SigN: SIMBox Activity Detection Through Latency Anomalies at the Cellular Edge

Anne Josiane Kouam, Aline Carneiro Viana, Philippe Martins, Cédric Adjih, Alain Tchana

SigN: SIMBox Activity Detection Through Latency Anomalies at the Cellular Edge

Anne Josiane Kouam
TU Berlin, Germany

Aline Carneiro Viana
INRIA, France

Philippe Martins
Telecom Paris, France

Cedric Adjih
INRIA, France

Alain Tchana
Grenoble INP, France

ABSTRACT

Despite their widespread adoption, cellular networks face growing vulnerabilities due to their inherent complexity and the integration of advanced technologies. One of the major threats in this landscape is Voice over IP (VoIP) to GSM gateways, known as *SIMBox* devices. These devices use multiple SIM cards to route VoIP traffic through cellular networks, enabling international bypass fraud with losses of up to $3.11 billion annually. Beyond financial impact, *SIMBox* activity degrades network performance, threatens national security, and facilitates eavesdropping on communications. Existing detection methods for *SIMBox* activity are hindered by evolving fraud techniques and implementation complexities, limiting their practical adoption in operator networks.

This paper addresses the limitations of current detection methods by introducing *SigN*, a novel approach to identifying *SIMBox* activity at the cellular edge. The proposed method focuses on detecting *remote SIM card association*, a technique used by *SIMBox* appliances to mimic human mobility patterns. The method detects latency anomalies between *SIMBox* and standard devices by analyzing cellular signaling during network attachment. Extensive indoor and outdoor experiments demonstrate that *SIMBox* devices generate significantly higher attachment latencies, particularly during the authentication phase, where latency is up to 23 times greater than that of standard devices. We attribute part of this overhead to immutable factors such as LTE authentication standards and Internet-based communication protocols. Therefore, our approach offers a robust, scalable, and practical solution to mitigate *SIMBox* activity risks at the network edge.

KEYWORDS

Cellular signaling, Network attachment, Cellular authentication

1 INTRODUCTION

Cellular networks provide digital communications for more than five billion people around the globe. However, their accessibility to the general public, inherent complexity, and integration of multiple advanced technologies have exposed these networks to numerous attacks, which have significantly increased over the past decades.

In this context, Voice over IP (VoIP) to GSM gateways, also known as *SIMBox*, are a significant source of security challenges within cellular networks. *SIMBox* appliances bridge two telecommunication technologies by converting VoIP traffic to traditional GSM cellular networks. This allows them to route calls initiated over the internet through cellular networks by re-originating them from one of their multiple SIM cards.

Although SIMBox appliances may have legitimate uses, such as reducing telecommunication costs or automating calls in dedicated companies, *this paper highlights that their potential for misuse poses significant security risks*. Indeed, *SIMBox* appliances are at the basis of international bypass frauds in cellular networks, recognized as one of the top four phone system frauds causing substantial losses to mobile network operators [10]. As depicted in Fig. 1, International bypass fraud, or simply *SIMBox* fraud, involves intercepting international mobile calls routing and diverting them through an internet flow (VoIP) to a *SIMBox* in the destination country. The *SIMBox* then re-originates the received VoIP traffic as a local mobile call from one of its SIM cards to the receiving party. Fraudsters bypass the regular interconnect operator, avoiding international termination charges by paying the lower local call termination charges, thus profiting from the difference.

Therefore, beyond the *growing revenue loss for operators*, estimated at 2.7 billion in 2019 [9] and 3.11 billion in 2021 [10][1], *SIMBox* usage *negatively impacts network quality* for legitimate consumers and *compromises national security*. Specifically, *SIMBox* fraud degrades the quality of experience for consumers due to call initiation delays and network unavailability, which in turn increases churn. Moreover, *SIMBox*'s re-originated calls introduce bias into operators' network usage records, distorting call origins and locations and affecting various analyses and research[27]. More critically, *SIMBox* usage enables international attackers to masquerade as national subscribers, a vulnerability that could be exploited for covert operations, including by terrorists. Furthermore, *SIMBox* appliances provide attackers with the ability to eavesdrop on international phone conversations [18], endangering user privacy and facilitating international espionage.

As a result, investigations into detecting *SIMBox* activity in cellular networks have gained the attention of researchers. The objective is to provide mobile operators with means to detect and regulate *SIMBox* usage on their networks by implementing legal registration for legitimate *SIMBox* operations (as exemplified in [36]) while blocking undeclared usage. Such investigations are typically conducted at the destination operator level, where fraud occurs. The most common approach involves analyzing network users' cellular activity traces to differentiate between *SIMBox* patterns and legitimate ones.

Authors' addresses: Anne Josiane Kouam, TU Berlin, Germany,; Aline Carneiro Viana, INRIA, France,; Philippe Martins, Telecom Paris, France,; Cedric Adjih, INRIA, France,; Alain Tchana, Grenoble INP, France,

[1]The CFCA's 2023 survey summary [11] indicates a 12% increase in global telecom fraud losses compared to 2021, highlighting the continuing rise in the economic impact of fraud. However, the full report is not publicly available for verification.

Most detection methods from the literature [20, 25, 26, 31, 32, 37, 38] extract the spatiotemporal communication behavior of each SIM card by relying on *Call Detail Records (CDR) traces*. CDRs are time-stamped and geo-referenced recordings of mobile device-generated events (i.e., call, text, data) collected by network operators. SIM cards used within *SIMBox* appliances tend to exhibit automated behavior, distinct from human or natural patterns, characterized by low mobility, repetitive calls at odd hours, or many contacts, as noted in [26]. Such literature contributions have demonstrated excellent detection performance (i.e., an average accuracy of 94.5%) and are implemented offline, leveraging historical data collected at the network core without impacting network performance.
However, *SIMBox* appliances currently available on the market offer functionalities that enable fraudsters to automatically mimic more advanced and human-like behavior, thereby evading CDR-based *SIMBox* activity detection [23].

Conversely, a few contributions focus on the cellular edge, proposing real-time monitoring of users' network activity to detect *SIMBox* patterns. These analyses include monitoring *call audio* quality [30] and speakers' voices [12] to identify potential degradation due to *SIMBox* routing. More recently, *cellular signaling data* has been leveraged [28] to create device model fingerprints and suggest an access-control-list prevention methodology.
Unfortunately, these approaches often overlook the computational challenges of cellular-edge-based deployment, which affects their practical relevance. Since these solutions must operate efficiently across the hundreds to thousands of cell towers comprising the cellular network edge, they must provide reliable indicators for detecting *SIMBox* patterns while minimizing the computational resources needed to process them network-wide. This scalability challenge remains unresolved and explains their limited practical adoption (cf. §3): e.g., *call-audio*-based solutions require monitoring all local calls across the network. Similarly, signaling-based fingerprinting necessitates maintaining an exhaustive list of device fingerprints at each base station for regular consultation.

This paper aims to bridge the gap posed by the limitations of the current solutions: *It addresses the detection of SIMBox patterns remaining undetected through CDR-based detection or overlooking computational costs, and proposes a novel and practical approach to unmask SIMBox activities at the cellular edge.* The proposed approach, i.e., named *SigN*, identifies and leverages an indicator of *SIMBox* activity: the *remote SIM card association. Remote SIM card association* is a ground technology used in the *SIMBox* to mimic human mobility pattern. It allows fraudsters to avoid the resource-intensive and easily-detected movements of *SIMBox* appliances by enabling the binding of a SIM card to a distant gateway (with cellular antenna), as depicted in Fig. 2. To the best of our knowledge, the *SIMBox* is the only system capable of physically separating the SIM card from the cellular antenna. *Remote SIM card association* is thus a distinct signature of *SIMBox* activity resulting in decoupled network devices, unlike traditional coupled devices (e.g., phones).

SigN precisely detects such *SIMBox*-decoupled network devices at their attachment to the network by analyzing their generated cellular signaling at the network edge. By especially characterizing the *latency of devices' signaling*, we provide empirical evidence of *a significant dissimilarity between SIMBox-decoupled devices and coupled ones during the network attachment.*

This includes the following contributions, outlined in Fig. 5:

- We set up inside a Faraday shield, realistic urban settings of an operator 4G/LTE radio access and core networks with specialized equipment. Indeed, 4G is the most widespread cellular technology, particularly in the developing countries where *SIMBox* activity is the most striking, with 5G still several years away. Our testbed provides real-time access to the cellular edge network attachment signaling from 12 phones and 7 LTE *SIMBox* appliances from two major manufacturers, collected at the base station (cf. **§5.1**).
- We make the first literature empirical characterization of network attachment latency, to the best of our knowledge (cf. **§5.2**). We report *SIMBox* decoupled devices generate at least 5 times more latency than standard phones, particularly during the *authentication phase* where their minimum latency is 23 times higher.
- We explain the latency overhead by analyzing the interactions between the SIM card and Mobile Equipment during the *authentication phase* for a SIMBox decoupled device compared to standard devices. Our investigation shows that the authentication latency in *SIMBox* decoupled devices is influenced by unavoidable factors, such as LTE authentication standards and Internet-based communication protocols and vagaries. Despite optimizations, this latency cannot match that of streamlined, legitimate devices, maintaining a clear distinction between *SIMBox* decoupled devices and their coupled counterparts (cf. **§5.3**).
- On the other hand, we show through data collection in an actual operator network that a standard phone cannot reach such high authentication latency peaks regardless of the network signal conditions (cf. **§5.4**). Our empirical findings confirm that *authentication latency* is a reliable, practical, and robust metric for distinguishing *SIMBox* decoupled devices from coupled ones.
- Based on this empirical evidence, we demonstrate in **§6** the practicality of *SigN* by introducing a novel method to monitor authentication latency at the cellular edge *without added overhead.* Through latency distribution analysis, *SigN* identifies devices with *consistently unusual latency values, enabling operators to promptly investigate potential threats.* Therefore, our statistical analysis shows that *SigN* achieves near-perfect accuracy in detecting *SIMBox*-decoupled device attachments.
- To ensure the reproducibility of our results, we have released the *SigN* datasets and code *at this anonymous link.*

Additionally, **§2** provides the background for our work, **§3** discusses the motivation, and **§4** outlines the threat model and our defense objectives. Finally, we conclude in **§8**. Readers can refer to the appendix for a list of acronyms used throughout the paper.

2 SETTING THE STAGE

This section outlines the background for our research, covering the cellular network ecosystem and *SIMBox* architecture.

2.1 Cellular networks

We overview several aspects of the 4G cellular networks, the most widespread cellular technology, particularly in the developing world where *SIMBox* activity is the most prevalent.

