propagation environment in which a drone Bob authenticates the sender Alice with the presence of a malicious agent Eve by exploiting his prior knowledge of the wireless channel statistic (fading, path loss, and shadowing), and moving to a set of positions on the map. By estimating the attenuations of the received signals, he will authenticate the sender. In this context, in [33] we formalize such an approach by modeling the problem as a zero-sum game against the intruder. Finally [42] considers a network where the user equipment is a drone and proposes a novel secret key generation solution when the eavesdropper is another node belonging to the network (curious device). We exploit drone mobility over realistic Rician fading channels to generate the key.

R5.5 - Work on trustworthy sensing, focusing on providing integrity guarantees for radar localization as a starting point.

Chapter 2 of D5.2 [34] addresses the vulnerabilities of Integrated Sensing And Communication (ISAC) in Non-Orthogonal Multiple Access (NOMA) systems, and the design of robust sensing signals. Chapter 17 studies adversarial attacks on ISAC systems; in particular, we consider a scenario where Alice and Bob cooperate to perform bistatic sensing of the environment while the attacker Trudy aims at disrupting such

a procedure by properly designing her transmitting beamformer using adversarial Machine Learning (ML) techniques.
Furthermore, Sections 19.2 and 19.3 of D5.2 present a new trust and reputation management model incorporating sensing information. In Section 19.5 of D5.2, we focus on the relation between CSI accuracy and sensing accuracy, and present a new channel estimation algorithm enhanced by radar sensing [43]. This algorithm combined environmental insights gathered from radar sensing with compressed sensing methods for channel estimation, enabling accurate assessment of channel states without relying on a large number of pilot signals. Key advantages of this approach included reduced pilot overhead in MIMO systems and improved estimation performance, both of which were crucial for optimizing spectral efficiency and minimizing system complexity. The proposed solution, which leverages radar sensing information, demonstrates a superior channel estimation accuracy compared to state-of-the-art algorithms with a normalized mean square error (NMSE) of less than 1 in dB, across various SNR scenarios.

R5.6 - Enhance accuracy, efficiency, and migration while minimizing updates and evaluating techniques under varying conditions for streamlining of RF fingerprinting.

We have contributed to the enhancement of accuracy, efficiency, and long-term robustness in RF fingerprinting by addressing two fundamental challenges that arise in RFFI.
The first challenge is receiver variability. In future deployments, receiver hardware may need to be replaced or upgraded due to maintenance or failures. When an RFFI model trained on one receiver is deployed on a different receiver, classification performance degrades significantly because the model inadvertently learns receiver-specific hardware characteristics together with the device fingerprints. As shown in Chapter 4 of D5.2 [34] and in [44], domain adaptation techniques can recover a significant portion of this lost performance, thereby improving the practicality and efficiency of RF fingerprinting.
The second challenge is temporal drift. Hardware characteristics of IoT transmitters can change over time due to various environmental and hardware-related factors. To investigate these effects, we collected the RFFI-Temporal Dataset, a longitudinal collection covering 30 transmitters (15 battery-powered and 15 DC-powered) with transmission intervals ranging from 15 seconds to 24 hours. Work on this dataset is ongoing. These two datasets have been publicly released as open research problems for the research community (see Chapters 5, 6, and the RFFI-Temporal Dataset A.3 in this deliverable).

R5.7 - Introduce a holistic trustworthiness approach, including lower layers to provide early identification of anomalies.

Chapter 19 of D5.2 [34] consolidates five contributions focusing on the role of the physical layer in fostering trust and trustworthiness in 6G systems [43, 45–48]. A unifying insight across these works is that future Cyber–Physical Systems (CPS) and multi-agent networks will necessitate the integration of *objective and quantifiable trust metrics* directly within the wireless substrate. Within this context, joint AoA and Time of Flight (ToF) sensing are identified as fundamental physical-layer enablers. By delivering verifiable spatial and temporal information, these techniques provide a robust foundation for enhancing trust, ensuring system integrity, and enabling accountability among autonomous entities.

