PROJECT “LOCUS”: LOCalization and analytics on-demand embedded in the 5G ecosystem, for Ubiquitous vertical applicationS

Grant Agreement Number: 871249
(https://www.locus-project.eu/)

DELIVERABLE D2.2

“Security and Privacy, preliminary version”

Deliverable Type:	R
Dissemination Level:	Public
Contractual Date of Delivery to the EU:	30/06/2020
Actual Date of Delivery to the EU:	30/06/2020
WP contributing to the Deliverable:	WP2 – Use cases, requirements, security, system architecture
Editor(s):	CNIT: Domenico Garlisi, Nicola Blefari Melazzi
Author(s):	CNIT: Stefania Bartoletti, Domenico Garlisi, Ivan Palamà, Francesco Mancini and Danilo Orlando Ericsson AB: Sara Modarres Razavi IMDEA: Domenico Giustiniano
Internal Reviewer(s):	Incelligent: Athina Ropodi

	Ericsson AB: Sara Modarres Razavi
Short Abstract:	The goal of this deliverable is to present a study of location/privacy issues as well as preliminary solutions to mitigate security attacks and ensure data privacy for the LOCUS system. Preliminary results on privacy attacks are also presented.
Keyword List:	Localization, 5G, security, privacy, IMSI catcher, jamming detector, integrity.

Executive Summary

Improved wireless connectivity and location awareness create enormous opportunities for new Location Based Services (LBS). These capabilities have a great potential for the operators, that can exploit them to offer new useful services to their customers. Examples of LBS include industry 4.0, intelligent transport systems, IoT applications, network planning and operation and maintenance support, flow monitoring and tracking systems, location-based advertising, social networking services. However, significant privacy and security concerns are raised by LBS.

The present document, D2.2, is the preliminary deliverable for the security and privacy aspects of the LOCUS project. LOCUS pays the highest attention to both location security and privacy to make sure that the technical solutions will process data in compliance with users' privacy rights. This Deliverable aims at addressing security and privacy as two critical and intertwined aspects of LOCUS, after the definition of the use cases (T2.1) and as overarching requirements for the LOCUS architecture (T2.3). T2.2 provides a robust and secure approach for data sharing and provision by users, services and internal LOCUS components with respect to various aspects, such as: endpoint and network security; data protection; application security; architecture security system resilience. T2.2 identifies security and privacy challenges that apply to objects, systems and networks and determine how to best address these challenges in LOCUS use cases and services.

VERSION CONTROL TABLE			
VERSION N.	**PURPOSE/CHANGES**	**AUTHOR (S)**	**DATE**
1.0	First draft and document structure	Domenico Garlisi, Bartoletti Stefania	11/05/2020
1.1	Add inputs contribute for privacy section	Domenico Garlisi, Ivan Palamà	18/05/2020
1.2	Add inputs contribute for security section	Danilo Orlando, Stefania Bartoletti	22/05/2020
1.3	Add inputs contribute for experimental results	Ivan Palamà, Francesco Mancini	03/06/2020
1.4	Add input contributes for security section	Domenico Giustiniano, Sara Modarres Razavi	15/06/2020
2.0	Add common parts document	Domenico Garlisi, Bartoletti Stefania	17/06/2020

2.1	Add input for location privacy and location security	Ioannis Chochliouros, Maria Belesioti	22/06/2020
2.1	Internal review	Athina Ropodi, Sara Modarres Razavi	25/06/2020
3.0	Complete version	Danilo Orlando, Stefania Bartoletti	27/06/2020
3.1	Final revision	Nicola Blefari Melazzi	30/6/2020

INDEX

1 Background

1.1 List of Abbreviations

ABBREVIATION	FULL NAME
3GPP	3rd Generation Partnership Project
AKA	Authenticated Key Agreement
AoA	Angle of Arrival
AN	Access Node
ARPF	Authentication Credential Repository and Processing Function
AV	Authentication Vector
AUSF	Authentication Server Function
BJ	Barrage Jammer
BS	Base Station
DoA	Direction of Arrival
DoS	Denial of Service
EM	Expectation Maximization
GDPR	General Data Protection Regulation
GLRT	Generalized Likelihood Ratio Test
GNSS	Global Navigation Satellite System
GUTA	Globally Unique Temporary User Equipment Identity
HN	Home Network
HSS	Home Subscriber Server
IMSI	Mobile Subscriber Identity
LBS	Location Based Service
MCC	Mobile Country Code
MIMO	Multiple Input Multiple Output
MNC	Mobile Network Code
MME	Mobility Management Entity
MS	Mobile Station
MSIN	Mobile Subscriber Identification Number

NR	New Radio
OAI	Open Air Interface
RRC	Radio Resource Control
RSS	Received Signal Strength
SBA	Service-based architecture
SDR	Software Defined Radio
SEAF	Security Anchor Function
SIDF	Subscription Identifier De-concealing Function
SN	Serving Network
SUPI	Permanent Identifier
SUCI	Concealed Identifier
TBD	Track Before Detect
TDoA	Time Difference of Arrival
ToA	Time of Arrival
TBS	Terrestrial Beacon Systems
TMSI	Temporary Mobile Subscriber Identity
UE	User Equipment
UDM	Unified data management
USRP	Universal Software Radio Peripheral

Table 1: Abbreviation List

2 Introduction

This Deliverable aims at exploring the security and privacy aspects supported by the LOCUS platform by relying also on the LOCUS architecture presented in Deliverable D2.1. Figure 1 shows the LOCUS architecture, where the security and privacy module is highlighted. The module (violet box in the figure) is positioned in between the data management and the location function modules. Security and privacy functions can be accessed by different modules of the LOCUS platform through the specific function interfaces.

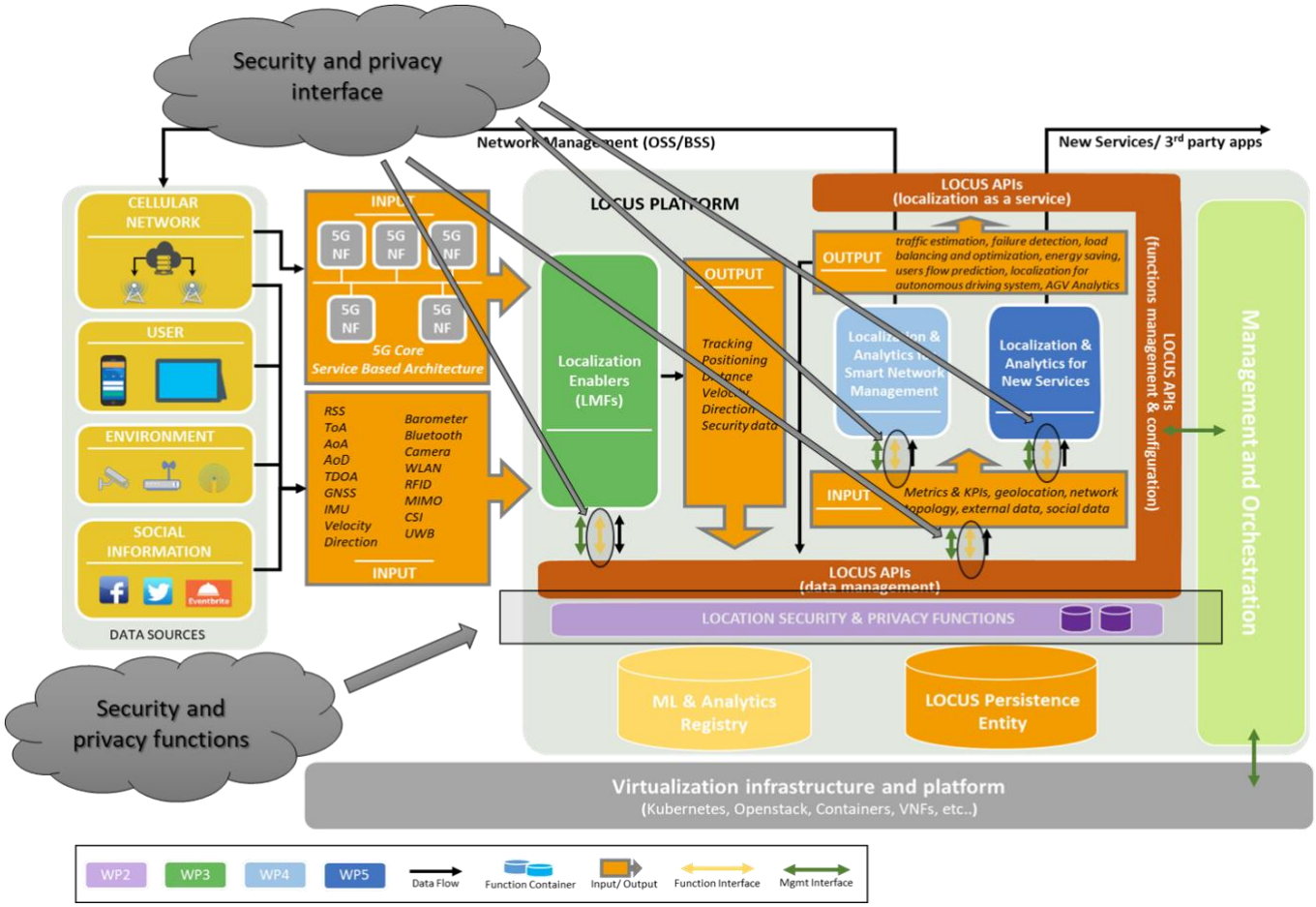

Figure 1 – Position of the security and privacy layer in the LOCUS platform

This Deliverable includes two main sections, after the introduction: Section 3 provides a complete study on the privacy issues in LBS, while Section 4 addresses the location security issues. Specifically, the privacy concern in LBS are: i) data authenticity, the reported location data from a user must be real and should not have been forged or modified by anybody, ii) privacy, the inability for an attacker to determine the presence, or track users in a specific area. Starting from these two key requirements, the authentication process is one of the base blocks of privacy measures in mobility technologies. An analysis of the authentication process in 4G and 5G cellular networks is provided in Section 3. From the analysis, we notice several weaknesses that only in part are addressed by the new 5G specifications. In particular, we

present the IMSI catcher attack, whose goal is to collect the identification code of the user terminal. We use an experimental set-up to implement a demonstrative IMSI catcher that exploits a Software Defined Radio (SDR) platform. We use this set-up to evaluate how different user terminals are robust with respect to the IMSI catcher attack. Finally, we present a preliminary study to improve 5G authentication issues.

Section 4 addresses the location security aspects of the LOCUS project. First, an overview of the countermeasures to face location security attacks in 4G and 5G networks is provided by considering both noise-like jammers as well as meaconing and spoofing attacks. Then, the main mitigation techniques that will be investigated within LOCUS are presented. The proposed techniques can act at two different levels: 1) raw data-based, i.e., they process directly the received waveforms when available; and 2) measurement-based, namely, they are fed by the range, angle, and signal strength measurements. Then, anti-spoofing techniques for non-3GPP signals are also investigated. Finally, the current work of the 3GPP standard related to positioning integrity is presented.

3 Location privacy

LBSs make users' life easier and more enjoyable but bring about several privacy issues. In some cases, individuals may be unaware of the potential risks implied by the use of these services and of who is being permitted access to their location information.

It is also known that telecommunication operators use location/position of their users both to optimize operations and to drive operational business opportunities [1]. For example, in 2014, Verizon built data centers in California to implement precision marketing with location data [2]. Thus, privacy has become a primary concern for LBSs.

In this section, we first provide an overview on the privacy implication for LBSs and cellular network, and afterwards, we propose possible solutions that can be applied to reduce system vulnerability.

Privacy concerns are legally addressed in the General Data Protection Regulation (GDPR); Article 25 state requirements for "data privacy by design" and "data privacy by default":

1. **Data privacy by design** means that appropriate organizational and technical measures to ensure personal data security and privacy must be embedded into the complete lifecycle of an organization's products, services, applications, and business and technical procedures. Technical measures can include, but are not limited to, pseudonymization and data minimization.
2. **Data privacy by default** means that (a) only necessary personal data is collected, stored, or processed and (b) personal data is not accessible to an indefinite number of people.

Following this provision, we consider aspects related to privacy in the complete lifecycle of the LOCUS platform, especially in the design part. Moreover, we consider the implementation of specific functionalities that reduce the amount of user personal data stored and prevents unauthorized access to the LOCUS persistence entity (repository where the data are stored).

Privacy issues arise both in the acquisition of data, and in the subsequent release from the repositories where the data is stored. If we consider that an adversary may be able to obtain one or more transactions at the time they are acquired, a defence technique must be applied at each transaction, as proposed in the case of requests to LBS. We call these techniques online privacy preserving techniques, as opposed to offline techniques that consider the privacy violations due to the release of a dataset of transactions acquired in a given time window. **There is a privacy threat whenever an adversary can associate the identity of a user to information that the user considers private** [3]. In the case of LBS, this sensitive association can be possibly derived from location-based requests issued to service providers. The identity and the private information of a single user can be derived also from requests issued by a group of users as well as from available background knowledge.

Privacy in general deals with the protection of information which may reveal or may hint at any details of personal information/activities about a specific user. Now, since the existing 4G/4G+ location technology and functionality will be significantly improved in 5G, it turns out that risk of privacy violation will be increased too. Otherwise stated, privacy in 5G networks will be an issue of primary importance because 5G will yield a great transformation in terms of new applications and access modes to digital services. Additionally, several new devices and gadgets will have positioning and tracking capabilities (location based services) [4] [5].

The knowledge of the specific cell tower or antenna with which communication is in progress, can disclose information about the end-user exact location. Each time an end-user connects to a cellular antenna, then mobile networks can identify that user location and can even determine the building where the user is. Threats such as semantic information attacks [6] (the use of incorrect information to cause harm) often target the location data of users. Security and privacy vulnerability are exhibited also by access point selection algorithms [7].

The 4G network technology has a broader coverage area than 5G and the signal is broadcast by a cellular tower over a wider angular sector with respect to 5G. As a matter of fact, 5G networks typically have a smaller coverage area due to the use of narrow beams and high frequencies. Thus, 5G allows for more accurate location and tracking of users even without deploying ad hoc and more sophisticated solutions and algorithms.

Location-based privacy requires anonymity-based techniques and systems where the users' true identity can be hidden. Messages should also be encrypted before their transmission to a location-based service provider. Obfuscation techniques where the quality of location information is reduced can also be used to protect location privacy.

For this reason, authentication is one the most important functionality for LBS. A possible mapping between databases and LBS requests or, more generally, location data acquisition consists in considering all the requests/acquisitions received in a given time slot as the tuples of a table, where each tuple has a respondent and a location as the value of a spatial attribute. Over the time, this mapping leads to a sequence of tables timestamped with a given time granule. Hence, the privacy problem in location data management is to release versions of these tables so that no sensitive associations can be derived. Finally, a privacy attack is a specific method used by an adversary to obtain the sensitive association. Figure 2 shows a graphical representation of this general privacy threat in LBS, where an adversary attack has been performed to bind user identity and private information.

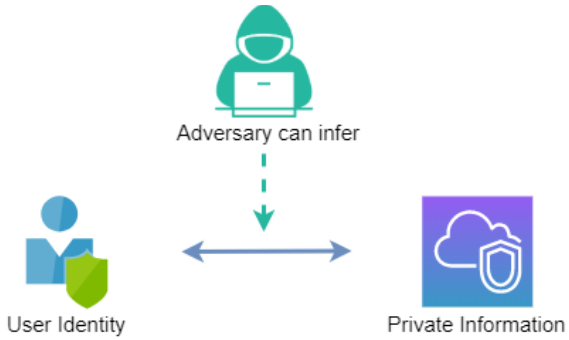

Figure 2 - General privacy threat in LBS

In the next subsections, we will present how the privacy aspects are addressed from cellular networks. We will present a study focused on the vulnerability of the authentication process in 4G and 5G network, and we propose possible solutions to reduce the weakness. **LOCUS activities focus on attacks concerning the acquisition of a cellular network user identifier. Experimental validation based on Software Defined Radio (SDR) is also performed.**

3.1 Identification and authentication in cellular network

In cellular networks, most devices rely on ubiquitous LBSs. A LBS uses location data, which are related to the smartphone and/or mobile device in order to deliver services to the users. Consequently, privacy attacks on the cellular network may lead to privacy issues for the user location. Indeed, from the user point of view, location data transmitted over the network are highly sensitive. If these data are revealed, the location, the trajectory, and the identity of the users can be leaked.

To this end, authentication is one of the building blocks of privacy measures in cellular networks. Authentication process has been a primary requirement in cellular networks since the old GSM (2G) and 5G has further invested in it by standardizing a novel public key-based approach to conceal the subscriber identity. Furthermore, the 3rd Generation Partnership Project (3GPP) has identified the following essential requirements related to user privacy [8]:

1. User **Identity Confidentiality**: The permanent identity of a user to whom a service is delivered cannot be eavesdropped on the radio access link.
2. User **Location Confidentiality**: The presence or the arrival of a user in a certain area cannot be determined by eavesdropping on the radio access link.
3. User **Untraceability**: An intruder cannot deduce whether different services are delivered to the same user by eavesdropping on the radio access link.

This section presents a comprehensive review of the core and enabling technologies that are used to build the 3GPP authentication model for cellular networks.

From the authentication perspective, a cellular network consists of three main components: a User Equipment (UE), a Serving Network (SN), and a Home Network (HN) (Figure 3).

Figure 3 - Cellular Network Architecture

In mobile telephony systems**, networks allocate to each SIM card a unique identifier**, referred to as International Mobile Subscriber Identity (**IMSI**) in **4G** and Subscription Permanent Identifier (**SUPI**) in **5G**. As the authentication between a user and its service provider is based on a shared symmetric key, it can only take place after user identification. The IMSI is a unique number residing in a SIM card of, usually, 15-digit.

The 3GPP defines an Authentication and Key Agreement (AKA) protocol as well as procedures that support entity authentication, message integrity, and message confidentiality in addition to other security properties [9]. A serving network in 4G consists of radio access equipment such as an Evolved NodeB (eNodeB) base station and a Mobility Management Entities (MMEs), among others. The UE communicates with a serving network through radio interfaces. In 4G, a home network usually consists of authentication servers as the Home Subscriber Server (HSS), which stores user credentials and authenticates users.

The 3GPP AKA protocol is a challenge-and-response authentication protocol based on a symmetric key shared between a subscriber and a home network. After the mutual authentication between a subscriber and a home network, cryptographic keying materials are generated to protect subsequent communications between a subscriber and a serving network, including both signalling messages and user plane data.

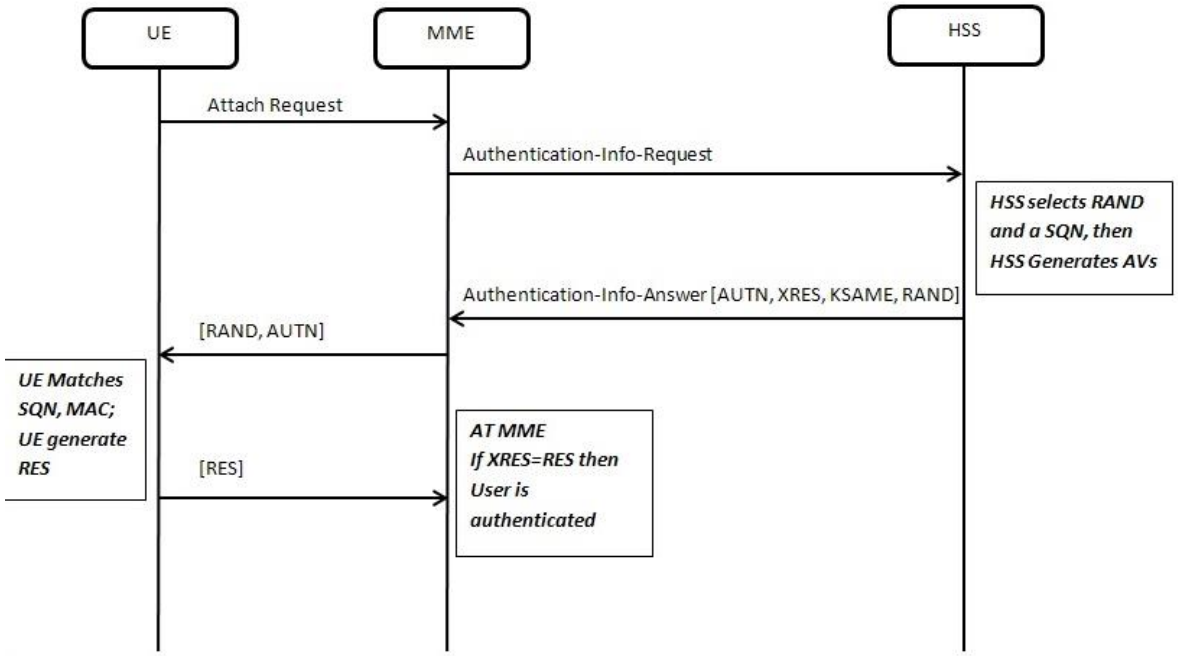

Figure 4 - Authentication process in 4G cellular network

Figure 4 shows the complete authentication process. The EPS-AKA is triggered after that the UE completes the Radio Resource Control (RRC) procedure with the eNodeB and sends an Attach Request message to the MME. The MME, in turn, sends an Authentication Request, including UE identity (e.g., IMSI) and the serving network identifier, to the HSS located in the home network. The HSS performs cryptographic operations based upon the shared secret key, Ki (shared with the UE), to derive one or more authentication vectors (AVs), which are sent back to the MME in an Authentication Response message. An AV mainly consists of an authentication (AUTH) token and an expected authentication response (XAUTH) token.