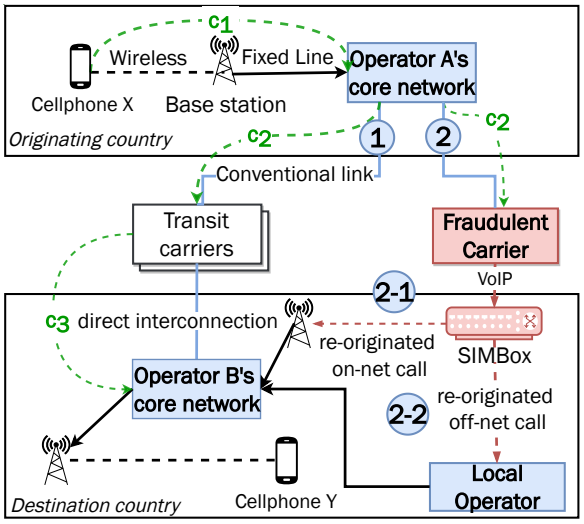


Figure 1: International call routing: (Flow 1) Legitimate scheme, (Flow 2) Fraudulent scheme.

Architecture. The cellular network infrastructure consists of end devices, also known as User Equipment or UE (e.g., phone, tablet), base stations, and the core network. A User Equipment (UE) is a mobile device registering to the network to receive access to communication services. It comprises two distinctive elements, the Mobile Equipment (ME) and the SIM card provided by the network operator. Base stations, called eNodeB in 4G networks, are intermediate connectors responsible for the radio transmission with the devices. At last, the core network handles administrative tasks such as the devices' authentication, security, and mobility management, intending to provide permanent service access.

The network attachment signaling procedure establishes a connection between end devices and the network. It occurs in four circumstances: when a device is powered on, when it moves into a new tracking area, when it loses connection with the network, or at a network trigger. In 4G, the network attachment (cf. Fig. 5, step 2) consists of several steps aiming (i) the acquisition of the device identity, i.e., International Mobile Subscriber Identity (IMSI), (ii) the mutual device and network authentication, (iii) the Non-Access Stratum (NAS) security setup, (iv) the device location update, and (v) the Evolved Packet System (EPS) session establishment [3].

The SIM card binds the mobile subscription to the network device. It securely stores the IMSI subscriber identifier and a secret symmetric key called the subscriber key (or K_i, in short) used in steps (ii) and (iii) of the network attachment. It also represents an environment protected from attackers where the network authentication and security algorithms are run following the AKA protocol [1].

2.2 *SIMBox* architecture and fraud

***SIMBox* fraud scheme.** As depicted in Fig. 1, *SIMBox* fraud interferes with the regular international voice call routing (*flow 1*). In a regular routing, the call traffic leaves the caller's mobile operator (Operator A) and is routed to the destination country through a set of transit operators. The traffic is received directly by the called party's operator (Operator B), who terminates it. Nevertheless, a transit carrier can be fraudulent. Indeed, transit carriers perform traffic interconnection between countries by buying and reselling international termination routes. A fraudulent carrier instead diverts the traffic it receives through a low cost VoIP trunk, as in the *flow 2* on Fig. 1. The diverted traffic is sent to a *SIMBox* (VoIP to GSM gateway) located in the destination country and re-originated as a national mobile call to its recipient. Once in the destination country, there are two possible fraudulent termination scenarios: (i) 2-1 is an on-net termination when the re-originated call is made using a SIM card of Operator B, the same operator of the called party, and (ii) 2-2 is an off-net termination when the fraudster uses a SIM card from a different local operator in the destination country.

The *SIMBox* operates as a VoIP GSM gateway. It receives a diverted call traffic as a VoIP client and terminates it by re-originating a cellular mobile call using one of its numerous SIM cards. The *SIMBox* continuously creates network devices by associating SIM cards and GSM modules (providing wireless link to the network). The *SIMBox* includes three kinds of hardware components:

- The *gateway* is a rack with a set of GSM modules maintaining the wireless communication inside a given cellular frequency range (i.e.,2G/3G/4G). It receives incoming VoIP traffic and distributes it to the GSM modules for termination as mobile calls. The gateway plays the role of Mobile Equipment for the formed *SIMBox* devices. Hence, the recorded network location of a *SIMBox* device is the location of its belonging gateway. Most gateways in the market include SIM slots for operation.
- The *SIMBank* is an appliance with numerous SIM slots that remotely holds a bundle of SIM cards (e.g., 128 in the SMB128 model [19]). It manages *SIMBox* SIM cards, including their addition, removal, and data transfer.
- The *control server* is a web server providing the *SIMBox* control functions, i.e., binding of SIM cards to GSM modules and architecture configuration. It can be hosted online to ease remote access from a web client.

Distributed *SIMBox* architecture involves the interaction of such appliances over an IP network using TCP or UDP protocols. Hence, as shown in Fig. 2, *SIMBox* devices formation can be done through *local SIM card association* if the SIM card is in the same appliance as its associated GSM module, or *remote SIM card association* if the SIM card is from another appliance, i.e., the SIMBank. *Local SIM card association* results in *coupled UEs*, while *remote SIM card association* yields *decoupled UEs*.

3 UNRAVELING LITERATURE GAPS

Despite the significant impact of *SIMBox* activity, it has received limited attention in the literature, with only 15 detection methods proposed since 2011. We categorize these contributions based on the type of cellular network data they use and how it is processed. Specifically, we distinguish between methodologies that rely on core network data (e.g., Call Detail Records/CDRs, §3.1) and those that use edge network data (e.g., call audio and cellular signaling, §3.2). This section highlights the strengths and weaknesses of current *SIMBox* detection approaches and positions *SigN* to address the

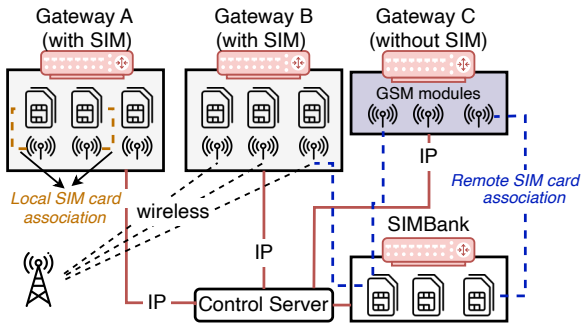


Figure 2: Example of a *SIMBox* distributed architecture.

remaining gaps. For a complete survey of *SIMBox* fraud solutions before 2021, refer to [22].

3.1 Network-core-based detection

Literature approaches operating at the network core [20, 25, 26, 31, 32, 37, 38] aim to identify *SIMBox* activity from SIM cards' communication and mobility behavior extracted from CDRs datasets. These methods rely on various features (e.g., #calls at night, #contacts, #incoming calls, #locations) to distinguish SIM cards used for *SIMBox* termination from SIM cards used by genuine consumers. *Such contributions have demonstrated high accuracy, averaging 94.5%, in detecting SIMBox activity characterized by unusual patterns in communication or mobility, as in [21, 25, 26]: numerous outgoing calls, few stay points, SIM card clusters, or no incoming calls.*

Unfortunately, *SIMBox* devices have evolved with functionalities that automatically mimic human behavior in CDR datasets, known as Human Behavior Simulation (HBS) [22]. HBS techniques enable *SIMBox* devices to operate while maintaining human-like behavior in terms of communication and mobility. In communication, this is achieved by thresholding the number and duration of initiated calls and controlling their contacts and timing. For mobility, fraudsters use *remote SIM card association*, binding a SIM card to a remote gateway (i.e., Mobile Equipment), resulting in an erroneous network recording of the SIM card location. Notably, the automatic binding of a *SIMBox* SIM card to gateways in different locations at various times creates human-like movements between network cells in CDRs at no cost to fraudsters.

Recent research [23] empirically examines the performance of HBS-generated *SIMBox* patterns compared to CDR-based *SIMBox* activity detection. The results indicate that the current HBS functionalities produce *SIMBox* patterns that closely mimic human behavior, enabling them to evade detection by CDR-based methods with a high degree of success. *This finding underscores the limitations of CDR-based approaches, which are insufficient to unmask all existing SIMBox patterns.*

Insight: *The wide adoption of HBS techniques in the SIMBox ecosystem limits the effectiveness of existing network-core-based SIMBox activity detection, justifying the need for detection techniques tailored to the fraud evolution.*

3.2 Network-edge-based detection

Network-edge-based detection methods operate at each base station within the cellular network, monitoring activity in real-time to detect *SIMBox* patterns. Unlike network-core-based solutions, these methods often face scalability challenges that affect their practical efficiency. Specifically, existing solutions analyze either *call audio* or *cellular signaling data*.

First, *call audio-based* solutions, which examine speakers' voices [12] or call quality [30], have demonstrated efficiency in lab settings but face real-world deployment challenges. Investigating all local calls across the network raises privacy concerns and significant scalability issues.

In contrast, Oh et al. [28] leverages *cellular signaling data* for *SIMBox* detection, proposing a fingerprinting-based approach, referred to as *ACLPrint*. This method compares the device's fingerprint and factory identifier code (i.e., TAC) to a pre-established database, rejecting devices with mismatched fingerprints. However, *ACLPrint* faces significant scalability issues. Frequent updates to the relying 3GPP LTE specifications (cf. Fig. 3), which occur roughly every three months, require constant manual monitoring of extensive documents and their numerous references, along with adjustment of the fingerprinting process. Additionally, *ACLPrint* relies on a pre-established database, making it vulnerable to new, unrecorded ME models and brute-force attacks, where fraudsters modify their *SIMBox* identifiers until they find a bypass. This also implies each network base station maintains and regularly consults such a vast database, complicating deployment.

These findings highlight the scalability challenges of *ACLPrint* and similar methods, limiting their real-world efficiency.

Insight: *The effectiveness of network-edge-based detection methods hinges on their scalability — their capacity to function across the entire network without compromising network performance. However, this challenge remains unmet in current literature contributions.*

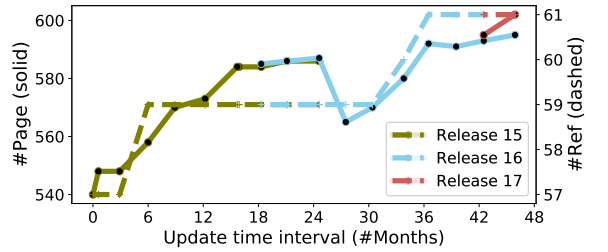
4 ESTABLISHING *SIGN* GROUNDS

In this section, we overview our investigation standpoint. In §4.1, we first outline the *SigN* threat model, defense objectives, and key insights. Then, in §4.2, we explain why signaling latency is central to the *SigN* methodology. Further discussions in §4.3 provide an overview of signaling procedures in LTE, justifying *SigN* focus on the network attachment procedure signaling.

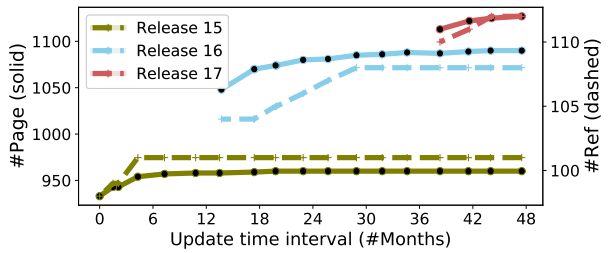
4.1 Threat models and defense goals

Threat model: *SigN* is designed in a complementary viewpoint to literature methodologies. It addresses "advanced *SIMBox* activity" that is undetectable with CDR-based approaches due to the use of Human Behavior Simulation (HBS) (cf. §3.1), and with network-edge-based methods due to privacy/scalability limitations (cf. §3.2).

