SIMBox bypass frauds in cellular networks: Strategies, evolution, detection, and future directions

Anne Josiane Kouam, Aline Carneiro Viana, Alain Tchana

SIMBox bypass frauds in cellular networks: Strategies, evolution, detection, and future directions

Anne Josiane Kouam
Inria, Ecole Polytechnique - IPP
Palaiseau, France
anne-josiane.kouam-djuigne@inria.fr

Aline Carneiro Viana
Inria
Palaiseau, France
aline.viana@inria.fr

Alain Tchana
Inria, ENS Lyon
Lyon, France
alain.tchana@ens-lyon.fr

***Abstract*—Due to their complexity and opaqueness, cellular networks have been subject to numerous attacks over the past few decades. These attacks are a real problem to telecom operators and cost them about USD 28.3 Billion annually, as reported by the *Communications Fraud Control Association*. *SIMBox* fraud is one of the most prevalent of these telephone frauds. It consists of diverting international calls on the VoIP network and terminating them as local calls using an off-the-shelf device, referred to as *SIMBox*. This paper surveys both the existing literature and the major *SIMBox* manufacturers to provide comprehensive and analytical knowledge on *SIMBox* fraud, fraud strategies, fraud evolution, and fraud detection methods. We provide the necessary background on the telephone ecosystem while extensively exploring the *SIMBox* architecture required to understand fraud strategies. Our goal is to provide a complete introductory guide for research on *SIMBox* fraud and stimulate interest for *SIMBox* fraud detection, which remains little investigated. In this vein, we conclude the paper by presenting insights into tomorrow's *SIMBox* fraud detection challenges and research directions.**

***Index Terms*—Telephony networks, *SIMBox* fraud survey, *SIMBox*, fraud detection.**

I. Introduction

Telephone fraud presents a considerable problem for Mobile Network Operators (MNO) around the world. According to the Communications Fraud Control Association, the global fraud loss is estimated to USD 28.3 Billion in 2019 [1]. Illegal bypass termination [2], also known as *SIMBox* fraud, is by far one of the most prevalent frauds affecting the telecommunication market [3]. In many countries, the international termination rate (ITR) is considerably higher than the local (retail) termination rate (LTR) within the country (e.g., 2.8 to 28 times higher in Cameroon [4]). This makes it profitable for fraudsters to bypass the regular interconnect operator when terminating calls in the country as they can pay the lower local rate instead of the ITR.

The simplest way of committing bypass fraud involves setting up a *SIMBox* (VoIP GSM gateway), a standard device that can be easily acquired via the internet and equipped with a bundle of SIM cards. The calls are typically routed via an internet flow (VoIP) to the *SIMBox* residing in the terminating country. The *SIMBox* then converts the VoIP call into a local mobile call to the receiving party.

SIMBox fraud is a significant problem for telecommunication operators and tax authorities of the affected countries, as international traffic taxes cannot be collected. Beyond direct revenue loss, bypass fraud also leads to poor customer experience. Examples of such call quality experience degradation are low voice quality due to latency issues, highly-compressed IP connections, longer call set up times, or still, missing or incorrect calling Line Identifiers. The latter results in many call rejections by the called party, while missed calls cannot be returned. Such degradation impacts the customer experience, which directly influences loyalty, lifetime value, and revenue. *SIMBox* fraud mainly affects developing countries. About 78% of African countries and 60% of Middle Eastern countries are fraud destinations [5], and as much as 70% of the incoming international call traffic is terminated fraudulently in some of these countries [6]. However, fraud is also present in America (e.g., interstate calls) and in Europe for international SMS termination [7].

SIMBox fraud has been around since at least 2011, but it continues to plague the telecom industry due to the following facts:

(1) There is very little research in the area; since 2011, only 14 articles have investigated this problem. It is mainly because research in this area requires the use of Call Data Records (and therefore partnerships with operators), private data as call audio records (e.g., [8]), or, in some cases, the purchase and set up of a *SIMBox* architecture (as done in [9]), which is not easily accessible. Indeed, in [8] and [10], the authors could not validate their detection model due to a lack of ground truth. It also explains why most of the information we identified on fraud is not from the scientific literature but from general research (articles from private anti-fraud companies, websites of *SIMBox* manufacturers, fraud businesses, association reports, and articles from news organizations) as categorized in Table I.

(2) Fraud techniques are evolving rapidly and are becoming more and more intelligent. Fraudsters adjust their strategies with each detection. As a result, the detection methods used by operators are limited, and some scientific contributions are outdated compared to the *SIMBox* advances at their arrival.

The purpose of this manuscript is to provide all the necessary elements to understand the *SIMBox* fraud problem in its entirety. It is intended to assist researchers in *SIMBox* fraud detection and it may spark new research interest in this currently under-explored area. To the best of our knowledge, this is the first paper in the literature presenting the current

TABLE I: Classification of surveyed information sources on *SIMBox* fraud

Category		Sources	Total
Scientific literature	Global information on *SIMBox* fraud	[2] [11–13]	4
	SIMBox fraud detection contributions	[8–10] [14–24]	14
Private anti-fraud companies		[6; 7] [25–31]	11
SIMBox manufacturers		[32–50]	19
SIMBox fraud businesses		[5] [51–53]	4
Association reports and news organizations		[1],[3] [54–57]	6

state-of-the-art of *SIMBox* fraud. To help the reader easily navigate through the paper, we provide in Figure 1 a road-map organization of all the sections discussed. A dashed line between two sections in the Figure indicates a direct link between these sections' ideas. Therefore, we provide a global understanding of *SIMBox* fraud and its context (Section II - III) which is necessary for the comprehensive review presentation of both the fraud strategies (Section IV - V) and the literature's detection contributions (Section VI). From these analysis, we point out open research issues in the area of *SIMBox* fraud detection (Section VII) and propose some directions (Section VIII).

More specifically, our contributions are as follows:

- We provide an overview of the telephony ecosystem around fraud, with the telephony networks and stakeholders involved, and discuss the main telephony functionality affected by *SIMBox* fraud, i.e., call routing (see Section II).
- We describe the *SIMBox* fraud ecosystem while explaining the fraud schemes, its financial benefits for fraudsters, and what facilitates its existence (see Section III).
- We deeply explore the system behind the *SIMBox* by commenting on its components and architecture (see Section IV). This work is based on extensive research into the specifications of different *SIMBox* models.
- We consider newer models of *SIMBox* having the advanced capability of simulating human communication behavior, which hardens detection. In this vein, we examine the temporal evolution of *SIMBox* fraud strategies related to human behavioral simulation (see Section V).
- After extensively exploring the *SIMBox* fraud ecosystem, we study the related detection methods in the literature. Here, a categorization of *SIMBox* detection methods is introduced and a qualitative comparison is presented (see Section VI).
- As in most security problems, *SIMBox* fraudsters are evolving at the same rate, if not faster, than the research community. In this perspective, we highlight some incoming challenges concerning the evolution of telecommunication technologies (such as 5G and 6G) and take the risk in prospecting the frauds of tomorrow (see Section VII).

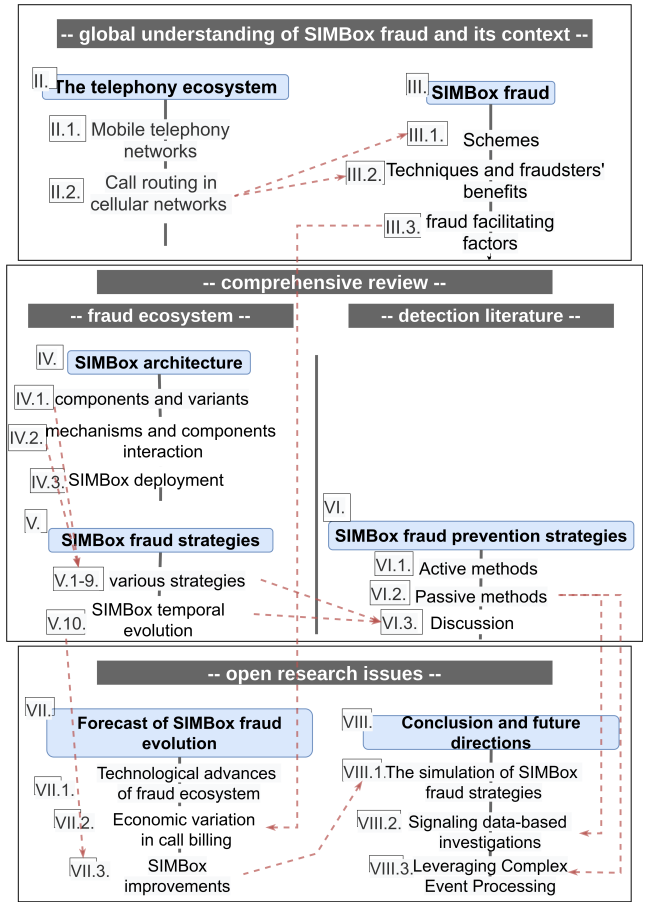

Fig. 1: Road-map organization of the manuscript

Finally, we draw conclusions of our work and comment on the future research directions to prevent *SIMBox* fraud in Section VIII.

II. The Telephony Ecosystem

In this Section, we present the key elements of telephone networks necessary for understanding *SIMBox* fraud. We focus on the *call service* by first discussing in Section II-A how calls are provided in mobile cellular and VoIP networks as well as the interface between these two networks. We then present in Section II-B legitimate call routing in cellular networks, including involved stakeholders, routing schemes, and money flow.

A. Mobile telephony networks

1) Calls in cellular networks: Call routing is a service provided by 2G, i.e., GSM networks, based on establishing a direct circuit between interacting subscribers. Although newer generations of wireless technologies offer access to a wide range of high-speed data services, they still rely on the 2G circuit-switched architecture to route telephone calls. Therefore, we mainly mention GSM networks throughout the manuscript, but more recent generations are also concerned.

For a mobile customer, the *Mobile Equipment* is the entry point to the cellular network. It is uniquely identified by its

International Mobile Equipment Identity (IMEI) and requires a *Subscriber Identity Module* (SIM) card to access the operator's network services. The SIM card is uniquely identified on the network by its *International Mobile Subscriber Identity* (IMSI) and contains a cryptography key assigned by the operator to encrypt communication.

Mobile Operators keep track of all services (i.e., calls, SMS, or data) transiting on their networks and resource use. This is done through the generation of *Charging Data Records* (CDR[1]) at the level of the core network switches. For instance, for a phone call, a CDR line may record information such as the caller and the called parties' IMEI and IMSI codes, their cell ids indicating their positions, the start timestamp of the call, and its duration. CDRs, collected and processed in a central location, are used for billing purposes or by fraud management units (referred to as *Revenue Assurance and Fraud Management* [14]) for telecommunication fraud prevention or detection.

2) Calls in VoIP networks: Voice over IP (VoIP) is the technology used to transmit voice over wired (cable/ADSL/optical fiber) or wireless (satellite, Wi-Fi, UMTS or LTE, etc.) IP networks. VoIP network is based on *VoIP servers*, providing authentication, management and routing services to *VoIP clients*, and optionally *VoIP gateways* allowing interconnection with other telephone networks (i.e. *Public Switched Telephone Network* (PSTN) and cellular). VoIP network components communicate and exchange data according to a signaling protocol such as *Session Initiation Protocol* (SIP) or H.323 [59].

VoIP clients are *IP hard-phones*, *IP soft-phones* and even analog PSTN phones combined with an *Analog Telephone Adapter*. In companies, they are managed by a central component called IP-PBX (IP Private Branch Exchange), which allocates an IP phone number to each station and connects internal calls (see Figure 2).

VoIP is provided as a cellular network data service through mobile applications referred to as *Over-The-Top* (OTT) apps (e.g., Skype, Discord, Whatsapp). These apps, developed on the top of VoIP protocols, provide cheap call services that attract more users and are seen as a threat by mobile operators [9]. VoIP calls are cheaper than cellular ones because they rely on an IP network's existing service and infrastructure (e.g., the Internet or an Intranet). Voice data is compressed and encapsulated as IP packets before its transmission over the network; this is done by specific algorithms called *codecs*, which determine the sound quality related to the bandwidth usage. On the other hand, VoIP calls quality is generally poor due to bandwidth sharing (for services other than VoIP) and IP network latency. As a result, they are affected by packet losses, delay, and jitter, which cause gaps in the audio flow.

3) VoIP to GSM gateway: Despite the growing trend of OTT applications, cellular phone calls are still widely used by customers and are, in some cases, requisite. Therefore, to exchange with the cellular network (e.g., a call to a customer

[1]Previously referred to *Call Detail Records* and later to *Charging Data Records* in the 3GPP specification [58]

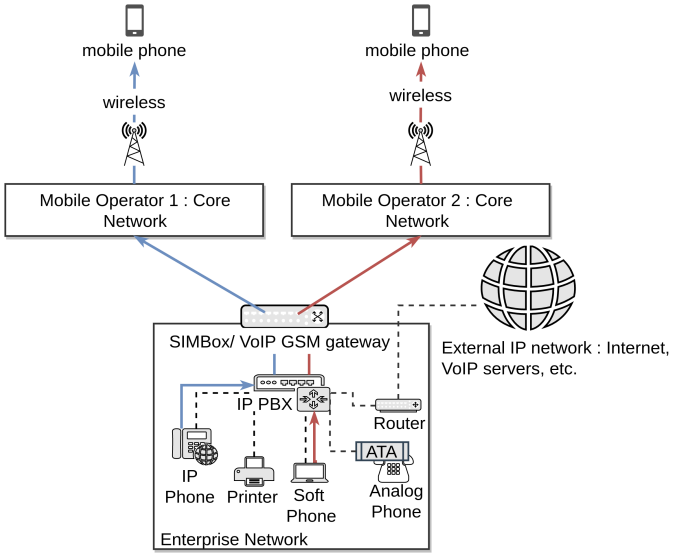

Fig. 2: Gateway from a company's internal VoIP network to the cellular network

or an employee on a mission), companies use VoIP GSM gateways, also known as *SIMBoxes*. A *SIMBox* manages a set of SIM cards from various mobile operators to ensure the live broadcast of the audio signal from the IP network to the cellular one and vice versa, which significantly extends the voice communication coverage (see Figure 2). Hence, each time a call is made from the company's IP hard/soft-phone to the cellular network, a SIM card is automatically allowed by the *SIMBox* to transmit the flow as a cellular phone call to the called party. For instance, in Figure 2, using two SIM cards, the *SIMBox* allows the simultaneous routing of two VoIP calls to subscribers of two different cellular networks.

In large companies and businesses, *SIMBoxes* have a strong market, for economic and operational reasons, but with a controlled usage by licensed telecommunication providers and regulators. Companies can significantly reduce their telephone expenses to the cellular network with these devices by avoiding roaming charges.

B. Call routing in cellular networks

Call routing schemes and involved stakeholders can vastly vary depending on if the call is domestic (within the country) or international; *on-network* or *off-network*. A phone call is said to be *on-network* (on-net) when the caller and the called parties are both customers of the same mobile operator, on the contrary of *off-network* (off-net) calls.

1) Stakeholders: The following parties may play a role in the termination of a phone call.

End-users emit and receive calls. Using multi-SIM devices, subscribers may have two or three SIM cards from different providers. This is common in developing markets as it helps subscribers always get the best offer from competing network operators.

Mobile Operators provide call routing through the traffic relay from the radio access network by the *base station* to the

core network. At the core network level, the *Mobile Switching Center* (MSC) establishes a route to the called party. In the case of an off-net call, it transfers the call traffic to the *Gateway MSC* for interconnecting with the destination mobile operator. The interconnection can be a direct link or through *intermediate carriers*.

Intermediate carriers are public (e.g., Tata Communications[2] in India) or private companies (e.g., Belgacom ICS[3] and Telia Carrier[4]) offering routes to termination or transit countries that they buy and acquire through partnerships and resell to others. They mainly intervene in international call routing when there is no direct link between the originating and the destination operators. Therefore, the route followed by the international call traffic is carrier-to-carrier hops from the originating mobile operator to the destination one.

The interconnection between carriers and operators is governed by agreements that provide the various terms and conditions, including the traffic measurement, the *Points of Interconnection* between carriers, and the quality of service standards [60]. There are numerous technologies of transport links a carrier can use to convey received traffic: satellite links, submarine communication cables, fiber rings, or such, impacting the pricing and the quality of the route. Therefore, a hop (in the international termination route) is considered *legal* if the carrier provides it with a license in its country to use a regulated transport link technology. Unlicensed carriers may use VoIP links, as they are difficult to regulate and control; formed routes/hops are thus considered *illegal*. Admittedly, there are three types of international termination routes: *white*, *grey*, and *black* [2] . A route is considered *white* when there is no illegal (black) hop all over the interconnection path. On the contrary, *grey* routes are arrangements where one hop is illegal, i.e., the originating operator sends the traffic to a legitimate carrier, but the traffic is terminated at the destination by an unlicensed carrier. This is the case of most calls from the USA to India [11]. In *black* routes, both source and destination use unconventional interconnections.

The telecommunication market is dynamic. A mobile operator usually has interconnections with several (maybe hundreds) carriers for each destination country and has to choose between them for the termination path. Besides, the quality and the price of these routes may vary weekly for the same carrier. To keep up with changes, *Least Cost Routing* algorithms [61] at the level of the *Gateway MSC* automatically select the most efficient route regarding quality and pricing to ensure the efficient use of the existing network infrastructure and maximize the operators' income.

Regulators, either public (e.g., ministries) or private, rule mobile operators' activities and partnerships in some countries. Indeed, governments usually consider telecommunications an essential public service and want to ensure services are supplied consistently with the national perception of the public interest.

2) International call: In Figure 3 an international call flow is depicted. Cellphone X calls Cellphone Y, a customer of operator B abroad. The originating operator (Operator A) transfers the call request through two intermediate carriers, with carrier 1 automatically chosen through least cost routing. Carrier 2, with a direct connection, sends the traffic to operator B's core network, which establishes the call route until cellphone Y. This example draws a white call routing as all carrier-to-carrier links are conventional, well-monitored connections specified by contractual engagements. The example assumes there are only two intermediate hops between operator A and operator B, but this value is unknown in practice as call routing is often opaque. Indeed, each carrier only knows the previous and the next hops of the termination route, as well as the originating and destination phone numbers [62]. Besides, the originating number can sometimes be missing or incorrect.