After receiving an Authentication Response message from the HSS, the MME sends an Authentication Request to the UE, including the AUTH token. The UE validates the AUTH token by comparing it to a generated token based on Ki. If the validation succeeds, the UE considers the network to be legitimate and sends an Authentication Response message back to the MME, including a response (RES) token, which is also generated based on Ki.

The MME compares the RES token with an expected response (XRES) token. If they are equal, the MME performs key derivation and sends a Security Mode Command message to the UE, which then computes the corresponding keys aimed at protecting subsequent NAS signalling messages. The MME will also send the eNodeB a key from which the RRC protection keys are derived. Once the corresponding keys are available at the UE, the subsequent communication between the UE and the eNodeB is protected.

From the analysis of 4G EPS-AKA two vulnerability aspects arise:

1. the UE identity is sent over radio networks without encryption. More importantly, the UE permanent identity may be sent in clear text in an Identity Response message when responding to an Identity Request message from a network.
2. a home network provides AVs when consulted by a serving network during UE authentication, but it is not a part of the authentication decision. Such a decision is solely made by the serving network.

If the IMSI/SUPI values are sent in plaintext over the radio access link, then users can be identified, located, and tracked using these permanent identifiers.

To avoid this privacy breach, temporary identifiers are assigned to the SIM card by the visited network. These identifiers are called Temporary Mobile Subscriber Identity (TMSI) in 3G systems and Globally Unique Temporary User Equipment Identity (GUTI) in both 4G and 5G systems. Although a temporary identifier may be used to hide a subscriber's long-term identity, researchers have shown that GUTI allocation is flawed: GUTIs are not changed as frequently as necessary [10], and GUTI allocation is predictable (e.g., with fixed bytes) [11].

Moreover, **there exist certain situations where authentication using temporary identifiers is not possible**. For instance, when a user registers with a network for the first time and is not yet assigned a temporary identifier. Another case occurs when the visited network is unable to resolve the IMSI/SUPI from the presented TMSI/GUTI.

An active man-in-the-middle adversary can intentionally simulate this scenario to force an unsuspecting user to reveal its long-term identity. These attacks are known as **IMSI catching** attacks [12] and persist in today mobile networks including the 4G LTE/LTE+ [13].

3.2 IMSI catching attack

IMSI Catchers are active attackers in cellular networks with the main goal of collecting IMSIs. The malicious devices usually act as a man-in-the-middle to perform more advanced attacks. IMSI Catchers can be used for (a) mass-surveillance of individuals in a geographical area, (b) linking a real person to his identity in the network, (c) tracing a person with a known IMSI, or (d) checking his presence in a building or area. The attack is based on the fact that the UE will use the IMSI as identifier when the TMSI is no longer available. This may happen when the TMSI is deleted by MME or UE due to timeout. The IMSI catching attack can be carried out in a passive or active way. A passive IMSI-catcher eavesdrops on the wireless traffic in its neighbourhood and collects all IMSIs captured.

A faster and active way to catch IMSIs exploits a "fake" base station which acts as a preferred base station in terms of signal strength. Mobile devices typically select base stations emitting the strongest signal. This fake base station can then be used to send an Identity Request

message to all mobile devices in the area, which will respond with their IMSIs since they assume that they are connected to a legitimate network which has lost access to the TMSI. In this way, IMSIs of all mobile device in the area can rapidly be captured.
The covered area varies according to the type of IMSI catcher and targeted network standard. For example, a semi-passive attack presented in [14] can locate an LTE compliant UE within a 2 km^2 area in an urban setting.
Based upon the way the IMSI catcher works, three categories of attacks can be listed:

1. **Passive:** a completely passive device sniffing packets from several frequency channels (i.e., capturing information without intruding in the communication), decoding, and possibly decrypting the captured information.
2. **Semi-active:** IMSI binders are semi-active and are the most common. Compared with passive, they are more complex devices that, under the false role of legitimate cells, attract victims by forcing the latter to establish a connection with the false node. This connection is released after obtaining the desired data.
3. **Active**: A fully active IMSI catcher is a refinement accessory of the previous concept. In this case, the device does not reject authentication, but maintains the connection with the victims also by redirecting packets from the victim to the legitimate network. Actually, these devices realize a real Man-in-the-Middle attack since they manage the traffic sent from mobile devices to the rest of the network.

IMSI Catchers were first built for 2G/GSM, and later extended to most recently protocols. The vulnerabilities in the recent protocols allow an adversary to trace the location of users with finer granularity, to enable DoS attacks, or to eavesdrop on the communication [15], [16]. General techniques to set up attacks against LTE include traffic capture, jamming, and downgrading to 2G. Academic works implement IMSI Catchers by modifying the code of open-source software projects, such as OpenLTE [17], srsLTE [18] or gr-LTE [19]. More recently, Rupprecht et al. have used Open Air Interface (OAI) [20] to check for the compliance of UE with the LTE standard encryption and to exemplify a man-in-the-middle attack. Most of existing works use OpenLTE code because managing and modifying the OpenLTE architecture is not difficult.
In this project, we use the OAI software in order to build an IMSI Catcher. We chose OAI because it supports in part the 5G architecture; a more detail description of the OAI framework is provided in Subsection 3.4.1.

3.3 Enhanced privacy protection in 5G networks

The Service-based architecture (SBA) has been proposed for the 5G core network in [21]. Some of the new entities relevant to 5G authentication are listed below.

- The Security Anchor Function (SEAF) is in a serving network and acts as a “middleman” during the authentication process between a UE and its home network. It can reject

an authentication from the UE, but it relies on the UE home network to accept the authentication.

- The Authentication Server Function (AUSF) is in a home network and performs authentication with a UE. It makes the decision for UE authentication, but it relies on a backend service for computing the authentication data and keying materials when 5G-AKA or EAP-AKA is used.
- Unified data management (UDM) is an entity that hosts functions related to data management, such as the Authentication Credential Repository and Processing Function (ARPF), which selects an authentication method based on subscriber identity as well as configured policy and computes the authentication data and keying materials for the AUSF when required.
- The Subscription Identifier De-concealing Function (SIDF) decrypts a Subscription Concealed Identifier (SUCI) to obtain its long-term identity, namely the Subscription Permanent Identifier (SUPI), e.g., the IMSI.

The SEAF may start the authentication procedure after receiving any signalling message from the UE. Note that the UE should send the SEAF a temporary identifier (a 5G-GUTI) or an encrypted permanent identifier (a SUCI) if a 5G-GUTI has not been allocated by the serving network for the UE. The SUCI is the encrypted form of the SUPI using the public key of the home network. Thus, a UE permanent identifier, e.g., the IMSI, is never sent in clear text over the radio networks in 5G. This feature is considered a major security improvement over prior generations such as 4G.

An Elliptic Curve Integrated Encryption Scheme (ECIES)-based privacy-preserving identifier containing the concealed SUPI is transmitted. A SUPI as defined in [21] is usually a string of 15 decimal digits. The first three digits represent the Mobile Country Code (MCC) while the next two or three form the Mobile Network Code (MNC) identifying the network operator. The remaining (nine or ten) digits are known as Mobile Subscriber Identification Number (MSIN) and represent the individual user of that particular operator.

The 3GPP standard for 5G security specifies two AKA protocols, Extensible Authentication Protocol AKA' (EAP-AKA') and 5G-AKA.

In 5G systems, SUCI is a privacy preserving identifier that contains the concealed SUPI. The UE generates a SUCI using a protection scheme with the public key of the HN that was securely provisioned to the USIM during the USIM registration. Only the MSIN part of the SUPI is concealed by the protection scheme while the home network identifier (MCC/MNC) is transmitted in plaintext.

The differences between 5G-AKA and 4G EPS-AKA are listed below.

- Entities involved in the authentication are different because of the new service-based architecture in 5G. Particularly, the SIDF is new and it does not exist in 4G.
- The UE always uses the public key of the home network to encrypt the UE permanent identity before it is sent to a 5G network. In 4G, the UE always sends its permanent identifier in clear text to the network, allowing it to be stolen by either a malicious network (e.g., a faked base station) or a passive adversary over the radio links (when communication over radio links is not protected).
- The home network (e.g., the AUSF) makes the final decision on UE authentication in 5G. In addition, results of UE authentication are also sent to UDM in order to create a log file. In 4G, a home network is consulted during the authentication only to generate authentication vectors; it does not make decisions on the authentication results.
- The 5G Key hierarchy is longer than that of 4G because 5G introduces two intermediate keys, KAUSF and KAMF (see Figure 5). Note that KSEAF is the anchor key in 5G equivalent to KASME in 4G.

Although the ECIES-based scheme is oblivious to the loss of the synchronization between the UE and HN and provides a robust key management, which leads to a significant reduction in connection failures, there are aspects that require further improvements and that represent vulnerability points for security and privacy attacks:

1. **Bidding Down Attacks**. An active adversary simulating a (false) base station can force the UE to use one of the previous generations (3G/4G) and then can catch the IMSI using an identity request message. Until all systems are not upgraded to 5G, nothing can be done against these bidding down attacks. In the current 5G security specifications [21], the SUPI is derived directly from the IMSI, so these bidding down attacks also compromise the SUPI.
2. **Chosen SUPI Attacks**. Any arbitrary third party can always select a SUPI of his choosing and send the corresponding SUCI to the HN. Thereafter, the adversary can look out for various responses from the HN, depending on whether or not the target user is present in that cell tower. Any noticeable variation in the perceived output would allow the adversary to either confirm or deny the presence of the target in that cell. There is no mechanism in the ECIES-based scheme to prevent this kind of attacks.
3. **Replay Attacks**. Note that the ECIES-based scheme does not have any inherent mechanism to ensure the information refresh to the HN and, hence, is susceptible to replay attacks. An adversary can always resend a previously encrypted SUPI to the HN and look out for various kinds of responses (an authentication challenge or a failure message). Based on the received response, a device whose SUPI is unknown to the attacker may be tracked with some confidence.
4. **Update of HN Public Key**. There could be situations requiring that the HN has a robust way to quickly update its public key. An example of this scenario could encompass a malware attack which tries to recover the home network private key. Such situations enforce the need to have a quick way of updating the corresponding public keys.

3.4 Experimental evaluation through a SDR platform

In this preliminary deliverable of the LOCUS project, we focus on IMSI catching attacks and we propose a possible solution to mitigate them. **We implement a 4G network through SDR and OAI, and we realize an IMSI catching to capture IMSI of legacy smartphones**. We test our approach on SDR Universal Software Radio Peripheral (USRP) devices from ETTUS research[1].

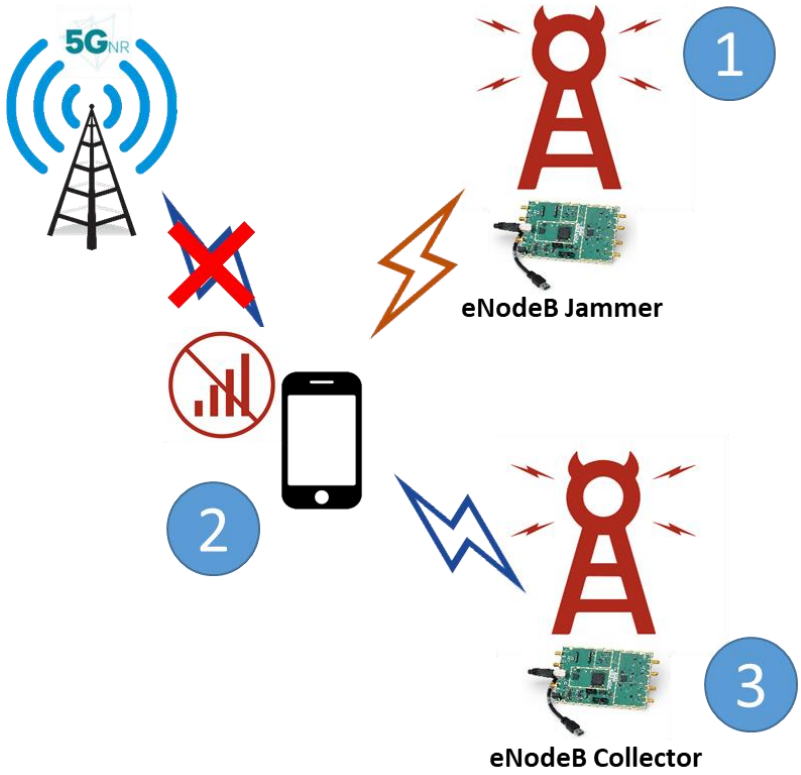

Figure 5 – Experimental set-up

Figure 5 shows the architectural structure of our set-up, it is composed of the following elements (numbered as shown in the Figure):

1. **LTE frequency hopping jammer** is used to jamming the operator eNodeB.
2. **UE victim** which represents the smartphone whose IMSI will be captured (as shown in the Figure, the smartphone loses the link connection after the activation of the jammer).
3. **eNodeB Collector** which is the malicious eNodeB that will retrieve the UE IMSI.

3.4.1 Open Air Interface (OAI) project

OAI is an open source project that implements both 3GPP radio access network and 3GPP core network and that run on general-purpose computing platforms together with Commercial

[1] http://www.ettus.com

Off-The-Shelf (COTS) SDR cards like the USRP. It allows to set up a compliant 4G/5G network and inter-operate with commercial equipment. The OAI team has been working to provide a standard compliant version of a 5G gNodeB that along with an updated version of the eNodeB will provide 5G Non-Standalone Access (NSA). Development of a 5G Core (5GC) network is also underway and in 2020 OAI is planning to support the Standalone Access (SA) mode of 5G. The current focus of OAI is on the development of a 5G NSA solution using the EN-DC architecture where the eNodeB handles all the CP and the gNodeB only needs to handle the user-plane traffic.

As for LTE, which supports channel bandwidths up to 20 MHz, the most commonly used platform is the USRP B210, even though different devices can be used in place of it. NR supports channel bandwidths up to 100 MHz in the sub-6GHz spectrum (called Frequency Range 1 - FR1 in 3GPP) but can also be configured for 80 MHz or 40 MHz wide channels. The usability of the device for NR mainly depends on the supported master clock rates and sampling rates, which have to be compatible with the ones specified by 3GPP.

The USRP B210 can be used for 40 MHz channel configuration with the limitation of 1 channel due to the limit of the Universal Serial Bus 3 (USB3) interface with the host. The X3x0 can be used for 80MHz channel configurations up to 2 channels. The N3x0 can be used for 100MHz bandwidth configuration up to 4 channels.

3.4.2 Experimental Setup

IMSI-catchers as StingRay [22] can be built up by using low-cost software defined radio. In 2010, Chris Paget has shown that it is possible to build a homemade active IMSI-catcher for about $1,500 using an SDR, two directional antennas, and a laptop running OpenBTS and Asterisk [23].

The experiments conducted here have been carried out in our wireless network security laboratory to avoid disturbing other normal UEs. The UE victim is maintained close to the IMSI Catcher system in each experiment in order to meet the radio signal power requirement of cell reselection procedure. An attacker needs to create a malicious network and needs to force the UE to connect to it. The IMSI Catcher system is therefore composed of two main components: the malicious network and the frequency LTE jammer.

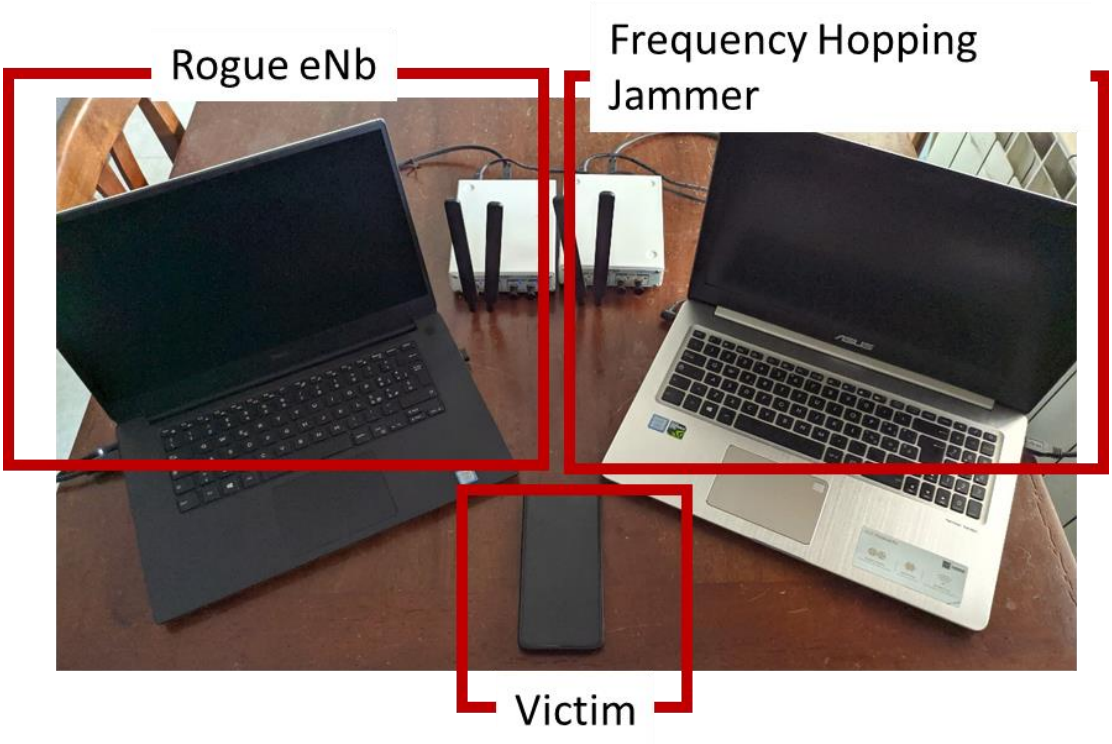

Figure 6 - IMSI catching experimental setup

All the hardware devices used for our experiment can be easily obtained from commercial markets.

As shown in Figure 6 we used two laptop computers (Dell XPS 15 7590, i7-9750H CPU@4.50 GHz × 6, and Asus VivoBook Pro N580G, i7-8750H CPU@4.10 GHz x 6). The operating systems of both computers are 64-bit Ubuntu 18.04 LTS with kernel version 5.3.0-53-lowlatency. Both computers were connected to the transceivers via USB 3.0. The transceivers are two USRP B210 devices, and we can program the USRP B210 to transmit and receive any radio signal we want over a wide radio frequency range, from 70 MHz to 6 GHz, covering all the LTE frequency bands. We used specific LTE antennas connected to the SMA of the SDR devices.

3.4.3 Malicious network

We ran **OpenAirInterface** to create the full malicious LTE architecture. In order to be able to deceive the UE and convince it to provide its IMSI it is necessary to configure the malicious network with the real parameters used by the operators in their networks, more precisely the parameters we need are: i) MCC, ii) MNC, iii) Physical Cell Identifier (PCI), iv) Tracking Area Code (TAC), and v) Inter-frequency cell re-selection priorities.

The **MCC** and **MNC** parameters are used to uniquely identify a mobile network operator in a specific country; MCC is the country code and always has 3 digits (some countries can use more than one MCC), MNC is the network code, it can have 2 or 3 digits. **PCIs**, or Physical Cell

Identifiers provide a pseudo-unique value to identify eNodeBs. The PCI value is created from two components: Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS). The PSS has the value 0, 1, or 2. The SSS can have a value between 0 and 167. The PCI value is [(3x SSS) + (PSS)], belonging to a range of values starting from 0 to 503, so the PCI is limited to 504 values.