Therefore, we focus on detecting advanced *SIMBox* patterns derived from HBS techniques implementation. As described in §3.1, these patterns involve *remote SIM card associations* to mimic human-like mobility behavior. We, therefore, assume that adversaries implement *remote SIM card association* using a distributed *SIMBox*



(a) NAS spef. updates from 27-03-2019 to 03-01-2023.



(b) RRC spef. updates from 19-02-2019 to 13-01-2023.

Figure 3: NAS[6] and RRC[5] protocol specifications size. Number of pages (#Page) in solid line. Number of references (#Ref) in dashed line.

architecture (cf. Fig 2) with lawfully-issued local SIM cards held within a SIMBank.

Defense objective: Our goal is the efficient, online prevention of such advanced *SIMBox* activity in a privacy-preserving and practical manner. We introduce *SigN*, *a network-edge-based SIMBox activity detection methodology based on cellular signaling data. SigN* aims to prevent fraudulent *SIMBox* activity on the mobile network surface. This is done through the reliable identification of network devices with advanced *SIMBox* patterns at the cellular edge *before any fraudulent calls* are made. Such an early detection effectively *prevents SIMBox owners from gaining financial advantage. SigN thus provides mobile operators with the means to detect and regulate SIMBox usage, enforcing legal registration for legitimate operations while blocking undeclared usage.*

We design *SigN* keeping in mind the open challenges of the network-edge-based *SIMBox* fraud detection literature (cf. §3.2). To ensure high real-world relevancy, we establish the following requirements: (i) *Privacy*: *SigN* should rely on network device features that operators can access without impeding privacy. (ii) *Practicality*: *SigN* implementation should be scalable and require minimal, non-constant effort from operators for wide deployment on the network surface.

Key insights: As advanced *SIMBox* patterns result from carefully crafted communications and movements to mimic human behavior, providing precise online detection indicators for mitigation at the cellular edge is a genuine challenge.

Our approach to addressing this challenge builds upon the indicator of advanced *SIMBox* activity: *remote SIM card association.* Specifically:

(1) *Indispensability of remote SIM card association*: *Remote SIM card association* is essential for *SIMBox* to mimic human behavior (cf. §3.1). Previous work [23] establishes that *SIMBox* activity without *remote SIM card association* results in distinctive communication behaviors efficiently detected through existing methods.

(2) *Uniqueness of remote SIM card association to SIMBox*: No other user end device (smartphones, tablets, laptops, IoT devices, modems, etc.) separates the SIM card from the Mobile Equipment (ME) during network operations, making *remote SIM card association* an explicit proxy for advanced *SIMBox* activity.

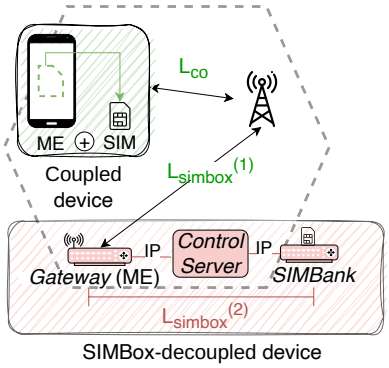


Figure 4: Signaling latency of coupled and *SIMBox*-decoupled devices.

Henceforth, by detecting the use of *remote SIM card association*, *SigN* effectively controls advanced *SIMBox* activity and prevents any malicious usage. *SigN* analyzes cellular signaling to determine if an attaching device is either *coupled* or *decoupled* via a *SIMBox*-operated *remote SIM card association.* This approach follows the intuition that signaling messages from *SIMBox*-decoupled devices exhibit higher latency compared to coupled devices, as below.

4.2 Preliminaries

Standard devices in cellular networks are a combination of a Mobile Equipment (ME) and a SIM card integrated within the ME, as depicted in Fig. 4: *it is a coupled ME-to-SIM combination.* They thus present a coupled signaling latency L_{co}, corresponding to the interaction time between the base station and the ME.

On the other hand, *SIMBox*-decoupled devices make a *logical IP-based* binding of a GSM module (inside the gateway operating as the ME) to a SIM card (inside the SIMBank) done at the level of the *control server*: *it is a decoupled ME-to-SIM combination.* Accordingly, the signaling latency of a *SIMBox*-decoupled device, i.e., L_{simbox}, includes:

- $L_{simbox}^{(1)}$ corresponding to the interaction time between the base station and the gateway (i.e., ME of the *SIMBox*-decoupled device), and
- $L_{simbox}^{(2)}$ corresponding to the interaction time between the gateway (i.e., ME) and the *SIMBox*-decoupled device's SIM card inside the SIMBank

such that $L_{simbox} = L_{simbox}^{(1)} + L_{simbox}^{(2)}$.

Therefore, compared to coupled devices' signaling latencies L_{leg}, *SIMBox*-decoupled devices' signaling latency L_{simbox}, will tend to be larger due to component $L_{simbox}^{(2)}$ involving one or more

Table 1: Summary of signalling procedure analysis

Signaling procedure	Description	#device processing	Moment of occurrence
Network attachment	Device connection and authentication to the network	4	- At the device power on - Device mobility dependent - Network initiated
Handover (X2)	Direct device connection's transfer between network base stations	1	Device mobility dependent
Handover (S1)	Device connection's transfer between base stations via the core network	1	Device mobility dependent
CQI update	Device's information to the network of channel quality	0	Network dependent (periodic or aperiodic)
Data bearer establishment	Setting up data transmission channel	2	Device communication dependent
Mobile originated SMS signaling	Texts transmission from the device to the network	1	Device communication dependent
CS Fallback call setup	Establishment of a traditional voice call circuit	7	Device communication dependent

exchanges of *SIMBox* components over the Internet during the signaling operation.

SigN methodology aims to identify such latency overhead, at the base station, to distinguish between coupled devices and *SIMBox*-decoupled ones. It, therefore, fulfills our mitigation requirements as follows:

(1) Privacy: Mobile operators have a natural access to signaling latency measurements as these do not relate to any specific individual or device model's identifier or data content and, thus, are not privacy-impeding.
(2) Practicality: Inspecting signaling latency is straightforward and is already implemented in LTE using *timers*. This practical approach ensures ease of deployment at no additional cost across the entire network edge.

4.3 Focusing on the network attachment

LTE standards provide several signaling procedures to deliver communication services to network devices. Aiming to distinguish *SIMBox*-decoupled devices by their latency overhead, we analyze the most common of such signaling procedures (cf. Table 1) based on two criteria: First, their *ability to involve device's processing*, i.e., the number of device processing necessarily occurring during the signaling procedure (*#device processing*), that maximizes the latency checking possibilities to detect latency anomalies of *SIMBox*-decoupled devices; Second, their *moment of occurrence* indicating when and how often a latency checking can done and whether such checking depends on the network or on the device behavior.

Our investigation relies upon the related 3GPP specifications and reports in Table 1 the uncovered *#device processing* and *moment of occurrence* per signaling procedure. We make the following observations:

- The number of device processing varies from one procedure to another, indicating that procedures with greater values are more suitable for our *SIMBox* activity detection goal. For instance, CQI updates, though fully network-controlled, do not involve any device processing, therefore, not allowing to uncover *SIMBox* latency overhead.
- Concerning the moment of occurrence, signaling procedures are triggered either by the device's communication or mobility behavior or by the network itself. Network-initiated or mandatory procedures are more relevant to guarantee minimal interference by fraudsters. For instance, data bearer establishment are only executed when a device starts a mobile data session, that can be avoided by fraudsters.

Based on these insights, *the network attachment procedure is the optimal choice for building SigN*, as it incurs a sufficient number of device processing compared to other signaling procedures. This procedure is mandatorily carried out by all network devices (coupled or *SIMBox*-decoupled) when they connect to an operator network upon being powered on. It, therefore, enables the implementation of a network access control that prevents any *SIMBox* activity-induced damage. Furthermore, it can be triggered by the operator (as a Tracking Area Update), independently of the device's behavior, to increase attempts to detect *SIMBox* activity.

In the following steps, we carry out an in-depth empirical study of the network attachment signaling latency to assess if this metric is satisfactory in distinguishing between coupled and *SIMBox*-decoupled devices.

5 LATENCY EMPIRICAL STUDY

In this section, we want to evaluate how well the network attachment signaling latency (referred to as *attachment latency*, for simplicity) can be used to differentiate between coupled and *SIMBox*-decoupled devices through experimental studies. To this end, we answer the following questions:

[Q1] *How different is the attachment latency between coupled and SIMBox-decoupled devices?*
[Q2] *What factors explain the attachment latency of SIMBox-decoupled devices, and is this latency reliable?*
[Q3] *What factors influence the attachment latency of coupled devices, and how might these variations compare to the latency observed in SIMBox-decoupled devices?*

Through extensive indoor and outdoor experiments, represented in Fig. 5, we make important observations relatively to the previous questions, which are summarized as follows:

[O1] *SIMBox-decoupled devices generate at least* 5× *more attachment latency compared to coupled ones in particular during the authentication phase where their minimum latency is* 23× *higher (cf. §5.2).*
[O2] *SIMBox-decoupled devices' latency is induced by both (i) SIMBox implementation and (ii) network protocols-imposed procedures (i.e., LTE standards and TCP/UDP correction and retransmission mechanisms). While the former allows for fraudsters improvement, the latter is beyond fraudsters reach and guarantee a minimum latency still* 2× *higher than coupled devices, in the authentication phase* (cf. §5.3).
[O3] *Regardless of the wireless network channel conditions, coupled devices' latency, in the authentication phase, cannot reach high values comparable to the one of SIMBox-decoupled devices (cf. §5.4). Therefore, the authentication latency unambiguously separates coupled and SIMBox-decoupled devices.*

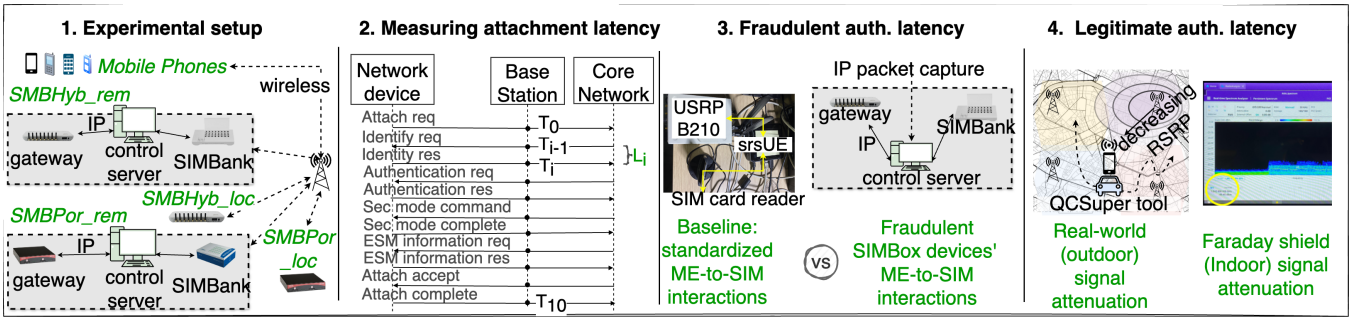


Figure 5: *SigN* attachment latency analysis methodology.