B.2 Quantifiable targets

- **Detection of jamming/interference denial of service attacks with accuracy higher than 90%.**

  In [35], a jamming detection scenario is proposed based on dynamic graphs, while in [36, 37], broadband jammers are detected using spectrograms and ML. Furthermore, two lightweight jamming

detectors with accuracy exceeding 90% were developed as part of the PLS-CL demonstrator [21]. All the aforementioned works achieve performance much higher than 90% accuracy.

- **Reach a detection accuracy higher than 70% for Sybil attack with the aid of source and device localization and RF fingerprinting.**

  We have demonstrated that rogue IoT transmitters attempting to impersonate legitimate devices can be detected using RF fingerprinting on the RF Fingerprinting Migration dataset. In the evaluated Sybil scenario, 5 transmitters are designated as legitimate devices, while rogue transmitters attempt to impersonate multiple legitimate identities. Detection performance on a held-out test set of real hardware captures exceeds the 70% target. Full evaluation results will be reported in a forthcoming paper. We have demonstrated in Security Challenge 2 that impersonation attacks can be identified with accuracy exceeding 90%.

- **Develop authentication and key agreement schemes with less than 5 msecs latency (static nodes).** In [27], we have presented a secret key agreement scheme which runs in less than 1 msec on a real demonstrator. Our protocols have a reconciliation rate greater than 99% [39, 40]. In [24], we have evaluated the time for authentication inference and showed that using Gradient Boosting Machine (GBM) classifiers, we achieve 90% accuracy with inference time 0.0018 sec. Combining [39, 40] and [24], we have Authentication and Key Agreement (AKA) schemes with less than 5 msec latency [30], and the works in Chapters 23 and 24 of D5.2 [34] have a detection latency equal to the devices' processing times.

- **6G resilience will be increased at least by** 20% **through the proposed mitigation techniques**. Under adversarial conditions, physical-layer resilience is primarily assessed through metrics that capture the system's ability to withstand, detect, and mitigate attacks such as jamming, spoofing, and eavesdropping. Key indicators include i) the jamming detection probability and possible adaptation to avoid degradation in throughput or reliability under attack scenarios, ii) spoofing detection probability, and iii) confidentiality against eavesdropping, although this list is clearly non-exhaustive. In our works in ROBUST-6G we achieved jamming detection probabilities exceeding 90%, spoofing detection probabilities approaching 100%, while we proposed privacy amplification techniques for SKG that are fast and robust against eavesdropping. Jointly, these results provide a first quantitative basis for understanding how effectively the physical layer can preserve secure and reliable communication in hostile environments. This aspect is to be further developed in [21]. In this work, we showcase that by using lightweight jamming identification algorithms in the monitoring stage of the proposed PLS-CL, we can subsequently drive power adaptation (closing the loop) in order to avoid link quality degradation. Overall, accounting for jamming attack detection accuracy > 90%, spoofing attack detection probability close to unity (as long as the legitimate and adversarial nodes are not colinear), and robustness against eavesdropping in SKG systems, there is a strong indication that the target of increasing resilience by at least 20% can be very comfortably achieved. Final results in this direction will be part of the Demonstrator 4 in D6.3.

B.3 WP5 Objectives

SO5.1 - Classification and identification of attacks in the radio, including jamming, attacks during network entry, false base stations, injection attacks, and malicious pilot contamination in MIMO, infrastructure safety for dMIMO and synchronization attacks, sensing integrity guarantees.

New solutions for jamming detection are presented in Chapters 6,7 and 8 of D5.2 [34]. As mentioned previously, two jamming detection approaches (distributed identification and local change point detectors)

will be integrated in Demonstrator 4. Jamming detection based on dynamic graphs and spectrograms and ML is also covered in [35–37]. Chapter 24 of D5.2 considers a cross-layer authentication mechanism which can be used to identify false base stations. Denial of service attacks are discussed in Chapter 9, which focuses on identifying irregular spectral signatures associated to threats using Convolutional Neural Network (CNN)-based classifiers. Classification of malicious users through radio frequency fingerprinting is also addressed in Chapter 10 of D5.2. PLA to prevent spoofing attacks in mMIMO systems is considered in Chapter 3; PLA techniques based on AoA [22–25] and CR mechanisms leveraging drones and RIS [30, 32] can be used to identify malicious users. Sensing integrity was addressed in Chapter 2 of D5.2 [34].