The **TAC** is the unique code that each operator assigns to each of their Tracking Areas (TAs), which are defined as logical groups of neighbour eNodeBs (this grouping is performed at the initial deployment of the network). The **Inter-frequency cell re-selection priorities list** is a list containing the available frequencies and their priorities, it is broadcasted by SIB5 and it is exploited by all UEs in the serving cell to perform inter-frequency cell reselection.

To retrieve these network cell parameters two different approach can be followed:

1. The first approach is more oriented to a test scenario; it exploits the UE to retrieve the connection parameters. For example -using an app called *NetMonster- it* is possible to retrieve all needed network cell parameters for the connection with the user equipment;
2. The second approach is more oriented to a real attack scenario in which a hypothetical attacker does not need UE usage and it is composed by 3 steps:
    a. **Autonomous Network Discover:** A hypothetical attacker can autonomously discover networks exploiting a OAI module called “Cell search” that is able to scan all possible frequencies with the aim of finding available network cells. Once network cells are found, an attacker can synchronize with them and decode broadcasted SIBs.
    b. **Obtaining SIBs scheduling information from SIB1:** We can decode System Information Block 1 which is carrying cell access related information. In particular, it also contains the SIBs scheduling information. SIB1, as opposed to others, uses a fixed schedule and for this reason it is possible to decode it exploiting the OAI module called “Cell measurement”, that analyses traffic at a given frequency and provides SIB1 payload. Once the SIB1 payload is retrieved, we can decode it by using an online free ASN.1 decoder.
    c. **Obtain inter-frequency cell re-selection priority list from SIB5:** Cell Reselection is a mechanism to change cell after a UE is camped on a cell and stays in IDLE mode. Thanks to this mechanism, the UE is able to connect to the cell that has the best condition among all the allowable cells. Cell Reselection follows some criteria and algorithms, in particular it uses 3 main criteria: i) Absolute Priority, ii) Radio Link Quality, iii) Cell Accessibility.

In our test, we exploit the second approach and the first criterion used in the Cell Reselection mechanism, the inter-frequency cell reselection priority list broadcasted in SIB5. In this way, we ensure that the UE will try to connect to the malicious eNodeB. Differently from SIB1, other

SIBs are being scheduled at the cycles specified by SIB scheduling information elements in SIB1. For this reason, once SIBs scheduling information is obtained from SIB1, we need to calculate the SIB5 scheduling transmission parameters to capture and obtain it.

We used a modified version of OAI's "*Cell measurement*" module to capture SIB5. Once SIB5 is obtained, we decode it with an online ASN.1 decoder and retrieve the Inter-frequency Cell Reselection priority list. In particular the information needed to calculate the SIB5 scheduling transmission are: i) Scheduling info list, ii) si-WindowLength, iii) si-Periodicity.

Once the Inter-frequency Cell Reselection priority list is obtained, the highest priority frequency is chosen for the malicious eNodeB configuration while all other frequencies are disturbed by the LTE jammer.

3.4.4 LTE jammer

The second main component of the IMSI Catcher system is the LTE jammer, it is needed to force the user to communicate with the rogue eNodeB using the prioritized frequency. In order to achieve this goal, it was decided to design a narrowband frequency hopping LTE jammer that would be able to disturb all the frequencies that can be used by the user terminal, except the prioritized frequency used for communications by the malicious eNodeB.

The narrowband frequency hopping LTE jammer was designed by using the open-source software development toolkit GNU Radio. It contains two main components: i) Noise transmitter chain, ii) probe signal chain.

The Noise transmitter chain is responsible for the creation of noise in the desired frequencies; in particular in our tests, the jammed frequencies are: band 1, band 3 and Band 20; leaving band 7 available to the IMSI catcher system. Thanks to the Probe signal chain, it is possible to dynamically change the jammed frequency every 0.5 s. Lower values have been tested but do not allow maximum power transmission during the frequency switch, consequently they do not allow to maximize the channel noise.

3.4.5 Experimental results

We use the previous described experimental setup to experiment an IMSI catching attack. This subsection presents the experimental results. In the first part of the experiment we retrieve SIBs scheduling information. They are the base information needed to perform the jamming operation. In other words, before start the jamming node, we need to know how it must be configured to jam the exact 4G cell. For this purpose, we run OAI to obtain SIBs scheduling information.

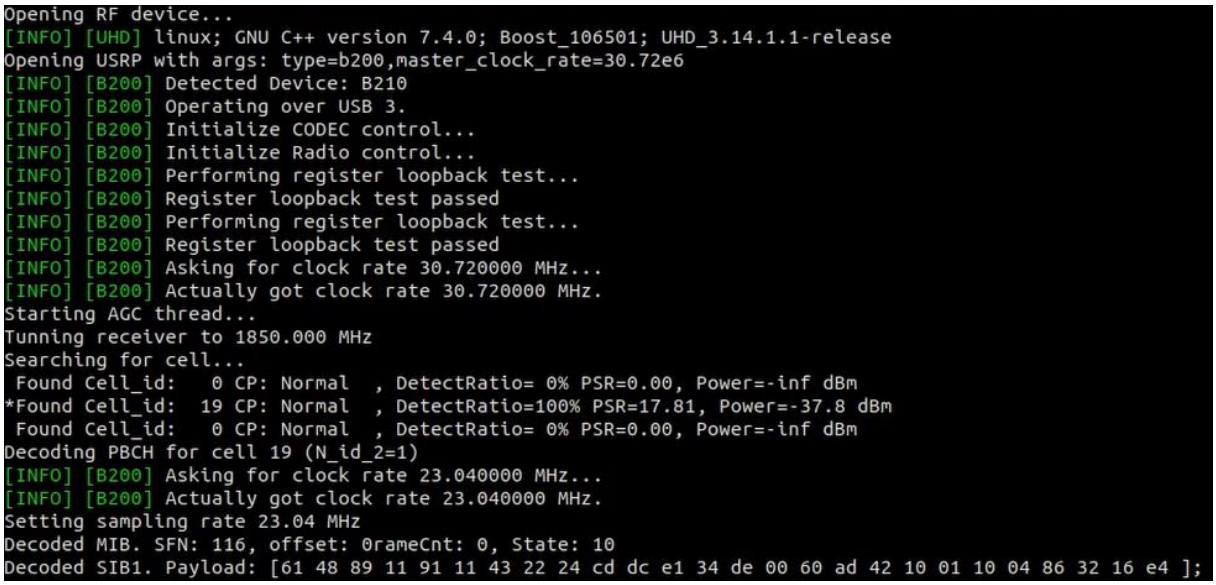

Figure 7 - SIB1 capture

Figure 7 shows the full steps performed to retrieve the SIB1. We used a Linux shell and ran OAI via command line. When the command is executed, the USRP B210 is recognized by the system and all the initialization operations are performed including the setting of the listening frequency. The program searches for the present cells. The cell with ID=19 is detected. For this cell the following information are retrieved:

1. PSR=17.81;
2. Power=-37.8 dBm.

As shown in Figure 7, the cell with ID=19 is selected and the respective SIB1 is decoded. SIB1 is shown in the last row of the figure.

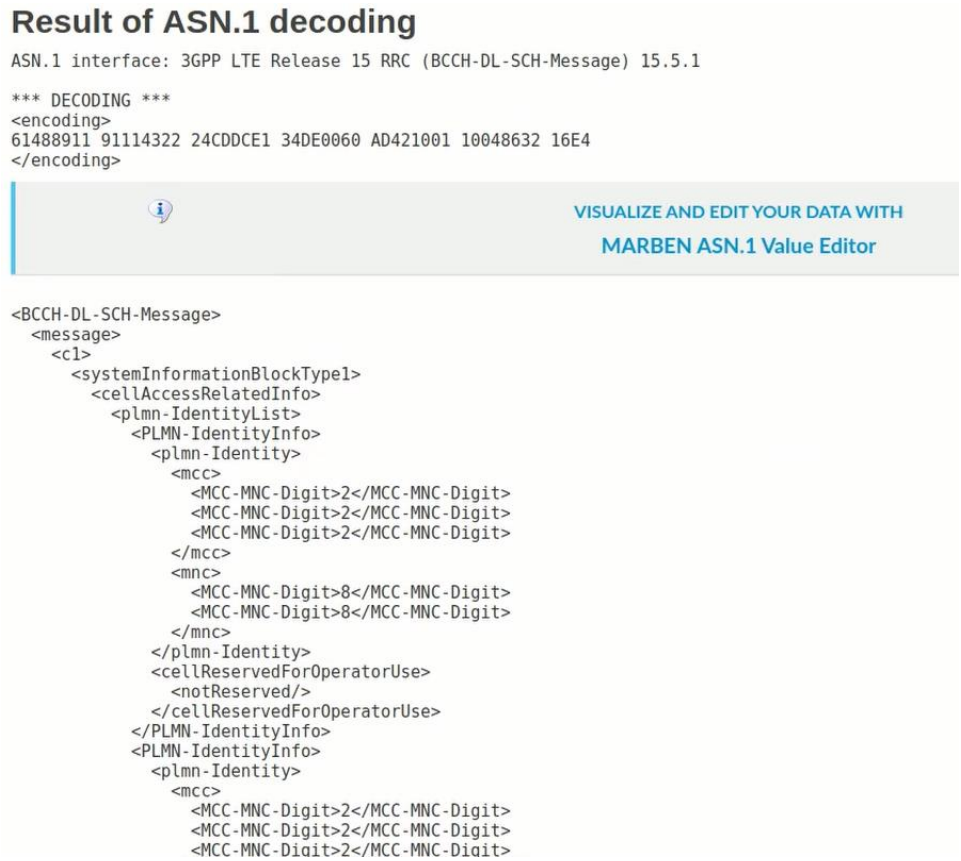

Figure 8 - SIB1 decoded with Marben ASN1 online decoder

Starting from the decoded SIB1, we need to extract the fields of interest. For this purpose, we exploit an online 3GPP LTE ASN.1 message extractor offered by Marben[2]. The decoded SIB1 payload is shown in Figure 8. Thanks to the Marben tool, we extract the following information:

1. MCC=2
2. MNC=8

Once the SIB1 payload is decoded we are able to retrieve the scheduling information needed to capture SIB5 and so to retrieve the inter-frequency cell re-selection priority list.

As stated before, the inter-frequency cell re-selection priority provides the list of the cells used by the UE to perform the connection procedure. Typically, the order of the cell in the list is based on the Radio Link Quality.

[2] https://www.marben-products.com/decoder-asn1-lte/

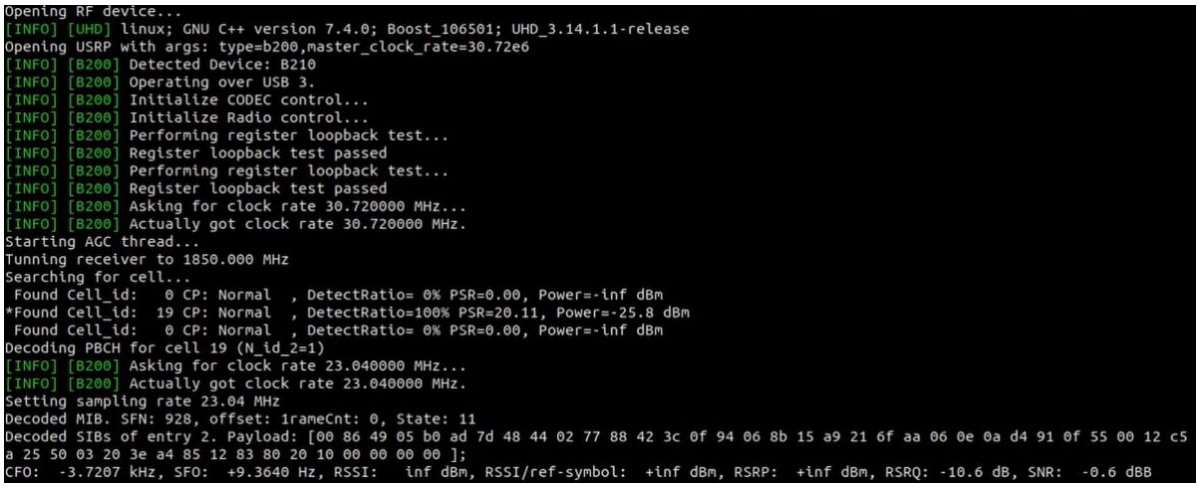

Figure 9 - SIB5 capture

Figure 9 shows the command line results when the OAI framework is used to extract the SIB5 payload. Again, the initialization phase of the USRP B210 is performed, afterward the cell (ID=19) is selected and finally the decoded SIB5 payload is shown.

Starting from the obtained SIB5 payload, we exploit the Marben tool to extract the E-UTRA Absolute Radio Frequency Channel Number (EARFCN) list and the relative priority. EARFCN uniquely identifies the LTE band and carrier frequency of the uplink and downlink channels, which ranges between 0 and 65535. For example, Band-1 and Band-4 can have the same Rx frequency 2110-2170 MHz, but their EARFCN are different. The EARFCN is independent of channel bandwidth.

EARFCN	Priority (1-7)
3350 (Band-7)	7
1650 (Band-3)	6
6200 (Band-20)	5
150 (Band-1)	5

Table 2 - Inter-frequency cell re-selection priority list from SIB5

Table 2 reports the Inter-frequency cell re-selection priority list broadcasted by the selected cell in the experiment (ID=19) through SIB5.

Band	Name	Bandwidth (MHz)	Mode	Earfcn DL	Downlink (MHz)	Earfcn UL	Uplink (MHz)
7	2600	70	FDD	3350	2680.00	21350	2560.00
3	1800+	75	FDD	1650	1850.00	19650	1755.00

20	800 DD	30	FDD	6200	796.00	24200	837.00
1	2100	60	FDD	150	2125.00	18150	1935.00

Table 3 - Technical specification for the complete list of EARFCN

We retrieve additional information about the EARFCN values through the online tool provided by Squimway[3]. Table 3 shows the complete technical specifications for the EARFCN list reported from the selected cell.

We exploit that list to configure the malicious eNodeB for the priority frequency and to configure the jammer to disturb all other frequencies. As shown in Table 2 the frequency related to EARFCN 3350 is the priority one, it will be used to configure the malicious eNodeB and all the other frequencies will be jammed.

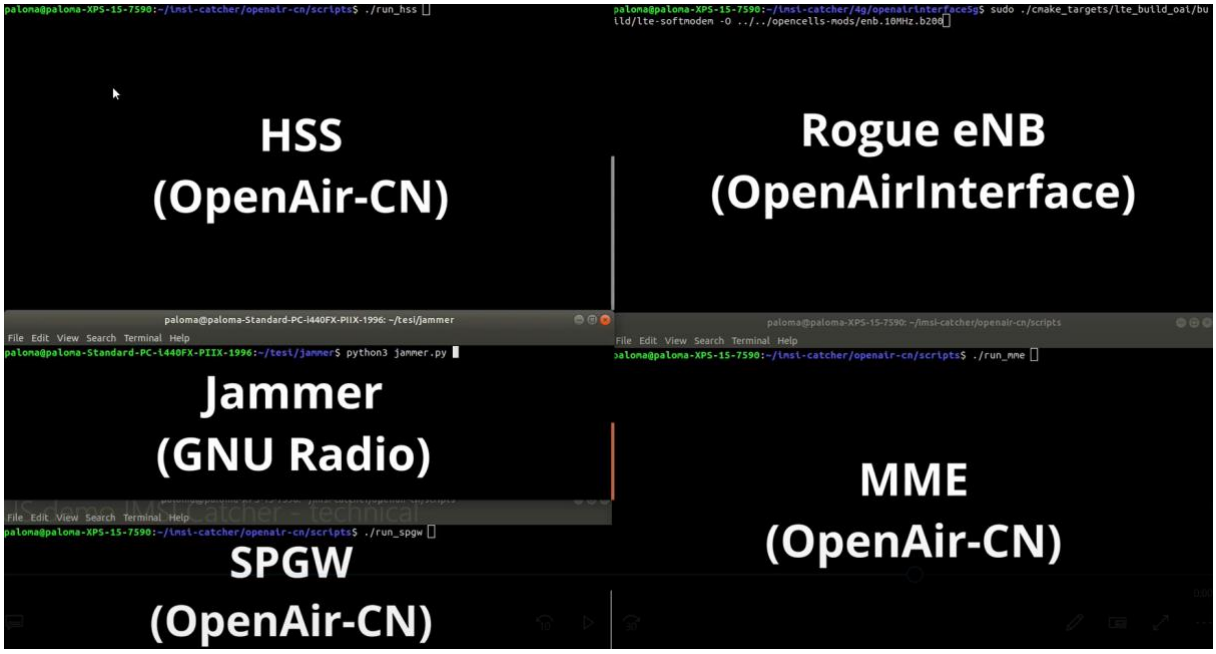

Figure 10 – IMSI catching experimental set-up

The next step of the experiment is the set-up of the jammer and of the malicious eNodeB. The jammer is performed via the GNU Radio framework[4]; it is configured to jam the frequency related to the band-3, band-20 and band-1, according to the capabilities of the Table 3. AOI Rouge EnodeB activation requires the 4G core network; it is composed of: i) HSS, ii) MME, and iii) SPGW. The SPGW (Serving and Packet Data Network Gateway) is the OAI module that provides the gateway operation between the 4G network and the IP network. For each of

[3] https://www.sqimway.com/lte_band.php
[4] https://www.gnuradio.org/

these elements we open a Linux shell and run the relative start command; the complete view of the screen with all the opened shell is showed in Figure 10.

Once all the entities (eNodeB and Jammer) are configured, as shown in figure 10, we are ready to perform the IMSI Catching attack. We active the OAI eNodeB and afterward the jammer process.

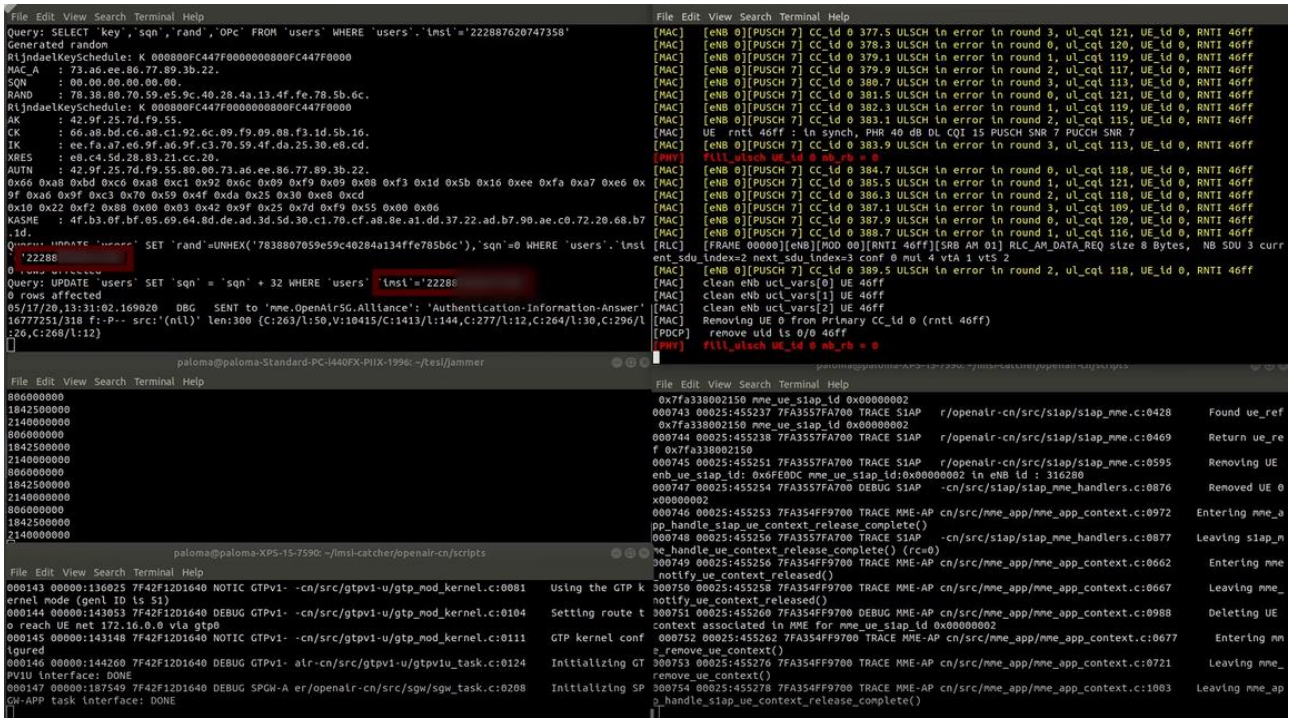

Figure 11 - IMSI catching attack

We use a smartphone as UE victim. When the jammer has been activated, the smartphone loses the connection with the operator eNodeB and performs a cell reselection according with the Inter-frequency cell re-selection priority list. This time the smartphone performs the attach procedure with the rouge eNodeB. The attach procedure is composed by 3 phases:

1. UE attempts to perform a Tracking Area Update (TAU)
2. The Rogue eNodeB replies rejecting the Tracking Area Update (specifying cause 9 *"UE identity cannot be derived by the network"*)
3. **UE tries to perform an Attach Request with its IMSI as plain text**

We then perform a successful IMSI catching attack to the UE victim; a screen shot of the last phase is showed in Figure 11, where we highlight with a box the received IMSI.