5.1 Experimental Setup

To ensure our study complies with regulations and avoids interference with live operator networks, we have designed a high-performance testbed that accurately simulates a real-world 4G cellular network. This setup utilizes Amarisoft's professional suite, a trusted solution in the wireless industry, and is housed within a $30m^2$ Faraday shield. Amarisoft's software-based technology is fully compliant with 3GPP standards and compatible with off-the-shelf hardware, including the physical layer [**?**]. With over 1,000 customers in more than 60 countries, including numerous public and private network operators, Amarisoft's solutions are widely adopted for both laboratory and field applications. This widespread adoption underscores the reliability and accuracy of Amarisoft's technology in replicating authentic network environments. Consequently, the signaling we capture in our testbed mirrors the exact procedures employed in operational networks, ensuring that our results are both precise and reflective of real-world conditions.

Our testbed employs a single PC to run both the base station and core network nodes, including the MME, IMS, and SGW. This PC handles baseband processing, while radio processing is managed by a software-defined USRP B210 connected to the PC, enabling seamless integration of the baseband and SDR systems. Detailed specifications of all testbed components, including the featured 4G cell and its radio parameters, are provided in Table 3 in the appendix. Signal quality, specifically RSRP, has been rigorously validated using a radio spectrum analyzer, showing excellent performance (around -71 dBm) consistent with real-world urban network conditions as documented in recent studies [24] (cf. Fig. 15a).

Our setup includes 12 mobile phones from five different vendors and 7 *SIMBox* devices from two manufacturers—Hybertone, the leader in the *SIMBox* market [17], and Portech. These devices are equipped with programmable SIM cards [35], ensuring they connect seamlessly to the LTE network inside the shield. Notably, Hybertone *SIMBox* appliances support both TCP and UDP protocols, while Portech devices only support UDP.

As illustrated in Fig. 5, step 1, we use the *SIMBox* appliances of both manufacturers to deploy (i) *remote SIM card association* and (ii) *local SIM card association*. The *remote SIM card associations* a *SIMBox* control server hosted on LAN-connected PCs (cf. Table 3). Thus, *SMBHyb_rem* and *SMBPor_rem* refer to devices using *remote SIM card association* from Hybertone and Portech, respectively, which are *SIMBox*-decoupled. Likewise, *SMBHyb_loc* and *SMBPor_loc* refer to devices using *local SIM card association* from the same manufacturers. These coupled devices, though potentially fraudulent, fall outside the scope of this research, as they involve already-addressed *SIMBox* activity (cf. §4.1)

5.2 Measuring attachment latency

Here we collect and analyze, for all the coupled and *SIMBox*-decoupled devices of our testbed, the latency at each step of network attachment procedure (cf. Fig. 5, step 2). We first detail the methodology for latency collection and computation and then present and discuss the obtained results.

Methodology. For each network device (i.e., phone model, *SIMBox* coupled and decoupled devices), we carry out 50 executions of the network attachment procedure. The resulting cellular signaling logs are recorded at the level of the base station. We consider only NAS-layer logs, as they provide information on the signaling between the device and the core network during the network attachment. These logs consist of 11 messages, as represented in Fig. 5, step 2. For each message we use the following associated fields for latency computation: `time, layer, direction, device_id, message`. The communication direction, i.e., uplink or downlink, indicates the message originator as the device or the network, respectively. Therefore, we compute the latency of each message as $L_i = T_i - T_{i-1}$, i.e, the delta time between the arrival of a message i and its previous one $i-1$. Depending on the message direction (uplink/downlink), L_i refers to the network's or the device's processing time along with the message transmission time to the base station. The total network attachment latency of a network device is thus $\sum_{i=1}^{n} T_i - T_{i-1}$ with $n = 10$ steps (cf. Fig. 5, step 2).

Results. Table 2 reports the obtained latency's mean and standard deviation values for each step of the network attachment. A particular interest is on the lines with *uplink* direction, enabling us to determine and compare the processing time per network device. We make the following observations:

- Regardless of the model, all phones have comparable latencies per step, similar to the *SIMBox* devices resulting from *local SIM association*. However, *distinguishing from coupled devices, the attachment latency for SIMBox-decoupled devices is significantly higher i.e.,* $\approx 9\times$ *for Hybertone and* $\approx 5\times$ *times for Portech.*

- *Such latency distinction of SIMBox-decoupled devices emerges at step 4 (authentication response)*, which consists of mutual authentication of the network and device, following the AKA procedure [1]. The authentication phase involves a computation internal to the SIM card (in the remote SIMBank) and, therefore, necessarily imputes IP-based interactions between *SIMBox* components, explaining the overhead. Particularly in the authentication phase, *SIMBox*-decoupled devices show a latency approximately 29× (for Hybertone) and 23× (for Portech) higher than coupled devices' latency.

Insight. *The previous results spotlight the authentication phase as the primary source of latency distinction for SIMBox-decoupled devices during the network attachment. Henceforth, we narrow the following investigations to understand such authentication latency.*

5.3 Decoupled devices' authentication latency

This section investigates the latency introduced by *SIMBox*-decoupled devices during authentication (i.e., Table 2, step 4) to uncover the cause and consistency of this overhead due to *remote SIM card association*. To this end, we capture ME-to-SIM interactions during authentication for both a coupled device, serving as a baseline reflecting 3GPP standards, and a *SIMBox*-decoupled device. Then by comparing these interactions, we explain decoupled devices' authentication latency, classifying its sources as either (i) specific to *SIMBox* implementation or (ii) imposed by standards and protocols.

5.3.1 Methodology. First, we describe the experimental process for capturing ME-to-SIM interactions during authentication for both a coupled device and a *SIMBox*-decoupled device.

Coupled devices. Aiming to capture standardized ME-to-SIM interactions during the authentication phase, we separate a coupled device's SIM card and radio processing, similarly to *remote SIM card association*: As depicted in Fig. 5, step 3 we set up a coupled network device combining (i) a srsUE softphone, i.e., a 4G phone implemented entirely in software, running on a Linux-system PC and connecting to the shielded LTE network, (ii) a physical SIM card within a SIM card reader connected to the softphone through an USB interface, and (iii) physical cellular antennas handled by a connected software-defined radio system (i.e., USRP B210). We then perform the network attachment of the formed device and align two sets of resulting timestamped logs: (i) signaling logs at the base station and (ii) SIM card logs at the softphone. *The use of the srsUE softphone, developed in the widely adopted srsRAN 4G framework [33], in our experiments attests to its generality and fidelity to 4G/LTE standards. Hence, our experiments thus provide insights into the actual implementation of 3GPP standards for the authentication phase (i.e., the AKA procedure [1]), publicly available in [34].*

***SIMBox*-decoupled devices.** Such devices disconnect the ME (i.e., the *SIMBox* gateway) from the SIM card (within the SIMBank), causing ME-to-SIM interactions to occur as packet exchanges over an IP network (cf. Fig 4). In order to capture these packet exchanges, we perform the network attachment of a *SIMBox*-decoupled device and monitor packets at the control server-hosting PC using Wireshark. This setup enables us to gain insights into the interactions between the SIMBank and the gateway, which are proprietary *SIMBox* appliances and thus typically concealed. Transport protocols (i.e., TCP or UDP) ensure reliability, order, and flow control in these IP-based interactions. We noted variations in the number of packets exchanged depending on the transport protocol used. By correlating the timing of these packets with network attachment signaling logs collected at the base station, we make specific observations for each transport protocol.

5.3.2 Observations. From the previous experiments we make the following observations, summarized in Table 4 in the appendix.

- First, confirming Table 2 insights, authentication is the primary phase involving ME-to-SIM interactions, making it the best context for identifying any latency overhead. For coupled devices, these interactions occur *only during authentication*, while *SIMBox*-decoupled devices also show *minor exchanges during the attach complete* phase. Specifically, TCP interactions involve 63 packets during authentication and 4 packets during the attach complete phase, while UDP interactions consist of 36 packets for authentication and 2 packets for attach complete (cf. Figs 16, 17). Shared by coupled and *SIMBox*-decoupled devices, interactions during the authentication split into ME-to-SIM *transfer* and *processing* phases. Transfers consist of *information transmission* from/to the ME/SIM card, while the processing phases are *internal computations* within the ME/SIM card following these transfers.
- *Transfers*: Logs from coupled devices reveal two physical layer round-trip transfers (four transfers in total) between the ME (i.e., softphone) and the SIM card, following the ISO/IEC 7816-4 protocol [16] illustrated on Fig. 14. These transfers are rapid, averaging 0.12 ms due to physical layer communication via UART serial interfaces.
  In contrast, *SIMBox*-decoupled devices show a significantly higher number of transfers. They are observable with TCP while UDP's unordered nature makes them less distinguishable. Fig. 6 illustrates 15 transfer sessions during authentication, each involving four packets (in total 60 packets) exchanged between the ME (i.e., gateway) and the SIM card in the SIMBank through the control server, which acknowledges and re-transmits the packets. On a local network, these transfers take on average 4.7 ms, totaling 70.6 ms, which is a lower bound compared to real-world *SIMBox* deployments that are Internet-based.
  This comparison highlights that transfer latency is a consistent indicator for distinguishing SIMBox-decoupled devices. Specifically, the 4 transfers mandated by standards result in a latency at least 39× higher than that of coupled devices, and *hardly controllable due to its dependence on (i) the transport protocol and (ii) Internet vagaries. Regarding the transport protocol, while TCP increases the number of packet exchanges, UDP poorly handles network congestion, leading to retransmissions and delays.* For instance, a comparison of the latency distribution over 50 authentications of *SMBHyb_rem* configured with TCP and UDP, as shown in Fig. 7, indicates that latency with UDP is significantly higher than with TCP. Additionally, *internet vagaries further contribute to transfer latency overhead.* We estimate this overhead by performing network attachment with the control server online, showing an average additional latency of 460 ms, and by measuring the