SO5.2 - Novel, PLS based security schemes for 6G leveraging mMIMO, RIS, dMIMO, focusing on providing measurable security guarantees, e.g., in terms of information leakage.

We have published several papers demonstrating the use of RISs in PLS based security schemes for challenge-response authentication, which are presented in Chapter 15 of D5.2 [34]. Furthermore, we discuss the impact of hardware impairments on RIS-based authentication in Chapter 18. We have investigated PLS-based Channel-Response Authentication (CRA) mechanisms leveraging the spatial processing capabilities of mMIMO and distributed MIMO (dMIMO) within ISAC systems. In particular, in D5.2, we developed a robust uplink NOMA-ISAC framework that jointly considers communication, sensing, and security in the presence of a passive eavesdropper with unknown location. Our approach leverages sensing-assisted system design to limit information leakage. Specifically, we model the uncertainty in the eavesdropper's location (angle and delay) using a Cramer-Rao bound (CRB)-based framework and explicitly propagate this uncertainty into the communication and eavesdropping channels. This enables the derivation of uncertainty-aware Signal-to-Interference-plus-Noise Ratio (SINR) expressions and the formulation of a joint beamforming and sensing-power optimization problem with explicit secrecy constraints. Regarding the ROBUST-6G objectives, our framework enables:

- OBJ 6.4 (attack detection < 5 minutes): The detection of adversarial activity can be achieved within a few channel estimation intervals by monitoring deviations from expected channel statistics and CRB-based estimation bounds, which directly impact the observed SINR and sensing metrics.
- OBJ 6.5 (false positives < 5%): By leveraging uncertainty-aware SINR modeling and post-processing of residual interference after NOMA Successive Interference Cancelation (SIC), the system can distinguish between normal channel variations and adversarial-induced anomalies, enabling controlled false alarm rates while maintaining high detection probability.

SO5.3 - Low latency and low footprint authentication and key agreement protocols for challenging use cases, such as for industrial IoT and identification of false base stations, that will form the basis for zero-touch Automatic Device Enrolment (ADE) solutions in WP4.

Authentication in industrial IoT is addressed in Chapter 24 of D5.2 [34], where the malicious transmitter is detected by comparing the known legitimate position with two estimates of it obtained from the CSI and traffic information. Furthermore, as part of WP4, ADE was proposed to consolidate the proposed AoA-PLA and fast SKG schemes. As an example, we assume an authenticated location is being used for ADE within a smart factory. Then, an ADE handshake between a new device placed in this location and a multi-antenna verifier has been conceptualized through a proposed authentication and key agreement handshake between the device and the authenticator [21].

SO5.4 - PHY enabled trustworthiness and resilience, including: i) localization privacy; ii) trustworthy sensing; iii) generalized anomaly detection collecting alerts from the physical and hardware layers, to be used in WP4.

Attacks on ISAC systems are presented in Chapter 17 of D5.2, where Alice and Bob cooperate to perform bistatic sensing of the environment, while Trudy aims at disrupting such a procedure by properly designing her transmitting beamformer to use adversarial ML techniques. Furthermore, as mentioned in earlier sections, we have addressed the trustworthiness of sensing in our works extensively in D5.2 [43, 45–48]. With respect to localization privacy, we propose the use of Channel Charting (CC) jointly with Differential Privacy (DP) to induce positioning privacy. CC is an unsupervised learning technique that maps high-dimensional CSI into a low-dimensional representation while preserving local spatial relationships, enabling location-based services without explicitly revealing true user positions. The quality of this representation is evaluated through the *continuity* and *trustworthiness* metrics [49], which respectively measure the preservation of neighborhood relationships and the avoidance of false proximities. However, despite its apparent privacy advantages, CC remains vulnerable to inversion attacks that could reconstruct actual locations, motivating the integration of DP. DP is introduced as a formal framework that ensures the output of a mechanism does not significantly depend on any single user's data by injecting calibrated noise based on sensitivity and privacy parameters ϵ and δ. The proposed approach, termed channel charting with differential privacy (CCDP), applies noise to carefully extracted and normalized CSI features before dimensionality reduction, with sensitivity bounded through clipping, thereby preserving local geometry while ensuring privacy guarantees. The work further extends to geo-indistinguishability, which adapts DP to spatial data by enforcing that closer locations are more indistinguishable than distant ones, enabling distance-aware privacy protection. Finally, different learning architectures (e.g., triplet loss, Siamese networks, PCA, and autoencoders) are evaluated, demonstrating a trade-off between privacy and utility: increasing noise improves privacy but degrades chart quality, as reflected in reduced continuity and trustworthiness, highlighting the fundamental balance between location privacy and representation fidelity in wireless systems. This work will be submitted in an upcoming paper.