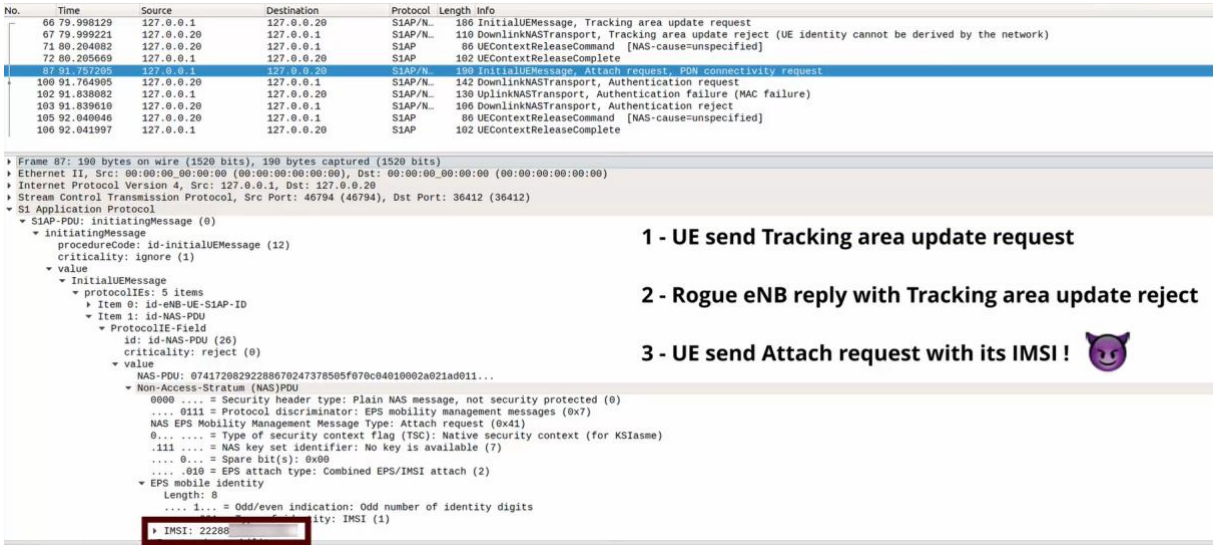

Figure 12 - Messages exchanged between UE and eNodeB during the IMSI catcher attack

Finally, Figure 12 reports the exchanged messages between UE and eNodeB during the IMSI catcher attack by using the Wireshark tool[5].

Model	OS	Modem	LTE Cat.	OpenAirInterface			
				without jammer		with jammer	
				Service Request	TAU Request	Service Request	TAU Request
Samsung Galaxy S9	Android 9	Exynos 9810	18	✓	✓	✓	✓
Samsung Galaxy A7 2018	Android 10	Exynos 7885	12	✓	✓	✓	✓
Samsung Galaxy Note Pro	Android 5	Snapdragon 800	4	✓	✓	✓	✓
Realme X2 Pro	Android 10	Snapdragon X24	20	✓	✓	✓	✓
Realme 6	Android 10	Helio G90T	13	✓	✓	✓	✓
Xiaomi Redmi Note 7	Android 9	Snapdragon X12	12	✓	✓	✓	✓
Xiaomi Mi A1	Android 9	Snapdragon X9	7	✓	✓	✓	✓
Huawei Mate 20 Pro	Android 9	HiSilicon Kirin 980	21	✓	✓	✓	✓
Huawei P30 Lite	Android 9	HiSilicon Kirin 710	12	✗	✓	✓	✓

[5] https://www.wireshark.org/

Huawei P8 Lite	Android 7	HiSilicon Kirin 655	6	✓	✓	✓	✓
Asus Zenfone 2	Android 5	Intel XMM 7260	4	✓	✓	✓	✓
iPhone 11	iOS 13	Intel XMM 7660	18	✗	✗	✓	✓
iPhone XS	iOS 13	Intel XMM 7560	16	✗	✗	✓	✓
iPhone 8	iOS 13	Intel XMM 7480	16	✗	✗	✓	✓
iPhone 7	iOS 13	Intel XMM 7360	9	✗	✓	✓	✓
iPhone SE	iOS 12	Qualcomm MDM9625M	4	✓	✓	✓	✓
iPhone 5S	iOS 12	Qualcomm MDM9615M	3	✓	✓	✓	✓
Huawei E3272 USB Stick	-	HiSilicon Balong 710	4	✓	✓	✓	✓
Huawei E392 USB Stick	-	Qualcomm MDM9200	3	✓	✓	✓	✓

Table 4 – IMSI catching attack results on different devices and operating system

We have tested many different devices in order to verify how their operating system could influence the behaviour of the device during an IMSI catching attack. Both Android and iOS devices were tested, the detailed list of tested devices and related results are reported in Table 4. For each device we report the success or failure of the IMSI catching attack in terms of SERVICE REQUEST and TRACKING AREA UPDATE, in presence and absence of the jammer module activated. The service request message is sent by the UE to the network to request the establishment of a connection.

The interesting result is that the great majority of devices immediately falls victim to our attacks, and often with no need of targeted jamming. Only the iPhone 7 and newer versions show more resilient behaviours, as they are more robust to IMSI catching and they provide IMSI only occasionally. We repeated multiple tests in order to better understand why iPhones behave differently. If we do not run the LTE jammer over non-priority frequencies, when we activate the malicious eNodeB, the iPhone switches to another available LTE cell. If we run LTE jammers over non-priority frequencies, often, when we activate the malicious eNodeB, the iPhone automatically downgrades to 3G or even GPRS without providing IMSI to our IMSI Catcher.

We also evaluate the impact of the operator on the IMSI catching attack. In fact, on one side, all the procedures and relevant parameters involved in our IMSI catching scenarios are actually managed by the Operator to which the UE is registered. But on the other side the vulnerabilities exploited by our IMSI catchers involve a level of detail which appears more related to the radio modem and device firmware, hence might not be influenced by the operators' configuration of the UE USIM.

To shed some light on this question, we bought commercial off-the-shelf USIMs/contracts from all four Italian Operators (Wind3, Vodafone, TIM, Iliad), and we ran experiments using one of the Android phones, specifically the Realme X2 Pro (Android 10).

Tim Wind3 Vodafone Iliad	X	Service Request Tracking Area Update Request	X	Without jammer With jammer	=	Positive

Table 5 - Analysis of operator impact on IMSI catching

Results suggest that the IMSI catching related behaviour of the victim does not depend on the operator. Indeed, as summarized in Table 5, the attack had success with all the possible combinations - highlighted in the table as the "cartesian" product among the four operators, the two different attack techniques, and the usage or not usage of jamming.

3.5 Preliminary study on IMSI catching attacks mitigation

A preliminary study on techniques to prevent IMSI catching attacks has been addressed by considering the possibility to use a jamming detector. Detection of smart-jamming attacks is feasible by monitoring any excess amount of energy on a specific physical channel (e.g. using masking) or any sudden change in the performance of the communication over this channel [24]. A common strategy in jamming detection is the use of a threshold with some performance metrics such as the packet delivery ratio (PDR), packet drop ratio (PDR), bit error rate (BER), and signal-to-noise ratio (SNR). These techniques monitor the level of these metrics during the absence and the presence of jamming attacks and set the threshold for detection. Threshold based detection are only efficient when we are dealing with constant jammers. In addition, because of the wireless environment dynamics, these methods have a high false alarm. Another detection category is statistically based [25]. These techniques use historical data and compute statistics to distinguish jammed signal from non-jammed signal. The last category is machine learning based. Several machine learning techniques such as random forest, decision tree, adaptive boosting, support vector machine, and expectation maximization are investigated in detecting jamming attacks [26] . Recently, deep learning which is a special case of machine learning is heavily investigated to detect jammers [27].

Base station should implement anti-jamming techniques. For example, if an exchange between the base station and user equipment is jammed, the base station should provide a spatial retreat, movement, time, and network reconfiguration.

We will explore deep learning approaches to anti-jamming. To train the deep learning algorithms, a large comprehensive dataset is needed, and the LOCUS system can be used for this purpose. The built dataset can be used to train and test deep learning techniques. Then, these techniques can be combined with sensing to detect the strategy of the jammer and actively select the communication channel that is not under jamming attacks. Deep learning models have to be trained to recognize a jamming signal from a legitimate user on the fly. Only if that can be performed, then the legitimate user can identify the pattern of the jammer and dedicate its strategy and accordingly define a mitigation hopping.

A different mitigation for the IMSI catching attack can be based on the study of the existing mechanism of 5G that provides keys between UE and core network. In [28], Authors proposed an authentication scheme which leverages the precomputation-based digital signature generation algorithms. However, their solution is a bit complicated and does not take full advantage of the existing private/public keys mechanism of 5G. We explore the possibility to increase the robustness of the identity/authentication mechanisms. The PKI mechanism in 5G network is designed to protect the user's permanent identity information SUPI. The mechanism exploits the ECC to generate a pair of public and private keys. The public key is stored in the SIM card of the 5G UE, and the private key is stored in the core network of carriers. When the network needs to request the UE identity information, the UE combines the public key stored in the SIM with the private key of the temporarily generated public-private key pair to form a shared key, which is used to encrypt the SUPI to generate the SUCI. The carrier core network then uses the temporary public key sent by the UE and the private key stored by itself to form the same shared key, which can be used to decrypt the SUCI and obtain SUPI. The certificate and signature are attached to the original SIB1 message and the UE determines whether to trust the gNodeB according to the signature information.

In order to prevent relay attacks, and avoid false announcement of the SIB2, we propose to use a delay parameter. According to the cell parameters, the operators can set a threshold T_h. The network side SIB message generation time is recorded as T_{gen} and the receiving time of SIB is recorded as T_{rec}. The UE calculates the difference between T_{gen} and T_{rec} and compares it with the threshold T_h. This can greatly reduce the attack window of malicious gNodeB. In order to prevent the attack from passively eavesdropping on the air interface information, we also consider an initial RRC and initial NAS layer message encryption scheme.

4 Location security

Location data in LOCUS will rely on metrics extracted from the uplink and down-link signals of the new radio (NR) (3GPP-technologies) as well as on non-3GPP technologies such as, for instance, Global Navigation Satellite System (GNSS), Terrestrial Beacon Systems (TBS), Bluetooth, WLAN, RFID, and other sensors.

In any case, the security requirements for localization techniques must include the capability of identifying and possibly mitigating any kind of deviation from the true locations. As a matter of fact, location data are highly vulnerable to both data-level and signal-level spoofing as well as meaconing attacks caused by malicious intruders. In order to face the aforementioned threats, the location security function should contain a preliminary stage responsible for the detection of an attack and a second stage (possibly shared with the location function) aimed at handling the fake measurements generated by the malicious platform. A high-level description of the location security function along with its link with the location function is shown in Figure 13. From the figure, it turns out that the location security function should manage measurements provided by heterogeneous sensors.

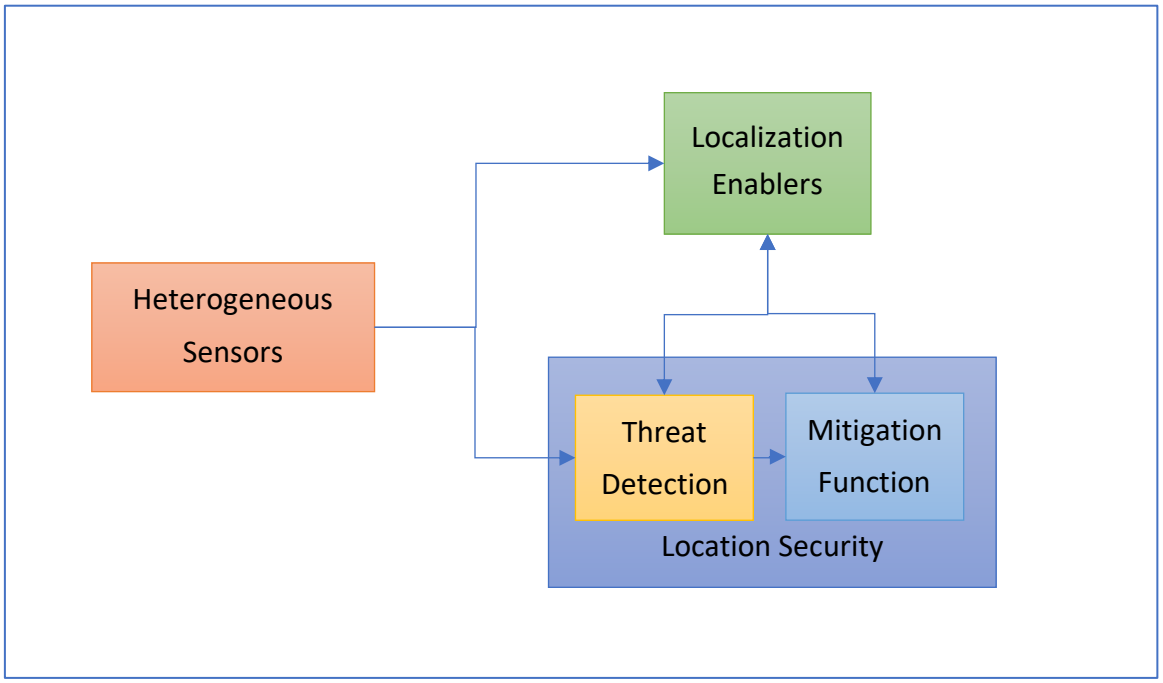

Figure 13. High-level description of the location security function and of its connections with the localization function and heterogeneous sensors.

4.1 Location security in 4G/5G networks

Location services in LOCUS mainly rely on the (possibly joint) exploitation of suitable measurements as the Time of Arrival (ToA), Time Difference of Arrival (TDoA), Angle of Arrival

(AoA), and Received Signal Strength (RSS) from 5G networks or other integrated non-3GPP technologies [29]. The localization process may occur at a network level (network-centric approach), where the computation of the UE position is performed by the network and transmitted to the latter, or at a user level (user-centric solution), where the UE collects information from the network and uses it to determine its position. In both cases, the location measurements can be combined with the information provided by other sources as, for instance, satellite positioning systems or other UEs in the proximity of that under consideration to enhance the localization accuracy.

In this highly connected context (see Figure 14), location security becomes of primary concern, especially in applications related to safety and liability. As a matter of fact, due to the large number of stages interacting towards the position estimation, the exposure to attacks is high indeed.

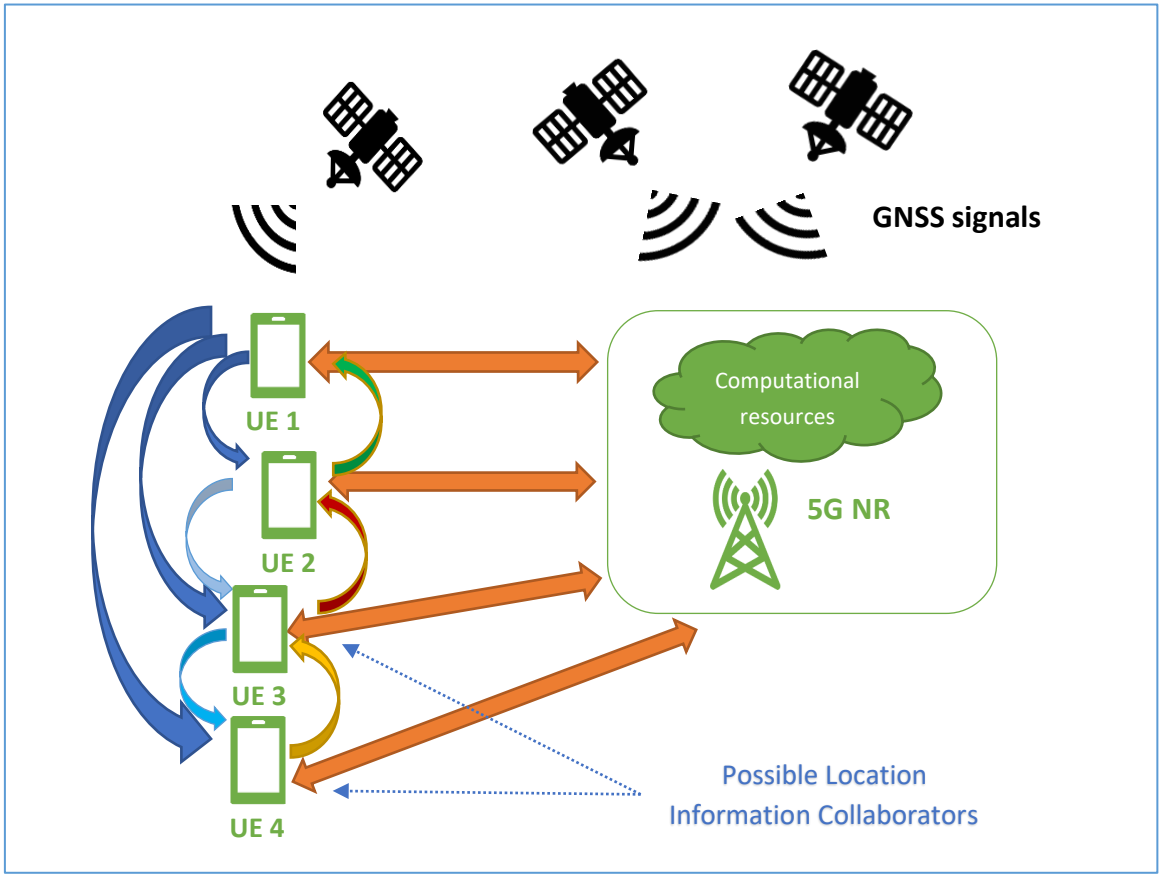

Figure 14. A highly connected scenario in 5G networks.

Otherwise stated, there exists a plethora of points of the distributed communication system or, equivalently, of the processing chain that can be subject to malicious actions. Therefore, new solutions aimed at preserving data integrity are strongly desirable. Among the possible

threats to location security, here we focus on intentional interference, which involves hostile platforms that target the UE and/or AN receivers in order to [30]:

- reduce the signal-to-noise ratio (noise-like jamming).
- inject false or erroneous information[6] (spoofing/meaconing).

Hostile platforms, that perform a Denial of Service (DoS) attack by transmitting high-power interfering signals to induce the disrupting of the receiver functionalities (see Figure 15), belong to the first class of threats and are referred to as Noise-like jammers (NLJs) [31] [32]. Signals transmitted by NLJs blend into the thermal noise of the receiver with the result of an increase of the noise power spectral density within the receiver bandwidth.

As for the second class of attacks, spoofing/meaconing threats intercept the positioning messages exchanged by two legitimate actors and alter them by synthetizing counterfeit information [30] [31]. These erroneous messages would prevent the location service from providing reliable position estimates (see Figure 16).

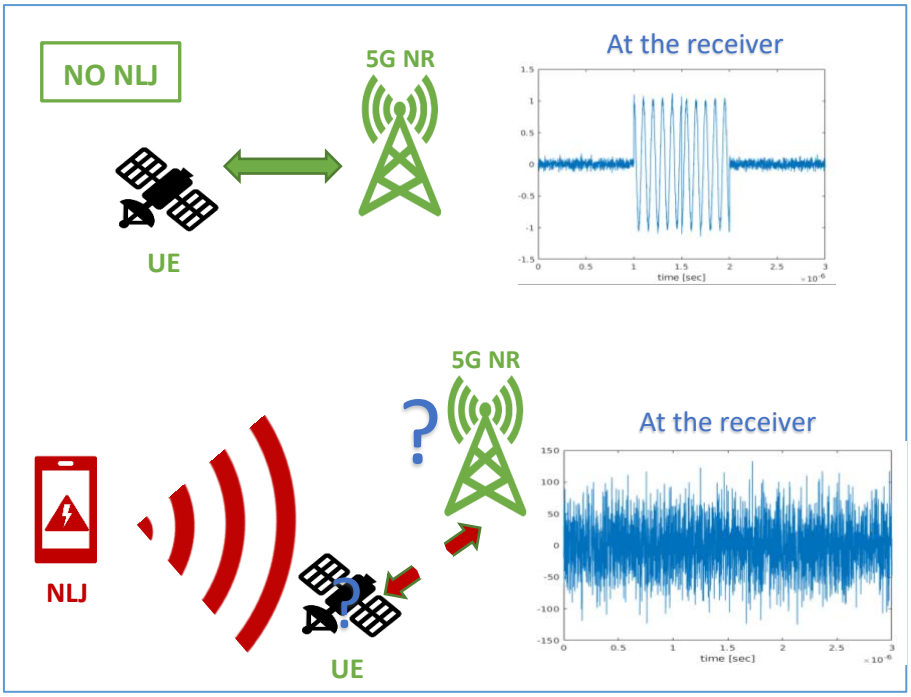

Figure 15. The effects of the NLJ attack.