3) Money flow: In all call routing schemes (domestic on-net/off-net and international), the caller pays the call termination fees; however, international calls are generally more high-priced than domestic calls. This is because an international call may travel over multiple intermediate operators before reaching its destination. Therefore, each transit operator gets a share from the call revenue for passing over the call traffic, referred to as *settlement rate*; and the destination operator receives the *call termination fee* for terminating the international call on its network. In Figure 3, the green dashed line represents an example of money flow. Operator A bills the end customer a *collection charge* c_1, including what it retains plus the sum c_2 it pays to carrier 1 for routing the call. Similarly, each transit carrier bill includes its fees and the sum required to ensure call routing to the destination. Lastly, the destination operator (Operator B) charges (c_4) represent the termination fees.

III. The *SIMBox* fraud

This Section provides a comprehensive overview of the *SIMBox* fraud from an in-depth review of the literature and specific research we carried out with *SIMBox* manufacturers. We rely on the 5-layer taxonomy of Figure 4 defined in [62] for telephone frauds. Hence we first present in Section III-A complete fraud schemes, then fraud techniques and benefits in Section III-B, and, finally, weaknesses favoring and facilitating the fraud existence in Section III-C.

SIMBox frauds are due to two main telephone systems' inherent characteristics. First, the possibility of *interconnection between VoIP and cellular networks* is the cornerstone for SIMBox fraud. This cannot be circumvented because, as discussed in Section II-A, it provides many benefits for companies besides extending the voice communication range. Also, the *variety of operators and services* offered (transit carriers and VoIP providers) in telephony makes it challenging to ensure each service provider/carrier has good purposes. This point is difficult to tackle without undermining competition and liberalization, thus slowing down service improvement.

[2]https://www.tatacommunications.com/
[3]https://bics.com/
[4]https://www.teliacarrier.com/

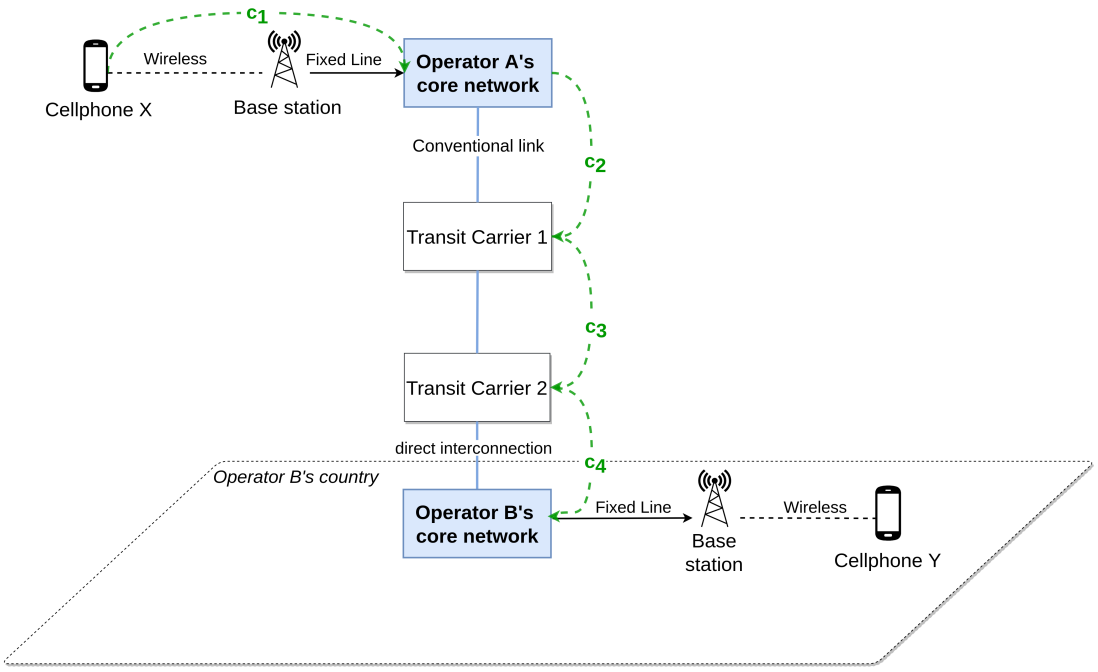

Fig. 3: International off-net call scheme

These root characteristics are the cause of weaknesses exploitable for the spread of fraudulent schemes and techniques.

A. *SIMBox fraud schemes*

SIMBox fraud consists of deviating call traffic from the conventional routing routes to a VoIP network using the appropriate gateway, i.e., the *SIMBox*. Its scheme can be broken down into four steps, summarized in Figure 5. *(1)* a call (mobile or landline) is emitted from one country to another and transits through regulated routes until a fraudulent carrier. *(2)* The fraudulent carrier uses a gateway to route the traffic through the VoIP network to a country where fraudsters' partners have a *SIMBox*, and the traffic is received at the *SIMBox* level. *(3)* The *SIMBox* reconverts the traffic to a mobile call using a SIM card as the call's origination. *(4)* The call re-originated by the *SIMBox* is terminated to the call recipient. It can be a domestic on-net call, a domestic off-net call, or an international call; the three cases are respectively distinguished in Figure 5.

1) Domestic on-net: This is the most recurrent case. The re-originated call is domestic, i.e., fraudsters partners with the *SIMBox* are located in the destination operator's (Operator B) country. An on-net termination indicates the SIM card used to re-originate the call is provided by the destination operator (Operator B). It is the most cost-effective case for fraudsters as on-net calls can be almost free charged; minimizing the cost of terminating calls through the *SIMBox* maximizes their revenues.

2) Domestic off-net: Fraudsters can use a SIM card from a competing operator to re-originate the call in domestic termination (see case 2 in Figure 5). It slightly reduces fraudsters' financial outcomes if off-net termination charges between the

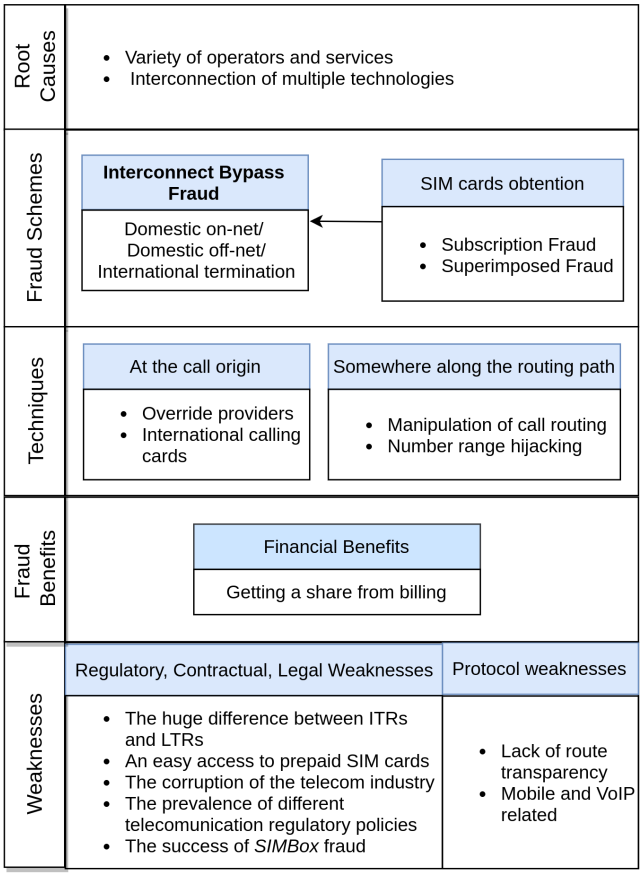

Fig. 4: *SIMBox* fraud overview according to [62]'s taxonomy

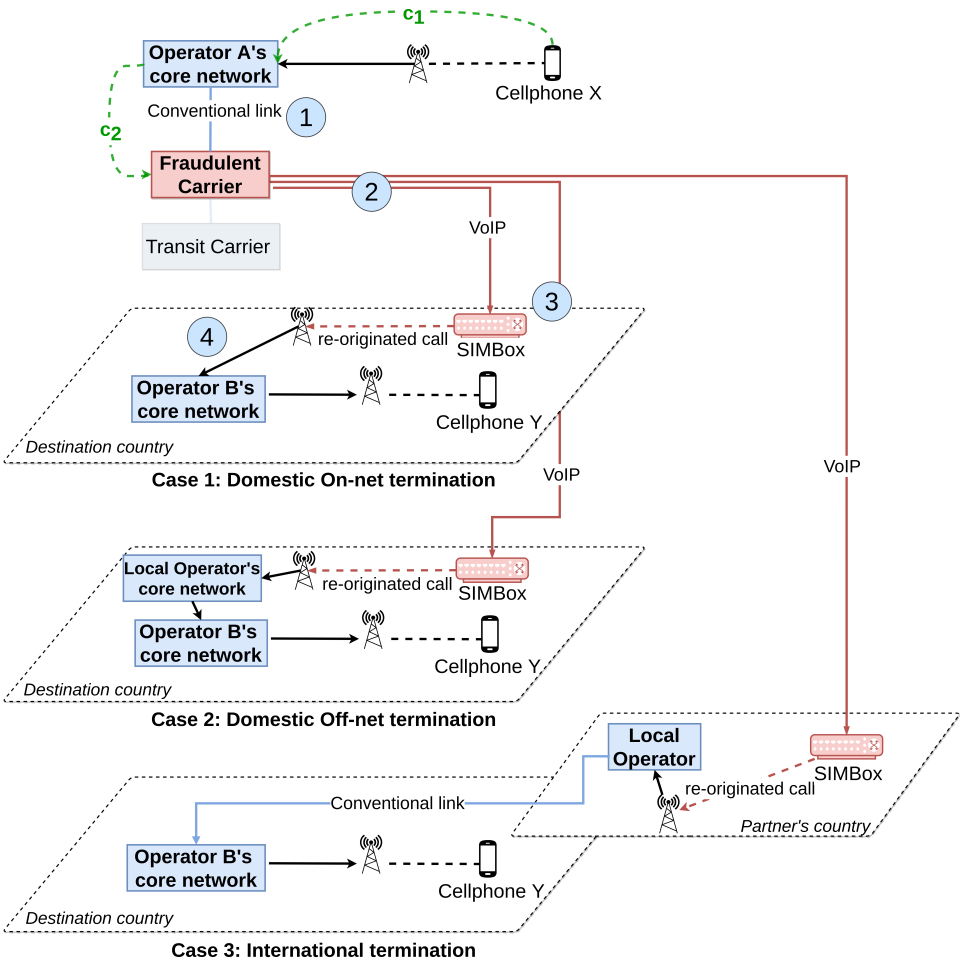

Fig. 5: Possible schemes of an international call flow in case of a *SIMBox* fraud

two operators (i.e., the destination operator and the competing one) are low enough; however, it makes it more difficult to detect the fraud in generated CDR.

Fraudsters obtain large amounts of low-charged or toll-free SIM cards from concerned mobile operators. They can act through theft or cloning, which is known as *superimposed fraud* [63; 64], or by impersonating existing subscribers' accounts, i.e., *subscription fraud* [15].

3) International: It is advantageous for fraudsters to have a partner in the destination country for the following three mains reasons: *(1)* They can easily and cheaply obtain prepaid SIM cards from both the destination and concurrent Mobile Operators; *(2)* Re-originated calls are disguised as low-charged on-net/off-net calls; and *(3)* The whole termination fees destined to the destination operator (operator B) are diverted to the fraudsters. Still, fraudsters can do the fraud even if they do not have a partner in the destination country; the re-originated call is, in this case, international. Indeed, suppose fraudsters are interested in a call with high termination charges but are limited as there is no partner in the termination country. They leverage partners' presence in countries where international call fees to the destination country are lower than the coveted termination fees. In such a way, fraudsters benefit from this termination cost discrepancy, although the generated call may be high-priced; this is known as arbitrage[5].

Usually, fraudsters compose an organization with one or a few international carriers and many partners with a *SIMBox* in different countries. Based on this, they divert much money using the different termination routes at their disposal from partners' locations. For countries where there is a partner, a domestic on-net/off-net termination is applied, and for others, an international termination is used if there is an arbitrage opportunity.

B. How do fraudsters benefit?

The primary motivation for the *SIMBox* fraud is financial. In order to obtain a *a share from the termination fees*, fraudsters insert themselves into the voice traffic termination route. We exemplify in Figure 6 fraudsters insertion in international call routing path. The Figure depicts different ways of routing international calls with actors likely to be fraudulent in a dashed-border circle and actors likely to be usurped by fraudsters in

[5]Arbitrage as a concept in economics is the manipulation of price discrepancies in different markets.

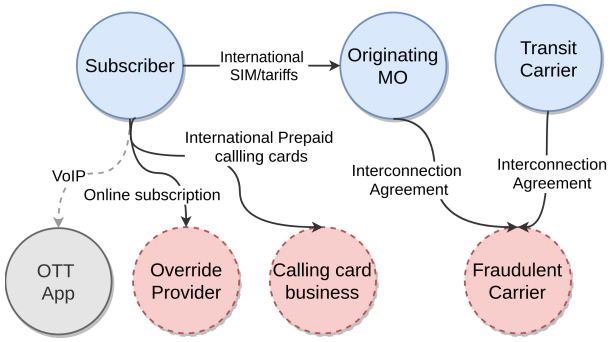

Fig. 6: Diagram of fraud intrusion in an international call routing path. Reddish circles with a dashed border represent actors who may be fraudulent. Others circles (except OTT app) represent actors whom fraudsters may usurp.

a blue solid-border circle. The insertion can be either *at the call origin* by inducing subscribers to send them the traffic instead of to their operator or *somewhere along the route* of the call as fraudulent carriers. In both cases, fraudsters' typical behavior is to pretend to route calls at costs low enough to be of interest to their audience (subscribers, mobile operators, or international carriers).

1) At the call origin: The call is diverted at its origin when the caller sends the traffic directly to fraudsters. Subscribers are motivated by the desire to make international calls at lower costs than those offered by operators. The main possibilities to achieve it are VoIP using OTT (Over The Top) applications and international calling services such as Override providers and International calling cards. *Override providers* (e.g., RebTel[6]) offer low-cost calls to mobile subscribers with an account on their application. They charge either monthly or sell call minutes. They are not explicit regarding the technology used to route calls and may use *SIMBox* devices. *International calling cards*, on the other hand, are prepaid cards that directly give international calling minutes to end consumers at discounted rates. As well, the routing method is not transparent and may be fraudulent.

The payment of fraudsters depends on the means used to insert into the voice traffic route. Calling card businesses, for instance, are paid by selling their cards on common commercial platforms, while override providers make monthly withdrawals from their subscribers or sell call minutes.

2) Somewhere along the termination route: Through the *number range hijacking* technique, fraudulent carriers insert themselves somewhere along the call routing path. *Number range hijacking* consists of dishonest carriers advertising very cheap rates for a destination number range by which they attract traffic from other operators. Usurped operators/international carriers thus transfer calls destined to such a number range to fraudulent carriers. Number range hijacking is facilitated by a *lack of due diligence* in partnership agreements between carriers. It is difficult to detect as there is no mechanism in telephony networks to directly authenticate the owner of a number range or check if an operator has the connectivity to route a call to a number range; this is known as *lack of route transparency*.

Fraudulent carriers are, in this case, paid according to the policies specified in the interconnection agreement. Therefore, the fraudulent carrier receives payment from usurped operator/carrier and shares with its partners managing *SIMBox* in destination countries.

C. What motivates and facilitates the fraud?

In this Section, we discuss regulatory, contextual, and contractual weaknesses favoring *SIMBox* fraud (see Figure 4); protocol weaknesses (lack of route transparency, Mobile and VoIP related) have already been mentioned in the previous sections. Regulatory, contextual, and contractual weaknesses explain the fraud prevalence in specific areas, i.e., Africa and the Middle East [25] compared to others, and therefore help better understand the whole fraud ecosystem. They are organized into five main factors.

(Factor 1): The huge difference between International Termination Rates (ITRs) and Local Termination Rates (LTRs). The *SIMBox* fraud benefits and main incentive lie in the fact that there is a difference between ITRs and LTRs.

Generally, international call charges are much higher than on-net/off-net calls'. This is mainly due to termination fees or destination country's ITRs having a considerable share in these charges. For instance, in Figure 3, the termination fees c_4 would generally be higher than the settlement rates obtained by transit carriers 2 (i.e., $c_3 - c_4$) and carrier 1 (i.e., $c_2 - c_3$), respectively as call traffic passes through all the switching levels in the destination country [4]. Other reasons explained below may justify this.

In their role of protecting service providers' and consumers' rights and interests within their countries, governments and National Regulatory Authorities would instead encourage high ITRs. Indeed, high ITRs serve as funding for the domestic network's development and leverage the internal economy. Besides, ITRs have no impact on domestic subscribers. Furthermore, most voice traffic originates from rich countries; money flows are thus from developed to developing countries, and high settlement rates favor the recipient countries. For instance, [54] presents the average price per minute United States' carriers pay to foreign carriers for call traffic termination. It reveals that Africa, the Caribbean, and the Middle East have the highest ITRs over time, increasing particularly for Africa.

On the other hand, to cope with the national competition, telecom companies have aggressive packages and promotions in bundles and unlimited access to calls for a given period, mostly on the same network, to lower the churn rates to attract new customers. This widens the difference between ITRs and LTRs for countries with high termination fees, particularly attracting fraudsters' attention. A primary *SIMBox* termination business, i.e., GoAntiFraud realizes in [51] a classification of the top 5 *SIMBox* fraud destination countries in 2020, in terms

[6]https://www.rebtel.com/en/

of mobile penetration, LTR, and population size. Nigeria, in West Africa, is the most prolific destination with 203 million people, 75% of mobile penetration, and an LTR range of \$0.042- \$0.048. In summary, the higher the mobile penetration and the lower the LTR, the better for fraudsters.

(Factor 2): An easy access to prepaid SIM cards. *SIMBox* fraud requires a considerable number of SIM cards used by fraudsters to re-originate bypassed international calls as local calls. A considerable amount of SIM cards is necessitated for the activity *(1)* to handle multiple calls increasing the fraud's potential and, therefore, revenues, *(2)* to perform SIM rotations in order to avoid suspicion, and *(3)* to continue providing service in case one or more SIMs are blocked. Indeed, the *SIMBox* allows to conduct several calls simultaneously and generally rotate the use of SIM cards so that ideally, a different SIM card is chosen for each call, limiting the detection of fraudsters (more detailed in Section V-A). Besides, at any time, one or more SIM cards may be blocked by anti-fraud activities and should be replaced to maintain the activity.
Moreover, fraudsters mostly use prepaid cards to limit their traceability as postpaid SIM cards require precise information on the SIM holder. Such information could facilitate the fraudsters' arrest if their SIM cards are intercepted.