Bibliography

[1] T. Damm, S. Freud, and D. Klein, "Dissecting the CHES 2018 AES challenge," *Cryptology ePrint Archive*, 2019.

[2] "Deliverable D6.1: Use case validation plan and testbed design," Horizon Europe Project ROBUST-6G, Tech. Rep. Grant Agreement No. 101139068, 2025. [Online]. Available: https://robust-6g.eu/wp-content/uploads/2025/01/ROBUST-6G-D6.1_v1.0.1-1.pdf

[3] A. van den Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," 2019. [Online]. Available: https://arxiv.org/abs/1807.03748

[4] "Study on channel model for frequencies from 0.5 to 100 GHz," 3GPP, Tech. Rep. TR 38.901, Mar. 2022, available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/.

[5] T. M. Pham, L. Senigagliesi, M. Baldi, R. F. Schaefer, G. P. Fettweis, and A. Chorti, "Leveraging angle of arrival estimation against impersonation attacks in physical layer authentication," *IEEE Transactions on Information Forensics and Security*, vol. 21, pp. 3226–3239, 2026.

[6] "Ultra-dense indoor MaMIMO CSI dataset," IEEE DataPort, 2023, available online: https://ieee-dataport.org/open-access/ultra-dense-indoor-mamimo-csi-dataset.

[7] A. Mayya, M. Mitev, A. Chorti, and G. Fettweis, "A SKG security challenge: Indoor SKG under an on-the-shoulder eavesdropping attack," in *Proc. GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2023, pp. 3451–3456.

[8] M. Shakiba-Herfeh and A. Chorti, "Comparison of short blocklength Slepian-Wolf coding for key reconciliation," in *2021 IEEE Statistical Signal Processing Workshop (SSP)*. Rio de Janeiro, Brazil: IEEE, Jul. 2021, pp. 111–115.

[9] A. K. A. Passah, R. C. de Lamare, and A. Chorti, "Channel state information preprocessing for CSI-based physical-layer authentication using reconciliation," 2026, under review in *IEEE Transactions on Signal Processing*.

[10] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3221–3227, Nov 2012.

[11] R. M. Oliveira and R. C. de Lamare, "Polar codes based on piecewise gaussian approximation: Design and analysis," *IEEE Access*, vol. 10, pp. 73 571–73 582, 2022.

[12] A. Mayya, L. Senigagliesi, and A. Chorti, "Theoretical and practical analysis of secret key rates based on design parameters and channel characteristics," *Submitted to IEEE IoT Journal*, 2026.

[13] J. Zhang, G. Shen, W. Saad, and K. Chowdhury, "Radio frequency fingerprint identification for device authentication in the Internet of Things," *IEEE Communications Magazine*, vol. 61, no. 10, pp. 110–115, 2023.

[14] L. Xie, L. Peng, J. Zhang *et al.*, "Radio frequency fingerprint identification for Internet of Things: A survey," *Security and Safety*, vol. 3, 2024.

[15] C. Ayyildiz, F. E. Yildiz, O. Ayyildiz, and V. C. Yildirim, "RF fingerprinting migration dataset," Zenodo, 2025, ROBUST-6G project, Version v2, CC-BY-4.0. [Online]. Available: https://doi.org/10.5281/zenodo.14801935

[16] "Deliverable D5.1: Library of known PHY attacks and PLS dataset," Horizon Europe Project ROBUST-6G, Tech. Rep. Grant Agreement No. 101139068, 2024. [Online]. Available: https://robust-6g.eu/wp-content/uploads/2025/01/ROBUST-6G-D5.1_v1.0.pdf

[17] S. AlHazbi, S. Sciancalepore, and G. Oligeri, "The day-after-tomorrow: On the performance of radio fingerprinting over time," in *Proc. Annual Computer Security Applications Conference (ACSAC '23)*. ACM, December 2023.