[6] Note that in the context of Electronic Warfare, such type of interference activity is referred to as Deception (Poisel, 2011).

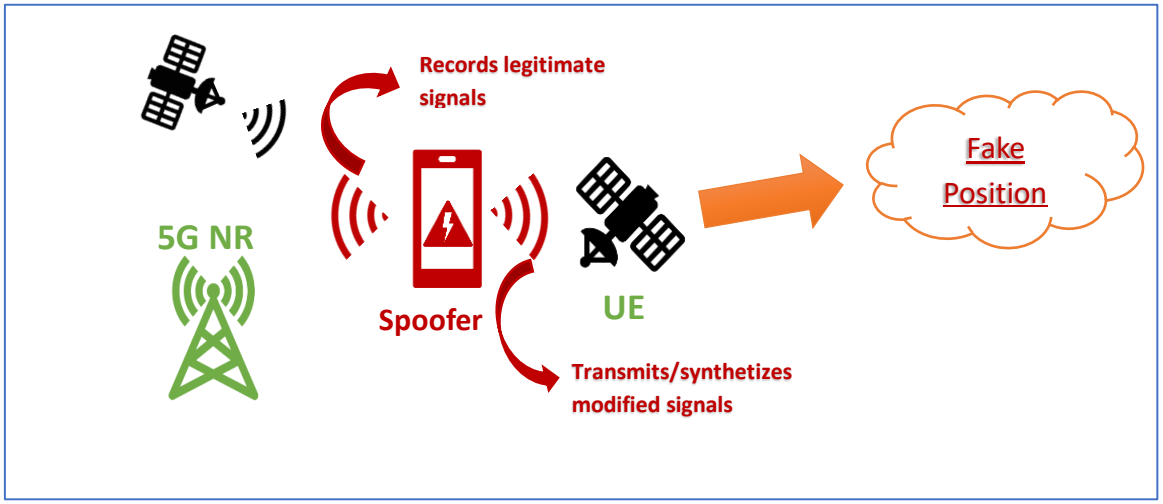

Figure 16. Spoofing/Meaconing operating scenario.

4.1.1 Noise-like jammers: a review of countermeasures

NLJs can be classified according to the bandwidth occupancy and the transmission duration [33] [34]. The simplest and, hence, most common form of noise-like interference is represented by a **Barrage Jammer** (BJ), which continuously transmits noisy signals regardless of the network waveforms and topology. Besides BJs, which are not efficient in terms of energy consumption, smart noise-like jammers are capable of optimizing the effectiveness of an attack and the energy use at the price of a more demanding complexity of the system (architecture and processing schemes). In fact, smart jammers are endowed with network sensing capabilities and can start their attack when a network activity on a certain channel is detected. Another important feature is that **Smart Jammers** can also corrupt part of a given message or transmit noise pulses which are concurrent with the communication signals (noise cover pulse). As a result, they are much more difficult to detect than BJs in practice. For these reasons, it is important that in this case a constant monitoring of the network must be performed [35].

Summarizing, NLJs could yield detrimental effects on the location functions of the LOCUS platform and, more generally, of any location systems by either significantly reducing the estimation accuracy or preventing the platform from obtaining the required measurements. Moreover, such attacks can be perpetrated on both the UE and AN sides. Indeed, the noise receiver level increases because of a NLJ attack and degrades the estimation quality of the AoA, ToA, TDoA, and RSS measurements at a generic AN, leading to inaccurate UE localization [36]. Remarkably, NLJs might also target the GNSS messages making weaker the user-centric positioning strategies or positioning algorithms based upon location information collaborators which are in the proximity of the UE to be localized [30].

Detection and mitigation strategies to cope with NLJ attacks are of primary importance to ensure reliable location services in wireless networks as corroborated by the huge amount of works existing in the open literature. In [37], the authors investigate the effect of noise-like interference on the Physical Uplink Control Channel in LTE and propose a mitigation strategy based upon the Radio Link Control protocol and, in the detail, the quality of the control messages. Anti-jamming techniques in the context of cognitive radio networks are addressed in [38], where several mitigation techniques are reviewed and a new anti-jamming protocol relying on probabilistic pairing and frequency tuning is proposed. In [39], the Q-learning algorithm is applied in the context of cognitive radio to learn the jammer strategy and, hence, to pro-actively avoid jammed channels. Other anti-jamming techniques in Wireless Networks or LTE exploit *Game Theory* [34] [40] [41]. Mitigation solutions conceived at the physical layer can be found in [33], where the received signal is classified by means of a deep convolutional neural network fed by the signal features in the wavelet domain. Alternative features or parameters to detect jamming attacks are represented by packet ratio, packet delivery ratio, received signal strength, and channel assessment [42] [43].

A breakthrough upgrade of 5G network with respect to 4G is the use of multiple antennas and, in particular, of massive multiple-input multiple-output systems which represent the key technology to increase the spectral efficiency in the new wireless networks generation. As a matter of fact, sensor arrays allow to achieve higher data rate and reliability by exploiting digital beamforming techniques. It is possible to steer the transmitted signal toward a desired direction and, at the same time, avoid receiving unwanted signals from undesired directions [44]. In this case, the anti-jamming techniques can take advantage of spatial diversity/angular resolution, flexibility associated to the processing of digital raw samples, and borrow techniques from *Electronic Warfare*. In [45], random matrix theory is exploited to conceive a multiple hypothesis test to detect the presence of jamming signals by estimating the jammer subspace through the sample covariance matrix [36]. Then, received data are projected onto the user subspace in order to mitigate the jamming components.

4.1.2 Spoofing and meaconing attacks: a review of countermeasures

This type of attack involves a platform that intercepts the positioning signals and suitably delays or modifies them (man-in-the-middle attack [30] [46]) in order to lead the victim system to compute fake positions. Remarkably, the malicious platforms can be also capable of synthetizing fake messages to deceive the legitimate actors. As already stated, the targets of these attacks may be either the UEs or the ANs or both. The attacker can also adulterate GNSS signals.

A taxonomy for these attacks can have the following structure [31] [32]:

- **Spoofing:** the attacker synthetizes and transmits counterfeit messages at different levels, namely
  - **Data-level**: data used to estimate the position are modified.
  - **Signal-level**: fictitious positioning signals are generated and delayed.
- **Meaconing:** the attacker records and replay the authentic positioning signals.

In addition, the malicious nodes may collect authentication credentials and other data [30] [46]. In [47], the Generalized Likelihood Ratio Test (GLRT) is applied to detect jamming signals contaminating user pilots, i.e. training sequences used to estimate channel state information. To this end, received raw data is first projected onto the subspace of the unused pilots in order to remove useful signal components and then the GLRT is derived by estimating the unknown parameters through the maximum likelihood approach. In [48] a pilot spoofing attack detection algorithm in massive MIMO systems is proposed, in which the information of channel statistics is unknown. First, users send pilots to the BS, then the latter transmits the conjugate of its received signal (which may contain spoofing signal) back to users, where the final decision on detection is made.

4.2 LOCUS Detection/Mitigation Techniques

The solutions reviewed in the previous section show that intentional interference detection/mitigation in wireless networks can be accomplished at different layers according to the desired trade-off between computational requirements and effectiveness. It is clear that processing raw data might lead to more reliable techniques at the price of an increased complexity, whereas algorithms fed by high-level data or measurements might be more efficient.

In the 3GPP work item on 5G positioning support [49], the investigation is conducted by applying algorithms based upon (possibly linearized) hyperbolic or least squares location methods which are fed by TDoA or AoA measurements. TDoA estimates are obtained by threshold-based algorithms, whereas AoA estimates are obtained by using direction-finding algorithms such as MUSIC. However, issues related to intentional (and also unintentional) interference are not addressed. As a consequence, it turns out that, without any additional protection methods, the vulnerability of the location function is high, i.e., noisy and/or counterfeit signals transmitted by a threat could dramatically decrease the measurement and, hence, estimation quality.

In Figure 17, we show the points of the LOCUS architecture which are exposed to possible attacks as well as functional blocks implementing the security features. Such functions deal with data integrity, which should be preserved within the cellular network or other networks

providing input to the LOCUS platform. In this context, data can refer to both the inputs and outputs of the localization algorithms.

Detection and mitigation of attacks can be performed at two different stages:

- before the ToA, TDoA, AoA, and RSS measurements are performed, i.e., at a waveform- or raw data level.
- after the ToA, TDoA, AoA, and RSS measurements are performed, i.e., at a measurement-level, through the statistical analysis of the measurements along with a censoring of the measurements labelled as unreliable.

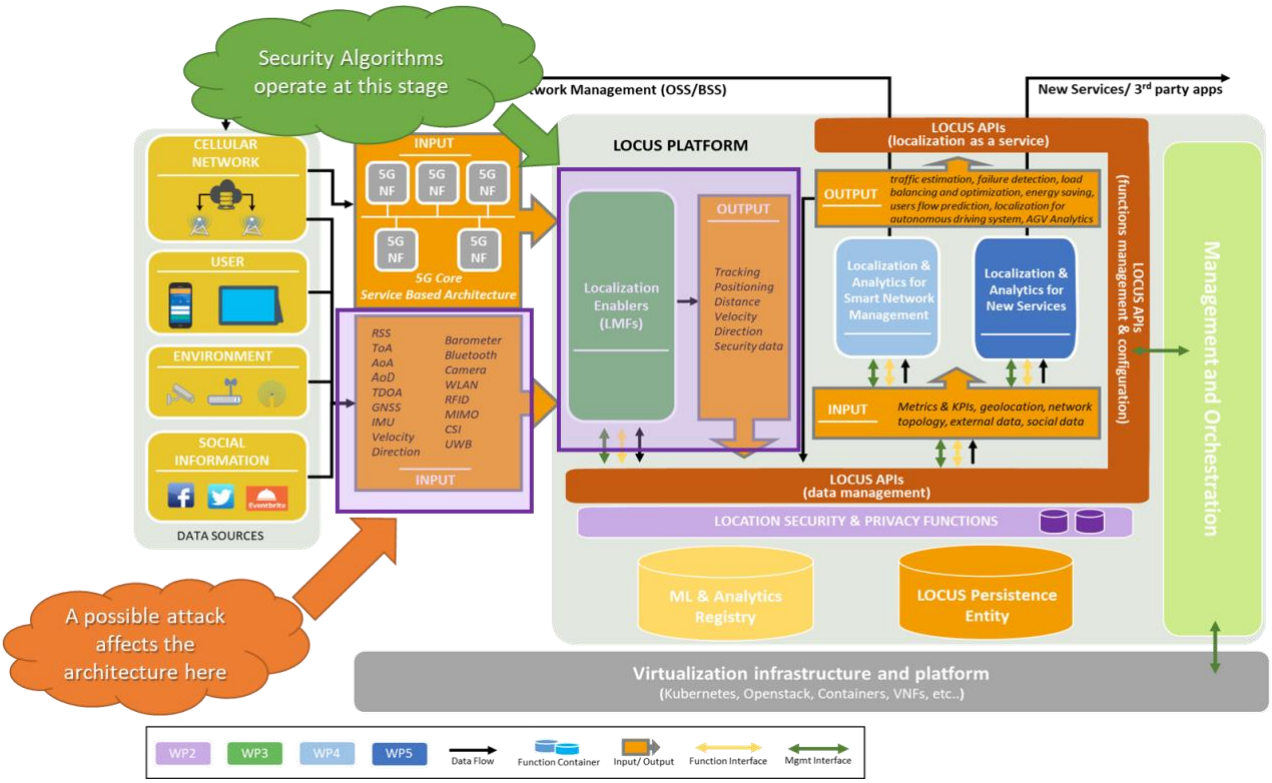

Figure 17. LOCUS Architecture and security functions.

In the following subsections, we address possible solutions for location security fed by either raw data or compressed data, namely the estimates provided by the network and/or localization enablers.

4.2.1 Raw Data-based Algorithms

This class of techniques process data before the ToA, TDoA, AoA, and RSS measurements (or estimation) are performed (i.e., data are the in-phase and quadrature samples). It is clear that since raw data contain all the available information, they might lead to more techniques at the price of an increased computational burden.

In what follows, we consider four general classes of algorithms relying on different paradigms ranging from Signal Subspace Decomposition to Compressive Sensing and Clustering Procedures.

4.2.1.1 Signal Subspace Decomposition

The first aspect related to vulnerability of the location service is that, in the presence of NLJ signals, the MUSIC algorithm should account for the actual signal subspace size in order to correctly identify the orthogonal complement. Thus, when beside the direction of interest there exist additional directions bearing energy, the risk of misidentifying the overall signal subspace is non-negligible. In fact, MUSIC algorithm selects the directions with minimum energy in the orthogonal complement of the signal subspace, namely, assuming to collect data from N sensors, the DoA estimate is obtained by solving

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left(\boldsymbol{v}_p^{\dagger}(\theta)\, \boldsymbol{P}^{\perp} \boldsymbol{v}_p(\theta) \right),$$

where $\boldsymbol{P}^{\perp} \in \mathbb{C}^{N \times N}$ is the projection matrix onto the orthogonal complement of the signal subspace and $\boldsymbol{v}_p^{\dagger}(\theta) \in \mathbb{C}^{N \times 1}$. Therefore, if the size of the signal subspace is underestimated, then there exist signal directions with significant energy in the subspace spanned by $\boldsymbol{P}^{\perp}$ (see Figure 18).

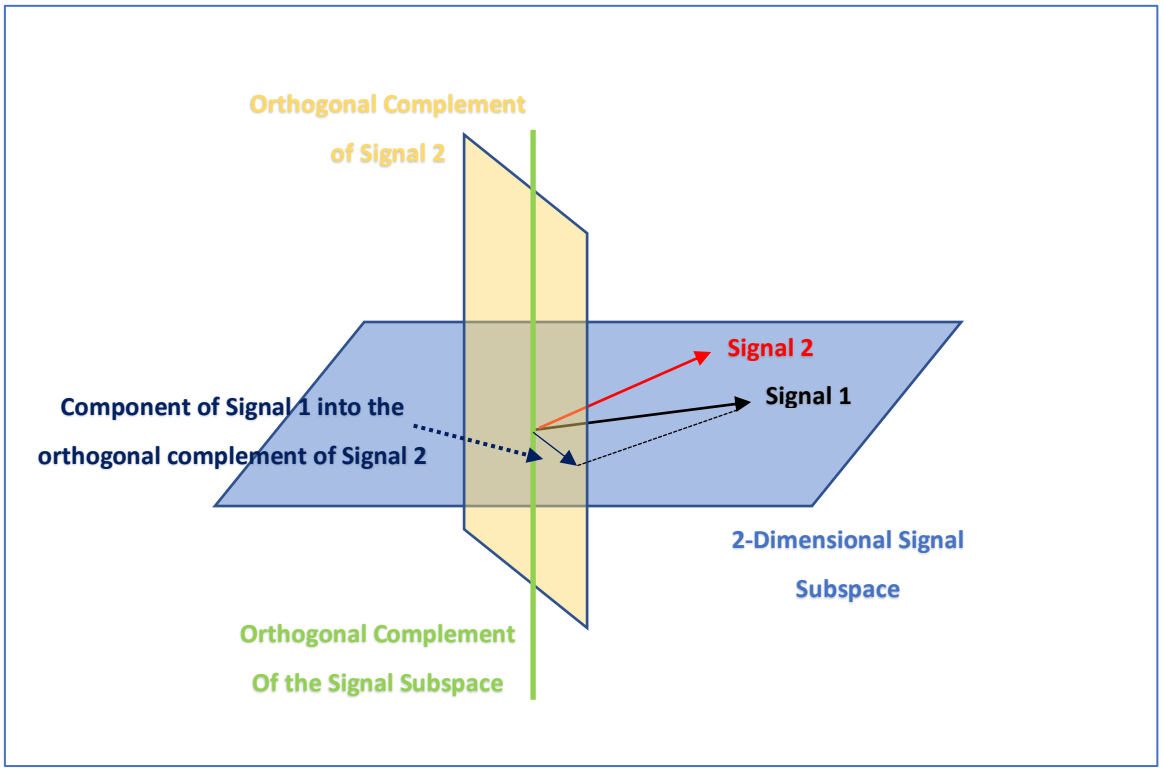

Figure 18. Underestimating the signal subspace leads to an erroneous orthogonal complement which contains unwanted signal components.

In order to cope with this situation, before applying direction-finding algorithms, it would be useful to estimate the size of the signal subspace by resorting to the Model Order Selection (MOS) rule as the Akaike Information Criterion, the Generalized Information Criterion, and the Bayesian Information Criterion [36], [21]. As an alternative, it would be possible to design a test on the eigenvalues of the sample covariance matrix to detect the energy jump due to the received signals.

Another drawback of subspace-based direction-finding algorithms as MUSIC is the sensitivity to the presence of unwanted signals that are correlated to that of the legitimate user (smart jammers). In this case, the estimation performance can be restored by means of algorithms that lower the correlation level as the so-called Spatial Smoothing, which builds up a sample covariance matrix based upon suitable subsets of sensors as shown in Figure 19. Such combination reduces the correlation level among received signals but, as drawback, decreases the angular resolution due to the lower vector size. In Figure 20, we show an outcome of the MUSIC algorithm in the presence of two highly correlated signals (at 0 and 10 degrees) whose peaks are not clearly identifiable. In Figure 21, we show that spatial smoothing is capable of restoring the discrimination performance even though the resolution is reduced.

Remarkably, notice that, in principle, the latter can be also used to detect the presence of correlated signals by comparing the subspace size of original data with that of smoothed data. A block scheme of this architecture is depicted in Figure 22.

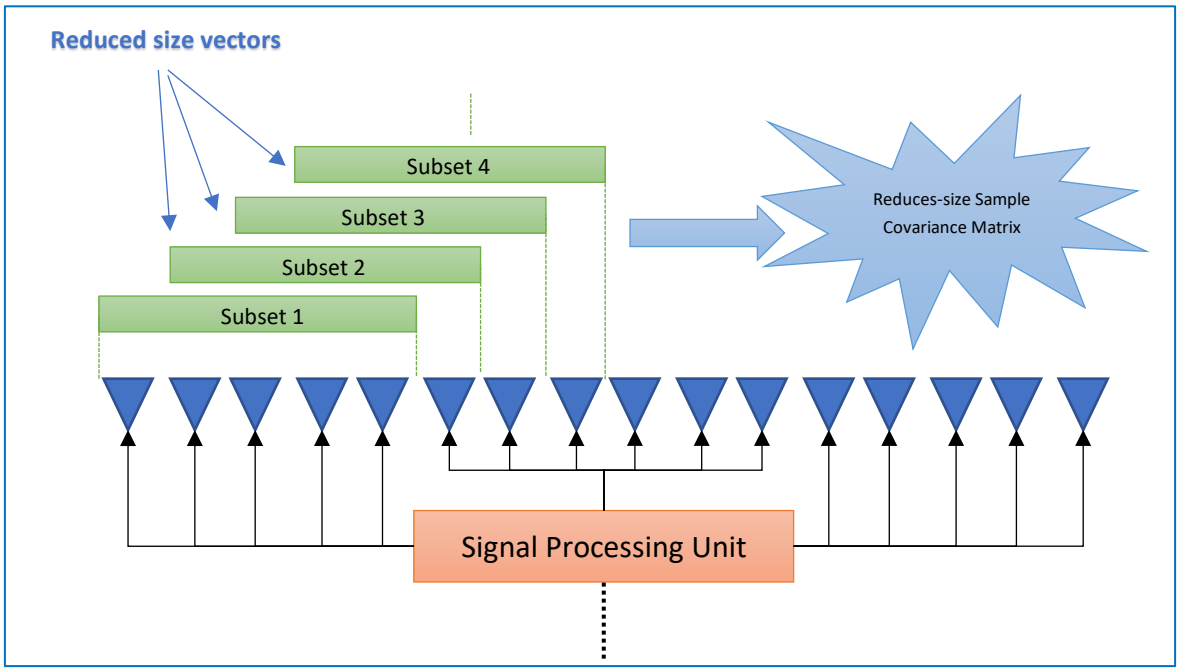

Figure 19. Main idea of spatial smoothing.