Consequently, areas allowing easy access to large quantities of prepaid SIM cards would be areas where fraudulent *SIMBox* activity could quickly spread. From this standpoint, Africa and the Middle East are the areas most concerned. Indeed, on average, 94% of mobile subscriptions in Africa are prepaid, and about 80% in the Middle East [55]. This can be explained by the flexibility prepaid SIM cards allow to their holders and the low financial inclusion rate in Africa[7], making it difficult for mobile operators to ensure regular and reliable payment of postpaid SIM cards.

(Factor 3): The corruption of the telecom industry.
SIMBox fraud has been a significant fraud issue for at least ten years partly because telecom operators are not entirely engaged in combating the fraud. Although operators are motivated by severe money loss, individualism and the search for personal profit give rise to corruption infiltrated at several levels in the fight against fraud [56]. We identified the following levels.

(1) In the wholesale telecom market, operators know that low-priced routes (20 to 70% off) are more likely to be grey routes. The operator's carrier team may be bribed to buy these routes and tangle them with legitimate ones. In some cases, there are conflicts within an operator because the carrier team is buying routes that the fraud prevention team is trying to detect.

(2) Fraudsters collude with re-seller kiosks to acquire SIM cards in markets where the legal identification of SIM cards is required. They pay re-sellers to turn a blind eye to fake identity cards (IDs) or sell SIM cards identified to other subscribers. Indeed, fraudulent re-sellers use clients' IDs to register SIM cards without their consent/knowledge and later provide them to fraudsters. Such collusion can occur with no suspicion from the re-sellers employer (i.e., the mobile operator).

(3) Vendors or anti-fraud private companies can also run bypass termination to their profit. To be hired by mobile operators, they boost *SIMBox* fraud traffic before a proof of concept to inflate the apparent size of the fraud and ensure instant results of their proposed solution. Similarly, senior managers of a target operator could have their own termination business and influence the fraud-prevention team by limiting their activities to avoid being detected.

(4) At last, even the fraud-prevention team can be corrupted. Fraudsters make enough money to influence fraud prevention team agents not to block their SIM cards when detected.

(Factor 4): The prevalence of different telecommunications regulatory policies. The telephony ecosystem comprises various regulation policies and laws, and the notion of legality can significantly differ depending on the country, and the communication medium [62]. That is how the grey routes came into being: on the one hand, there is a broad regulation that permits to send traffic to a VoIP carrier legally, and on the other hand, a strict regulation that only considers the traffic of a few legacy operators as legitimate (e.g., USA and India as described in [11]).

Hence, *SIMBox* fraud is fuelled by many actors legally operating in their country. It is the case of the *SIMBox* manufacturers who produce, advertise, and sell these devices. A *SIMBox* device can be easily ordered online on Amazon or Alibaba by any individual. Some GSM gateways providers such as Sysmaster [32] explicitly provide methods to counteract detection strategies from operators and authorities. Similarly, some international transit carriers openly propose and encourage partnership for *SIMBox* termination to internauts. It is the case of Antrax [33], a company based in Latvia, which has been operating termination in 74 countries since 2017.

On the other hand, fraud is outlawed and actively combated by many countries where the economic impact is negative. For example, in Pakistan, the rapid advance of technologies developed by gateway manufacturers is seen as an "arms race" against which regulators have reacted by introducing additional fraud prevention and detection. Pakistan reportedly went as far as installing deep packet inspection technology to block all unauthorized virtual private networks in the country [54]. In some other countries, bypassing the official termination can be considered a violation of local laws. The National Communications Authority of Ghana regularly arrests offenders and imprison some. Some countries ban VoIP usage to protect their revenue from bypass termination [57].

Furthermore, *the lack of cooperation* of law enforcement authorities makes identification of fraudsters difficult, even when the fraud is detected [26]. Despite international organizations' presence, there is a *lack of joint industry initiative* to fight fraud. Due to privacy concerns and competition, operators are usually unwilling to share their pricing terms, routing options, or fraud-related findings [62]. Besides, not all operators have the same incentives to fight fraud. Indeed,

[7]According World Bank Group [65] there would be an average of about 40% of adults with a mobile-money account or in a financial institution.

TABLE II: *SIMBox* architectural variants and some providers

	Gateway		SIMBank		Control server		Some providers
	function(s)	number	function(s)	number	function(s)	number	
1	Voice server SIM client SIM manager (present or not)	1 to n	SIM manager	1 to n	Config. unit SIM server	1	Hybertone Antrax Ejoin 2N VoiceBlue Portech Dinstar
2	Voice server SIM client SIM server Config. Unit	1	none	0	none	0	Hybertone Antrax Ejoin 2N VoiceBlue Portech Dinstar Hypermedia
3	Voice server SIM client	1 to n	Config. unit SIM server (max. 32 SIM cards)	1	none	0	Hybertone
4	Voice server SIM client SIM manager (at least one)	1 to n	none	0	Config. unit SIM server	1	Hybertone Antrax Ejoin 2N VoiceBlue Portech Dinstar

sometimes a competing operator can profit from the losses and the bad reputation induced by the fraud. In other cases, fighting small-scale fraud can be more expensive than the losses due to the fraud itself.

(Factor 5): The success of the SIMBox fraud. The success of the *SIMBox* fraud acts as a factor in its perpetuation. The Communications Fraud Control Association reports the losses caused each year by this fraud as enormous, i.e., in the order of billions of dollars. These losses are a profit from the fraudsters' point of view, a real motivation that would be difficult to drop overnight. Therefore *SIMBox* fraud activity can be seen as a business opportunity: fraudsters invest their time and money and expect to receive remuneration.
SIMBox fraud is "safe"; fraudsters are rarely arrested as anti-fraud teams mainly act by blocking SIM cards [52]. When one or more SIM cards are blocked, fraudsters simply identify the flaw, adjust, and replace these SIMs to continue the activity. Besides, *SIMBoxes* are getting cheaper and more featured; more cheap-international-calling Apps are present on the App Stores (for IoS and Android users) to easily get traffic from users. It motivates other individuals (e.g., unemployed people) to take up the activity and have benefits.

In sum, the ecosystem around fraud is complex and involves factors challenging to control and contribute to fraud prevalence.

IV. *SIMBox* ARCHITECTURE

Understanding the *SIMBox* architecture and internal functioning is a significant and required step to define and comprehend the various fraud strategies, as presented in Section V. There are various *SIMBox* models on the market, depending on the manufacturer. These models may differ according to the appellation, the size, or the functionalities of the components. Yet, from an investigation performed on models offered by the five most popular *SIMBox* manufacturers [53], we could observe that they present similar organization and functional architecture. Therefore, in this Section, we first present the different components of the *SIMBox* and their organization in architectures (Section IV-A). In Section IV-B, we discuss the interaction between all these components for routing a call request. Finally, we present how *SIMBox* components are deployed and organized at a termination country or city in Section IV-C.

A. SIMBox components and architectures

A *SIMBox* is made up of three main components: the *gateway*, the *SIMBank*, and the *control server*, with each a primary role(s). A *SIMBox* architecture is defined by the quantity of each of these components, which can vary from zero (in the case the component is not present) to several, as well as the functions handled by each component. Table II summarizes the different architectural variants of the *SIMBox* with for each variant some popular manufacturers who provide it. Figure 7 depicts the first *SIMBox* architecture as presented in Table II; it is the standard *SIMBox* architecture for being the most widespread and allowing to have the most traffic capacities. Thus, we present below each of the three *SIMBox* components; respectively (1) its main roles, (2) its physical description, and (3) its functions in the first *SIMBox* architecture. At last, we discuss *SIMBox* architectural variants of Table II.

1) The gateway: Its main role is the actual conversion of VoIP traffic to cellular traffic and vice versa. As such, the gateway has a *VoIP interface* through which it receives VoIP traffic flow according to signaling protocols; it then transcodes it using supported codecs and ensures its routing on the wireless cellular network through its *cellular interface*.

The gateway's cellular interface is constituted by on-board *cellular modules* allowing the creation and maintaining of *radio channels* to establish communication with the cellular network through surrounding base stations. The gateway's cellular modules can operate at different frequencies corresponding to different generations of wireless technologies (GSM, CDMA,

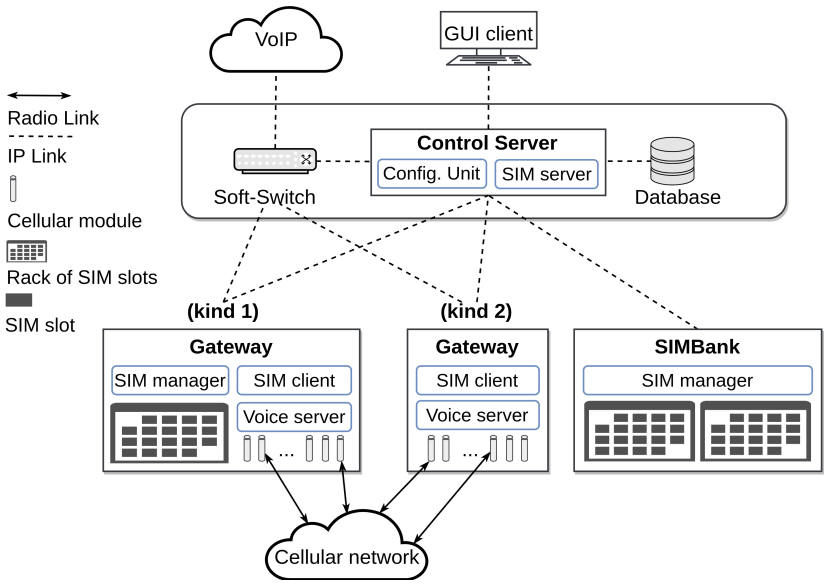

Fig. 7: *SIMBox*/VoIP GSM gateway installation functional architecture

WCDMA/UMTS, or LTE). However, for convenience, we will use the term *GSM module* to refer to a cellular module of any generation and the term *GSM channel* to refer to the corresponding established radio channel. At the physical layer, a GSM module has an antenna managed by an integrated firmware and is identified by an IMEI code.

Physically, the gateway is off-the-shelf equipment, with many GSM antennas allowing the simultaneous routing of many calls. Indeed, each antenna corresponds to a GSM module and provides the termination of one call at a time; therefore, the number of GSM modules determines the number of simultaneous calls that can be routed through the gateway. There are two kinds of GSM gateways, as shown in Figure 7. One of them (kind one) contains, in addition to GSM modules, slots for inserting SIM cards. The gateway interacts with the other components of the *SIMBox* architecture through the IP network; it, therefore, has Ethernet ports intended for intranet or internet connection and network sharing. For other connectivity, it has a USB port to which a modem or flash drive can be plugged and a *Direct Current* port for the power supply.

The basic functions of the gateway are handled by the *SIM client* module, the *voice server* and the *SIM manager* for kind one gateways.

The *SIM client* is responsible for creating GSM channels by assigning a SIM card to every GSM module. The GSM channel represents a communication node on the radio network through which mobile services (calls, SMS, or data) are provided. It is a combination of a GSM module that allows signal transmission and reception on the physical layer through antenna emitter and receiver to a SIM card that contains data and protocols for interfacing with the cellular network. GSM channels are continually created within the *SIMBox*: as soon as a GSM module is released (i.e., not bound to any SIM card), the *SIM client* requests the *SIM server* module of the control server for a new SIM card in order to proceed the binding. A GSM channel is characterized by the following: the IMEI code of the corresponding GSM module, the IMSI code of the SIM card, a status indicating the current connection state to the cellular network (i.e., *mobile registered* or *unregistered*), the current base station Identifier if mobile registered, the current call status[8], and quality indicator parameters such as the *Received Signal Strength Indicator* (RSSI) of the current cell, the *Answer-Seizure Ratio* (ASR), the *Average Call Duration* (ACD), or the *Post Dial Delay* (PDD).

The *Voice server* module manages formed GSM channels; it receives call routing requests from the *Soft-Switch* and selects the suitable GSM channel to handle each request according to configured policies. Gateways of kind one also have a *SIM manager* module in charge of the transfer of SIM card data and status to the *control server*.

2) The SIMBank: Its role is to hold a bundle of SIM cards used by the *SIMBox* for call routing. The SIMBank has the advantage of offering remote operation capability, which eases management tasks, minimizes maintenance expenses, and eliminates displacement needs to change blocked SIM cards. A SIMBank allows the installation and management of SIM cards of different mobile operators; depending on its model, it can comprise 32 to more than 256 SIM slots. Moreover, several SIMBanks can connect to the *SIMBox* providing the ability to use practically unlimited SIM cards.

Physically, the SIMBank is either a sub-rack with one or many SIM boards, each SIM board having independent control of many SIM cards, or a lighter unit easily transportable (the

[8]The call status is *Idle* if there is no call on the channel, *Processing* if a call is connecting, *Alerting* if the destination is ringing, *Active* if a call is connected, and *Calling Waiting* if the gateway is receiving another call during a running conversation.

latest models).

SIM cards within the SIMBank are controlled by the *SIM manager* module that ensures the transfer of SIM cards' data, status, and protocols to the *control server*, in order to allow the creation of GSM channels.

3) The Control server: It is the central component of the architecture and has three main roles:

(1) It coordinates the whole system's functioning by leading the establishment of GSM channels. A *SIMBox* architecture is divided into two parts: a set of gateways providing GSM modules and a set SIMBanks (or gateways of kind one) providing SIM cards. For call routing, there is a continuous creation of GSM channels within the *SIMBox*, i.e., interconnections between GSM modules and SIM cards. The control server, specifically the *SIM server* module, coordinates this process. It obtains from each SIMBank's SIM manager information on plugged SIM cards, allocates an available SIM card to each SIM client's request according to configured policies, and sets communication between connected SIM card and GSM module.

(2) The control server, specifically the *Config Unit* module, provides centralized remote administration of the *SIMBox* architecture, including visualization of the system's general state and its components' configuration. To this end, the control server stores all necessary data (settings, statistics, or logs) in a database. It has a user interface through which monitoring features are allowed. It also synchronizes all functional units integrated into the system and coordinates the interaction between them.

(3) Finally, the control server manages the VoIP traffic incoming into the architecture. To this end, it communicates with the Soft-Switch to allow the routing of traffic to prior created SIP accounts or through a SIP trunk[9]. Therefore, A SIP account is associated with either a unique GSM module or a set of GSM modules of the architecture. In the latter case, an incoming call request will be handled by an available GSM module of the SIP account, chosen according to one of the following routing policies :

- *In-Turn:* traffic is routed to the first released GSM module.
- *Balance:* traffic is routed to the fewest historical calls GSM module.
- *Sequence:* traffic is routed to the next GSM module of a sequence.
- *Random:* the GSM module is randomly chosen among the available ones.

A GSM module can be configured to route only calls to a defined number range determined by a phone number prefix. For example, if the prefix +237 is set on a GSM module, it will only terminate calls to Cameroon.

Physically, the control server is usually a web server with a database component, hosted on a *Dedicated Server* or a *Virtual Private Server* (VPS) on the cloud for availability and eased access through a web client.

4) SIMBox architectural variants: As represented in Table II, each architecture has a *traffic capacity* determined by the potential number of GSM modules and SIM cards of the architecture, which indicates the number of calls the *SIMBox* can route simultaneously, and an *operational capacity* determined by the functionalities that the architecture provides.

As above-mentioned, Architecture 1 has the most significant traffic capacity; it can support multiple GSM modules (from different gateways) and multiple SIM cards (from both SIMBanks and kind one gateways) centrally managed by a hosted control server. Also, in terms of functionality, this architecture is the most complete. Each gateway can be installed at a different location that allows simulating SIM cards' mobility (further discussed in Section V-C); the presence of one or more SIMBanks provides operational facilities, such as easy addition or removal of SIM cards from the system. Finally, the control server allows centralized management of all connected devices through advanced configurations.

Compared to Architecture 1, Architecture 4 is limited in terms of operational capacity; the architecture does not include SIMBank(s), and SIM cards are inserted in at least one kind one gateway. Therefore, to remove or add a SIM card from/to the *SIMBox*, the fraudster must move to the gateway(s) holding the SIM card, which can be tedious if many gateways hold SIM cards.

In architecture 2, there is only one device: A kind one gateway. Indeed a kind one gateway can operate standalone; it has SIM slots to insert SIM cards in the system and GSM modules for interacting with the cellular network. The architecture is light, but the traffic capacity is limited to the number of SIM slots and GSM modules of the gateway used, and only basic configurations can be performed.

Architecture 3 is a specific *SIMBox* architecture provided by Hybertone, where the SIMBank has the Config unit and SIM server functions. It makes use of a specific SIMBank model, which contains 32 SIM slots and provides SIM cards to remotely connected gateways of kind two only, i.e., gateways with no SIM slots. The traffic capacity is thus limited to 32 calls at a time, and as well configuration possibilities are limited compared to what is offered by a control server.

B. Interaction

In this Section, we describe interaction flows between *SIMBox* components; first, for the continuous creation of GSM channels and, secondly, for the termination of an incoming VoIP call.

1) Binding of GSM modules and SIM cards: The bindings are done manually, i.e., by selecting a specific SIM card for a GSM channel or using SIM and GSM grouping. For the latter method, there are two methods of binding through groups.

(First) This method is the most supported by existing *SIMBox* models. GSM modules are combined into GSM groups and SIM cards into SIM groups at the *Config Unit* level. A GSM group is a set of GSM modules, with shared

[9]A SIP trunk or SIP Peer refers to two direct static IP connections between the client's router and SIP server.

configurations, that can be located on different *gateways* and managed by different *voice servers*. Similarly, a SIM group is a set of SIM cards with shared configurations and a unique identifier. SIM cards of a SIM group can be located on different *SIMBanks* (or gateway of kind one) and managed by different *SIM managers*. A SIM group has various parameters that determine the behavioral pattern of its SIM cards.

To create the bindings, an administrator links a GSM group and one or more SIM groups. This way, SIM cards from the SIM group(s) can be bound only to GSM modules from the linked GSM group, according to SIM groups' parameters, as follows:

- When a *SIM client* requests a SIM card for a released GSM module (included in a GSM group), the Control server considers potential SIM cards, i.e., available SIM cards from the linked SIM group(s).
- These SIM cards are ordered according to a criterion to ensure SIM cards rotation (see Section V-A). The first SIM card is selected for the GSM module, and further checks are made.
- The control server checks the SIM card's availability according to its activity limitation parameters (see Section V-B). For example, it will check if the SIM card can operate at that time (time limitation) or if the SIM card has not reached the day's call threshold (parameter limitation per period). If the SIM card is unavailable, the next SIM card is selected, and the same check is made until a suitable SIM card is found.
- The control server then checks whether the selected SIM card can be connected to the GSM module according to the SIM card's GSM module selection configurations for migration (see Section V-C). The control server chooses the first SIM card that validates both last checks. If no SIM card validates, the GSM channel creation procedure is canceled and can be resumed later.