Table 2: Latency (in ms) per device model reported per network attachment step

Step	Direction	Fair Phone5G	Galaxy A90	Galaxy Note4	Galaxy S3	Galaxy ZFold25G	OnePlus Nord	Sony XPERIA	Xiaomi10 Lite5G	Xiaomi9 Pro5G	SMBHyb_loc	SMBHyb_rem	SMBPor_loc	SMBPor_rem
0. Attach request	Uplink	0	0	0	0	0	0	0	0	0	0	0	0	0
1. Identity request	Downlink	1± 0	1± 0	1 ± 0	0.9± 0.3	1± 0	1± 0	1± 0	1± 0	1± 0	1 ± 0	0.9 ± 0.2	/	/
2. Identity response	Uplink	31±0	27±6	38.3 ± 2.3	31.0± 10.4	31±0	31.8± 2.4	25.0± 6.4	31± 0	31± 0	31.8 ± 3.5	31.0 ± 4.3	/	/
3. Authentication request	Downlink	1±0	1 ± 0	1.0 ± 0.3	1±0	1±0	1± 0	1± 0	1± 0	1 ± 0	1 ± 0	0.9 ± 0.3	0.9± 0.1	1 ± 0
4. Authentication response	Uplink	57.6 ± 11.4	74.1 ± 22.1	84.5 ± 36.5	67.9± 12.2	70.2± 18.2	69.8 ± 10.0	69.1± 5.9	69.9± 8.2	67.9± 16.2	71.7 ± 10.8	**2122.7± 309.9**	71.2± 10.7	**1640.2± 286.7**
5. Security mode command	Downlink	1±0	1± 0	1±0	1±0.1	1±0.1	1± 0	1± 0	1± 0	1 ± 0	1 ± 0	0.9 ± 0.3	1 ± 0	1 ± 0
6. Security mode complete	Uplink	20.5± 3.2	19.3± 1.6	37.0± 6.3	33.0± 9.5	21.8± 14.3	31.3± 12.5	19.6± 2.6	21.9± 4.6	21.8± 10.5	22.4 ± 5.9	20.1± 3.7	19 .7 ± 2.7	21.1± 5.8
7. ESM information request	Downlink	1±0	1 ± 0	0.9 ± 0.2	/	1 ±0	1± 0	1± 0	1±0	0.9± 0.1	1 ± 0	1.0± 0	/	/
8. ESM information response	Uplink	19±0	19.7 ± 2.6	37.3 ± 5.5	/	22.8± 21.4	26.2± 9.3	19.6± 2.3	22.6± 5.6	20.7± 4.2	22.9 ± 5.8	20.6 ± 3.9	/	/
9. Attach accept	Downlink	50.4 ± 4.8	48.7± 2.5	66.2 ± 6.8	56.9 ± 8.7	50.0± 4.4	66.5 ± 14.3	50.9± 5.9	48.8± 3.9	49.3± 4.3	46.9 ± 10.3	43.7± 9.3	50.7± 6.8	57.9 ± 26.1
10. Attach complete	Uplink	32.4± 1.9	32.8± 3.4	49.7 ± 7.1	60.1± 1.1	34.3± 6.0	35.5 ± 8.2	54.5± 6.4	38.8± 3.9	33.5± 4.1	57.3 ± 10.1	53.2 ± 9.5	78.5± 6.8	52.2± 4.7
Total		215.0± 21.3	225.6± 38.2	316.8± 65.5	251.9± 42.4	234.2± 64.6	265.1± 56.9	242.8± 29.5	237.1± 33.5	228.2± 38.5	257.1± 42.7	**2295.5± 341.7**	253.9± 31.3	**1773.5± 323.3**

median RTT of internet communication within the same country based on an empirical distribution of 1000 RTTs (cf. Fig. 8). The median value of 57.4 ms suggests that *the 2 RTTs imposed by the standard guarantee a transfer latency overhead of 114.8 ms for SIMBox-decoupled devices, which is almost twice the avg. auth. latency of coupled devices (cf. Table 2).*

- *Processing*-induced latency is much higher than transfer latency. Analysis of coupled devices' logs reveals two standard-imposed processing phases on the SIM card and three on the ME, averaging 15.6 ms and 9.4 ms respectively (cf. Table 4). In contrast, *SIMBox*-decoupled devices exhibit as many as 14 processing phases (8 for the SIM card and 6 for the ME), averaging 218 ms and 211 ms respectively with TCP, and 12 (6 each for the SIM card and the ME), averaging 236.1 ms and 139.3 ms respectively with UDP. *This comparison highlights that the elevated processing latency in SIMBox-decoupled devices is likely due to their implementation involving a higher number of processing phases than necessary and significantly longer average times. Although this could be optimized by SIMBox manufacturers, some overhead may be inevitable due to the simultaneous control of multiple SIM cards and GSM modules, as well as the encapsulation/decapsulation of information exchanged during authentication into IP packets. This overhead is challenging to quantify for proprietary devices.*

Insight. *In essence, our investigations show that SIMBox-decoupled devices' authentication latency is influenced by factors beyond SIMBox owners' control. Even with optimization efforts, i.e., reducing the transfer count and processing time, this latency cannot match that of stripped-down, coupled devices, maintaining a consistent distinction between SIMBox-decoupled devices and their coupled counterparts.*

5.4 Coupled devices' authentication latency

Here, we assess the feasibility of instances where a coupled device's latency could be high enough to be mistaken for a *SIMBox*-decoupled one. To this end, we scrutinize the latency of coupled devices during the authentication phase by breaking down a coupled device's auth. latency into two components: (i) *A transmission latency* that includes the wireless propagation time along with any delay related to the wireless network channel condition (§5.4.1) and (ii) *A processing latency* in which the device locally runs the authentication algorithm until a response is generated (§5.4.2). Note that these latencies differ from the ones detailed in §5.3.2, which focused on intra-device interactions. The current context instead aims to analyze the latency in the device's communication with the network through the base station.

5.4.1 Transmission latency of the authentication. Transmission latency refers to the back-and-forth communication time between the device and the base station during an authentication request and response. Predicting this latency is challenging due to various factors that affect the quality of each device's experience. In particular, although the *wireless signal propagation* time is negligible as the signal moves at the speed of light, and LTE employs an admission control mechanism [13] to prevent delays during the attachment caused by *network congestion*, the impact of *signal quality* on authentication latency remains to be determined.

Indeed, poor signal quality, indicated by an LTE RSRP of less than -110 dBm, often results from significant distance to the cell tower, interference, or device sensitivity issues. This can cause retransmissions and signaling delays. In the following, we assess this impact on authentication latency by measuring it in an outdoor

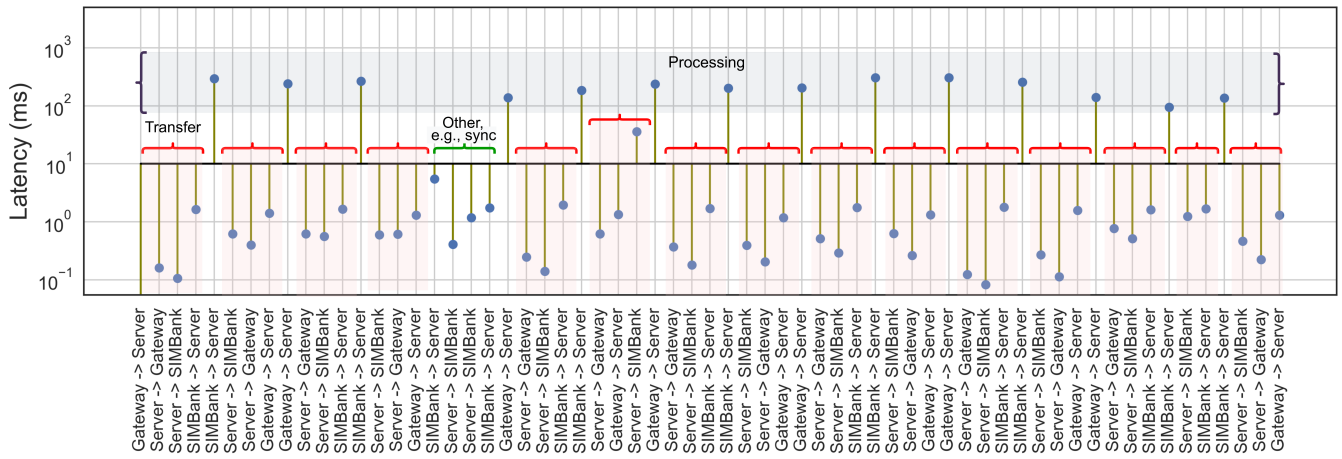


Figure 6: Hybertone *SIMBox* components TCP interactions during the authentication phase.

network and generalizing the results in a controlled indoor testbed (cf. Fig. 5, step 4).

Methodology. Regarding *outdoor signal attenuation*, we measure the outdoor signal quality over three days with varying weather conditions (sunny, rainy, and windy) using a Samsung Galaxy Note4 (referred to as *GalaxyNote4*) in a vehicle covering over 80km of city roads in a central urban area in the Paris region. As open phone signaling data is not public, we use the QCSuper open-source diagnostic logging tool [29], which is compatible only with Qualcomm-based phones, to capture the network attachment signaling messages. Data collection is done with a QCSuper-installed laptop connected to a *GalaxyNote4* phone, the only compatible device in our testbed. The collected outdoor dataset comprises 2287 network attachments of the *GalaxyNote4* phone. Each network attachment marks the phone's entrance into a new network cell and induces an authentication process. From the collected logs, we extract signal quality measurements (i.e., RSRP) within the respective cell for each network attachment.

To extend these findings across various phone models, we replicate the network *indoor signal attenuation* in a controlled environment (cf. §5.1) using static attenuators. Specifically, we apply two attenuation levels to the initial network signal quality of -71 dBm (excellent quality), resulting in measurements of (i) -90 dBm (medium quality) and (ii) -100 dBm (poor quality) (cf. Fig. 15, in the appendix). For each signal quality condition, we conduct 50 network attachments for every phone model (described in Table 3), including the *GalaxyNote4* phone used in the outdoor measurement scenario, and record the corresponding authentication latency.

Observations. In Fig. 9, we break down the attachment latency distribution for the *GalaxyNote4* in the outdoor scenario, detailing each step of the network attachment process. Since latency data is collected at the device level, our focus is on the downlink messages (UE←BS), which include the transmission latency of interest. Specifically, we examine the impact of signal quality (i.e., RSRP) on the latency of the *"security mode command"* message, which inherently captures the transmission latency of the *"authentication response."* To streamline interpretation, we approximate the *"security mode command"* latency as the transmission latency of the authentication process in Fig. 10. With signal quality measurements ranging from -65 to -119 dBm, our results cover all radio frequency conditions, from "Cell Edge" to "Excellent" [2], ensuring the representativeness and depth of our findings.

Notably, Fig. 10 shows that the upper limit for the auth. response's transmission latency is negligible. Although some outliers around 200 ms, the majority of the values indicate low latency, irrespective of meteorological conditions (denoted by the days) or signal quality (denoted by the RSRP). *This result convincingly shows that signal quality has a minimal impact on authentication transmission latency. Furthermore, linear regression of latency and signal quality confirms this trend across a broader signal quality spectrum.*

Figure 11, shows the authentication latency under indoor signal attenuation, extending our findings to other phone models. The results indicate that signal quality has a negligible impact on the authentication latency for the *GalaxyNote4, GalaxyS3,* and *GalaxyZFold5G*, confirming our earlier interpretations from outdoor scenarios. However, the six remaining phone models did not initiate network attachment under medium to high signal attenuation, likely due to the lower sensitivity of their receivers. In a carrier network, these phones would have connected to nearby cells with stronger signals, thereby avoiding any latency issues related to signal quality.