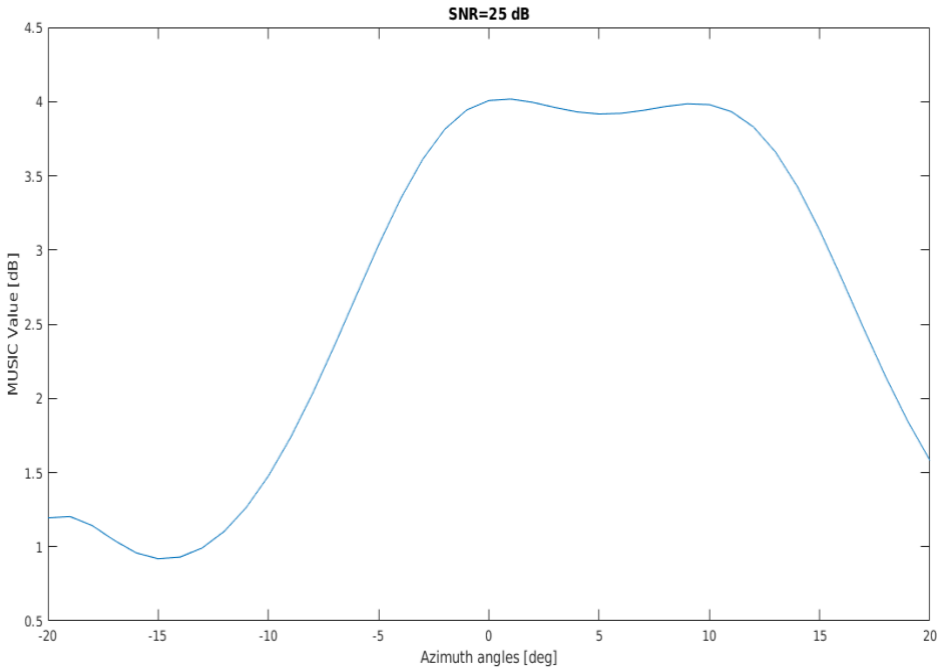

Figure 20. MUSIC values versus the angle of arrival assuming SNR = 25 dB and two fully-correlated targets at 0 and 10 degrees.

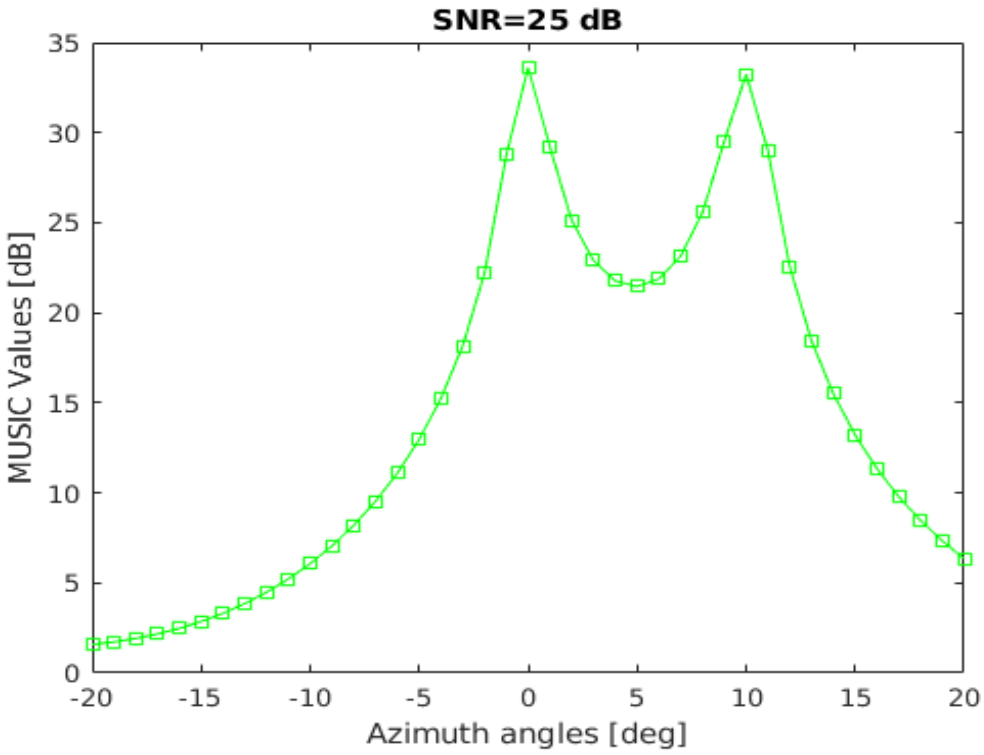

Figure 21. Spatial Smoothing- based MUSIC performance in the presence of two fully-correlated signals at 0 and 10 degrees assuming SNR=25 dB.

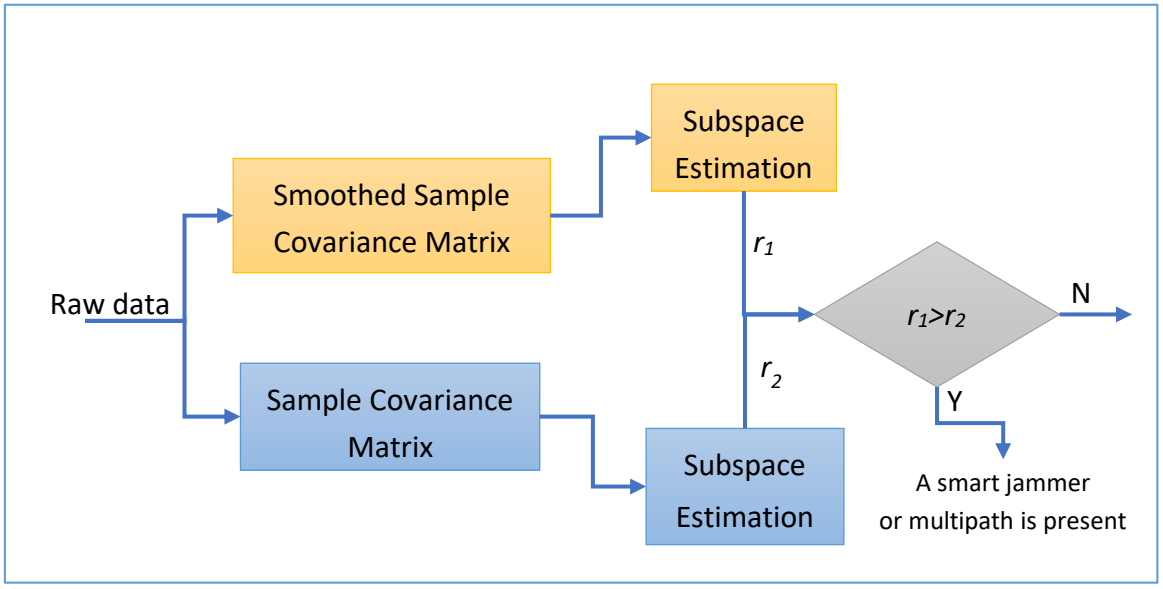

Figure 22. Detector of correlated signals based upon raw data.

4.2.1.2 Hypothesis Testing

In 5G systems, antenna arrays will be intensively exploited allowing for digital beamforming techniques or other sophisticated processing schemes. The problem of jammer detection in the presence of useful communication signals can be formulated in terms of a binary hypothesis test based upon raw data vectors

$$\begin{cases} H_0: & data\ contain\ useful\ signals, \\ H_1: & data\ contain\ useful\ signals\ and\ jammer\ components, \end{cases}$$

and that can be solved resorting to decision theory [50]. Specifically, since a uniformly most powerful detector, generally speaking, rarely exists, we focus on suboptimal tests such as the GLRT, the Rao test, and the Wald test, that are asymptotically equivalent [50]. Among them, the GLRT is the most commonly employed in statistical signal processing, even though it does not share any known optimality property for a finite number of observations. Therefore, it would be worth investigating the remaining design criteria since they might

- lead to decision schemes more robust than the GLRT in the presence of mismatched environments (usually present in real operating situations);
- be less time demanding than the GLRT.

Finally, we underscore that the presence of nuisance parameters could make the detector designs knotty in terms of mathematical tractability, leading to decision rules of difficult implementation. For this reason, we can resort to heuristic modifications of the above criteria which simplify the resulting detection architectures.

4.2.1.3 Compressive Sensing

The inherent sparse nature of the signal model can be exploited since the involved actors occupy a number of angular positions which cover a very low portion of the overall angular sector associated with the cell. In order to make clear this concept, let us consider the following scenario (see Figure 23):

- a **phased array** on the AN receiver pointing the UE and using a specific communication channel.
- the **UE** sharing the same channel as the AN.
- a **NLJ** transmitting within the UE channel.

It turns out that if we uniformly sample the angular sector covered by the 5G AN, only two directions are representative of the considered entities, whereas the other angular positions are free of any signal component except for the thermal noise which is ubiquitous. Thus, by properly modelling the above scenario, it is possible to shed light on its inherent sparse nature (see Figure 23 and Figure 24) and taking advantage of the compressive sensing (CS) approach. From an analytical point of view, denoting by $\boldsymbol{z} \in \mathbb{C}^{N\times 1}$, with N the number of sensors, the vector of received data, the model for a structured interferer is

$$\boldsymbol{z} = \boldsymbol{V}\boldsymbol{\alpha} + \boldsymbol{n},$$

where $\boldsymbol{V} \in \mathbb{C}^{N\times L}$ is the dictionary matrix with $L \gg N$ and $\boldsymbol{\alpha}$ is a sparse vector accounting for the magnitude associated with each direction.

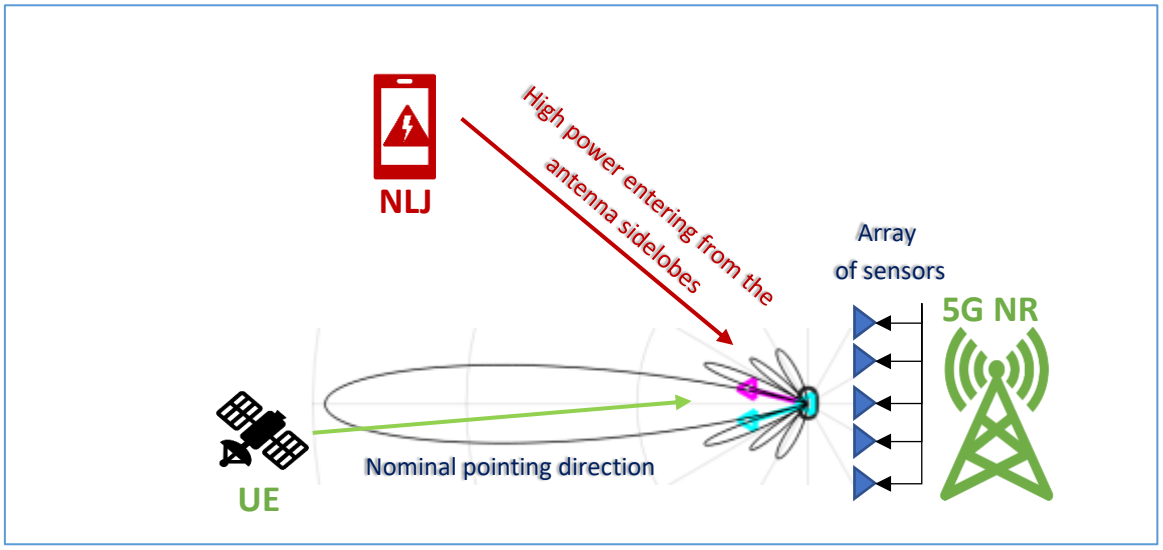

Figure 23. NLJ attack assuming an array of sensors at the AN receiver.

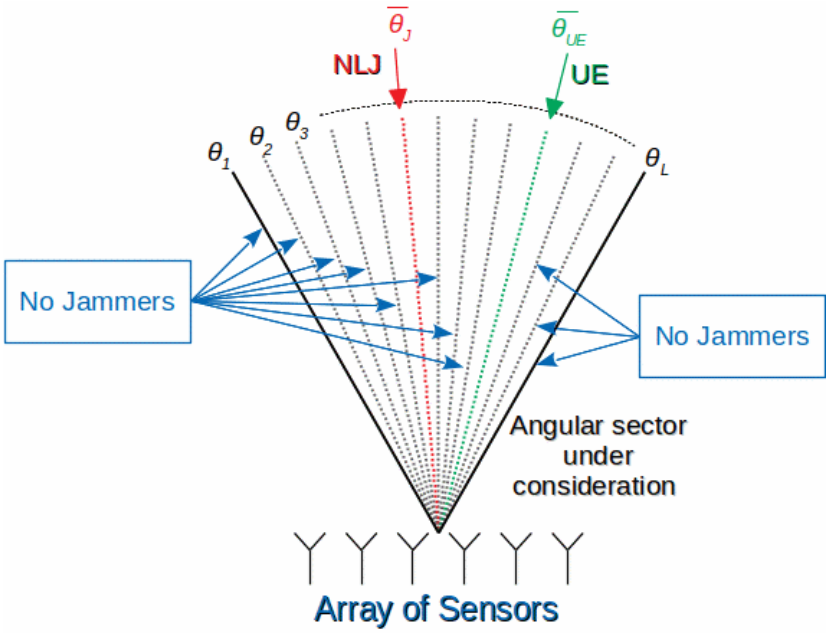

Figure 24. Sparse model for the scenario of Figure 23. NLJ attack assuming an array of sensors at the AN receiver.

In the case of noise jammer, the sparse nature moves to the structure of the data covariance matrix, namely

$$\boldsymbol{M} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{\dagger} + \sigma_n^2\boldsymbol{I},$$

where $\boldsymbol{D} \in \mathbb{R}^{L\times L}$ is a sparse diagonal matrix accounting for the power received from each direction and σ_n^2 is the thermal noise power. In this context, compressed sensing algorithms can be applied/devised to jointly perform jammer detection and, as a byproduct, jammer angular estimation. Specifically, sparsity-promoting priors can be assumed at the design stage and a joint Bayesian-Fisher estimation approach can be pursued as proposed in [51] with respect to the radar scenario. Once the jammer has been detected and its angular position estimated, the latter information can be used to design a set of weights for the receive antennas in order to place nulls along the interference directions and mitigate their effects (signal-processing-based mitigation techniques).

4.2.1.4 Clustering

Clustering algorithms can be used for anomaly detection or change detection [50] working in the original domain or in a transformed domain of the received waveform (Wavelet Transform, Wigner-Ville transform, etc.). As a matter of fact, these transformations may lead to an increase of the separability between the useful signal component features/parameters and those of the interference signal.

As for the change detection, this problem can be formulated in terms of the following binary hypothesis test [52]:

$$\begin{cases} H_0: & \text{data are statistically homogeneous,} \\ H_1: & \text{data can be partitioned into two homogeneous subsets,} \end{cases}$$

where, under H_1, the first subset has different statistical properties with respect to the second subset. In Figure 25, we show a case adhering the above problem and related to an NLJ attack which increases the power of the noise affecting the useful signal.

4.2.2 Measurement-based

Intentional interference detection/mitigation algorithms can be designed resorting also to compressed data, namely that actually represent the estimates provided by earlier stages in the LOCUS architecture (localization enablers) or the 5G network equipment. It is clear that, since a preliminary processing is performed on data, the available information is limited and, more importantly, affected by the drawback of the latter. Nevertheless, this reduced amount of information allows for computational efficiency and, hence, algorithms that can run fast without expensive hardware requirements.

In the next subsections, we describe three design strategies dealing with compressed data.

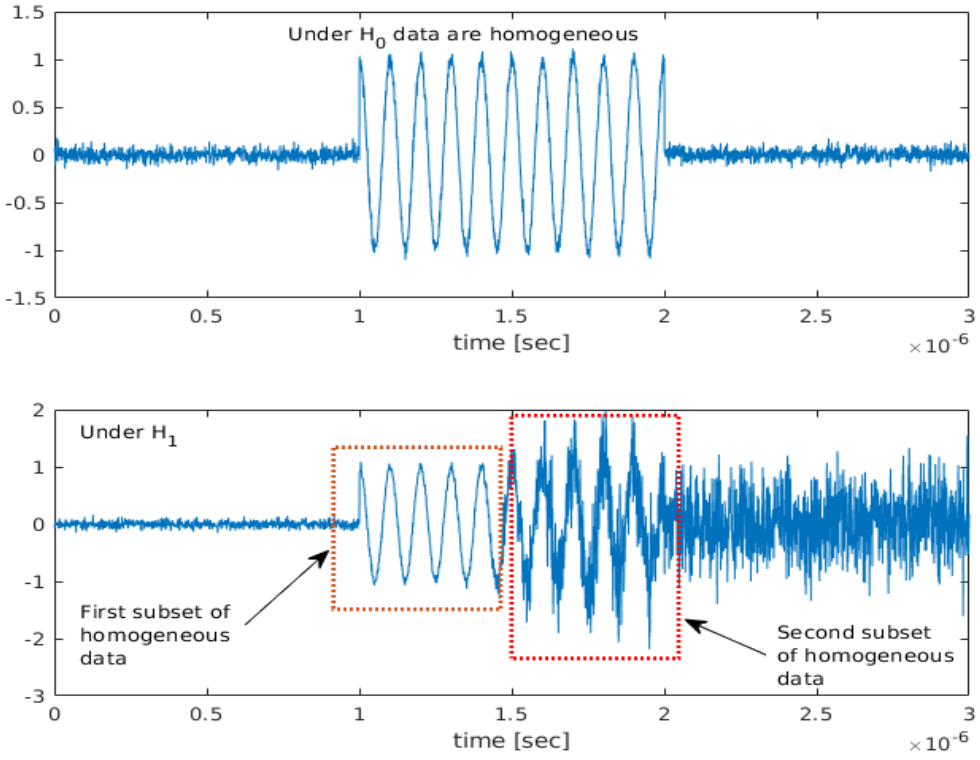

Figure 25. Received signal before an attack (subfigure on the top) and in the course of an attack (subfigure on the bottom).

4.2.2.1 Summary statistics

Monitoring estimation degradation can be effective to detect the presence of a malicious attack. Specifically, the estimates provided by the network (ToA/TDoA/AoA/RSS/Frequency measurements) could be analysed from a statistical point view in order to extract summary statistics (e.g., mean, standard deviation, kurtosis, skewness) which can represent the initial value of a tracking algorithm (e.g., Kalman filters and particle filters). The output of the tracking algorithm can be used to detect the presence of an interferer and infer the introduced bias in order to mitigate its effect by correcting the measurements. For instance, the covariance matrix of the estimation error can be used to define an uncertainty ellipsoid according to a given confidence level. Then, if the next measurement lies outside this region, it most likely represents an anomaly (see Figure 26).

An effective strategy to deal with low-accuracy measurements is the Track-Before-Detect (TBD) paradigm that processes batch of measurements allowing for energy integration [53]. The idea behind the TBD paradigm is sketched in Figure 27. However, the main drawback of TBD concerns the computational load or the latency before providing the first estimate. Moreover, additional information as environment maps can be incorporated into the parameter tracking algorithms in order to predict/correct the next measurements. Once an attack is detected and possibly mitigated, the location security function provides the localization functions (i.e., the localization technique developed in WP3) with an alarm of attack detection or directly with a corrected measurement.

Another source of information is represented by compressed data covariance matrix, that can feed a change detection algorithm [52]. In fact, the malicious platforms transmit high power inducing a discontinuity in the eigenvalues of the data covariance matrix. Such discontinuity can be caught through suitable tests on the sample covariance matrix [54].