(Second) Here, bindings are based on a *scheduling group* which includes several SIM cards and GSM modules. SIM cards (respectively GSM modules) can be located on different SIMBanks (resp. gateways) and managed by different *SIM managers* (resp. Voice servers). Within a scheduling group, SIM cards and GSM modules are randomly bound together according to the group scheduling parameters. The scheduling group has two main parameters, namely the *re-allocation interval* referring to the working duration and the *sleep time* relating to the break duration after each work session. After each *re-allocation interval*, SIM cards and GSM modules cancel their current binding and turn into a hibernation state, which lasts for the *sleep time*. After hibernation, there are new bindings in the group, and the cycle starts again.

In the remainder, i.e., Section V we consider only the first mode of group binding as it is the most supported and offers more configuration capabilities to fraudsters.

2) Call termination flow: The communication diagram of Figure 8 summarizes a call termination flow in the *SIMBox*. The following steps can be identified in the Figure. ⓪ As a prerequisite step, GSM channels are continuously created within the system through *SIM clients*' requests to the *control server* (see Section IV-B1). ① A call comes from the fraudsters' VoIP network to the Soft-Switch. ② The Soft-Switch sends the routing request for the call to a registered SIP account of the Control server. ③ After processing the request (which includes anti-spam rules, see Section V-I1), the control server responds with the selected GSM channel, if there is any available. If not, the call is dropped. The control server selects the GSM channel according to the policy defined in the control server's *Role 3* (see Section IV-A3). ④ The Soft-Switch routes the call to the *voice server* of the selected GSM channel. ⑤ The voice server terminates the call to the cellular network. ⑥ At last, the call connection is completed.

C. SIMBox deployment

In practice, fraudsters usually follow some rules when deploying a *SIMBox* architecture in a country/ city for fraudulent termination. First, they have to choose where to locate the different gateways. Gateway locations must be crowded places such as city centers, high-rise buildings in the center of the market, station districts, market areas, densely residential districts, or call-center areas. Crowded places enable the camouflage of *SIMBox* calls by the massive flows of calls made in these areas. Fraudsters then rent offices or apartments with stable power and Internet access at these locations to have continuous functioning of gateways and connect them to the architecture network.

On the other hand, the system's SIMBank(s) do(es) not emit a cellular signal and can be located anywhere in the country or abroad. Still, the IP connection's quality should be good beyond a certain threshold to allow smooth communication with the gateways. For instance, in the Hybertone architecture, the network connecting the control server, the SIMBanks, and the GSM gateways should meet a packet delay of less than 300ms and a packet loss rate of less than 1%. The network bandwidth required depends on the number of SIM cards used simultaneously, and reaches the peak of 11 Kbps during a gateway registration.

The control server, as mentioned above, is usually hosted on a private server, and its visualization and configuration interface is accessible online. Finally, the Soft-Switch emitting the VoIP traffic to the system is managed by the fraudulent carrier.

V. *SIMBox* FRAUD STRATEGIES

The newer models of *SIMBoxes* are not just limited to terminating calls but also provide advanced features that help fraudsters in their activities. Indeed, as mentioned in Section III-C, fraudsters invest enough to obtain large quantities of SIM cards, and it would be a substantial financial blow if their SIM cards were blocked just after some time of activity. To avoid this, they optimize their strategies by simulating user behavior (e.g., traffic habit) [56]. This is done automatically through various features (strategies) integrated into most of the new *SIMBox* models. It falls within the framework of what is called in the literature *Human Behavior Simulation*

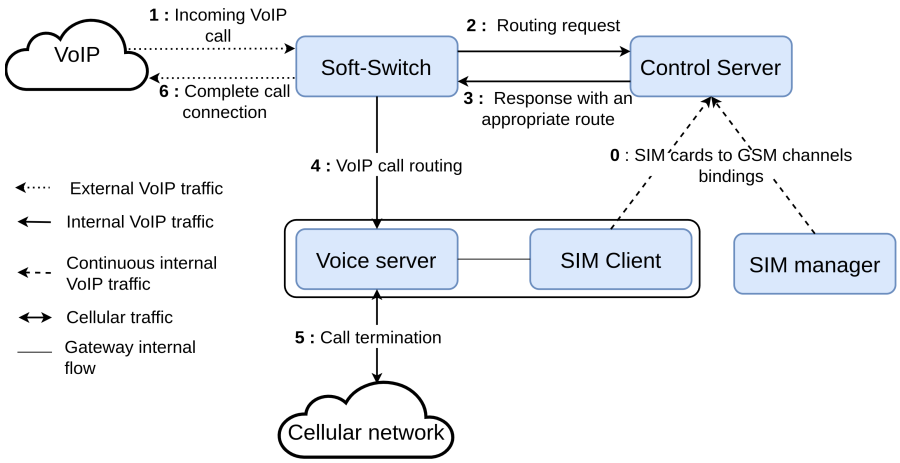

Fig. 8: Communication diagram : main stages of a call flow through the *SIMBox* functional components.

(HBS) [13]. HBS features are thus designed and improved to counteract fraud detection techniques (more details are provided in Section VI where these detection techniques are described).

We present below the different functionalities included in HBS and their current mode of operation in the light of studies we carried out on *SIMBox* models from the most popular suppliers. Table III summarizes such functionalities. Furthermore, we present some features of the *SIMBoxes* not related to HBS, but still relevant to consider, as they help fraudsters avoid detection. Finally, in the last sub-section, we aim to build a history of the temporal evolution of the *SIMBox* from the point of view of these different functionalities.

A. SIM card rotation

SIM card rotation allows the variation of the connected SIM card for a GSM module, i.e., a GSM module will repeatedly change the SIM card with which it operates to form new GSM channels. It ensures the distribution of termination traffic among the system's SIM cards, preventing one SIM card or a small group of cards from being used excessively during working hours or beyond a certain threshold. SIM card rotation, therefore, makes SIM cards operate in limited hours a day, which simulates regular customers' behavior. Its operating principle is based on the linkage of one or many group(s) of SIM cards to a GSM module (through *SIM and GSM groups links*, see Section I). At the SIM client's request, available linked SIM cards are candidates for forming the new GSM channel. The method for selecting the next SIM card for a GSM module is one of the following:

- The *round-robin or constant method* puts all available SIM cards in a circular loop and selects the next SIM card in sequence according to a defined step.
- The *random method* selects the next SIM card randomly from the SIM cards available (not including the active SIM card).
- The *statistic based factor* selects the next SIM according to the ascending or descending values of a particular defined factor. Examples of factors are SIM cards' remaining talk time, the cumulative call duration, the total calls count, and the total SMS count.
- The *statistic based factor per period* method follows the same principle as the previous method, except that factors are aggregated by period (e.g., the day or the month).

The switching of the current connected SIM card for a GSM module consists of canceling the current GSM channel and the SIM client's request for a new SIM card. It's triggered according to one of the two following methods.

(Method 1) The threshold method allows setting limit values for specific factors beyond which the switching is triggered. These parameters are indicated in the column *tresholding for switching* of Table IV. When at least one of these parameters exceeds the defined value, the SIM switching process begins when there is no active call in progress. If there is, the action to be taken should be defined, either hang up the call or wait for its end.

(Method 2) The activity script method allows setting scriplets (simple scripts) combining several parameter limitations with the operands *and* and *or*. Two parameter limitations linked by the logical operand *and* implies both limitations should be reached to end the active GSM channel session. As well, the logical operator *or* implies at least one limitation must be reached to end the GSM channel session. These parameters are indicated in the column *script-lets composition for switching* of Table IV.

B. SIM card activity limitations

Most *SIMBoxes* incorporate usage restrictions to distribute the termination traffic between the different SIM cards of the system and prevent them from being used beyond a certain threshold. Configured limitations automatically block SIM cards that have reached the limit values for defined parameters. The unblocking and reset of locked SIM cards is done either manually or automatically after a defined period. Limitations are classified as follows:

TABLE III: *SIMBox* HBS functionalities, parameters, and illustrations

HBS feature	parameter	value	Illustration
SIM rotation	Method for selecting the next SIM card	round-robin method	Hybertone [50], Antrax[35]
		random method	Hybertone [36], Portech [37]
		statistic-based factor	Hybertone [36], Antrax [35], Dinstar [38]
		statistic-based factor per time period	Antrax [35]
	Trigger to switch SIM cards	Threshold method	Hybertone [36], Antrax [39; 40], Portech [37]
		Activity script method	Antrax [39; 40]
SIM activity limitations	Type of limitation	Parameter limitation	Hybertone [36], Dinstar [41], Ejoin [42], Portech [37]
		Parameter limitation per time period	Antrax [39; 40] Dinstar [41]
		Time limitation	Hybertone [36], Antrax [39; 40], Ejoin [42], Dinstar [41], Portech [37]
SIM migration	Method for selecting the next GSM module	Manually fixed method	Hybertone [36]
		Any except previous	Antrax [43], Hybertone [36]
		Any except previous zone ID	Hybertone [44]
		Any gateway	Antrax [43]
		Specified order	Antrax [43]
Base station switching/locking	Method for selecting a base station	Manually	Portech [37]
		Default	Hybertone [36], Ejoin [42], Portech [37], Dinstar [41]
		Fixed	Hybertone [36], Dinstar [41]
		Random	Dinstar [41]
		Poll	Hybertone [36], Ejoin [42]
		Advanced	Hybertone [36] Dinstar [41]
Changeable IMEI	Method for changes at GSM module	Manual editing	Hybertone [36]
		IMEI auto change	Hybertone [36]
	Method for changes at SIM slots	Manual editing	Ejoin [42] Antrax [45]
		Random IMEI	Antrax [45], Dinstar [41]
		Prefix IMEI	Ejoin [42], Antrax [45]
		Registry IMEI	Antrax [45]
		IMEI based on TAC	Antrax [45], Dinstar [41]
The usage of other network services	Services used	Internet	Ejoin [42], Dinstar [41], Antrax [46]
		USSD commands	Hybertone [36], Ejoin [42], Dinstar [41], Antrax [47], Portech [37] 2N voiceblue [48]
		SMS	Hybertone [36], Ejoin [42], Dinstar [41], Antrax [49], Portech [37] 2N voiceblue [48]
Family list	Services used	SMS inter-sending	Ejoin [42]
		Inter-calling	Ejoin [42]
Call forwarding	Forwarding conditions	Unconditional	Hybertone [36], Dinstar [41]
		Busy	Hybertone [36], Dinstar [41]
		No reachable	Hybertone [36], Dinstar [41]
		No reply	Hybertone [36], Dinstar [41]

1) Parameter limitation: Numerous parameters related to SIM cards' call and SMS behaviors are used to limit their activity. They are distinguished in terms of consecutive occurrence, total number, or total duration. These factors are listed in the column *parameter limitation* of the Table IV.

2) Parameter limitation per time unit: Some parameters are aggregated and evaluated by period (day, week, or month). They are listed in the column *parameter limitation per time period* of the Table IV. The column contains for each parameter the potential aggregation periods. Therefore, if a SIM card reaches the limit value for a defined parameter in a configured period, it is automatically blocked until the end of this period.

3) Time limitation: The *SIMBox* makes it possible to set working periods, break times per day, or delays between each use for a SIM group. In some *SIMBox* models, the administrator can configure the week's days on which a SIM group can operate.

In some *SIMBox* models, these different limitation types can be combined using the *"and"* and *"or"* operands to form more accurate control scripts.

C. SIM card migration

The purpose of SIM card migration is to simulate human mobility. Indeed, *SIMBox* equipment (particularly the gateway) is static because, as aforementioned, its installation requires setting up a specific environment. It does not correspond to a regular customer's behavior, as people move around and may make calls at many different places. To avoid being detected *SIMBoxes* allows simulating SIM cards' movement by migrating their bindings with GSM modules from gateways to gateways (the system's gateways being located at different places in a city/country).

Therefore, the *SIMBox* allows setting a policy for choosing the next GSM module to which a SIM card will be connected. It is one of the following:

1) Manually fixed: The administrator can manually bind between a SIM card and a GSM module. The binding is stopped either at removing one of the two elements (GSM module or SIM card) or manually by the administrator.

TABLE IV: Recorded features for *SIMBox* configuration

Parameter	Tresholding for switching	Script-lets composition for switching	Parameter Limitation	Parameter limitation per time period	SMS and call inter-sending	Base station balancing
Number of outbound SMS	x	x		day		
Duration of a SIM card binding	x				x	
No cell service duration	x					
Number of call attempts	x	x		day	x	x
Total call duration	x	x	x	day, month	x	x
Total balance used	x		x	day	x	
Number of successful calls		x		day		
Consecutive successful calls					x	
Consecutive failed calls			x		x	x
Consecutive non-answered calls			x			
Consecutive calls of short duration			x			
Consecutive fast alerting calls			x			
Consecutive fast answered calls			x			
Consecutive SMS successfully sent			x			
Consecutive failed SMS			x			
Total number of failed to send SMS			x			
Total number of SMS received			x			
Consecutive network attachment failures			x			
Consecutive SIP release			x			
Consecutive GSM channel release			x			
Consecutive outbound calls with no ringback tone			x			
Answer Seizure Ration (ASR)			x			
Minimum ASR			x			
Average Call Duration (ACD)			x			
SIM card balance			x			
Post Dial Delay (PDD)			x			
Number of USSD command sending failures			x			

2) Any except previous: A SIM card can be linked to any GSM module, except to a defined amount of the previous ones to which it was connected. The amount of last GSM modules is known as *previous gateway depth*.

3) Any except previous zone ID: A SIM card cannot be linked to a GSM module with the same zone ID as the previously connected GSM module. This way, SIM cards automatically switch GSM module and location at each re-allocation.

4) Any gateway: There is no restriction; SIM cards can be linked to any gateway.

5) Specified order: SIM cards are linked according to a defined sequence list of GSM modules. The list is an ordered selection of available GSM modules.

D. Base station switching/locking

Some *SIMBoxes* can simulate mobile phones' smaller movements, which are conveniently reflected by connecting one surrounding base station to another depending on the signal strength. Therefore the *base station switching and locking* functionality configures the selection and connection to a base station for a GSM module.

First, the *SIMBox* provides a list of all surrounding base stations, each featured by: *Mobile Country Code*, *Local Area Code*, Cell Identifier, *Base Station Identity Code*, *Broadcast Control Channel*, and *Received Signal Level*. It is as well possible to have a spatial arrangement of these base stations around the GSM module. From this list, the selection of a specific base station is made according to the following modes:

1) Manually: The administrator manually selects the base station to which the GSM module should be connected;

2) Default: This mode uses the default GSM base station selection mechanism;

3) Fixed: This mode locks the GSM module to be registered to a specified base station or to switch between up to three fixed base stations;

4) Random: The base station is randomly chosen by the system in accordance to some conditions, including minimum signal strength, the frequency for base station switch, and whether the system can make a switch during a running call;

5) Poll: This mode enables the device to switch to the next base station of an ordered list at a specified frequency. The list is called base station *polling list*. It gathers all the surrounding base stations in the descending order of their signal strength, as recorded by the GSM module. The *maximum polling channel* defines the maximum number of base stations in the polling list. The *channel switching interval* establishes the frequency of base station switching occurrence. The frequency value can be randomly chosen at each switching, between a range set by the administrator. A white-list and a blacklist of base stations can be edited to define base stations that are going to be used in the polling and base stations that will not be part of the polling list, respectively;

6) Advanced: This mode triggers base station switching when the GSM module reaches the threshold value defined

for some parameters in the column *base station balancing* of the Table IV. The administrator can fix a minimum allowed signal strength as well.

E. Changeable IMEI

A GSM gateway has one IMEI per GSM module. Therefore, all SIM cards connected to a GSM module would typically match the IMEI of this GSM module in the CDRs of mobile operators. It is an obvious way to detect *SIMBoxes* as using many SIM cards in a single mobile device is quite unusual. The *SIMBox* provides the ability to set an IMEI to any used SIM card to overcome this weakness and simulate a regular customer's behavior. Depending on the *SIMBox* model, IMEI changes are made either to GSM modules or to SIM cards.

Changes related to the GSM module are done manually by the *SIMBox* administrator or automatically. In the latter case, the IMEI is modified according to one of the following: a specified frequency (by default once per hour and not more than once per 10 minutes), a threshold on the number of calls made by the channel (not less than 10), or each time a SIM card changes.

IMEI changes to the SIM card imply the IMEI code is related to the SIM card; therefore, the formed GSM channel's IMEI code is the SIM card's regardless of the GSM module to which it is connected. Changes at the SIM card level are either manual or automatic. In the latter case, they are applied to a SIM group according to one of the following patterns: randomly, prefix-based (similar to random, but with a prefix), *Type Allocation Code*[10]-based, or registry-based (the full IMEI code comes from a registry). Changes are triggered according to a *generation rule* which is either *null* – meaning that IMEI codes will be generated only once –, *periodic* – meaning that IMEI codes will be generated every period (hour, day, week, or month) – or *registration count* – meaning that the IMEI codes will be generated after a certain amount of bindings of the SIM card.

Changeable IMEI can, in some cases, become a weakness because, with an intensive GSM termination, a SIM-card can change its IMEI code at high recurrences, which is suspicious for mobile operators.

F. The usage of other network services

The primary use of *SIMBox*'s SIM cards is to terminate voice traffic through phone calls. This is a weakness as regular customers use other mobile services, namely SMS, USSD commands, and data. Some *SIMBoxes* allow to use SMS, data services, and to send USSD commands to simulate this human behavior.

1) Data services: A SIM group is configured to use a specified amount of data (generally less than 2048 MB) within a time interval. The main data services supported are web browsing and e-mails; for the former, the administrator has to define some website URLs and the Access Point Names (APN) for the connection.

[10]The Type Allocation Code (TAC) is the initial eight-digit portion of the 15-digit IMEI and 16-digit IMEISV codes used to uniquely identify wireless devices.

2) USSD commands: Unstructured Supplementary Service Data (USSD)[11] commands are sent by GSM channels to get the phone number of used SIM cards, their balances, or to make top-ups. For this purpose, the administrator selects one or more GSM channel(s) or all[12] from the interface, enters the USSD command in the appropriate field, and sends it. It is also possible to automatically send USSD commands under certain conditions set by schedule, call duration, or GSM module connection time.

3) SMS: SMS are sent similarly to USSD commands, i.e., by GSM channels. They can be used to make top-ups, get SIM card phone numbers, or SIM balances. Thus, one, several, or all GSM channels are selected for sending an SMS. A list of recipients, as well as the message content, are configurable. The encryption mode of the message can be chosen between ASC7/8 (ASCII 7/8 bit) and UCS2 (Unicode 16 bit), and the maximum length of the SMS varies accordingly. Finally, it is possible to view messages received by a GSM module and the sent messages' history.