[18] A. Elmaghbub and B. Hamdaoui, "No blind spots: On the resiliency of device fingerprints to hardware warm-up through sequential transfer learning," in *Proc. 17th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '24)*. ACM, May 2024.

[19] C. Ayyıldız, F. E. Yıldız, V. C. Yıldırım, and D. Çakmak, "RFFI-Temporal: A long-term RF fingerprinting dataset for temporal drift analysis," Zenodo, 2026, version 1.0.0. [Online]. Available: https://doi.org/10.5281/zenodo.18952487

[20] M. Piana, G. A. Alghisi, A. V. Guglielmi, G. Perin, F. Gringoli, and S. Tomasin, "BRISC: A dataset of channel measurements at 5 GHz with a reflective intelligent surface," 2026. [Online]. Available: https://arxiv.org/abs/2602.21102

[21] A. Chorti, S. Skaperas, M. Delamou, L. Chen, L. Senigagliesi, P. Giardina, C. Ayyildiz, S. Tomasin, S. Berri, and L. Luzzi, "Automating trust: Closed-loop physical layer security for 6G RAN," in preparation.

[22] S. Skaperas and A. Chorti, "On the robustness of AoA as an authentication feature under spoofing: Fundamental limits from misspecified Cramer Rao theory," *in arXiv:2603.21219 and under review in IEEE Wireless Communications Letters*.

[23] M. Srinivasan, L. Senigagliesi, H. Chen, A. Chorti, M. Baldi, and H. Wymeersch, "AoA-based physical layer authentication in analog arrays under impersonation attacks," in *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 496–500.

[24] B. Trinh-Nguyen, S. Berri, S. G. Teo, T. Truong-Huu, and A. Chorti, "High-accuracy AoA-based localization using hierarchical ML classifiers in outdoor environments," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Taipei, TW, Dec. 2025.

[25] L. Senigagliesi, A. V. Guglielmi, M. Baldi, and S. Tomasin, "Security analysis of RIS-assisted physical-layer authentication over multipath channels," in *Proc. IEEE 17th IEEE International Workshop on Information Forensics and Security (WIFS)*, Perth, Australia, Dec. 2025.

[26] A. Mayya, A. Chorti, R. F. Schaefer, and G. P. Fettweis, "Secret key generation rates for line of sight multipath channels in the presence of eavesdroppers," in *Proc. 27th International Workshop on Smart Antennas (WSA)*. IEEE, 2024.

[27] A. Mayya, Y. Richhariya, A. K. Boroujeni, S. Vorberg, M. Matthé, R. Vinz, L. Senigagliesi, K. Klamka, and A. Chorti, "Context-aware secret key generation demonstrator based on physical layer security," in *Proc. 2025 IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2025.

[28] "Deliverable D6.2: Intermediate validation results," Horizon Europe Project ROBUST-6G, Tech. Rep. Grant Agreement No. 101139068, 2025.

[29] "Deliverable D2.3: Final ROBUST-6G architecture and ROBUST-6G dataspace," Horizon Europe Project ROBUST-6G, Tech. Rep. Grant Agreement No. 101139068, 2026.

[30] S. Tomasin, T. N. M. M. Elwakeel, A. V. Guglielmi, R. Maes, N. Noels, and M. Moeneclaey, "Analysis of challenge-response authentication with reconfigurable intelligent surfaces," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 9494–9507, 2024.

[31] L. Crosara, A. V. Guglielmi, N. Laurenti, and S. Tomasin, "Divergence-minimizing attack against challenge-response authentication with IRSs," pp. 1986–1991, 2024.

[32] F. Ardizzon, D. Salvaterra, M. Piana, and S. Tomasin, "Energy-based optimization of physical-layer challenge-response authentication with drones," in *Proc. 2024 IEEE Globecom Workshops (GC Wkshps)*, 2024, pp. 1–6.