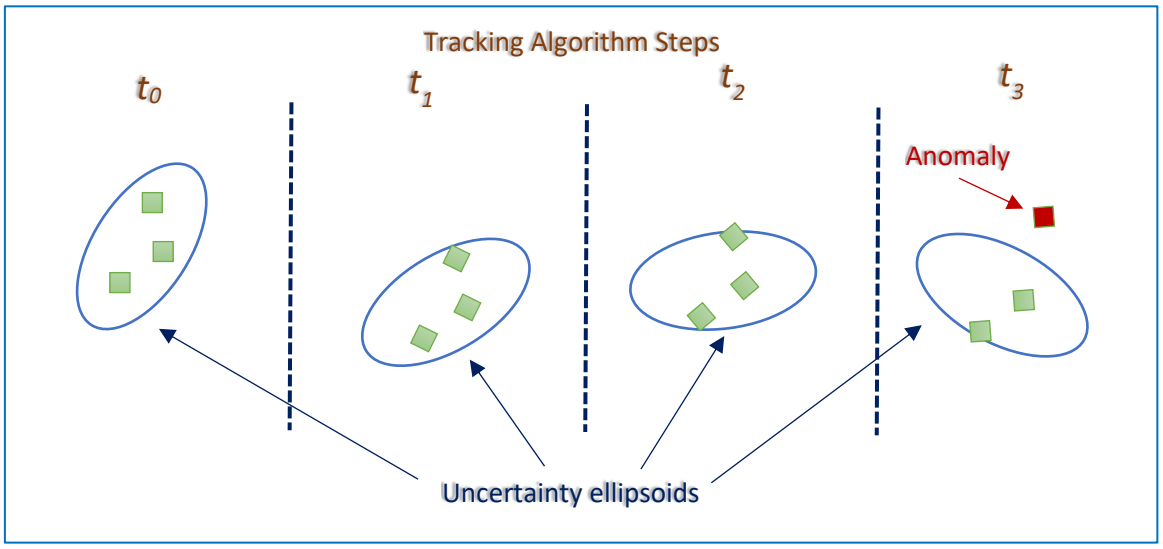

Figure 26. Steps of a tracking algorithm and anomaly detection.

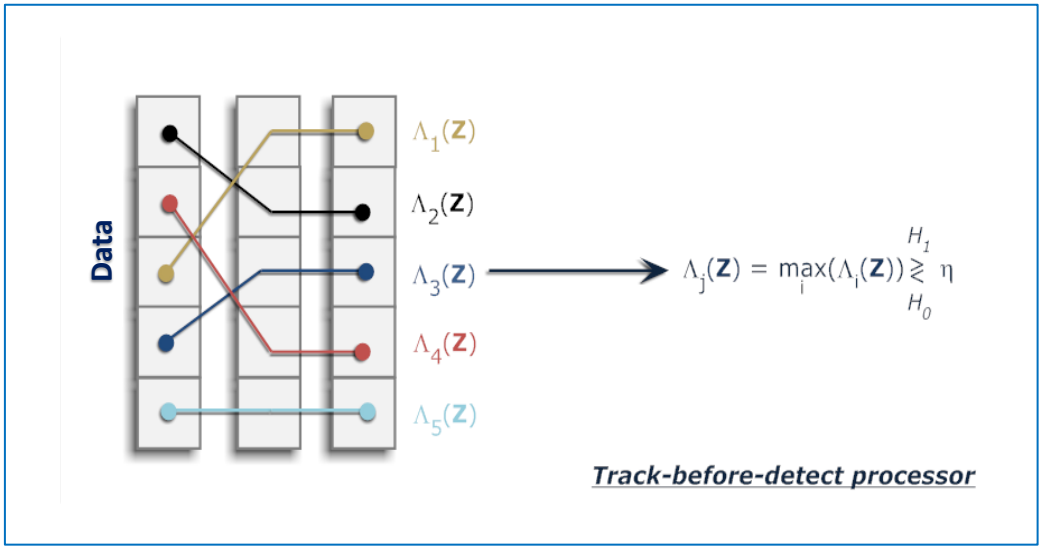

Figure 27. Track-Before-Detect paradigm: the algorithm exploits several temporally consecutive measurements.

4.2.2.2 Clustering

Under the hypothesis that the action of a malicious platform is not concurrent with the request of location for a given UE, the identification of degraded/erroneous measurements can be accomplished by partitioning the available measurements into subsets according to specific parameters. To this end, it would be possible to resort to the so-called latent variable model [55] that introduces hidden random variables representing the classes to which each data sample belongs. Then, the Expectation-Maximization (EM) algorithm [56] can be exploited to come up with suitable estimates of the unknown parameters.

Further techniques as Support Vector Machine, K-means, K-nearest neighbour, etc. are also available [57] possibly after transferring data into a suitable domain (see also Subsection 4.2.1.4). In this context, the number of clusters might be a priori obtained or estimated through the MOS rules, which overcomes the limitations of the maximum likelihood approach in the case of multiple nested hypotheses.

4.2.2.3 Information Theory

Localization techniques that rely on a set of possible values rather than on single value estimates will be developed in WP3 and employed in LOCUS. Such techniques belong to the paradigm of soft-information, where the localization process consists of a training phase and an online phase [58]. During the training phase, ToA, AoA, and RSS measurements are processed through unsupervised machine learning to perform density estimation (see Figure 28). During the online phase, the density estimate is evaluated based on the online measurements. This results in a likelihood of possible range and angle values. One or multiple

likelihoods can be jointly processed and updated with contextual information to obtain the final position estimate.

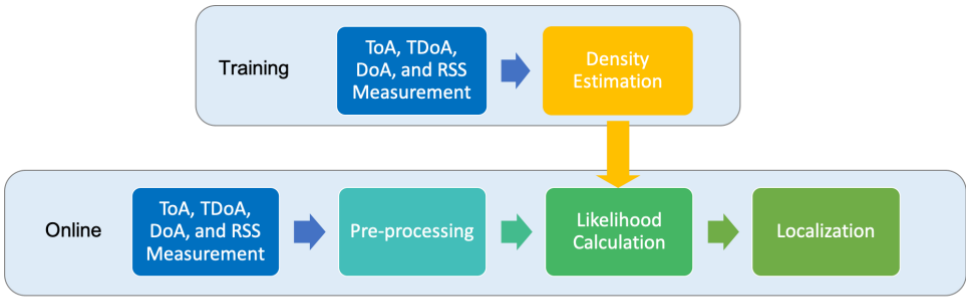

Figure 28 – Training and Online Phase for Localization via Soft-Information

In this context, the presence of a malicious attack on raw data will change the distribution of the measurements with respect to what was collected during the training case. Therefore, the density estimation can be performed multiple times during the online phase (see Figure 29). The comparison between the density estimate obtained during the training phase and the one obtained during the online phase can reveal if any significant distortion is experienced. In such a case, a test based upon information distance between the training density function and the density function computed from online data can be provided. Several metrics can be used as, for instance, the Kullback-Leibler divergence or the Bhattacharyya distance.

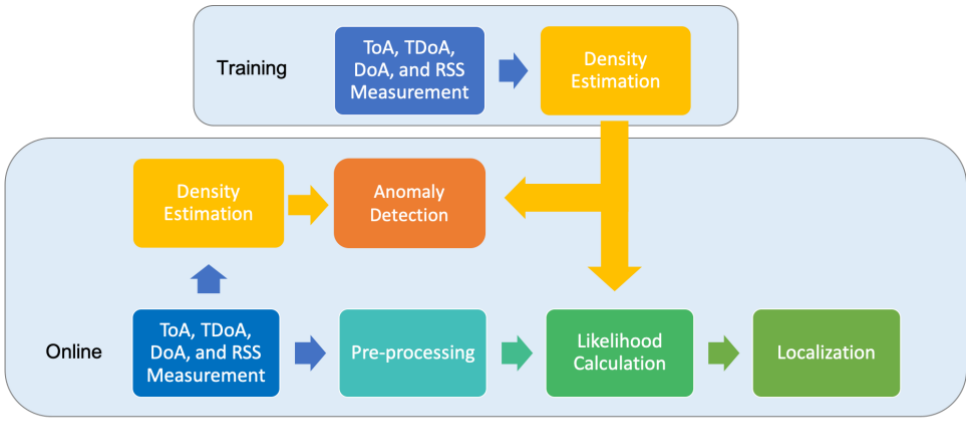

Figure 29 – Training and Online phase for Localization via Soft-Information with Anomaly Detection

4.3 Anti-spoofing (Non-3GPP)

There exists ongoing activity in 3GPP for security and privacy in 5G localization. Although the technical reports mention some key issue, they leaves space for contribution. In particular, TR-33.814 lists potential security and privacy issues in RAT-independent technologies, including Bluetooth and WLAN. However, it does not provide so far a study for Global Navigation Positioning System (GNSS)/Global Positioning System (GPS), which is instead presented in other reports and specifications from 3GPP.

GNSS/GPS is a positioning system widely used nowadays in our lives for real-time localization on Earth. This technology is highly vulnerable to spoofing/jamming attacks caused by malicious intruders. In recent years, there has been high concern about the security of GPS signals, since it has been demonstrated that GPS is vulnerable to spoofing and jamming attacks where attackers can introduce fake GPS [59] signals in the channel. Sophisticated and planned attacks to the GPS receivers can destabilize the economy of a country, besides endangering human lives.

Figure 30 illustrates at high level how a GPS attack can be performed. The malicious attacker only has to generate and transmit a fake GPS signal in the channel so that the power received in the target area is sufficiently higher than one received from the GPS satellite(s). In this way, the GPS receiver will take the fake signal as the legitimate one (GPS spoofing). This is easily plausible nowadays thanks to the quick rise of commodity and low-cost Software-Defined Radio (SDR) transceivers. For less than 300 euros is possible to acquire a full equipment of antenna and SDR with transmission capabilities being able to transmit a signal in a range of 200 meters. For all these reasons, GPS attacks are real, and must be taken as a real threat in wireless communications.

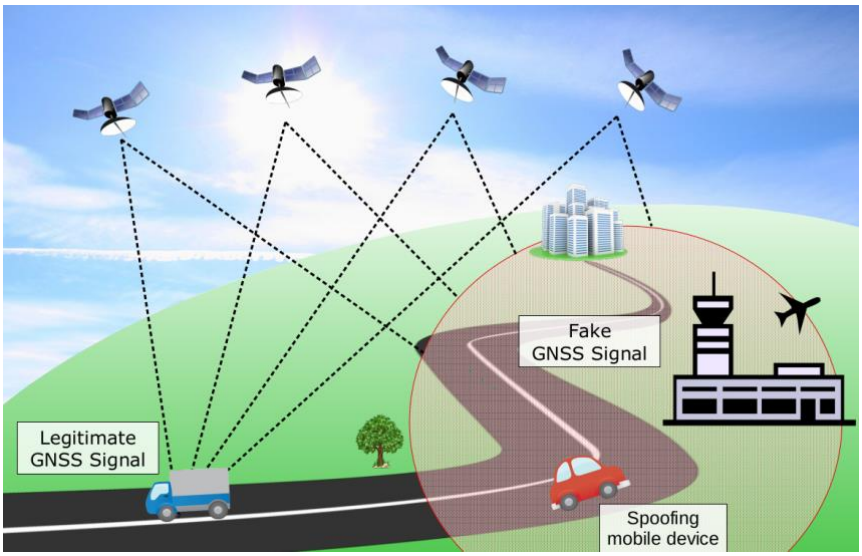

Figure 30 - A GNSS spoofing attacker can compromise wireless communications by impersonating the legitimate signal and affecting nearby infrastructures.

One typical GPS attack is jamming by transmission of any arbitrary signal at a relatively high power. This is relatively easy to detect, and it is usually done with the involvement of a hostile national authority. Another category of attacks is intentional spoofing of GPS signals. This can be done by giving a false location deliberately. In both cases attackers can be static or moving around the targets. We exploit the effect of imperfections in the spoofing signal to determine the presence of an anomaly.

A subset of the aforementioned attacks are not accurately detected by only looking at the SNR of the specific signal if the spoofing is sophisticated. The common factor among all GNSS technologies that we exploit to design a spoofing detection system is that *satellites are constantly orbiting the Earth along known trajectories*. This generates Doppler shift fingerprint of the signal.

The Doppler shift is the change in frequency of a signal caused because the transmitter (satellite in the legitimate case or spoofer in case of an attack) is moving. GNSS satellites are constantly moving along an orbital speed of about 14.000 km/hour. The latter creates a characteristic Doppler shift that can be observed in the GNSS receiver (as shown in Figure 31). Therefore, for every location on Earth at every time we do know how many satellites are visible and how their Doppler shift should look like. All this information can be used to fingerprint somehow the signal received for every satellite. A potential mobile GNSS attacker will try to inject a non-legitimate signal in the channel with the proper characteristics of a GNSS signal, but it will also add the Doppler shift created by its own movement towards or

around the sensors. We exploit the Doppler difference between the legitimate GNSS signal and the spoofing signal to determine the presence of a non-authorized transmitter.

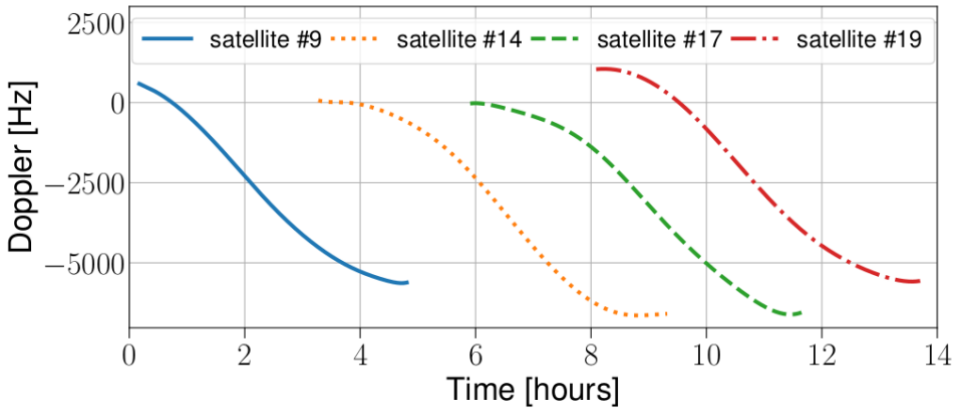

Figure 31 - Doppler shift detected by a GPS receiver observing different satellites of the constellation. [Madrid - (40.336994,-3.770459)]

We propose to use a deep learning approach to detect anomalies in the GNSS channel rather than traditional signal processing techniques. The latter requires the knowledge of the exact trajectories for the GNSS deployment under consideration, while a deep learning approach can directly exploit past Doppler measurements and the periodicity of the trajectories to train the system.

We rely on a LSTM RNN architecture to build a GNSS anomaly detector [60]. LSTM fits properly in our scenario since the Doppler shift of different satellites have similar patterns (see the past Figure) and the same satellite does not always have the same Doppler shift at the same time in consecutive days (as shown in Figure 32) due to the inclination of GPS orbits of about 55 degrees with the Earth's equator. Our model's input is the Doppler shift of the visible satellites. This input can be obtained by taking measures with a GNSS receiver or by using historical data provided by the GNSS satellite networks, e.g. NASA[7] releases daily data of GPS broadcast. This data can be used by specific software[8] to generate the simulated signal and Doppler for the GPS satellites in view.

[7] ftp://cddis.gsfc.nasa.gov/gnss/data/daily
[8] https://github.com/osqzss/gps-sdr-sim

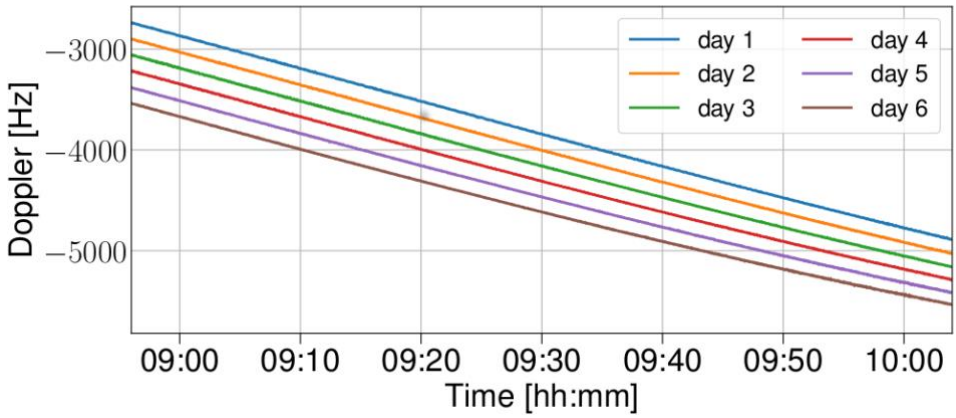

Figure 32 - Doppler shifts detected by a GPS receiver from the same satellite at same time interval along 6 adjacent days [Madrid - (40.336994, -3.770459)]

We build a model based on 2 LSTM layers with 32 neurons, 1 dropout and 1 dense layer. The prediction of the model is built on the knowledge of the previous 50 measurements (look-back) to predict the next state. We train the network in batches of 128. We rely on Keras framework to train and evaluate our model [61]. For the training dataset we use GPS Doppler shift measurements with 5 seconds of granularity recorded in L1 channel (1575.42~MHz) for 100 days using GNSS SDR software [62] and RTL-SDR. The training dataset could be also obtained by looking at the historical daily data provided by NASA.

The work exploits the predictability of the Doppler characteristics of the received GPS signals to determine the presence of anomalies or malicious attackers. As GNSS is an essential technology for 5G localization systems, it is deemed of paramount importance that its measurements can be considered reliable. The proof of concept for the spoofing detection is performed using Nvidia Jetson Nano and RTL-SDR. However, we envision that smartphones that support a similar framework could be able to infer the anomalies while the infrastructure would be responsible for learning the parameters of the neural network. In particular, tensor flow lite can run on smartphones, but it does not support LSTM model yet. However, this is on our roadmap.

Figure 33 shows the Doppler shift of one satellite for one-single day in a specific position in Madrid (Spain). The same figure also shows the Doppler shift forecast computed by our model. The error of our forecasting model over the testing dataset is in the order of 0.0006 %, which implies a Mean Absolute Error (MAE) of 4 Hz in the Doppler shift prediction. Therefore, this model is able to forecast the Doppler shift of a GPS satellite with a very small error which allows us to determine very accurately the presence of an anomaly in the GPS channel.

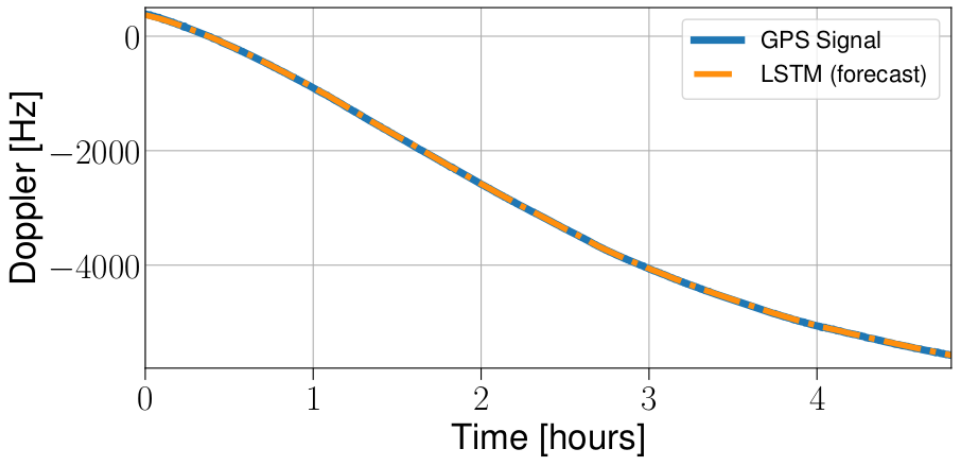

Figure 33 – LSTM forecast of Doppler shift of the satellite #3 of GPS constellation. [Madrid – (40.336994,-3770459)]

Then, we simulate the presence of a spoofing device that is moving at 20 km/h towards the sensor transmitting a non-legitimate GPS signal to slightly change the location decoded by the receiver. Figure 34 shows that the model is able to forecast the Doppler using the previous knowledge of the acquired data with a very small error. Then, we activate the simulated spoofing device (at minute 425 in the same Figure) which intends to fake the GPS signal. We can observe how this spoofing signal generates a slightly different Doppler shift due to its own movement and speed. This Doppler variation can be detected by our model, which immediately starts making Doppler forecasting using only predicted data (generated by the model) since the real Doppler shift acquired by the receiver is not trustable anymore. The lower the speed of the transmitter, the lower is the difference between the Doppler generated by the legitimate transmitter and the spoofing device. Our model is able to detect GPS anomalies by looking at the signal imperfections, detecting static spoofers with Doppler shift deviations higher than 4~Hz, and catching mobile spoofing devices at higher speeds of 2.5 km/h.