G. List of family

The family list consists of building a virtual family network for each *SIMBox*'s SIM card. Indeed, *SIMBox*'s SIM cards make several outgoing calls to many non-related network consumers as part of the voice traffic termination. It is an abnormal behavior because most regular consumers only call and receive calls from a restricted group of network consumers, named *family list*, who, in some cases, also make calls to each other. Therefore, some *SIMBoxes* allow exchanging SMS and voice traffic between inserted SIM cards to constitute a family list.

1) SMS inter-sending: The administrator can manually schedule an SMS sending by the active SIM cards of a SIM group. He selects the phone number(s) of one or several inserted SIM cards as SMS recipients and edits the SMS content. Also, SMS recipients can be automatically selected, either randomly or according to defined threshold values on parameters listed in the column *SMS and call inter-sending* of the Table IV. Therefore, when a SIM card reaches the defined threshold for a parameter, it is sent a predefined message.

2) Inter-calling: Calls between SIM cards are featured by a *minimum and maximum duration* and a *message sending* option indicating whether an SMS should be sent by the called SIM to the caller one, just before the call. They are triggered from an idle GSM channel (i.e., registered to the cellular network but not in a call) to other GSM channels of the same GSM group. This is done randomly or according to threshold conditions on the parameters listed in column *SMS and call inter-sending* of the Table IV. If the *message sending* option

[11]USSD is a service that is provided by telecom operators and allows GSM/WCDMA mobile phones to interact with the telecom operator's computers. USSD messages travel over GSM/WCDMA signaling channels and are used to query information and trigger services. Unlike similar services (SMS and MMS), which are stored and forwarded, USSD is a session-based service. It establishes a real-time session between mobile phones and telecom operators' computers or other devices.

[12]With the option "select all" one can choose all the GSM modules available on the GSM gateway.

is activated, the called party will, before each call, select a message from a predefined message list and send it to the caller.

In addition to the here-above traffic exchange mechanism between SIM cards, some providers offer call routing based on the *SIMBox* activity history. As a result, a SIM used by the *SIMBox* to route a call from a specific phone number will be preferred to route future calls from that phone number when the SIM is available. In this way, the fraudulent SIM inserts itself into the family list of the phone number whose calls it routes.

H. Call forwarding

The call forwarding feature allows a call intended for a SIM card used in the *SIMBox* to be forwarded to another phone number so that a human agent can reply to the call. It is done according to the following policies :

- *Unconditional:* it allows to forward all incoming calls unconditionally;
- *Busy:* it allows to forward incoming calls only when the called number is busy;
- *Not reachable:* it allows to forward incoming calls when the called number is not reachable or cannot register to the mobile operator network;
- *No reply:* it allows to forward incoming calls when there is no reply from the called number.

I. Other relevant SIMBox features

Most *SIMBox* models incorporate features that are not part of human behavior simulation but are still relevant to explore as they help achieve and maintain fraudulent activity. In the following, we discuss some of these features.

1) Anti-spam: Anti-Spam consists of setting up lists of SIM numbers from which calls can be filtered. Fraudsters use this feature to filter test calls coming from detection teams, as explained in Section VI-A1. For this purpose, there are *white*, *gray*, and *black* lists of numbers.

The *white list* represents the phone numbers authorized to make outgoing calls when the outgoing call authentication mode is set to "white list." A phone number is added to this list either manually or automatically, according to certain conditions related to the number of calls and their duration over a certain period.

The *grey list* represents the phone numbers listed for a period to be reviewed by the system. Phone numbers are automatically added to the grey list based on filtration algorithms. Indeed, if, over a certain period, the amount of calls made by a phone number exceeds a specific value established for the grey list, the phone number is added to the grey list. Similarly, if, over a defined period, the number of calls of short duration exceeds a maximum allowed, the phone number is added to the grey list. A phone number remains in the grey list for a defined time, and, depending on its behavior, the blocking will be extended or not.

The *black list* represents the list of numbers that are not allowed to make outgoing calls through the system. Phone numbers are added to the blacklist either manually or based on filtration algorithms. In this latter, thresholds are set for phone numbers in the grey list. Such limits are related to the number of calls made and their duration over a given period. If a phone number exceeds the allowed values, it is transferred from the grey list to the blacklist.

2) Voice and codec configuration: Configurations to voice and codecs used in the *SIMBox* can improve the call quality. It is useful for fraudsters as some detection methods are based on audio call pitfalls (packet losses and jitter) identification. The *SIMBox* supports a variety of codecs. The administrator can activate and order them according to his preferences. Here is a non-exhaustive list of codecs supported by the *SIMBox*: G.711 a-law, G.711 μ-law, G.723, G.723.1, G.729, G.729-16, G.729-24, G.729-32, G.729-40, G.729A and G.729AB. Besides, the user can make voice configurations. He can define a minimum length for each VoIP packet received and activate the jitter buffer. The jitter buffer is designed to remove the effects of jitter from the decoded voice stream, buffering each arriving packet for a short interval (called *jitter buffer delay*) before playing it out. It substitutes additional delay and packet loss (discarded late packets). If a jitter buffer is too small, an excessive number of packets may be discarded, leading to call quality degradation. If a jitter buffer is too large, then the additional delay can lead to conversational difficulty. A fixed jitter buffer maintains a constant size, whereas an adaptive jitter buffer has the capability of adjusting its size dynamically to optimize the delay/discard trade-off. Three modes of jitter buffer are supported:

- *Fixed:* The fixed mode, which is the default mode, is a simple *First In First Out* (FIFO) mode for arriving packets, with a fixed jitter buffer delay.
- *Sequential:* The sequential mode is a fixed jitter buffer delay mode in which the jitter buffer function looks at the packets' timestamps for dropped or out-of-sequence packet problems. The data packets are sorted based on the packets' timestamps.
- *Adaptive:* The adaptive mode optimizes the size of the jitter buffer delay and depth in response to network conditions, in addition to the sequential mode functions.

Some *SIMBoxes* allow the activation of audio silence suppression through *Voice Activity Detection* (VAD) in combination with the *Comfort Noise Generator* (CNG). The purpose of VAD and CNG is to maintain an acceptable perceived quality of service while simultaneously keeping transmission costs and bandwidth usage as low as possible. In conjunction with VAD algorithms, CNG quickly determines when periods of silence occur and inserts artificial noise until voice activity resumes. The insertion of artificial noise allows a constant transmission flow with a consistent background sound throughout the call so that the listener does not notice if the line is cut. Echo cancellation mechanisms are also supported.

TABLE V: *SIMBox* Voice CDR fields

CDR Field	Description
Port	Identifier of the *SIMBox* GSM module which terminated the call
SIM card IMSI	Identifier of the SIM card used by the GSM module for the call termination
Originating number, Original destination number	Caller and called phone numbers respectively
Filtered destination number	Called number after the appliance of a filter (where relevant). Otherwise identical to the Original destination number.
Dial time	The date and time the gateway received the call request
Alert time	The date and time the called party has begun ringing (if reported by the remote party and supported by the outgoing resource)
Hangup time	The date and time any of the parties had ended the call
Call type	Either mobile originated or terminated call
Source hangup direction	0 = caller party had ended the call; 1 = call was ended by the *SIMBox* (due to cancellation by the called party, no route to the destination, unavailable destination resource, etc.)
Destination hangup direction	0 = called party had ended the call; 1 = call was ended by the *SIMBox* (due to cancellation by the caller party, no route to the destination, unavailable destination resource, etc.)
Hangup reason cause	Code as reported by the party that had first ended the call
Call duration	Time difference between answer time and hangup time

3) CDR management: The *SIMBox* provides CDR generated by its activity for traffic and accounting management. A line of CDRs as collected by the *SIMBox* can contain items reported in Table V.
CDRs are saved either on an external disk or on a server to which requests can be made. Requests aim to obtain CDR records that meet certain conditions on call duration, caller and called identifiers, call start time, call end time, and call type. They enable fraudsters to identify SIM cards/GSM modules that may behave suspiciously and refine their activity.

J. SIMBox temporal evolution

The SIMBox has evolved, both in terms of hardware and functionality. We try in this Section to build a chronological sequence to this evolution which is threefold useful: (1) it gives better visibility on the pace at which the *SIMBox* (and therefore, fraud) evolves (discussed in Section V-J3), (2) it provides insights on the motivations for this evolution, which helps us guess how the fraud can further improve (discussed in Section VII) (3) it allows getting a global view of the potential of the *SIMBox* (and therefore, fraud) positioned in time, which can be opposed to the detection potential (depicted in Table VI).

Very little information is available on the evolution of *SIMBoxes*. *We have been able to collect the data presented in Figure 9 by reporting news about the addition or update of components on the websites of the leading SIMBox suppliers.* The Figure gathers the upgrades of four different manufacturers (*Hybertone*, Dinstar, Ejoin, and Antrax). Each line represents upgrades from one category (hardware or functional) provided by one manufacturer. The name and the category of the upgrades are indicated below the line. For the [Antrax: functional] line's updates, the text is right-aligned for space reasons.

1) Hardware evolution: The SIMBank appears to be the first update for each supplier. Its first occurrence is in 2012 by Hybertone, with the capacity of only 32 SIM slots. Over time its size, i.e., the number of simultaneous SIM cards, evolved very quickly, to the peak of 256 SIM slots from 2015 by Ejoin. The presence of the SIMBank allows the deployment of *SIMBox* of architectures 1, 4, and 3 for Hybertone (see Table II). *Therefore, since 2012, fraudsters can manage their gateways remotely and realize the SIM migration strategy (ref V-C) with at least two gateways*.

On the other hand, gateways' evolution is seen in the number of modules and SIM slots and the supported cellular network technologies. Similarly to SIMBank, the number of GSM modules evolves from the value eight by Hybertone in January 2012 to the peak of 32 in 2012 by Dinstar provider. *Therefore, since 2012, it is possible to terminate 32 calls simultaneously using a gateway.*

We notice that despite the presence of the SIMBank, throughout the evolution, manufacturers still provide gateways incorporating SIM slots, i.e., kind one gateways. The number of slots is a multiple of the number of modules and gradually evolves to a maximum of 512 ports in March 2017 by Ejoin. It proves that kind one gateways remain popular because they allow easy management as done with *SIMBox* architecture 2. *Nevertheless, we think that this SIMBox architecture is primarily used by companies (ref Section II-A3) as it offers limited (fraud) functionalities, but a significant percentage of fraudsters could use them as well.* The physical evolution of this *SIMBox* model has been made in size: from sub-rack with many GSM and SIM boards to lighter and easily transportable units. Only Antrax provides gateways without SIM slots, making it clear that it promotes a distributed architecture.

Over time gateways support new generations of cellular network technologies. In 2013 we had GSM and CDMA gateways, WCDMA in 2015, and LTE in 2016. The more cellular technologies a gateway supports, the more features it offers. For instance, GSM gateways cannot enable data services contrarily to CDMA and WCDMA gateways, and LTE gateways allow higher data rates. *We deduce that since 2013 with the providing of CDMA gateways by Dinstar, fraudsters can make usage of other mobile services (ref V-F).* Hence manufacturers are providing gateways that support multiple cellular technologies (e.g., Dinstar and Ejoin).

2) Functional evolution: Information on functional evolution is more challenging to obtain. We were only able to collect them for two manufacturers: *Hybertone* and *Antrax*. Both providers make a point of *changeable IMEI* and *limitations of activity of the SIM card*. Indeed, since 2011 with *Hybertone*,

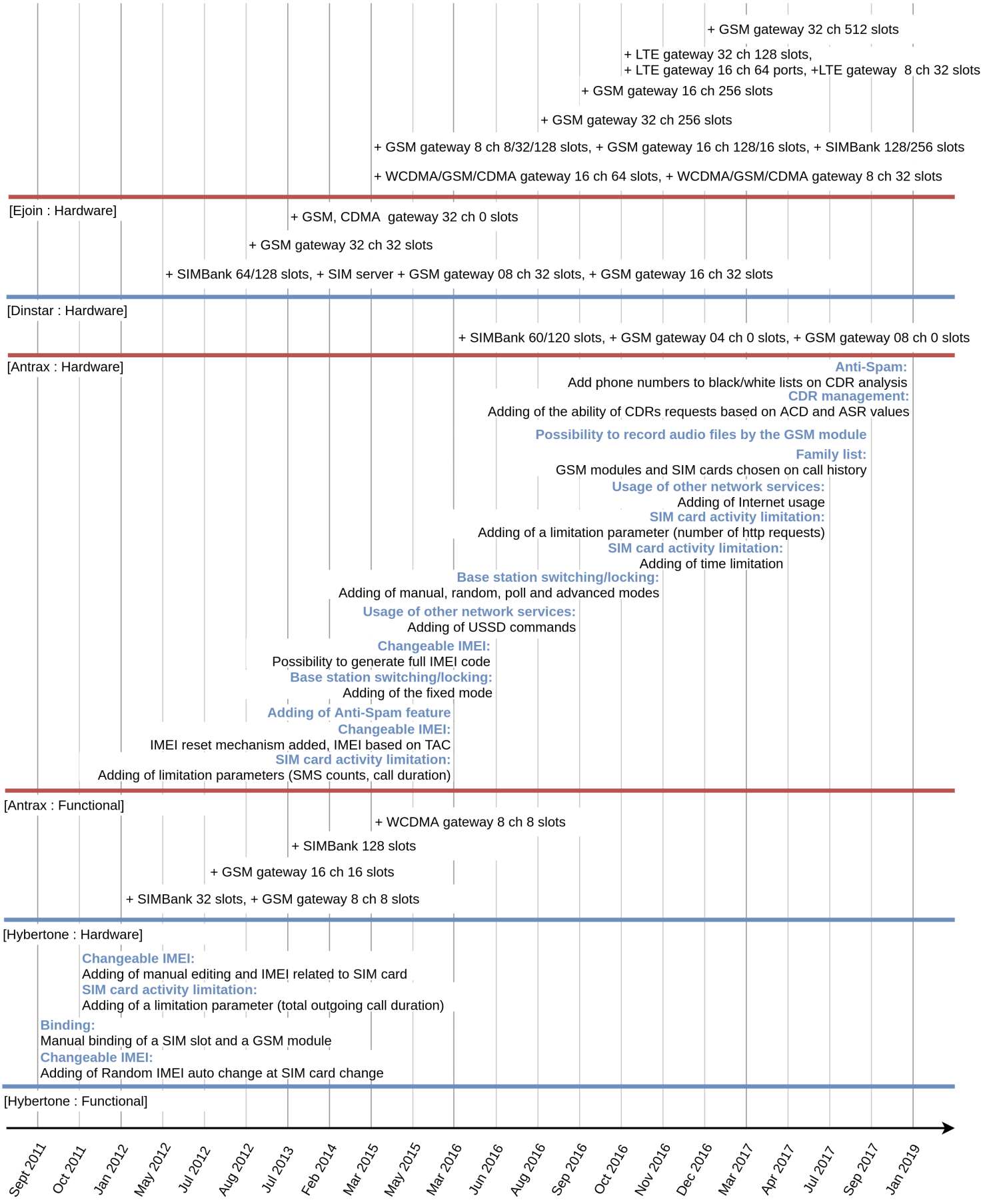

Fig. 9: Timeline of the *SIMBox* evolution from material and functional points of view.

the *changeable IMEI* feature is already supported, as well as the SIM card call activity limitation. This is expected as the IMEI, if not changed, makes it easy to determine all SIM cards operating in a *SIMBox*' GSM module. Similarly, if the call traffic is not regulated and distributed to the different SIM cards of the architecture, the SIM cards will be easily identifiable. *Antrax* started its activity later in time (in 2015); it added in 2016 the *Anti-Spam* feature to register and block numbers used by operators for test calls and developed many other features above-mentioned. Furthermore, since 2017, the *Antrax SIMBox* model can record audio tracks of calls routed through the *SIMBox*. It represents a real intrusion because fraudsters can eavesdrop on bypassed calls without the call sender and recipient's permission.

3) Evolution pace: The evolution of the *SIMBox*, both from a hardware and functional point of view, is significant. We notice that during the period we were able to collect information from each provider, the updates are frequent. There is on average one update every six months within a provider, with a maximum of 2 updates per month (e.g., Hybertone from Sept 2011 to Oct 2011) and a minimum of one update per year (e.g., Antrax from Sep 2017 to Jan 2019). With a comparable evolution rate, we assume that *SIMBox* nowadays (and therefore, fraud) supports many more features, more accurate than what is available in the literature (i.e., provider websites). For instance, one of the latest functional updates made by Antrax is the possibility to analyze CDRs generated by the *SIMBox*. It paves the way for a wide range of advanced HBS features based on current AI advances.

VI. *SIMBox* FRAUD-PREVENTION STRATEGIES

This Section presents all existing solutions for detecting and preventing *SIMBox* fraud in the literature. They are summarized in the Table VI. The Table only lists the major detection contributions of the literature. Regarding their operation mode, we organize in Figure 10 the detection solutions into two categories: active and passive solutions. Also, we classify the passive solutions into three sub-categories according to the analyzed data type. In each category/sub-category in Table VI detection solutions are in chronological order. For each detection solution, we also present the countermeasures adopted by fraudsters (previously discussed in Section V for most) with the year of their first usage according to the *SIMBox* temporal evolution (see Section V-J).

A. Active methods

The active methods require a permanent action by one or more entities. These methods are commonly considered as *classical methods* because they represent the first response of telecommunication companies to *SIMBox* fraud. They require significant material resources to be implemented.

1) Test Call Generation (TCG): The principle of TCG consists of setting up test phone numbers in a target mobile network and make calls to those test numbers from different countries through many different interconnect voice routes around the world. This way, a local *Calling Line Identification* (CLI) indicates a *SIMBox* number and can be acted upon accordingly. Once routes having a high volume of *SIMBox* terminations are detected, the call campaign focuses on them to maximize detection. A large number of test calls are generated in a very short time and may use anti-white list services [66] and specialized SIM cards with *CLI Restriction Override* (CLIRO) to overcome the hiding of the CLI in calls routed through the *SIMBox*.

TCG is all about probability, i.e., the more test calls cover routes, the more likely *SIMBox* fraud cases are to be detected. TCG is known for not making false positives, which explains its wide adoption by anti-fraud services [27–31]. Yet, it is expensive because it requires making several calls, and every test call is associated with cost/network resource consumption.