Insight: *Our findings demonstrate that the round-trip transmission latency during the authentication phase is negligible and robust, showing little sensitivity to fluctuations in network signal quality. Even in the worst-case scenario, poor signal quality will cause mobile devices to switch network cells rather than increase transmission latency.*

5.4.2 Processing latency of the authentication. The processing latency denotes the time required for authentication computations within coupled devices, specifically within the SIM cards provided by the operator. As a result, it barely varies across different phone models or depends on phone features (i.e., processor, battery, or RAM), since modern phones are designed to handle much more

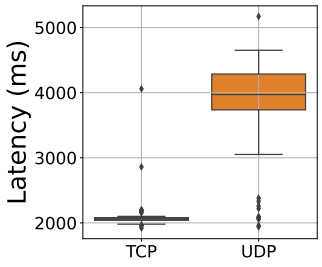


Figure 7: *SMBHYB_rem* TCP vs UDP auth. latency.

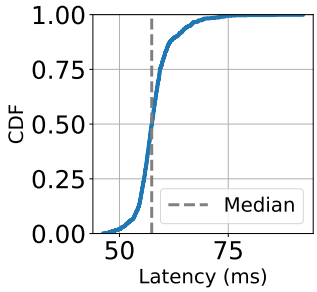


Figure 8: RTT latency distribution over Internet

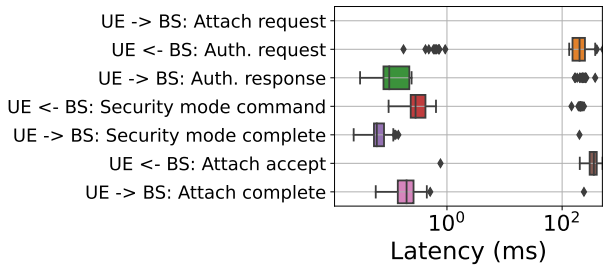


Figure 9: Latency (in log-scale) in different attachment steps for *GalaxyNote4*'s outdoor scenario

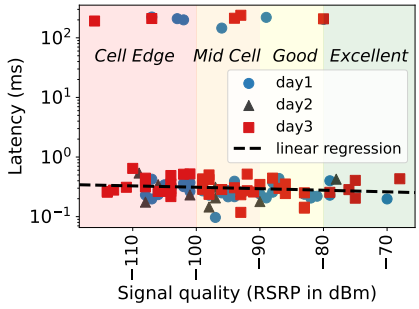


Figure 10: Outdoor transmission latency (in log-scale) of the authentication response w.r.t. the signal quality.

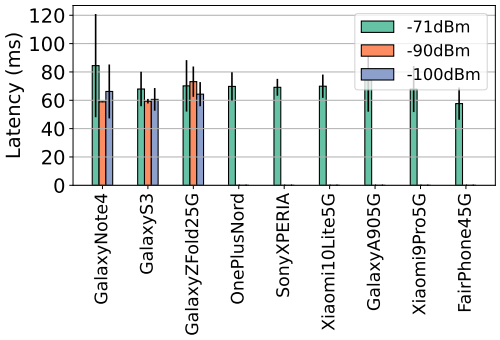


Figure 11: Indoor network authentication latency per phone model w.r.t. the signal quality.

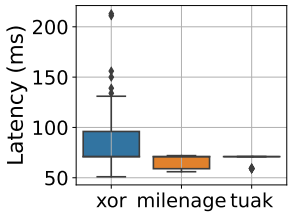


Figure 12: *GalaxyNote4*'s authentication latency variation w.r.t. the SIM authentication algorithm.

resource-intensive applications. Thus, the authentication processing latency is primarily determined by the *SIM card dauthentication algorithm*, which runs inside the SIM card to compute the expected network Authentication Response (RES). Chosen by each operator and kept secret to prevent account impersonation, these algorithms are typically variants of standardized algorithms like XOR [4], Milenage [14], or Tuak [7].

We evaluate the impact of the standardized SIM card authentication algorithms on the authentication latency. Specifically, we configure inside our indoor testbed (cf. §5.1) such different authentication algorithms on the *GalaxyNote4* phone and conduct 50 network attachments for each algorithm, recording the resulting authentication latency values. The findings in Fig. 12 confirm that the authentication algorithm impacts both the processing latency and its variability. *Thus, mobile operators can achieve lower processing latency by carefully selecting their SIM authentication algorithm.*

6 SIGN IMPLEMENTATION

This section delves into utilizing the authentication latency metric for the practical implementation of *SigN* for *SIMBox* activity detection at the mobile edge.

Supporting insights. The experiments conducted in §5 underscored a *significant distinction* in attachment latency between coupled and *SIMBox*-decoupled devices, specifically *during the authentication phase* (cf. Table 2). Acknowledging its variability influenced by internal/external factors, we further investigated the authentication latency of both coupled and *SIMBox*-decoupled devices. First, *coupled devices authentication latency in outdoor networks remains within a consistent range and is markedly lower than observed for SIMBox-decoupled devices in indoor settings* (cf. Fig. 9). Second, our investigations reveal that *a non-negligible portion of the observed authentication latency in SIMBox-decoupled devices is imposed by factors beyond fraudsters' control*, such as LTE standards and Internet-based communication protocols and vagaries (cf. §5.3). Moreover, like coupled devices, *SIMBox-decoupled devices experience additional transmission latency* in the real-life conditions of operator networks. Given these findings, ***monitoring authentication latency at the network edge proves to be a reliable and practical method for distinguishing SIMBox activity from regular one.***

Monitoring the authentication latency. In LTE, the authentication procedure initiates the monitoring of the induced latency through a logging mechanism. Consequently, the latency of each authentication is automatically logged at each base station, and is useful for network performance monitoring, optimization, and Quality of Service (QoS) management.

However, *experiments in this paper highlight that authentication latency is not only an indicator of network QoS but also a robust metric for identifying vulnerabilities related to SIMBox activity.* Our findings demonstrate that, while the latency is generally acceptable

(i.e., within the 6s timer range [15]) for both coupled and *SIMBox*-decoupled devices, it creates a clear distinction between *SIMBox* activity and regular one.

Therefore, *SigN approach suggests a new monitoring usage of authentication latency in cellular networks, allowing the detection of SIMBox activity through its distribution.* According to 3GPP standards, **network operators have the flexibility to initiate at any time an authentication procedure when a signaling connection with a device exists** [15]. This flexibility enables operators to capture the distribution of mobile devices' authentication latency by initiating multiple authentications at different times throughout the day. Randomly timed authentications are essential to accurately reflect a device's behavior, as passive collection could be manipulated by *SIMBox* operators who might switch between remote and local SIM card associations to skew the latency distribution.

Relying on the distribution rather than a single measurement is crucial for ensuring robustness against high-value outliers that may occur for coupled devices (cf. Fig. 10). Hence, analyzing a full day's data yields key metrics such as the mean, median, and standard deviation of authentication latencies, highlighting *devices with consistently unusually high values and prompting further investigation by the operator.*

Such monitoring is lightweight, leveraging existing automatic functions in cellular networks, i.e., logs collected by each base station. As a result, it allows operators to make informed decisions without the network edge's overhead.

Statistical support. We statistically validate the effectiveness of the *SigN* approach from our network measurements. Fig. 13 plots realistic distributions of authentication latency for: (i) coupled devices with outdoor transmission latency, (ii) Current measurements of *SIMBox*-decoupled devices with outdoor transmission latency, and (iii) Optimized *SIMBox*-decoupled devices (cf. 5.3.2) consisting of *SIMBox* coupled device with outdoor transmission latency and simulated reduced ME-to-SIM transfers (2 RTTs as on Fig. 8). We employed a *t-test*, a key statistical tool, to compare the means of coupled and *SIMBox*-decoupled devices (both current and optimized). Details are provided in the appendix (cf. §A).

The *t-ratio* of 15.29 (with $t = 25.27$ and *critical value* $= 1.65$) reveals a significant statistical difference between the coupled and *SIMBox*-decoupled device groups. The *p-value* of 1.3×10^{-102} indicates an almost zero chance of overlap between the two groups and a high probability of correctly identifying the attachment activity even of the most optimized *SIMBox*-decoupled device.

7 LIMITATIONS

Acknowledging the limitations of this study is crucial to contextualizing its findings and guiding future research efforts. While we believe our approach and results provide valuable insights into monitoring *SIMBox* activity in mobile networks, there are areas where constraints, assumptions, and specific conditions may have influenced the outcomes.

First, due to the proprietary nature of *SIMBox* devices, we were unable to perform an in-depth reverse engineering of their hardware and software. This limitation constrained our ability to fully uncover the specific implementation techniques used by *SIMBox* manufacturers. Instead, we adopted a standards-based approach to evaluate potential baseline measures that fraudsters must adhere to. While our approach sheds light on baseline security assumptions, future research could explore advanced logic analysis techniques to extract and study SIM card secrets from these devices, providing deeper insights into their operation.

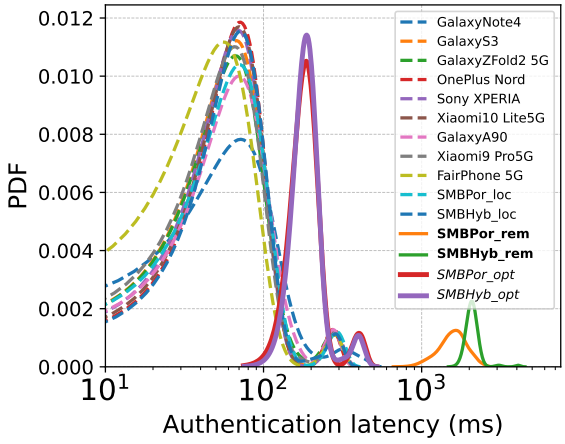


Figure 13: Distribution of authentication latency of (i) coupled devices (*dashed*), (ii) current (*bold label*) and (iii) optimized (*italic label*) *SIMBox*-decoupled devices

Second, while our experimentation was conducted on a simulated 4G network powered by the Amarisoft suite, we acknowledge that this does not fully replicate the complexities of a live operator network. However, the testbed design and the use of Amarisoft's professional-grade software, widely trusted by Mobile Network Operators (MNOs), ensure results that closely approximate those observed in real-world conditions. Additionally, testing in a controlled environment allowed for precise measurements and analysis without the risk of interfering with live operator networks. Despite this, future validations on live networks could provide additional insights into practical deployment scenarios.

Lastly, the TCP and UDP latency measurements presented in Figure 8 are derived from a single network configuration, which may not fully capture the variability across different operator networks. While these measurements are representative and align with general expectations, they may overestimate latency in certain cases. We believe this limitation can be mitigated by adapting our detection method to each operator's specific network conditions through a straightforward calibration process. Future work should consider extending these measurements across diverse network environments to enhance the generalizability of our findings.

8 CONCLUSION

This paper introduced *SigN*, an online prevention solution to uncover *SIMBox* activity at the cellular edge. Based on an empirical study of network attachment latency in coupled and *SIMBox*-decoupled devices, we found that *SIMBox*-decoupled devices exhibit higher authentication latency. *SigN* optimizes existing cellular monitoring, improving *SIMBox* activity detection efficiency.

SigN's significance lies in its effectiveness and practicality, enabling easy integration into operator networks. This offers substantial economic benefits and resolves challenges faced by current

network-edge solutions, making *SigN* a key advancement in securing networks against *SIMBox* activity.