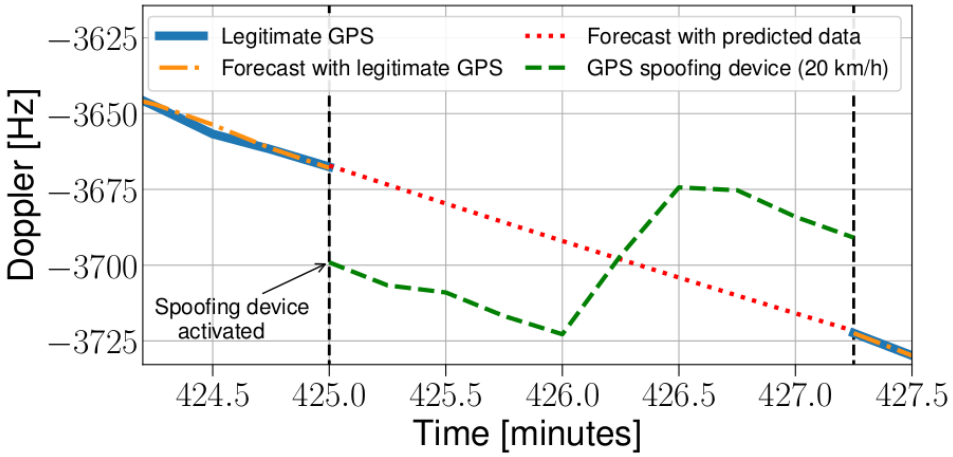

Figure 34 – Our model tracks the Doppler shift of the legitimate GPS signal. In the presence of a spoofing signal the model uses predicted data for forecasting since Doppler acquired by the receiver is not trustable anymore

5 Integrity in 3GPP Release 17

Until now, accuracy has been the main positioning performance metrics discussed and supported in 3GPP. However, emerging applications relying on high-precision positioning technology, demand high integrity and reliability, in addition to high accuracy.

Integrity is the measure of trust that can be placed on the correctness of information supplied by a navigation system. Integrity includes the ability of a system to provide timely warnings to user receivers in case of failure.

The 5G service requirements specified in TS 22.261 include also the need to determine the reliability, and the uncertainty or confidence level, of the position-related data [63] [64].

In RP-193237 [65], a new SI on "New SID on NR Positioning Enhancements" has been approved in which one of the objectives is to:

• *Study solutions necessary to support integrity and reliability of assistance data and position information:*

- *Identify positioning integrity KPIs and relevant use cases.*
- *Identify the error sources, threat models, occurrence rates and failure modes requiring positioning integrity validation and reporting.*
- *Study methodologies for network-assisted and UE-assisted integrity.*

Any use-case related to positioning in Ultra Reliable Low Latency Communication (URLLC) naturally requires high integrity performance. AS a matter of fact, use-cases in URLLC comprise V2X, autonomous driving, UAV (drones), eHealth, rail and maritime, emergency and mission critical scenarios. In use-cases in which errors can lead to serious consequences such as wrong legal decisions or wrong charge computation, etc., the integrity reporting becomes crucial.

In Figure 35, we try to illustrate the definition of accuracy, precision, validity, reliability, consistency, confidence level and integrity. Basically, we can conclude that accuracy is the same term as validity in positioning. Also, terms such as reliability, precision, certainty and confidence level can be used interchangeably. However, integrity requires the evaluation of both accuracy and reliability.

The integrity study in 3GPP Rel.17 would not be limited to RAT-independent positioning methods, mainly GNSS, in which the discussion has rooted from; the intention is that the integrity support would be also requested for RAT-based positioning methods.

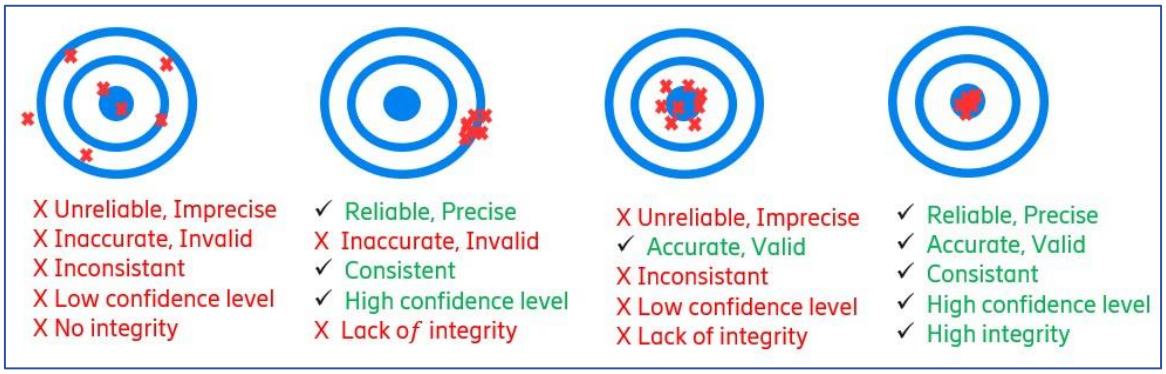

Figure 35 – Positioning integrity definition

There are already some integrity KPIs defined in the satellite navigation area which are likely to be soon discussed in the postponed 3GPP Rel.17 meeting in August. Below these KPIs are summarized:

- **Alert Limit (AL):** is the largest allowable error for safe operation.
- **Time to Alert (TTA):** is the maximum allowable elapsed time from the onset of a positioning failure until the equipment announces the alert.
- **Integrity Risk (IR):** is the probability of providing a signal that is out of tolerance without warning the user in a given period of time. The integrity risk is often assumed to include the Time to Alert, a sort of latency time left to the system before it detects a failure after it has occurred.
- **Protection Level (PL):** is the statistical error bound computed to guarantee that the probability of the absolute position error exceeding the said number is smaller than or equal to the target integrity risk.

In the LOCUS project, we intend to closely follow up the standardization progress of integrity support in Rel.17 and to hopefully contribute to the topic.

6 Conclusions

This document has presented the initial activities carried out in LOCUS for what concerns localization security and location privacy.

Location privacy has been concretely addressed in terms of ability for an attacker to disclose the real identity of nearby users. We experimentally proved the feasibility of IMSI-catching attacks using low cost software-defined radios, thus highlighting the various weaknesses that the 3GPP authentication process does expose. We concluded the privacy section with a preliminary study regarding mitigation techniques for privacy attacks. Our next activities will extend the experimental analysis to 5G systems, which, despite the new IMSI concealment solutions being standardized, we expect, will not clear the issues raised in this deliverable.

Then we addressed location security in terms of candidate countermeasures to noise-like jammers and meaconing attacks. A few selected solutions will be further investigated and studied continuing the on-going work of WP3 on signal processing techniques. LOCUS will leverage techniques for data encryption, data anonymization, and privacy preservation, so that all relevant actors can deal with private and secure data without compromising confidentiality, identities, and privacy of the users.

Finally, we briefly introduced the task of location or positioning integrity. Even if not originally foreseen in the project proposal (written before 3GPP Release 17 was chartered), this is a new concept which will be brought up in 3GPP Release 17, and the intention is to add other metrics on reliability of the positioning estimation, in addition to the well-known accuracy metrics which have been considered in previous releases on 4G and 5G 3GPP positioning issues. LOCUS will monitor and may contribute to this topic in the upcoming 3GPP Study Item in positioning.

The work initiated in this deliverable will be also harmonized with the specific approaches and solutions being proposed and investigated in other LOCUS work packages. In addition to analysing and improving security and resilience of the WP3 localization algorithms, we will focus also on robust, secure and privacy-preserving approaches for the management of location data management within the LOCUS platform, thus including the integration of privacy enhancement technologies within the WP5 analytics related tasks. We refer the reader to the WP5 deliverables for details on preliminary work already carried out in this direction (to avoid duplication of material, privacy aspects in location data analytics will be entirely documented in WP5).

7 References

[1] R. L. Villars, C. W. Olofson, M. Eastwood., “Big data:What it is and why you should care [J].,” *White paper, IDC,* 2011.

[2] X. F. Zheng, “Big data application and revelation offoreign telecom operators.,” *Mobile communications,* pp. 29-33, 2015.

[3] C. Bettini, S. Mascetti, D. Freni, X. Wang, and S. Jajodia, “Privacy and Anonymity in Location Data Management,” *Privacy-Aware Knowledge Discovery: Novel Applications and New Techniques.,* no. 10.1201/b10373-13, 2010.

[4] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, “Location-aware Communications for 5G Networks: How Location Information can Improve Scalability, Latency and Robustness of 5G,” *IEEE Signal Processing Magazine,* vol. 31(6), pp. pp.102-112, 2014.

[5] M. Koivisto, A. Hakkarainen, M. Costa, P. Kela, K. Leppanen and M. Valkama, “High-Efficiency Device Positioning and Location-Aware Communications in Dense 5G Networks,” *IEEE Communications Magazine,* vol. 55 (8), pp. pp.188-195, 2017.

[6] Schneier, B., “Inside Risks: Semantic Network Attacks,” in *Communications of the ACM*, 2000.

[7] S. Farhang, Y. Hayel and Q. Zhu, “PHY-layer location privacy-preserving access point selection mechanism in next-generation wireless networks,” in *IEEE Conference on Communications and Network Security (CNS)*, Florence, Italy, 2015.

[8] H. Khan, B. Dowling, and K. M. Martin, “"Identity Confidentiality in 5G Mobile Telephony Systems",” in *"Security Standardisation Research"*, Springer International Publishing, 2018.

[9] ““3GPP System Architecture Evolution (SAE)—Security Architecture” (Release 15),” technical specification (TS) 33.401, v15.2.0, September 2018.

[10] Zhenhua Li, Weiwei Wang, Christo Wilson, Jian Chen, Chen Qian, Taeho Jung, Lan Zhang, Kebin Liu, Xiangyang Li, and Yunhao Liu,, ““FBS-Radar: Uncovering Fake Base Stations at Scale in the Wild”,” Proceedings of the Internet Society Symposium on Network and Distributed System Security (NDSS), February 2017.

[11] Byeongdo Hong, Sangwook Bae, and Yongdae Kim,, ““GUTI Reallocation Demystified: Cellular Location Tracking with Changing Temporary Identifier”,” Proceedings of the

Internet Society Symposium on Network and Distributed System Security (NDSS), February 2018.

[12] A. Lilly, "IMSI catchers: hacking mobile communications.," *Network Security,* vol. 2, no. https://doi.org/10.1016/S1353-4858(17)30014-4, pp. 5-7, 2017.

[13] Shaik, A., Seifert, J., Borgaonkar, R., Asokan, N., Niemi, V., "Practical attacks against privacy and availability in 4g/lte mobile communication systems," in *23rd Annual Network and Distributed System Security Symposium NDSS*, San Diego, California, USA,, 2016.

[14] R. Borgaonkar, A. Shaik, N. Asokan, V. Niemi, and J.-P. Seifert., "LTE & IMSI Catcher Myths. In BlackHat," in *Europe'2015*, 2015.

[15] Jover, R.P., "LTE security, protocol exploits and location tracking experimentation with low-cost software radio," in *CoRR abs/1607.05171*, 2016.

[16] Rupprecht, D., Jansen, K., P¨opper, C., "Putting LTE security functions to the test: A framework to evaluate implementation correctness," in *10th USENIX Workshop on Offensive Technologies (WOOT 16)*, 2016.

[17] "OpenLTE: An open source 3GPP LTE implementation," https://sourceforge.net/projects/openlte/.

[18] Gomez-Miguelez, I., Garcia-Saavedra, A., Sutton, P.D., Serrano, P., Cano, C., Leith, D.J., "srsLTE: an open-source platform for LTE evolution and experimentation," arXiv preprint arXiv:1602.04629, 2016.

[19] "gr-LTE: GNU Radio LTE receiver," https://github.com/kit-cel/gr-lte.

[20] "Open Air Interface: 5G software alliance for democratising wireless innovation," http://www.openairinterface.org.

[21] "3GPP, "Security Architecture and Procedures for 5G System" (Release 15)," technical specification (TS) 33.501, v15.5.0., September 2018.

[22] Ravishankar Borgaonkar, Andrew Martin, Shinjo Park, Altaf Shaik, Jean-Pierre Seifert., "White-Stingray: Evaluating IMSI Catchers Detection Applications.," https://ora.ox.ac.uk/objects/uuid:15738ed0-c144-49e9-a4fa-466362cf7754, 2017.

[23] Paget., C., "Practical cellphone spying," in *DEFCON*, https://media.defcon.org/DEFCON18, 2010.

[24] Y. Arjoune and S. Faruque, “Smart Jamming Attacks in 5G New Radio: A Review,,” in *10th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA,, 2020.

[25] O. Osanaiye, A. Alfa, and G. Hancke,, ““A statistical approach to detect jamming attacks in wireless sensor networks”,” *Sensors,* Vols. vol. 18, no. 6,, pp. p. 1691,, 2018.

[26] O. Punal, I. Aktas¸, C.-J. Schnelke, G. Abidin, K. Wehrle, and J. Gross,, “Machine learning-based jamming detection for ieee 802.11: Design and experimental evaluation,” in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2014.

[27] T. Erpek, Y. E. Sagduyu, and Y. Shi, “Deep learning for launching and mitigating wireless jamming attacks,” *IEEE Transactions on Cognitive Communications and Networking,* vol. 5, no. 1, p. pp. 2–14, 2018.

[28] S. R. Hussain, M. Echeverria, A. Singla, O. Chowdhury, and E. Bertino,, “Insecure connection bootstrapping in cellular networks: The root of all evil,” in *Conf. Secur. Privacy Wireless Mobile Netw.,*, May 2019.

[29] Gustafsson, Fredrik; Gunnarsson, Fredrik, “Mobile Positioning Using Wireless Networks,” *IEEE Signal Processing Magazine,* vol. 22, no. 4, pp. 41-53, 2005.

[30] M. Liyanage, I. Ahmad, A. Bux Abro, A. Gurtov and M. Ylianttila, A Comprehensive Guide to 5G Security, John Wiley & Sons, 2018.

[31] D. Margaria, B. Motella, M. Anghileri, J.-J. Floch, I. Fernandez-Hernandez and M. Paonni, “Signal Structure-Based Authentication for Civil GNSSs,” *IEEE Signal Processing Magazine,* vol. 34, no. 5, pp. 27-37, 2017.

[32] L. Heng, J. J. Makela, A. D. Domínguez-García, R. B. Bobba, W. H. Sanders and G. X. Gao, “Reliable GPS-based timing for power systems: A multi-layered multi-receiver architecture,” in *2014 Power and Energy Conference at Illinois (PECI)*, Champaign, IL, USA, 2014.

[33] O. Alp Topal, S. Gecgel, E. M. Eksioglu and G. K. Kurt, “Identification of smart jammers: Learning-based approaches using wavelet preprocessing,” *Elsevier Physical Communication,* 2020.

[34] F. M. Aziz, J. S. Shamma and G. L. Stüber, “Jammer-Type Estimation in LTE With a Smart Jammer Repeated Game,” *IEEE Transactions on Vehicular Technology,* vol. 66, no. 8, pp. 7422-7431, 2017.

[35] K. Grover, A. Lim and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: a survey," *International Journal of Ad Hoc Ubiquitous Computing,* vol. 17, no. 4, p. 197–215, 2014.

[36] H. L. V. Trees, Optimum Array Processing: Part IV Detection, Estimation, and Modulation Theory, John Wiley & Sons, 2004.

[37] M. Lichtman, T. Czauski, S. Ha, P. David and J. H. Reed, "Detection and Mitigation of Uplink Control Channel Jamming in LTE," in *2014 IEEE Military Communications Conference*, Baltimore, MD, 2014.

[38] R. D. Pietro and G. Oligeri, "Jamming mitigation in cognitive radio networks," *IEEE Network,* vol. 27, no. 3, pp. 10 - 15, 2013.

[39] F. Slimeni, B. Scheers, Z. Chtourou and V. Le Nir, "Jamming mitigation in cognitive radio networks using a modified Q-learning algorithm," in *2015 International Conference on Military Communications and Information Systems (ICMCIS)*, Cracow, 2015.

[40] K. Firouzbakht, G. Noubir and M. Salehi, "On the Performance of Adaptive Packetized Wireless Communication Links Under Jamming," *IEEE Transactions on Wireless Communications,* vol. 13, no. 7, pp. 3481-3495, 2014.

[41] X. Liu, G. Noubir, R. Sundaram and S. Tan, "PREAD: Foiling Smart Jammers Using Multi-Layer Agility," in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, Barcelona, 2007.

[42] Y. Arjoune, F. Salahdine, S. Islam, E. Ghribi and N. Kaabouch, "A Novel JammingAttacks Detection Approach Based on Machine Learning for Wireless Communicatio," in *The 34thInternational Conference on Information Networking (ICOIN 2020)*, Barcelona, 2020.

[43] M. Spuhler, D. Giustiniano, V. Lenders, M. Wilhelm and J. B. Schmitt, "Detection of Reactive Jamming in DSSS-based Wireless Communications," *IEEE Transactions on Wireless Communications,* vol. 13, no. 3, pp. 1593-1603, 2014.

[44] S. M. Razavizadeh, M. Ahn and I. Lee, "Three-Dimensional Beamforming: A new enabling technology for 5G wireless networks," *IEEE Signal Processing Magazine,* vol. 31, no. 6, pp. 94-101, 2014.

[45] J. Vinogradova, E. Björnson and E. G. Larsson, "Detection And Mitigation Of Jamming Attacks In Massive Mimo Systems Using Random Matrix Theory," in *IEEE 17TH Int. Workshop on Sig. Proc. Adv. in Wireless Communications (SPAWC)*, Edinburgh, 2016.

[46] Project, 3rd Generation Partnership, "Technical Specification Group Services and System Aspects Study on 5G Security Enhancement against False Base Stations," 2019.

[47] H. Akhlaghpasand, S. M. Razavizadeh, E. Björnson and T. T. Do, “Jamming Detection in Massive MIMO Systems,” *IEEE Wireless Communications Letters,* vol. 7 , no. 2, pp. 242-245, 2018.

[48] W. Xu, “Detection of Pilot Spoofing Attack in Massive MIMO Systems,” in *IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019.

[49] T. S. G. R. A. Network, “Study on NR positioning support (Release 16),” 3GPP TR 38.855 V16.0.0 , 2019.

[50] S. Kay, Fundamentals Of Statistical Processing, Volume 2: Detection Theory, Pearson Education, 2009.

[51] L. Yan, P. Addabbo, Y. Zhang, C. Hao, J. Liu, J. Li and D. Orlando, “A Sparse Learning Approach to the Detection of Multiple Noise-Like Jammers,” *IEEE Transactions on Aerospace and Electronic Systems,* 2020.

[52] D. Orlando, “A Novel Noise Jamming Detection Algorithm for Radar Applications,” *IEEE Signal Processing Letters,* vol. 24, no. 2, pp. 206-210, 2017.

[53] F. Ehlers, D. Orlando and G. Ricci, “A Batch Tracking Algorithm for Multistatic Sonars,” *IET Radar, Sonar and Navigation,* vol. 6, no. 8, pp. 746-752, 2012.

[54] R. J. Muirhead, “Aspects of Multivariate Statistical Theory,” John Wiley & Sons, 2009.

[55] P. Addabbo, S. Han, D. Orlando and G. Ricci, “Learning Strategies for Radar Clutter Classification,” *arXiv : 2004.08277.*

[56] A. P. Dempster, N. M. Laird and D. B. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 39, no. 1, pp. 1-38, 1977.

[57] S. Theodoridis, Machine Learning: A Bayesian and Optimization Perspective, Academic Press, 2015.

[58] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey and M. Z. Win, “ Soft Information for Localization-of-Things,” *Proceedings of the IEEE,* vol. 107, no. 11, pp. 2240-2264,, 2019.

[59] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, “Unmanned aircraft capture and control via gps spoofing,” *Journalof Field Robotics,* Vols. vol. 31, no. 4, p. pp. 617–636, 2014.

[60] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, “Long short term mem-ory networks for anomaly detection in time series,” in *Presses universitaires de Louvain*, 2015.

[61] Geron, A., “Hands-On Machine Learning with Scikit-Learn, Keras, andTensorFlow,” *O’Reilly Media,* 2019.

[62] C. Fernandez-Prades, J. Arribas, P. Closas, C. Aviles, and L. Esteve, “Gnss-sdr: An open source tool for researchers and developers,” in *24th International Technical Meeting of TheSatellite Division of the Institute of Navigation (ION GNSS)*, 2011.

[63] ESA, “RP-192981: Motivation to study integrity in the position domain,” RAN#86, 2019.

[64] Deutsche, Telekom Swift Navigation and, “RP-192750: Motivation for Study on Positioning Integrity,” RAN#86, 2019.

[65] Qualcomm, “RP-193237: New SID on NR Positioning Enhancements,” RAN#86, 2019.