TCG method worked very successfully for many years. Yet, around 2012 and 2013, its effectiveness dropped off significantly because of the following reasons. First, the *SIMBox* fraudsters figured out how to avoid detection by test calls. For instance, they perform an analysis on the voice call traffic coming toward their *SIMBoxes*. Based on usage patterns (e.g., as discussed in Section V-I1), they can differentiate calls to real subscribers from those originating from a TCG campaign. They can then either block test calls and prevent them from reaching the *SIMBox* or reroute them to a legitimate route. Second, fraudsters can allocate pools of SIM cards to be sacrificed. Therefore, they allow the detection of these SIM cards by TCG to make the mobile operators feel confident of their results. It deceives anti-fraud teams' vigilance and will enable fraudsters to conduct the activity with other SIM cards. Moreover, the sacrificed SIM cards are chosen not to use HBS; they, therefore, have an obvious fraudulent profile (i.e., high call traffic and no mobility) that cannot be leveraged (using other methods) to identify other SIM cards.

2) Rule-based methods in Fraud Management System (FMS): Rule-based methods consist of establishing basic rules for subscriber profiling to identify fraudulent SIM cards (e.g., done in [23]). This involves the analysis and monitoring of call patterns (e.g., outgoing call count, cell ids counts, incoming to outgoing call ratio, SMS originating/terminating counts, etc.) of a set of subscribers by experts looking for an abnormal behavior originating from an operator's SIM card or terminating over it. Any case identified and validated (through a call or a similar action) can then be used to profile and uncover other similar SIM cards.

This approach is less costly and has better coverage than TCG because once a profile is established, it can be extended over all available subscribers for a wide detection range. However, it has several limitations. First, it causes a non-negligible rate of false positives and requires continuous monitoring and field expert intervention. Second, through time, the whole process of analyzing data gets more complex as rules are added to the system. It increases the detection latency, allowing fraudsters to make enough profit before being blocked. Finally, this method can not scale as it requires human intervention.

FMS has been effective in detecting *SIMBox* fraud prior the integration of *Human Behavior Simulation* (HBS) (discussed

TABLE VI: Summary of existing detection work on *SIMBox* fraud detection classified by category

Category		Detection method	Year	Principle			Countermeasure	Year
Active		TCG	∼2010	Generation test calls from abroad to a target network and checking of the CLI for each call			HBS - AntiSpam - Call forwarding Obtention of a country's phone numbers white list Sacrifice of obvious SIM cards profiles	2012
		Rule-based methods	/	Basic rules establishment for subscriber profiling			HBS - SIM card activity limitations - SIM rotation - SIM migration	2011
Passive	CDR-based			**Data Preparation**	**Model building and evaluation**	**Detection time**		
		[16]	2013	Call behavior features	Artificial Neural Network	A day	HBS - SIM card activity limitations - SIM rotation	2011
		[17]	2014	Call behavior features	Support Vector Machine	A day	HBS - SIM card activity limitations - SIM rotation	2011
		[15]	2014	- Call behavior - Mobility - Entity properties features	- Detection based on IMEI code - Linear combination of Random Forest, ADTree and Functional Tree	A week	HBS - Changeable IMEI - SIM migration - SIM card activity limitations - SIM rotation	2011
		[10]	2015	- Call behavior features - Mobility features - Mobile services usage features	Fuzzy logic	Not mentioned	HBS 1. SIM migration 2. SIM card activity limitations 3. SIM rotation 4. Usage of other network services	2011 (1-3) 2013 (4)
		[18]	2015	- Call behavior features - Mobility features	Complex Event Processing	Real-time	HBS - SIM migration - SIM card activity limitations - SIM rotation	2011
		[14]	2016	- Call behavior features - Entity properties features	- Artificial Neural Networks - Support Vector Machine - Boosted Trees - Logistic Classifier	A day	HBS - SIM card activity limitations - SIM rotation Fraudulent obtention of SIM cards	2011
		[19]	2018	- Call behavior features - Mobility features - Mobile services usage features - Entity properties features	- Artificial Neural Networks - Random Forest - Support Vector Machine	- 4 hours - A day - The month	HBS 1. Changeable IMEI 2. SIM migration 3. SIM card activity limitations 4. SIM rotation 5. Usage of other network services Fraudulent obtention of SIM cards	2011 (1-4) 2013 (5)
		[20]	2019	Not mentioned	- Artificial Neural Networks - Support Vector Machine	Not mentioned	/	/
		[21]	2020	- Call behavior features - Mobility features - Mobile services usage features	- Artificial Neural Networks - Random Forest - Support Vector Machine	- An hour (using a sliding window) - Fixed 4 hours	HBS 1. Changeable IMEI 2. SIM migration 3. SIM card activity limitations 4. SIM rotation 5. Usage of other network services Fraudulent obtention of SIM cards	2011 (1-4) 2013 (5)
		[22]	2020	Call behavior features	Complex Event Processing - Lossy Counting with fast forgetting algorithm	Real-time	HBS 1. SIM card activity limitations 2. CDR-based analysis	2011 (1) 2019 (2)
	Audio-based	[9]	2015	Detection of call audio degradation (packet losses and jitters) indicative of routing through the VoIP network		Real-time	Variety of codecs Voice Activity Detection (VAD) Comfort Noise Generator (CNG)	/
		[8]	2017	- Analysis and recognition of call speakers voices - A SIM card used by a number of speakers beyond a defined threshold is considered fraudulent			/	/
	Signaling-based	/	Mentioned in 2015	- Intuition: SIMBox devices generate a fingerprint in signaling messages exchanged with the cellular network - Not yet exploited in the literature - Described as promising (real-time and accurate)			None to the best of our knowledge	/

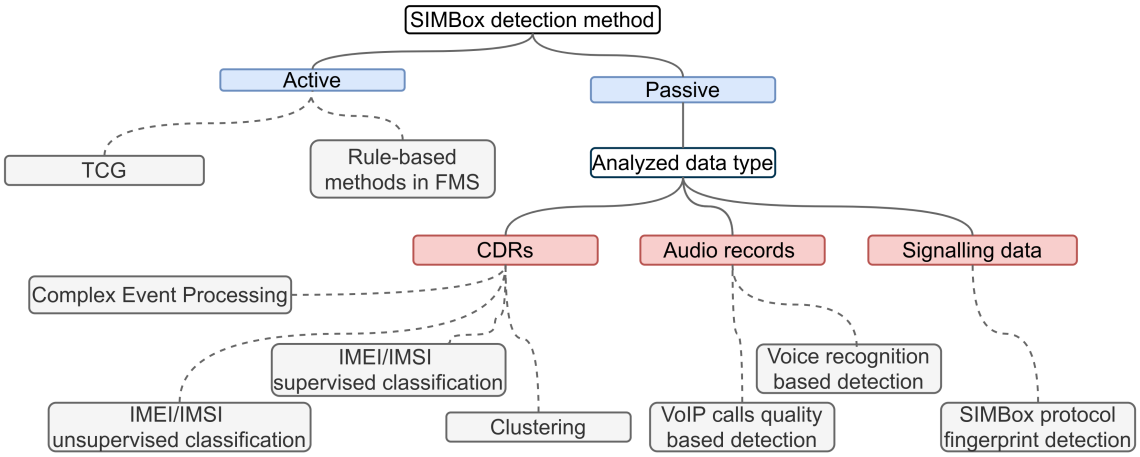

Fig. 10: Categorization of existing *SIMBox* fraud detection methods.

in Section V) into *SIMBoxes*. HBS significantly increases the false positive rate to paralyzing the decision to block a detected SIM card.

B. Passive methods

Passive methods don't require a permanent human action: they are deployed in a mobile operator's system to detect fraudulent entities automatically. We categorized in Figure 10 passive methods based on the data type analyzed throughout the detection process. This classification allows us to distinguish three sub-groups discussed in the following: CDR analysis-based approaches, audio analysis-based approaches, and signaling data analysis-based approaches.

1) CDR analysis-based approaches: This is a passive method consisting of analyzing both the content and the occurrences of CDRs, unlike Rule-based methods that only focus on the latter.

CDRs gather all traces of events carried out on the network, whether it is voice, text message, or data usage, as listed in Table VII. As described in Table VI, the large majority of *SIMBox* fraud detection solutions use CDR content data to shed light on anomalies of all kinds [67], through Machine Learning or *Complex Event Processing* (CEP). Machine Learning CDR-based solutions correspond to a *classification problem*. The entity to be classified is either a *SIM card* (identified by its IMSI) or a *mobile device* (identified by its IMEI).

A common methodology [68] is applied in CDR-based *SIMBox* fraud solutions. It can be summarized as a data preparation step followed by model building and evaluation.

Data preparation: Data preparation includes data understanding, feature selection as well as any form of data preprocessing. A CDR has several fields, some of which may not be meaningful for *SIMBox* fraud detection. Table VII summarizes all the fields used in the literature for *SIMBox* detection, their common usage, and all the works, to the best of our knowledge, using them. We organize the commonly used features (built by aggregated fields) into four types.

(Feature type 1) Call behavior. They highlight how calls are generated within the *SIMBox*. Because of its traffic termination role, the *SIMBox* is known to generate tremendous outgoing calls than incoming calls. Therefore, a feature such as the ratio of the number of outgoing calls to the number of incoming calls over a given period is usually considered to detect this pattern. Besides, SIMs within the *SIMBox* are considered to be more or less heavily used and sometimes at irregular hours. This usage behavior is evaluated by features such as the total and the average number of calls, the cumulative or the average duration of calls made during a period (day, week, etc.), and at irregular hours. Finally, the *SIMBox* is known to make outgoing calls to many different individuals; thus, the total number of individuals called is generally considered. Early fraud detection works [16; 17] are solely based on features of this category. [22], recent work on *SIMBox* fraud detection, considers only the total number of calls made by a SIM card during a small time window W and tries to capture the repetitiveness of this pattern. Authors in this [22] claim a fraudulent SIM card repeatedly makes a large number of calls during short periods, and this pattern is reproduced by other SIM cards of the same country, named mirror behaviors.

(Feature type 2) Mobility. As far as mobility is concerned, the *SIMBox* is known to be not very mobile and makes calls through the few surrounding base stations to its location. As a result, movement during a call will be practically non-existent, and a significant amount of traffic will be generated in these cells to the point of qualifying them as *hot cells*. To detect this, the total number of different cell IDs where a SIM or device is located over a given period, the event load (number of calls, SMS, etc. per unit of time) on a cell or a set of geographically close cells, the number of subscribers making voice calls within the same location, the average or cumulative distance traveled by an individual (SIM) during a call, and the ratio of calls made without displacement, etc. are examples of attributes that can be considered. Murynets et al. [15] exploited this category of features.

TABLE VII: CDR fields used in the literature for *SIMBox* fraud detection grouped by feature type

Feature type	CDR Field	Description	Usage	Illustrative papers
All features type	Originating number	Phone number of the caller	Gives the number of calls made per user during a period of time	[15],[10],[16], [17],[14],[24], [18],[22],[21] [19]
	Terminating number	Phone number of the called party	Allows to obtain the number of calls received per user during a period as well as to calculate ratios in relation to the number of calls sent.	[15],[10],[16], [17],[14],[24], [18],[21],[19]
	Originating IMSI	Calling subscriber unique SIM card identifier	Gives the number of unique subscribers calling a given subscriber during a period of time.	[15],[10],[16], [17],[24],[18], [21],[19]
	Terminating IMSI	Called subscriber unique SIM card identifier	Gives the number of unique subscribers called by a subscriber during a period of time.	[15],[10],[16], [17],[24],[18], [21],[19]
Call Behavior	Call type	Mobile originated/terminated call	Distinguishes calls made to the mobile network from calls received by the network	[15],[10],[24], [18],[21],[19]
	Originating country code	Country code of the caller	Detects calls from international destinations	[15], [24],[18]
	Terminating country code	Country code of the called party	Detects calls destined for international destinations	[15],[24],[18]
	Event type	Local or international destination	Similarly to originating and terminating country codes, it allows the detection of calls made or intended for international calls.	[14],[18]
	Time	Date and time of the call	Allows to make districts, sorting and selection according to the period of time in which the events occurred. The time period can be weekday, daytime, peak hours during the day, time period during the night etc.	[15],[10], [16], [17],[14],[24], [18],[22]
	Duration	Call duration	Gives the average or cumulative call time over a period of time.	[15],[10], [16], [17],[14],[24], [18]
	Cause for termination	Cause for the end of the call: *caller/called call drop*, *normal call release*, *call timeout* or *network cause*	Gives the total number of calls answered and not answered by a party	[18]
Mobility	LAC-CID at the origination of the call	Local Area Code and Cell Id (base station location identifier) at the start of the call	Allows the study of a subscriber's mobility for a given period of time	[15],[10],[24], [18],[21],[19]
	LAC-CID at the termination of the call	Local Area Code and Cell Id (base station location identifier) at the end of the call	Allows to study of a subscriber's mobility during a call	[10],[18],[21], [19]
Mobile services Usage	Service type	Call, SMS, MMS or mobile data	Allows to distinguish between the different types of service used by a subscriber and to study the frequency of use	[10],[14],[21] [19]
Entity properties	IMEI	Device identifier	Allows to identify the mobile devices acting on the network and to study the number of SIM cards per device	[15],[10],[14]
	Account age	Time since account activation	Allows you to study the average duration of SIM cards and SIM cards associated with a given device	[15],[24],[21], [19]
	Customer segment	Prepaid/postpaid/corporate account	Serves as an indicator as to which accounts may or may not be fraudulent	[15],[24]

(Feature type 3) Mobile services usage. SIMBoxes are known to be specialized in terminating calls. As a result, the SIMs used in *SIMBoxes* make little or no use of other mobile services such as SMS, MMS, GPRS, or other mobile internet services. The number of voice calls to the total number of other services can be studied globally and per individual for a SIM card to detect this usage behavior. Further analysis can be carried out on a service-specific study (SMS, data upload and download traffic, etc.).

(Feature type 4) Entity properties. Finally, some features are used to perform detection based on the account information and properties and not on the entity (SIM card or Mobile equipment) activity. For example, in the literature, we count features such as the number of SIM cards per IMEI, which is usually high for *SIMBoxes* because they are designed to hold hundreds of SIM cards. The age of a customer account is also considered by [15; 19; 21], as fraudulent SIMs operate for less time than regular SIMs because they are usually blocked by operators as soon as they are detected. The customer segment also helps identify fraudulent entities since prepaid accounts are more likely to be deceitful than postpaid or corporate accounts (as discussed in Section III-C).

Data preparation often includes data sampling to reduce the input data's size to ensure a proper proportionality between fraudulent and non-fraudulent entities, leading to better results [69]. In this vein, the proportionality of 66% normal cases versus 34% fraudulent cases is commonly adopted as in [15–17]. Also, 75% of normal cases versus 25% of

fraudulent's is considered in [19; 21].

Model building and evaluation: Several classification models have been used in the literature for *SIMBox* fraud detection. The most recurrent are described in the following.
Artificial Neural Network (ANN). ANN [70] is the first classifier used for *SIMBox* fraud detection. ANN is composed of several neuron layers: an input layer, zero of several hidden layers, and an output layer. Each layer forwards its outputs as input to the next layer until a final result is delivered by the output layer. The optimal weights associated with the nodes are calculated during the neural network training. This is usually done according to an optimization algorithm that aims to minimize training errors. The backpropagation algorithm is commonly used.

In the context of *SIMBox* fraud, each neuron of the input layer represents a feature of the entity to be detected, and the output layer comprises two nodes, one indicating a fraudulent entity when it is activated and the other indicating a normal entity. It is possible to have a single node indicating one of the two possibilities depending on its output value.

Sallehudin et al. [16] tested 240 NN models for the detection varying the values of 4 parameters to choose the optimal architecture: the number of hidden layers, the number of hidden nodes per hidden layer, the learning rate, and the momentum[13]. The authors used the Sigmoid function as the activation function. The best results (accuracy of 98.7% and RMSE of 0.10380) are obtained with a low value of momentum and a relatively high value of learning rate. Kashir et al. [20] proposed a similar NN architecture for fraud detection but with the Sign function as the activation function. The authors tested five NN variants by varying the model optimization algorithms for a dataset with 8695 normal subscribers and 50 fraudulent subscribers. They obtained very good performances: an accuracy of 99.87% with the *Bayesian regularization algorithm* and an RMSE of 0.01654.

Support Vector Machine (SVM). SVM classifies cases by finding a hyperplane separator in the feature space between two classes to maximize the distance between the hyperplane and the closest data points of each class (referred to as support vectors). When the dataset is not linearly separable, the training samples are mapped to a higher dimensional space by applying a kernel. The most commonly used are the linear kernel, the polynomial kernel with a p parameter as the polynomial degree, the Radial Basis Function (RBF) kernel with a ∂ parameter, and the Sigmoid kernel.

Sallehudin et al. [17] tested out 40 SVM models considering three kernel functions and varying for each kernel the values of parameters. The authors used the same features as in their previous work [16] (based on ANN, described above) and obtained a maximum accuracy of 98.9% and a minimum RMSE of 0.105. The authors also provided an extensive comparison between the ANN model and the SVM model for fraud detection. They showed evidence that the SVM model is more efficient in classifying fraudulent subscribers with a lower false-negative rate overall than the ANN model as the training data percentage increases. The ANN model is better suited to classify regular subscribers; however, it presents dramatically increased outliers for specific percentages of training data considered. Regarding classification accuracy, the SVM model has the highest value regardless of the percentage of data used for training. Finally, evaluating the time taken to build the model showed that the SVM model is about three times faster than the ANN model and requires less computational power.

Kashir et al. [20] tested out 5 SVM kernels compared by classification accuracy and regression. The authors reported a lower performance than the model based on ANN (presented above), which contradicts Sallehudin et al. [17].

Albougha et al. [14] compared SVM to ANN, *Boosted Tree* [71] and *Logistic* classifiers for *SIMBox* fraud detection. The *Boosted Tree* classifier shows the best performance (i.e., 91.12% of accuracy) and the ANN the worst (i.e., 60.37% of accuracy).

Fuzzy logic. Unlike SVM, a binary and non-probabilistic classifier, Fuzzy logic deals with approximate reasoning rather than fixed and exact. Therefore, an element belongs to a fuzzy set according to a *Membership Function* (MF), whose value is between 0 and 1. If this value is 0 for an element x for a fuzzy set A then, it has no membership to this set, and if it is 1, there is a full membership. From this logic, fuzzy rules are defined as conditional statements based on elements belonging to fuzzy sets. Therefore, a fuzzy rule is valid at a specific rate, which results from calculations on MF values of the different fuzzy sets included in the conditional statement. The collection of fuzzy sets with their MFs and fuzzy rules constitutes a fuzzy system. Fuzzy logic brings the reasoning closer to humans with linguistic expressions that refer to fuzzy sets.