Note that while this paper measurements focus on smartphones, the findings apply to other devices with a physical or embedded SIM card, including tablets, laptops, and IoT devices. The key distinction is the separation of the SIM card from the Mobile Equipment in *SIMBox*-decoupled devices.

REFERENCES

[1] 3GPP. 2020. *3G Security; Security architecture (Release 16).* Technical Specification TS 33.102. European Telecommunications Standards Institute (ETSI). https://www.etsi.org/deliver/etsi_ts/133100_133199/133102/16.03.00_60/ts_133102v160300p.pdf

[2] 3GPP. 2021. *3GPP TS 36.133: Requirements for support of radio resource management.* Technical Report. 3GPP. https://www.3gpp.org/ftp//Specs/archive/36_series/36.133/36133-960.zip

[3] 3GPP. 2022. *General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access.* Technical Report TS 23.401. 3GPP. https://www.3gpp.org/ftp/Specs/archive/23_series/23.401/23401-i00.zip

[4] 3GPP. 2023. *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Packet Core (EPC); Common test environments for User Equipment (UE) conformance testing.* Technical Specification 36.508. 3GPP.

[5] 3GPP. 2023. *LTE RRC Protocol Specification.* Technical Report TS 36.331. 3GPP. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2440

[6] 3GPP. 2023. *Non-Access-Stratum (NAS) Protocol for Evolved Packet System (EPS); Stage 3.* Technical Report TS 24.301. 3GPP. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1072

[7] 3GPP. 2023. *Specification of the TUAK algorithm set: A second example algorithm set for the 3GPP authentication and key generation functions f1, f1*, f2, f3, f4, f5 and f5*; Document 1: Algorithm specification.* Technical Specification 35.231. 3GPP.

[8]]amarisoftWebsite Amarisoft. [n. d.]. Amarisoft Official Website. https://www.amarisoft.com. Accessed: 2025-01-13.

[9] CFCA. 2019. *CFCA 2019 Fraud Loss Survey.* Technical Report. Communications Fraud Control Association. https://cfca.org/document/cfca-2019-fraud-loss-survey-pdf/

[10] CFCA. 2021. *Communications Fraud Control Association 2021 Fraud Loss Survey.* Technical Report. Communications Fraud Control Association. https://cfca.org/wp-content/uploads/2021/12/CFCA-Fraud-Loss-Survey-2021-2.pdf

[11] Communications Fraud Control Association. 2023. Telecommunications Fraud Increased 12% in 2023 Equating to an Estimated $38.95 Billion Lost to Fraud. https://cfca.org/telecommunications-fraud-increased-12-in-2023-equating-to-an-estimated-38-95-billion-lost-to-fraud/ Accessed: 2025-01-13.

[12] Osama Mohamed Elrajubi, Ali Mustafa Elshawesh, and Mustafa Ali Abuzaraida. 2017. Detection of bypass fraud based on speaker recognition. In *2017 8th International Conference on Information Technology (ICIT).* 50–54. https://doi.org/10.1109/ICITECH.2017.8079914

[13] ETSI. 2018. *LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access; Stage 2.* Technical Specification 123 401. ETSI. https://www.etsi.org/deliver/etsi_ts/123400_123499/123401/15.04.00_60/ts_123401v150400p.pdf

[14] ETSI. 2021. *Universal Mobile Telecommunications System (UMTS); LTE; 3G Security; Specification of the MILENAGE algorithm set: An example algorithm set for the 3GPP authentication and key generation functions f1, f1*, f2, f3, f4, f5 and f5*; Document 2: Algorithm specification (3GPP TS 35.206 version 14.0.0 Release 14).* Technical Specification 35206. ETSI.

[15] ETSI. 2021. *Universal Mobile Telecommunications System (UMTS); LTE; Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3.* ETSI Technical Specification 124 301 V15.17.0. European Telecommunications Standards Institute. https://www.etsi.org/deliver/etsi_ts/124300_124399/124301/15.04.00_60/ts_124301v150400p.pdf

[16] European Telecommunications Standards Institute (ETSI). 2007. *ETSI TS 100 977 V8.14.0 (2007-06).* Technical Specification. ETSI. https://www.etsi.org/deliver/etsi_ts/100900_100999/100977/08.14.00_60/ts_100977v081400p.pdf

[17] GoAntiFraud. 2018. Top 5 Popular GSM Gateway Manufacturers. https://goantifraud.com/en/blog/818-top-5-popular-gsm-gateway-manufacturers.html.

[18] GoAntiFraud. accessed 2023. Call Recording. https://goantifraud.com/en/ejointech-skyline-gsm-termination-solution#call-recording.

[19] Hybertone. Accessed 2023-04-27. Remote SIM Bank, Model: SMB128. http://www.hybertone.com/en/pro_detail.asp?proid=57.

[20] Hagos Kahsu. 2018. *SIM-Box Fraud Detection Using Data Mining Techniques: The Case of ethio telecom.* Ph. D. Dissertation. School of Electrical and Computer Engineering Addis Ababa Institute of Technology.

[21] KGDC Kehelwala, HMND Bandara, RA Yasaratne, P De Almeida, IKKS Ilesinghe, and PDKE Wickramasinghe. 2015. *REAL-TIME GREY CALL DETECTION SYSTEM USING COMPLEX EVENT PROCESSING.* Technical Report. IET, Sri Lanka. http://theiet.lk/wp-content/uploads/2017/10/22-p7.pdf

[22] Anne Josiane Kouam, Aline Carneiro Viana, and Alain Tchana. 2021. SIMBox bypass frauds in cellular networks: Strategies, evolution, detection, and future directions. *IEEE Communications Surveys Tutorials* 23, 4 (2021), 2295–2323. https://doi.org/10.1109/COMST.2021.3100916

[23] Anne Josiane Kouam, Aline Carneiro Viana, and Alain Tchana. 2024. Battle of Wits: To What Extent Can Fraudsters Disguise Their Tracks in International bypass Fraud?. In *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security* (Singapore, Singapore) *(ASIA CCS '24).* Association for Computing Machinery, New York, NY, USA, 366–382. https://doi.org/10.1145/3634737.3657023

[24] Nicholas Krawczeniuk. 2019. *Analysis of LTE network RF performance in a dense urban environment.* Undergraduate thesis. Pace University. https://digitalcommons.pace.edu/honorscollege_theses/269

[25] Hussein M. Marah, Osama Mohamed Elrajubi, and Abdulla A. Abouda. 2015. Fraud detection in international calls using fuzzy logic. In *International Conference on Computer Vision and Image Analysis Applications.* 1–6. https://doi.org/10.1109/ICCVIA.2015.7351891

[26] I. Murynets, M. Zabarankin, R. P. Jover, and A. Panagia. 2014. Analysis and detection of SIMbox fraud in mobility networks. In *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications.* 1519–1526. https://doi.org/10.1109/INFOCOM.2014.6848087

[27] Diala Naboulsi, Marco Fiore, Stephane Ribot, and Razvan Stanica. 2016. Large-Scale Mobile Traffic Analysis: A Survey. *IEEE Communications Surveys & Tutorials* 18, 1 (2016), 124–161. https://doi.org/10.1109/COMST.2015.2491361

[28] Beomseok Oh, Junho Ahn, Sangwook Bae, Mincheol Son, Yonghwa Lee, Minsuk Kang, and Yongdae Kim. 2023. Preventing SIM Box Fraud Using Device Model Fingerprinting. In *Network and Distributed Systems Security (NDSS) Symposium.*

[29] P1sec. accessed 2023-04-23. QCSuper: An open-source tool for capturing and decoding data transmitted over Qualcomm-based cellular devices. https://github.com/P1sec/QCSuper.

[30] Bradley Reaves, Ethan Shernan, Adam Bates, Henry Carter, and Patrick Traynor. 2015. Boxed Out: Blocking Cellular Interconnect Bypass Fraud at the Network Edge. In *Proceedings of the 24th USENIX Conference on Security Symposium* (Washington, D.C.) *(SEC'15).* USENIX Association, USA, 833–848. https://doi.org/10.5555/2831143.2831196

[31] Roselina Sallehuddin, Subariah Ibrahim, Azlan Zain, and Abdikarim Elmi. 2013. Detecting SIM Box Fraud Using Neural Network. In *IT Convergence and Security 2012*, Kuinam J. Kim and Kyung-Yong Chung (Eds.). Springer Netherlands, Dordrecht, 575–582. https://doi.org/10.1007/978-94-007-5860-5_69

[32] Roselina Sallehuddin, Subariah Ibrahim, Azlan Zain, and Abdikarim Elmi. 2015. Detecting SIM Box Fraud by Using Support Vector Machine and Artificial Neural Network. In *Jurnal Teknologi*, Vol. 74. 137–149. https://doi.org/10.11113/jt.v74.2649

[33] SRS. accessed 2023. SRS 4G Documentation. https://docs.srsran.com/projects/4g/en/latest/index.html.

[34] srsran/srsRAN_4G. 2023. srsRAN 4G - pcsc_usim.cc. GitHub. https://github.com/srsran/srsRAN_4G/blob/master/srsue/src/stack/upper/pcsc_usim.cc File: pcsc_usim.cc.

[35] Sysmocom. Accessed: March 8, 2023. SysmoUSIM. https://www.sysmocom.de/products/lab/sysmousim/index.html.

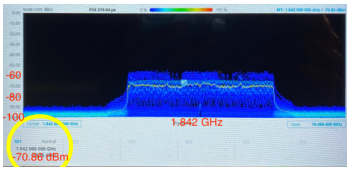
[36] T-mobile. 2024. Bring your own phone when you switch to T-Mobile. https://www.t-mobile.com/resources/bring-your-own-phone

[37] Fitsum Tesfaye. 2020. *Near-Real Time SIM-box Fraud Detection Using Machine Learning in the case of ethio telecom.* Ph. D. Dissertation. School of Electrical and Computer Engineering Addis Ababa Institute of Technology.

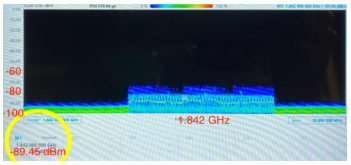
[38] Bruno Veloso, Shazia Tabassum, Carlos Martins, Raphael Espanha, Raul Azevedo, and João Gama. 2020. Interconnect bypass fraud detection: a case study. *Annals of Telecommunications* 75 (Oct. 2020), 583–596. https://doi.org/10.1007/s12243-020-00808-w

A T-TEST PROCEDURE AND STATISTICAL ANALYSIS

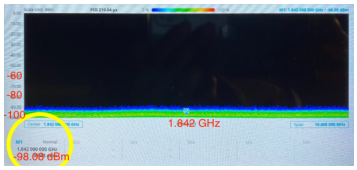
To perform a t-test, we first calculate the means of the two groups (μ_l and μ_f), their respective sizes (n_l and n_f), and the pooled standard error (SE) of the two groups, as shown in equation 1. The t-statistic (t) is then computed as the ratio of the difference between the group means to the pooled standard error, as detailed in equation 2.