[33] M. Piana, F. Ardizzon, and S. Tomasin, "Challenge-response to authenticate drone communications: A game theoretic approach," *IEEE Transactions on Information Forensics and Security*, vol. 20, pp. 4890–4903, 2025.

[34] "Deliverable D5.2: Report on the use of PLS in 6G," Horizon Europe Project ROBUST-6G, Tech. Rep. Grant Agreement No. 101139068, 2024.

[35] A. Hossary, L. Crosara, and S. Tomasin, "Jamming detection in cell-free MIMO with dynamic graphs," in *2025 IEEE 36th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2025, pp. 1–6.

[36] M. Varotto, S. Valentin, and S. Tomasin, "Detecting 5G signal jammers using spectrograms with supervised and unsupervised learning," in *Proc. IEEE Int. Conf. on Commun. Work. (ICC Work.)*, 2024, pp. 767–772.

[37] M. Varotto, S. Valentin, F. Ardizzon, S. Marzotto, and S. Tomasin, "One-class classification as GLRT for jamming detection in private 5G networks," in *Proc. 2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 201–205.

[38] L. Luzzi and V. Bioglio, "Finite-blocklength performance of polar wiretap codes under a total variation secrecy constraint," in *IEEE International Symposium on Information Theory (ISIT)*, 2026.

[39] A. K. Angélo Passah, R. C. De Lamare, and A. Chorti, "Physical layer authentication using information reconciliation," in *Proc. 2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, 2024, pp. 1–5.

[40] A. Kokuvi Angélo Passah, A. Chorti, and R. C. de Lamare, "Enhanced multiuser CSI-based physical layer authentication based on information reconciliation," *IEEE Wireless Communications Letters*, vol. 14, no. 2, pp. 544–548, 2025.

[41] A. K. Angélo Passah, R. C. De Lamare, and A. Chorti, "Adaptive CSI preprocessing for physical layer authentication," in *Proc. 2026 IEEE Wireless Communications and Networking Conference (WCNC)*, under review.

[42] M. Piana and S. Tomasin, "Secret key generation on aerial rician fading channels against a curious receiver," in *Proc. 2025 IEEE 26th International Workshop on Signal Processing and Artificial Intelligence for Wireless Communications (SPAWC)*, 2025, pp. 1–5.

[43] D. Wang, L. Chen, and F. Nait-Abdesselam, "SA-SWOMP: Radar-assisted sparse channel estimation for joint sensing and communication," in *Proceedings of IEEE International Conference on Communications, ICC 2025*, 2025.

[44] C. Ayyıldız, F. E. Yıldız, and B. E. Süzek, "Data-efficient domain adaptation for receiver-invariant radio frequency fingerprinting identification," in *Proc. IEEE European Conference on Networks and Communications (EuCNC) & 6G Summit*, 2026, accepted for publication.

[45] S. Gil, M. Yemini, A. Chorti, A. Nedić, H. V. Poor, and A. J. Goldsmith, "How physicality enables cy-trust: A new era of trust-centered cyber–physical systems," *Proceedings of the IEEE*, vol. 113, no. 10, pp. 1121–1154, 2025.

[46] B. Trinh-Nguyen, S. Berri, S. G. Teo, T. Truong-Huu, and A. Chorti, "A framework for global trust and reputation management in 6G networks: Position paper," in *Machine Learning for Networking: 7th International Conference, MLN 2024, ACM Digital Library*, 2024.

[47] M. Delamou, L. Chen, E. M. Amhoud, and A. Chorti, "Enhanced physical layer authentication via robust and trustworthy sensing," in *Proceedings of IEEE International Conference on Communications, ICC 2026*, Glasgow, UK, 2026.

[48] R. Khanzadeh, F. Ademaj-Berisha, S. Berri, L. Senigagliesi, A. Chorti, A. Springer, and H.-P. Bernhard, "Trustworthiness-aware resource allocation in network slicing via hierarchical reinforcement learning," in *Proceedings of IEEE International Conference on Communications, ICC 2026*, Glasgow, UK, 2026.

[49] L. Ribeiro, M. Leinonen, H. Al-Tous, O. Tirkkonen, and M. Juntti, "Channel charting aided pilot reuse for massive MIMO systems with spatially correlated channels," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 2390–2406, 2022.