Marah [10] proposed a fuzzy system for *SIMBox* fraud detection based on five fuzzy sets. Each fuzzy set is established based on a fraud detection pattern related to mobility, call behavior, or the use of mobile services by subscribers. For each input SIM card, the authors find to what extent it conforms to each fraudulent pattern; this means determining the MF value of each fuzzy set for that SIM card. The MFs are calculated in a triangular way by identifying for each pattern the maximum and minimum values of a dataset and applying the ratio $\frac{Value-Min}{Max-Min}$ for an input value. The SIM card detection process is based on the average of the MF values for the five patterns. If it is above a certain threshold, the SIM is considered fraudulent or to be watched. Evaluations of the model were not performed to validate the model due to a lack of ground truth data.

We note that contrary to the models presented above, this model is not much adjustable; only the threshold's value determines the detection, which could cause some limitations.

[13]The momentum in an ANN helps in the early stages of the algorithms, by the increasing rate at which the weights approach the neighborhood of optimality

Random Forest (RF). It is a classification method using many decision trees. It uses bagging and feature randomness when building each tree to create an uncorrelated forest of trees whose committee's prediction is more accurate than that of any individual tree. The number of trees to build is selected during the training phase. The prediction made by random forest is determined by the majority rule of the generated decision trees.

Using ten-fold cross-validation and a 60-40 percentage split validation, Hagos et al. [19] compared the RF model with ANN and SVM for three dataset profiles obtained by aggregation of features on a 4-hours, daily, and monthly basis. The results suggest that all models have comparable performance measures. However, the RF model achieves a little better than the two others in accuracy and training time for the 4-hour dataset. Similarly, the confusion matrices' outcomes verify that the RF model has a slightly lesser false positive rate than the other models. However, for the daily and the monthly datasets, the RF performance is slightly lower in accuracy. We also note that generally, the RF model's false-negative rate is higher than that of the other models.

Fitsum et al. [21] compared the same three models with a dataset aggregated on a fixed 4-hours basis, i.e., a day split in 6 fixed part as done in [19], and aggregated on a Sliding Window (SW) of 4 hours which adds one hour to the cumulated results of the previous 3 hours. Authors obtained comparable results as [19] in terms of training time, accuracy, false-positive and false-negative rates of the RF model compared to the two others for both datasets. According to their results, RF achieves much better accuracy (i.e., a minimum difference of 5% with ANN) and training time (i.e., a minimum difference of 126 minutes with ANN and a maximum difference of 5651 minutes with SVM) than the others.

Murynets et al. [15] obtained a similar outcome as [19] regarding false-negative and false-positive rates of the RF model. The authors mitigated this issue by combining RF with an *Alternating Decision Tree* (ADTree) model [72; 73] and a *Functional Tree* model [74]. For each record to be classified, a linear combination of the predictions of these three classifiers (in the form 0 and 1) is carried out based on coefficients ($\in [0.1]$ and whose sum is equal to 1) chosen to minimize the classification error. The obtained value is compared to a threshold α (obtained by minimizing the classification error) for decision making. The resulting model has a lower false-negative rate and better accuracy.

Complex Event Processing (CEP).

CEP [75], also known as event stream processing, uses appropriate tools to query a data stream before its storage in a database. It helps aggregate a lot of different information and identifies and analyzes cause-and-effect relationships among events in real-time. CEP continuously matches incoming events against defined patterns which allow taking effective actions proactively (e.g., block a fraudulent SIM card).

The authors in [18] defined seven patterns for recognizing a SIMBox fraudulent phone number depending on if it is on-net or off-net. They based the detection on a set of 26 features obtained through real-time CEP queries, using the free-licensed WSOA CEP tool [76]. These features are established based on analysis of the ground truth (known fraudulent and regulars phone numbers). They are of the mobility and call behavior categories. With features obtained from one day CDR, they achieved F1-scores of 0.786 and 0.789 for on-net and off-net numbers detection, respectively. The used database counts 25 on-net and 15 off-net fraudulent numbers out of 1,573,680 customers.

[22] also leverages real-time CEP to detect fraudulent calling numbers producing a high volume of calls during a small time window. The authors in [22] used the Lossy Counting algorithm [77] which computes frequency counts exceeding a defined threshold over data streams. They improved it with a Fast forgetting algorithm they designed to capture this fraudulent pattern's repetitiveness. Additionally, they applied the Opus Miner frequent set mining algorithm [78] to detect the same fraudulent pattern distributed on different numbers (what they called mirror behaviors). They were able to detect some fraudulent phone numbers from the same country sharing international traffic to avoid peaks (as described in V-A and V-B).

Some remarks: First, the choice of the entity to classify, i.e., IMSI or IMEI (see Table VII), has a meaningful impact on the results as it guides feature determination and, therefore, detection patterns. On the one hand, by using HBS features, fraudsters strive to make SIM cards look identical to regular cards, making IMSI-based detection complex. On the other hand, IMEI codes are regularly changed according to V-E. Therefore, an IMEI code is very volatile; it can appear just once in CDRs, making it challenging to detect relevant patterns.

Second, the *detection time* of the entity is also relevant. It is the time required to obtain the entity's features before the classification. In fact, the longer it takes to detect a fraudulent SIM card or *SIMBox*, the more revenue it can generate. Nevertheless, most of the contributions in the literature do not consider optimizing the *detection time*. This information is either not mentioned (see Table VI) or is hidden and difficult to identify in some cases. In the former case, the methodologies are presented as effectively applicable regardless of the data collection interval, which should be validated. CEP is a good option for real-time CDR-based fraud detection, as seen in [18; 22]. However, it remains challenging to identify relevant patterns for fraud detection in real-time, which can cause a high false-negative rate.

2) Audio analysis-based approaches: Audio records hold valuable information to obtain attributes such as the origin of a call, the types of telephone networks a call has traversed, and to perform analysis and profiling on packet loss and noise as in [79]. Two research works used audio record analysis to detect *SIMBox* fraud to the best of our knowledge.

[9] leveraged the fact that calls performed over the VoIP network suffer from audio degradation in terms of packet losses and jitters (as discussed in Section II-A2). Therefore,

the authors try to detect calls with such degradations, whether they are concealed or not, as an indication of a bypass over the VoIP network using a *SIMBox*. The system is designed for deployment at the base station level for real-time detection of calls characterized by a stream of GSM audio frames. For evaluations, the authors used audio recordings from the TIMIT Acoustic-Phonetic continuous Speech corpus [80] to generate 1960 calls from a set of 98 randomly chosen speakers. They simulated the use of three codecs: G.711, GSM-FR, and GSM-FR with PLC. By considering a SIM to be fraudulent when at least 25% of the calls it makes are deemed fraudulent, 100% of fraudulent SIMs are identified with a 5% VoIP packet loss rate for the three simulated codecs%. Fewer SIM cards are identified as loss rates decrease, and in the case of 1% VoIP packet losses, 43% of G.711 SIMs and 28% of GSM-FR SIMs could be identified. Finally, experiments conducted with a real *SIMBox* showed that the solution could detect 87% of fraudulent SIMs with no false positives.

Elrajubi et al. [8] proposed a voice-recognition-based approach to fraud detection. The solution is based on the fact that fraudulent SIMs are used to terminate traffic that may originate from several different callers to local numbers. However, these calls appear in the CDR traces as coming from a single SIM which can be identified by analyzing the voices of the different speakers using it to differentiate between them. Therefore, the methodology adopted is to identify, from calls audio samples periodically extracted, the speakers using originating and destination SIM cards. A recognized speaker is added in a database as a new speaker for the used SIM card, if not yet existing. Else, the system increments the number of calls made by the corresponding speaker. From the database, two variables are defined for each SIM card: M the greatest number of calls made by a unique speaker using this SIM card, and T the total number of calls made with this SIM card. Additionally, F, a threshold value between 0 and 1, is experimentally chosen so that if M is less than $F \times T$, a SIM is considered fraudulent, and else the SIM is considered normal. Although the idea is very promising, the system was not implemented due to privacy issues in telephone calls, raising several questions about its effectiveness. Will the audio signal quality in a real case allow us to recognize the speakers' voices? After how long will the system detect, and can it not be circumvented by fraudsters? Is it possible to alter the voice of the users at the level of the *SIMBox*? If so, is it within reach of fraudsters, and how can the system be extended to prevent this?

3) Signalling data analysis approaches: The analysis of signaling data to detect *SIMBox* fraud is a recent and not very exploited technique. It was mentioned by LATRO Services [81] in 2015 and is described as highly effective.
Indeed, signaling messages are exchanged between *User Equipment* (UEs) and the core mobile network according to well-defined protocols. They aim to control the UEs, manage access to the network and services, and monitor terminals in case of mobility. The protocols are distributed according to the *Access Stratum* (AS) and the *Non Access Stratum* (NAS). The AS (RRC, PDCP, RLC, MAC, and PHY) manages the signaling between UEs and base stations for radio resource management, handover, and data encryption/compression. The Non-Access Stratum (EMM and EPS in LTE) manages the signaling between UEs and the core network, including establishing data or call sessions and mobility.

Network Attachment for example, is a procedure that involves the AS and the NAS. It is carried out when the UE is switched on, after a loss of network coverage or a change of *Mobile Management Entity* (MME)[14]. It consists of the authentication of the SIM card to the network, the authentication of the network to the SIM card[15], the identification of the UE, and the update of the mobile subscriber's location on the core network, through specific information exchanges (IMSI, IMEI, authentication vector, etc.). Similarly, signaling information is exchanged when a call service is set up or for SMS transfers.

The authors in [81] argue that *SIMBox* components (Gateways, SIMBanks, and the control server) generate a specific set of these signaling messages that constitute a fingerprint allowing fraudulent devices to be distinguished from other devices on the mobile network. The analysis of these messages' data and parameters can be performed in real-time. For example, the *SIMBox* signature can be detected when SIM cards get attached to the network, preventing any use.

Therefore, this technology has several advantages, including the fact that it is passive (no permanent human action is needed) and stops fraud before any revenue is lost. High motivation is thus attached to the exploration of its possibilities. However, there are some difficulties related to the accessibility to this type of data from mobile operators. Their processing as well might be challenging because data volume is much higher than with CDRs.

C. Discussion

First, we notice that CDR-based detection methods are limited. They are based on prior detection knowledge provided by the mobile operator, i.e., ground truth used to train machine learning models. However, the ground truth is limited; fraudulent SIM cards are identified through active detection methods that the fraudsters know to counter, and non-detected SIM cards are considered non-fraudulent. This explains why most CDR-based detection solutions do not consider HBS features (yet existing long before, as shown in Table VI) but still achieve excellent detection accuracy. *Therefore, CDR-based solutions currently detect only the SIM card profile with apparent fraudulent behaviors* (e.g., limited mobility, a large number of outgoing calls, etc.)

Second, *Audio-based solutions are tricky and challenging to explore because they deal with private data*. Although they allow real-time detection (e.g., [9]), the *SIMBox* temporal evolution shows that fraudsters already have access to calls audio recordings and could modify them to avoid detection.

[14]The MME in LTE replaces the *Visited Location Register*(VLR) in 2G and 3G mobile network standards.

[15]From 3G, mutual authentication is realized. The SIM card verifies that the terminal is connected to a legitimate serving network to prevent 'fake base station' attacks.

Moreover, recent *SIMBoxes* support various codecs (see Section V-I2), which makes these approaches difficult to scale.

Third, although virtually unexplored, signaling data analysis-based solutions promise more efficient and accurate *SIMBox* fraud detection regardless of the strategy used. *Indeed, it is difficult for fraudsters to access and modify this data type because it mainly involves the operator's components (i.e., SIM cards and base stations)*. Although challenging, this provides a great incentive to investigate solutions based on this type of data.

VII. Forecast of the *SIMBox* fraud evolution

SIMBox temporal evolution discussed in Section V-J shows that *SIMBox* fraud evolves over time. As a result, today's challenges in fraud detection will not be the same in a few years.

In this Section, we identify the various factors that may influence *SIMBox* fraud, and on this basis, we forecast what tomorrow's fraud may be. The purpose of this exercise is to allow readers interested in fraud detection to propose detection solutions that will not be limited to today's challenges and quickly become outdated but that will be able to adapt and face tomorrow's possible challenges.

We distinguish three categories of factors that can influence *SIMBox* fraud : *(1)* Technological advances of fraud ecosystem elements; *(2)* Economic variations in the billing of calls; *(3)* and Improvements to the *SIMBox* for more efficient and accurate fraud.

A. Technological advances of fraud ecosystem elements

Cellular networks in which the *SIMBox* fraud is deployed are rapidly evolving. Indeed, since 2014 several studies [82–84] focus on 5G, whose deployment is ongoing, while scientists are already working on beyond 5G specifications, with the definition of the next-generation 6G wireless system [85; 86]. Through technologies such as high-speed connectivity, *Internet of Things* (IoT), augmented virtual reality, and so on, these new standards will considerably increase the quality of VoIP communications by allowing high data rates and low latency [82]. Hence, phone calls routed through the *SIMBox* will be indiscernible from a quality perspective and may even be of better quality than cellular-only voice calls. As a result, the efficiency of audio quality-based detection methods [9] will be significantly reduced, if not nullified, because VoIP's current pitfalls (packet losses and jitters) will be practically indistinguishable.

Furthermore, the subject of virtual SIM cards [87] is topical [88–91] and a good option for mobile operators to handle the emergent massive mobile connectivity. Virtual SIM cards technology could revolutionize the way fraud is currently carried out and give rise to pro-fraud (e.g., easily obtaining large quantities of SIM cards, advanced *Human Behavior Simulation*-capabilities architecture) or counter-fraud (e.g., more precise control of SIM card distribution) possibilities depending on how mobile operators will implement it.

B. Economic variations in the billing of calls

The use of OTT applications for voice, audio, video, or other media delivery services is expanding. This tendency will be accentuated with the democratization of 5G, which will give rise to a high bandwidth allowing the birth of new types of OTT applications (e.g., tactile internet applications [92]). This growing trend is seen as a credible and measurable threat to mobile operators [93] because OTT apps provide services (voice calls and messaging) that can substitute their own relatively more expensive ones. Besides, OTT apps use the mobile operators' infrastructure and network to deliver services without directly[16] contributing to the mobile operators' revenue. In order to balance this, efforts are being made to regulate and tax OTT and VoIP services in some countries [94; 95], while in others, VoIP usage is banned [57]. The former remains challenging because of the difficulty of finding a consensus on what digital content is and how tax should be applied [96]. However, this will likely come into effect in some countries in a few years, and consequently, VoIP calls will be billed. We believe that this could increase *SIMBox* fraud as current users of OTT apps for international calls might instead use *cheap-international-calling apps* to get a better quality/cost compromise. As discussed in Section III-B, these *cheap-international calling apps* are a way for *SIMBox* fraudsters to get international call traffic.

C. SIMBox's improvements

Fraudsters create/refine their fraud strategies to (1) adapt to existing detection solutions and be able to evade them or (2) to have higher traffic termination capacity (more GSM channels for simultaneous calls, more SIM slots on a single device, support for CDMA, LTE or other new unpredictable functionalities).

To respond to CDR analysis-based detection solutions, we believe that *SIMBox* fraudsters can leverage ML algorithms to control the behavior of SIM cards. Indeed, fraudsters have CDRs generated by *SIMBox* activity (see Table V) and have access to the publicly available ML-based detection algorithms developed by researchers (discussed in Section VI-B1). By replicating these algorithms, for instance, they could check in real-time if a SIM card is detectable to limit its usage. Concerning SIM card migration, [54] mentions a type of GSM gateway that can be mounted in a vehicle and powered by a car battery to simulate mobile traffic. It suggests that fraudsters may develop more of this model to limit detection based on mobility behavior. They could go further and design mobile GSM gateways with integrated batteries to power themselves and allow any movement.

As a counteract to audio analysis-based detection solutions, fraudsters could incorporate a *SIMBox* feature to modify the call audio characteristics (which they can already eavesdrop, as discussed in Section VI-B1) uniquely during each call. Many methodologies exist to do such modification [97–99] and this

[16]We still have an indirect contribution because the use of OTT services requires the purchase of mobile data, which adds to the operators' revenues.

will considerably limit the efficiency of methods based on voice and audio features recognition [8].

VIII. Conclusion and Future directions

The *SIMBox* fraud is tricky as it involves economic, technical, and even character factors (people's mentality). Moreover, it evolves by adapting to existing detection solutions and is, therefore, a real challenge. This document surveyed both the *SIMBox* manufacturer's community to highlight the fraud schemes and various strategies and the scientific literature regarding fraud detection. We identified the limitations of existing solutions and why fraud continues to be rampant.

Our review provides the key elements to understand and tackle this not-recent but not-enough-studied security issue, which may become more challenging in the future with the forthcoming technological developments and possible economic variations. In this vein, we list below some potential research directions to better address *SIMBox* fraud detection and evolution.

A. The simulation of SIMBox fraud strategies

There is a need for a *SIMBox* fraud simulator to unleash research in the field of *SIMBox* fraud detection. Such a simulator would include all existing fraud strategies and *SIMBox* architectures based on our survey, thus allowing thorough evaluations of current and new fraud detection methods. With implementations made in a scalable way, it would enable easy adding new fraud strategies from the *SIMBox* evolution or realistically designed ones as a forecast. A notable contribution of such a simulator would be to provide access to the different data types currently used to investigate *SIMBox* fraud (CDR, audio, and signaling), thus overcoming the lack of data and privacy limitations.

B. Signaling data-based investigations

As discussed above, the analysis of signaling data is very promising for *SIMBox* fraud detection. Signaling on cellular and VoIP networks is made according to well-defined and secure protocols (e.g., SIP and H. 323 in VoIP networks, RRC, PDCP, and RLC in the LTE Network Stratum). Hence, signaling messages are usually encrypted, and their alteration is beyond the reach of fraudsters. For instance, in cellular networks, signaling messages are exchanged between the SIM card, *a smart card provided by the Mobile Operator designed to secure programs and data stored inside it*, and the base station, *owned by the Mobile Operator*. Therefore, although challenging, exploiting signaling would allow developing solutions that fraudsters could not counter and possibly end the fraud.

C. Leveraging Complex Event Processing

CEP is a tangible asset to be used to detect *SIMBox* fraud. Indeed, time is a critical metric in detecting *SIMBox* fraud but is not enough considered in the literature's contributions. The longer fraudulent SIM cards operate, the more revenue the fraudsters make. Moreover, fraudsters need little return on their investment, which keeps them motivated to continue the fraud. CEP tools and algorithms allow for real-time detection (as in [22], and [18]), but their effectiveness is limited because they are based on static (rule-based) patterns. We argue that these tools could be used in Signaling-based and Audio-based solutions where patterns may be more exploitable than in CDR-based solutions.