(a) Reference state: -71 dBm



(b) -90dBm RSRP attenuation



(c) -100dBm RSRP attenuation

Figure 15: Reference Signal Received Power (RSRP) measurement inside the testbed: x-axis: RSRP (dBm), y-axis: Frequency (Hz)

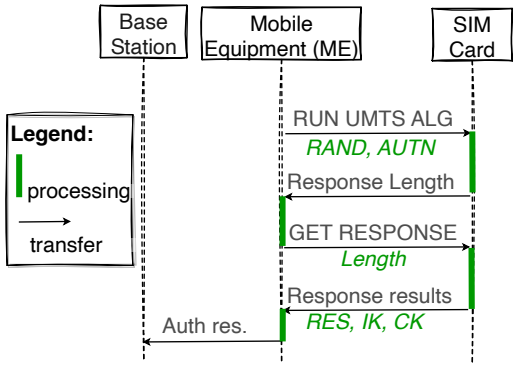


Figure 14: coupled ME to SIM card interactions during the authentication

$$SE = \sqrt{\left(\frac{\sigma_l^2}{n_l}\right) + \left(\frac{\sigma_f^2}{n_f}\right)} \quad (1)$$

$$t = |\frac{\mu_f - \mu_l}{\sqrt{SE^2(\frac{1}{n_f} + \frac{1}{n_l})}}| \quad (2)$$

Unlike the classical standard deviation (σ_l or σ_f), which measures dispersion within each group, the pooled standard error (SE) incorporates variability from both groups to provide a more accurate estimate of the population standard deviation.

A higher t-statistic value suggests a more significant difference between the groups. To determine the statistical significance of this difference, we compare the computed t-statistic to a *critical value*, which is typically based on a predetermined confidence level (e.g., 95%). If the computed t-statistic exceeds this critical value, it indicates that the observed difference is statistically significant and unlikely due to random chance, thereby supporting the validity of the findings.

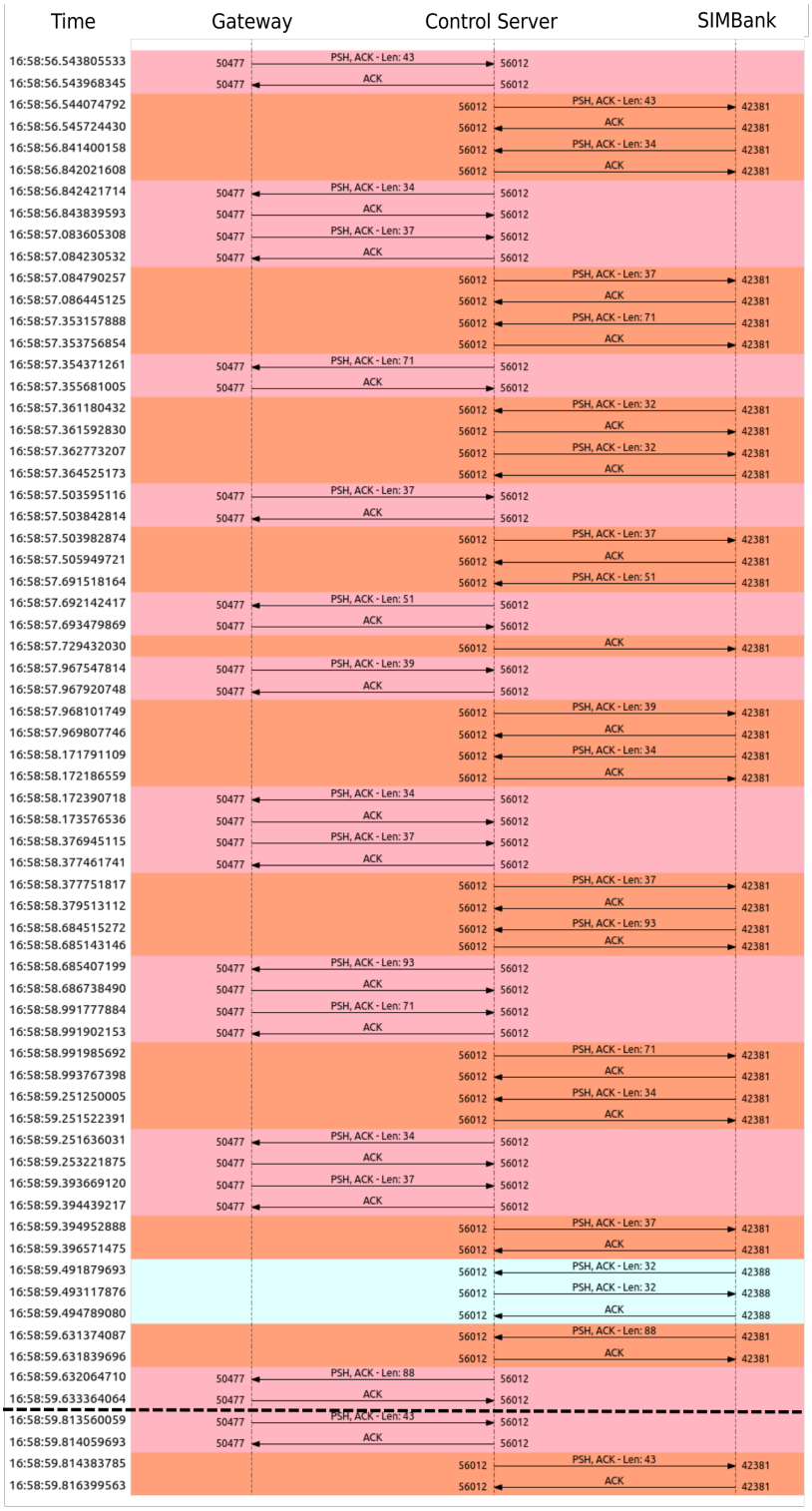


Figure 16: Hybertone *SIMBox* components TCP interactions during authentication

Table 3: Testbed component specifications

Parameters		**Values**
Host PC (BS, MME, SGW)		Intel(R) Core(TM) i9-10900K CPU@3.70GHz, 16GB RAM, GB Ethernet controller
Cell	Bandwidth	5MHz FDD
	Configuration	SISO (Single Input Single Output)
	Frequency	Downlink center frequency: 1845 MHz, Band 3
Programmable SIM cards		Sysmocom SysmoSIM-SJS1
Mobile Phones		Samsung Galaxy Note 4 (x3)
		Samsung Galaxy S3
		Xiaomi Redmi Note 9
		Xiaomi 10 Lite 5G (x2)
		FairPhone 4 5G
		OnePlus Nord Model 5G
		Sony XPERIA
		Samsung galaxy Z Fold2 5G
		Samsung Galaxy A90 5G
SIMBox appliances		Hybertone - SIMBank: SMB32 - Gateway: GoIP8 (x2) - Control server v. 2022-5-11 (Host PC: Intel(R) Core(TM) i5-4590 CPU @ 3.30GHz, 8GB RAM, GB Ethernet controller)
		Portech - SIMBank: SBK-32 - Gateway: MV-374 - Control server SS-128 (Host PC: Intel(R) Core(TM) i7-4610M CPU @ 3.00GHz, 16GB RAM, GB Ethernet controller)

Table 4: Finegrained analysis of authentication latency (in ms)

		SMBHyb_rem						*SMBPor_rem*			*srsUE softphone*		
		TCP			*UDP*			*UDP*					
Step	**Dir.**	**latency**	**Transfer**	**Processing**	**latency**	**Transfer**	**Processing**	**latency**	**Transfer**	**Processing**	**latency**	**Transfer**	**Processing**
4. Authentication response	Uplink	3259	15 sessions 4.7 ± 9.2 total: 70.6	14 occurences: - SIMBank (8) 218±8 total: 1744.3 - Gateway (6) 211±6 total: 1265.8	2379	Not clearly identified	12 occurences: - SIMBank (6) 236.1±116.3 total:1416.4 - Gateway (6) 139.3±73.6 total: 835.9	1199	1 session total: 4.2	Not clearly identified - Before transfer 774.4 - After transfer 420.4	59	4 sessions 0.12±0.15 total: 0.58	5 occurences - SIM card (2) 15.6±14.5 total: 31.1 - ME (3) 9.4±10.8 total: 28.1
10. Attach complete	Uplink	40	1 session total: 2.8	/	60	1 session total: 0.2	/	52	/	/	48	/	/
Total latency		3411 ms			2521 ms			1320 ms			259 ms		

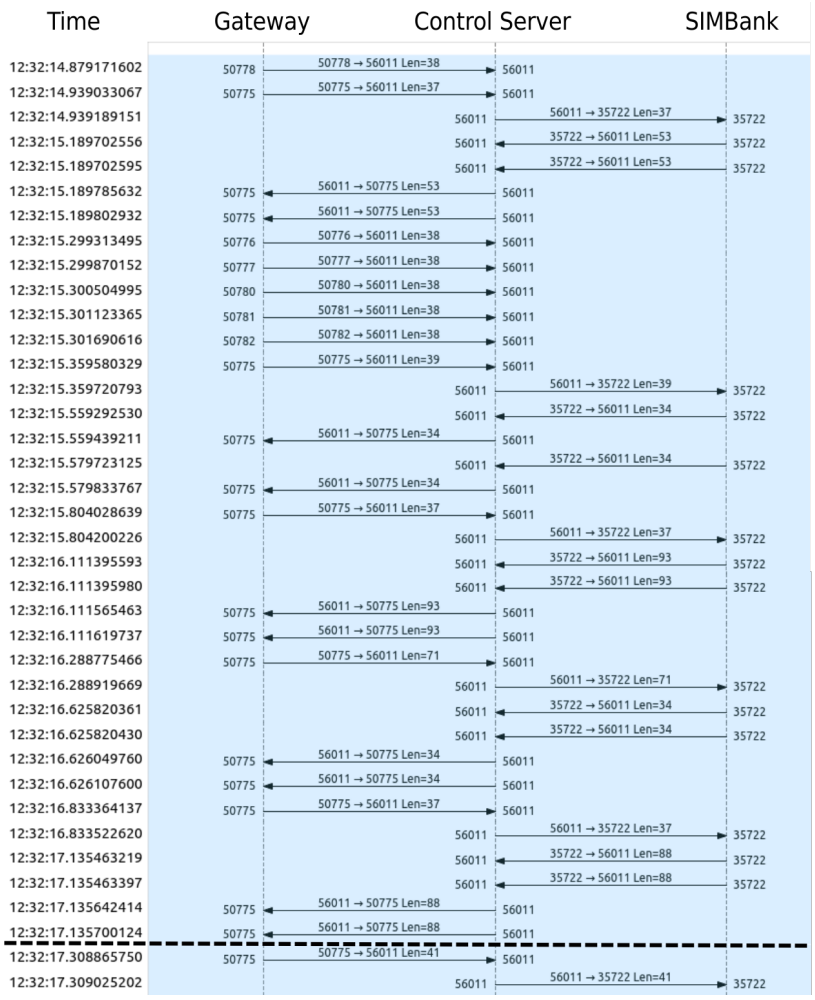


Figure 17: Hybertone *SIMBox* components UDP interactions during authentication