References

[1] CFCA, "CFCA 2019 fraud loss survey," Report, 2019. [Online]. Available: https://cfca.org/document/cfca-2019-fraud-loss-survey-pdf/

[2] K. Baskar, "A study on internet bypass fraud: national security threat," *Forensic Research & Criminology International Journal*, vol. 7, p. 31–35, Sep. 2019.

[3] CFCA, "CFCA 2017 fraud loss survey," Report, 2017. [Online]. Available: https://cfca.org/document/cfca-2017-fraud-loss-survey-pdf/

[4] NCC, "An assessment of international voice traffic termination rates," Report, Jul. 2015. [Online]. Available: https://www.ncc.gov.ng/docman-main/industry-statistics/policies-reports/681-the-principles-of-international-termination-rate/file

[5] GoAntiFraud, WebPage, 2021, accessed: 2020-09-03. [Online]. Available: https://goantifraud.com/en/blog/categories/article

[6] Revector, "Simbox fraud and ott bypass biggest threats to mobile operator revenues," Article, Nov. 2016.

[7] Black Swan Telecom Journal, "Mapping the interconnect resale routes of fraudsters: How a global robot network detects voice and sms bypass," Article, Jun. 2015. [Online]. Available: http://bswan.org/interconnect_fraud_routes.asp

[8] O. M. Elrajubi, A. M. Elshawesh, and M. A. Abuzaraida, "Detection of bypass fraud based on speaker recognition," in *2017 8th International Conference on Information Technology (ICIT)*, 2017, pp. 50–54.

[9] B. Reaves, E. Shernan, A. Bates, H. Carter, and P. Traynor, "Boxed out: Blocking cellular interconnect bypass fraud at the network edge," in *Proceedings of the 24th USENIX Conference on Security Symposium*, ser. SEC'15. USA: USENIX Association, Aug. 2015, pp. 833–848.

[10] H. M. Marah, O. M. Elrajubi, and A. A. Abouda, "Fraud detection in international calls using fuzzy logic," in *International Conference on Computer Vision and Image Analysis Applications*, 2015, pp. 1–6.

[11] H. Kumar, "Technical note on illegal international long distance telephone exchange in india," ITS, Meerut, India, Tech. Rep., Aug. 2012.

[12] M. Sahin, A. Francillon, P. Gupta, and M. Ahamad, "Sok: Fraud in telephony networks," in *2017 IEEE European Symposium on Security and Privacy (EuroS P)*, Apr. 2017, pp. 235–250.

[13] F. Okumbor N. Anthony and A. A. J. Olokunde, "Grappling with the challenges of interconnect bypass fraud," *IOSR Journal of Mobile Computing and Application (IOSR-JMCA)*, vol. 6, pp. 35 – 41, Jan. 2019.

[14] M. R. AlBougha, "Comparing data mining classification algorithms in detection of simbox fraud," Master's thesis, St. Cloud State University, Dec. 2016.

[15] I. Murynets, M. Zabarankin, R. P. Jover, and A. Panagia, "Analysis and detection of simbox fraud in mobility networks," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, Apr. 2014, pp. 1519–1526.

[16] R. Sallehuddin, S. Ibrahim, A. Zain, and A. Elmi, "Detecting sim box fraud using neural network," in *IT Convergence and Security 2012*, K. J. Kim and K.-Y. Chung, Eds. Dordrecht: Springer Netherlands, 2013, pp. 575–582.

[17] ——, "Detecting sim box fraud by using support vector machine

and artificial neural network," in *Jurnal Teknologi*, vol. 74, Apr. 2015, pp. 137–149.

[18] K. Kehelwala, H. Bandara, R. Yasaratne, P. De Almeida, I. Ilesinghe, and P. Wickramasinghe, "Real-time grey call detection system using complex event processing," IET, Sri Lanka, Tech. Rep., 2015. [Online]. Available: http://theiet.lk/wp-content/uploads/2017/10/22-p7.pdf

[19] H. Kahsu, "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, Nov. 2018.

[20] M. Kashir and S. Bashir, "Machine learning techniques for sim box fraud detection," in *2019 International Conference on Communication Technologies (ComTech)*, Apr. 2019, pp. 4–8.

[21] F. Tesfaye, "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, Feb. 2020.

[22] B. Veloso, S. Tabassum, C. Martins, R. Espanha, R. Azevedo, and J. Gama, "Interconnect bypass fraud detection: a case study," *Annals of Telecommunications*, vol. 75, pp. 583–596, Oct. 2020.

[23] Y. Alraouji and A. Bramantoro, "International call fraud detection systems and techniques," in *Proceedings of the 6th International Conference on Management of Emergent Digital EcoSystems*, ser. MEDES '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 159–166.

[24] N. A. Ibrahem Soliman Alsadi, "Study to use neo4j to analysis and detection sim-box fraud," *Journal of Pure & Applied Sciences*, vol. 17, no. 4, pp. 31–35, Jan. 2019.

[25] S. Limited, "Bypass fraud - are you getting it right?" White Paper, 2011. [Online]. Available: https://www.subex.com/pdf/Bypass_Fraud.pdf

[26] ——, "Subex wholesale fraud management survey 2013 is the industry ready to tackle a growing issue?" Survey, 2013. [Online]. Available: https://billingviews.com/wp-content/uploads/delightful-downloads/2013/09/Subex-Wholesale-Fraud-Management-Survey-2013.pdf

[27] CSGi, Webpage. [Online]. Available: https://www.csgi.com/portfolio/digital-wholesale/assure/assure-sim-box-detection/

[28] Araxxe, Webpage. [Online]. Available: https://www.araxxe.com/p/our-services/global-transaction-verification/global-transaction-verification/test-call-generator-outsourcing.html

[29] Pixip, Webpage. [Online]. Available: https://www.pixip.net/index.php/solutions/test-call-generation.html

[30] Calltic, Webpage. [Online]. Available: https://www.calltic.com/

[31] MediaFon, Webpage. [Online]. Available: https://www.mediafonts.lt/

[32] Sysmaster, Webpage. [Online]. Available: http://www.sysmaster.com/products/gsm_termination.php

[33] Antrax, Webpage. [Online]. Available: https://en.antrax.mobi/

[34] L. Shenzhen HyberTone Technology Co., Webpage. [Online]. Available: http://www.hybertone.com/en/

[35] Antrax, *Sim Server Factor Script*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/sim-server-factor-script.md#sim-server-factor-script

[36] L. Shenzhen HyberTone Technology Co., *GoIP User Manual*, Shenzhen HyberTone Technology Co., Ltd., Jun. 2016, version 1.5. [Online]. Available: http://www.hybertone.com/uploadfile/download/20140304125509964.pdf

[37] Portech, *SIM Server User Manual*, Portech, version 2.0.1. [Online]. Available: https://www.portech.com.tw/data/SIM%20Server%20User%20Manual%20V2.pdf

[38] Dinstar, *Instructions for Using Multi-SIM Function of DWG*, Dinstar. [Online]. Available: https://www.dinstar.com/WEB/files/48826/2019-05-22/Multi-SIM%20of%20DWG%20Instruction.pdf

[39] Antrax, *Session Period Script*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/session-period-script.md

[40] ——, *Activity Period Script*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/activity-period-script.md

[41] Dinstar, *UC2000-VE/F/G GSM/CDMA/WCDMA VoIP Gateway User Manual*, Dinstar, Feb. 2017, version 2.2. [Online]. Available: https://www.dinstar.com/WEB/files/47154/2019-04-30/UC2000-VE&VF&VG%20GSM&CDMA&WCDMA%20VoIP%20Gateway%20User%20Manual.pdf

[42] Ejoin, *EJOIN ACOM5xx VoIP Gateway User Manual*, Ejoin, Oct. 2019, version 1.4. [Online]. Available: https://fr.scribd.com/document/484621350/ACOM5xx-User-Manual-V1-4-1

[43] Antrax, *Gateway Selector Script*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/gateway-selector-script.md

[44] L. Shenzhen HyberTone Technology Co., *Gateway Selector Script*, Shenzhen HyberTone Technology Co., Ltd., version 1.01.1. [Online]. Available: http://www.hybertone.com/uploadfile/download/20171222180904804.pdf

[45] Antrax, *IMEI Generator Script*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/imei_gen_script.md

[46] ——, *HTTP request*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/business-activity-scripts/http-request.md

[47] ——, *USSD*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/business-activity-scripts/ussd.md

[48] N. VoiceBlue, *2N VoiceBlue Enterprise User Manual*, 2N VoiceBlue, version 1.11. [Online]. Available: http://www.mpinetworks.com/images/catalogue/id_13/images/22_VoiceBlue_Enterprise_-_User_Manual_EN.pdf

[49] Antrax, *SMS*, Antrax, Jul. 2017, commit hash ee1d40cf. [Online]. Available: https://gitlab.com/flamesgroup/antrax/-/blob/master/doc/manual/scripts/business-activity-scripts/sms.md

[50] L. Shenzhen HyberTone Technology Co., *GoIP32-X4 Quick Setup Manual*, Shenzhen HyberTone Technology Co., Ltd. [Online]. Available: http://www.hybertone.com/uploadfile/download/20180913163352145.pdf

[51] GoAntiFraud, "Goantifraud gsm termination in africa: Top 5 destinations in 2020," Article, Feb. 2020, accessed: 2020-02-20. [Online]. Available: https://goantifraud.com/en/blog/1186-gsm-termination-in-africa-top-5-destinations-in-2020.html

[52] ——, "10 misconceptions about gsm termination," Article, Nov. 2016, accessed: 2020-04-24. [Online]. Available: https://goantifraud.com/en/blog/368-10-misconceptions-about-gsm-termination.html

[53] ——, "Top-5 popular gsm gateway manufacturers," Article, Feb. 2018, accessed: 2020-02-23. [Online]. Available: https://goantifraud.com/en/blog/818-top-5-popular-gsm-gateway-manufacturers.html

[54] OECD, "International traffic termination," *OECD Digital Economy Papers*, no. 238, 2014.

[55] Y. Theodorou, K. Okong'o, and E. Yongo, "Access to mobile services and proof of identity 2019: Assessing the impact on digital and financial inclusion," GSMA, Tech. Rep., 2019. [Online]. Available: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/02/ProofOfID_R_WebSpreads.pdf

[56] D. Morrow, "Telco corruption fuels simbox frauds," Article, Jul. 2017, publisher: Comms Risk. [Online]. Available: https://commsrisk.com/telco-corruption-fuels-simbox-frauds/

[57] S. Guerraoui, "Morocco banned skype, viber, whatsapp and facebook messenger. it didn't go down well," Article, Mar. 2016, publisher: Middle East Eye.

[58] 3GPP, "Telecommunication management; charging management; charging data record (cdr) parameter description," Specification. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1915

[59] S. Karapantazis and F.-N. Pavlidou, "Voip: A comprehensive survey on a promising technology," *Computer Networks*, vol. 53, no. 12, pp. 2050 – 2090, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128609001200

[60] C. Blackman and L. Srivastava, *Telecommunications Regulation Handbook : Tenth Anniversary Edition*, ser. World Bank Publications. The World Bank, Jun. 2011, no. 13278. [Online]. Available: https://ideas.repec.org/b/wbk/wbpubs/13278.html

[61] P. Cholda, M. Kantor, A. Jajszczyk, and K. Wajda, "Ngl01-1: Least cost routing in inter-carrier context," in *IEEE Globecom 2006*. IEEE, 2006, pp. 1–5.

[62] M. Sahin, "Understanding telephony fraud as an essential step to better fight it," Ph.D. dissertation, TELECOM ParisTech, Sep. 2017.

[63] U. Murad and G. Pinkas, "Unsupervised profiling for identifying superimposed fraud," in *Principles of Data Mining and Knowledge Discovery*, J. M. Żytkow and J. Rauch, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 251–261.

[64] S. Rosset, U. Murad, E. Neumann, Y. Idan, and G. Pinkas, "Discovery of fraud rules for telecommunications—challenges and solutions," in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '99. New York, NY, USA: Association for Computing Machinery, 1999, p. 409–413.

[65] Demirguc-Kunt, Asli, L. Klapper, D. Singer, S. Ansar, and J. Hess, "The global findex database 2017 measuring financial inclusion and the fintech revolution," World Bank Group, Tech. Rep., 2018. [Online]. Available: http://documents1.worldbank.org/curated/en/332881525873182837/pdf/126033-PUB-PUBLIC-pubdate-4-19-2018.pdf

[66] Black Swan Telecom Journal, "Araxxe on the art of deception and analysis in sim box fraud warfare," Article, Feb. 2019. [Online]. Available: http://bswan.org/araxxe_art_of_deception.asp

[67] P. Ferreira, R. Alves, O. Belo, and L. Cortesão, "Establishing fraud detection patterns based on signatures," in *Advances in Data Mining. Applications in Medicine, Web Mining, Marketing, Image and Signal Mining*, P. Perner, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 526–538.

[68] C. Shearer, "The crisp-dm model: The new blueprint for data mining," *Journal of Data Warehousing*, vol. 5, no. 4, 2000.

[69] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial Intelligence*, vol. 97, no. 1, pp. 245 – 271, 1997, relevance.

[70] S. Haykin, *Neural Networks: A Comprehensive Foundation (3rd Edition)*. USA: Prentice-Hall, Inc., 2007.

[71] Z. Tu, "Probabilistic boosting-tree: learning discriminative models for classification, recognition, and clustering," in *Tenth IEEE International Conference on Computer Vision (ICCV'05)*, vol. 2, 2005, pp. 1589–1596.

[72] Y. Freund and L. Mason, "The alternating decision tree learning algorithm," in *Proceedings of the Sixteenth International Conference on Machine Learning*, ser. ICML '99. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999, p. 124–133.

[73] G. Holmes, B. Pfahringer, R. Kirkby, E. Frank, and M. Hall, "Multiclass alternating decision trees," in *Machine Learning: ECML 2002*, T. Elomaa, H. Mannila, and H. Toivonen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 161–172.

[74] J. Gama, "Functional trees," *Machine Learning*, vol. 55, no. 3, p. 219–250, Jun. 2004.

[75] E. Wu, Y. Diao, and S. Rizvi, "High-performance complex event processing over streams," in *Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 407–418.

[76] S. Traut and P. Purich, *Getting Started Guide for Oracle Complex Event Processing*, Oracle, Aug. 2012, version 11.1.1.6.3. [Online]. Available: https://docs.oracle.com/cd/E23943_01/doc.1111/e14476/overview.htm#CEPGS106

[77] G. S. Manku and R. Motwani, "Approximate frequency counts over data streams," in *Proceedings of the 28th International Conference on Very Large Data Bases*, ser. VLDB '02. VLDB Endowment, 2002, p. 346–357.

[78] M. Riondato and E. Upfal, "Efficient discovery of association rules and frequent itemsets through sampling with tight performance guarantees," *ACM Transactions on Knowledge Discovery from Data*, vol. 8, no. 4, Aug. 2014.

[79] V. A. Balasubramaniyan, A. Poonawalla, M. Ahamad, M. T. Hunter, and P. Traynor, "Pindr0p: Using single-ended audio features to determine call provenance," in *Proceedings of the 17th ACM Conference on Computer and Communications Security*, ser. CCS '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 109–120.

[80] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, N. L. Dahlgren, and V. Zue, "Timit acoustic-phonetic continuous speech corpus," *Philadelphia: Linguistic Data Consortium*, 1993.

[81] Technology Research Institute and LATRO Services, "White paper: Network protocol analysis: A new tool for blocking international bypass fraud before revenue is lost," LATRO, Tech. Rep., 2015. [Online]. Available: http://bswan.org/research/network_protocol_analysis_white_paper.pdf

[82] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[83] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[84] A. Gupta and R. K. Jha, "A survey of 5g network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[85] K. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. Zhang, "The roadmap to 6g: Ai empowered wireless networks," *IEEE Communications Magazine*, vol. 57, pp. 84–90, 08 2019.

[86] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[87] M. Richarme, "The virtual SIM - a feasibility study," Master's thesis, Technical University of Denmark, Department of Applied Mathematics and Computer Science, Apr. 2008.

[88] S. Zhao and B. Smeets, "Virtual sim card cloud platform," US Patent 10,349,272, Jul. 2019.

[89] D. L. Polehn, P. R. Chang, F. Weisbrod, and C. J. Christopherson, "System and method for virtual sim card," US Patent 10,123,202, Nov. 2018.

[90] J. Liu, X. Qin, and B. Du, "Method and system for international roaming using virtual sim card," US Patent App. 11/746,493, Jan. 2008.

[91] G. Shi, V. Tangirala, T.-Y. Siu, J. Durand, and S. A. Sprigg, "Virtual sim card for mobile handsets," US Patent 8,811,969, Aug. 2014.
[92] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5g-enabled tactile internet," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 460–473, 2016.
[93] J. Sujata, S. Sohag, D. Tanu, D. Chintan, P. Shubham, and G. Sumit, "Impact of over the top (ott) services on telecom service providers," *Indian Journal of Science and Technology*, vol. 8, no. S4, pp. 145–160, 2015.
[94] Telecom Regulatory Authority of India, "Regulatory framework for over-the-top (ott) services," eSocialSciences, Working Papers, Apr. 2015. [Online]. Available: https://EconPapers.repec.org/RePEc:ess:wpaper:id:6687
[95] N. A. Wasmi, "TRA says viber is 'unlicensed' in the UAE," Article, Sep. 2014. [Online]. Available: https://www.thenationalnews.com/business/technology/tra-says-viber-is-unlicensed-in-the-uae-1.240739
[96] R. Katz, "The impact of taxation on the digital economy," ITU, Tech. Rep., 2015. [Online]. Available: https://www.itu.int/en/ITU-D/Conferences/GSR/Documents/GSR2015/Discussion_papers_and_Presentations/GSR15_session4_Katz.pdf
[97] A. H. Zhou, T. T. Zhou, and D. T. Zhou, "Systems and methods for digital multimedia capture using haptic control, cloud voice changer, and protecting digital multimedia privacy," US Patent 8,968,103, Mar. 2015.
[98] F. Horikawa, "Telephone with voice changer and control method and control program for the telephone," US Patent App. 11/438,834, Nov. 2006.
[99] P. Bonnard, I. Bourmeyster, X. Fourquin, and P. Ladouce, "Telecommunication terminal able to modify the voice transmitted during a telephone call," US Patent 7,796,748, Sep. 2010.