SoK: Fraud in Telephony Networks

Merve Sahin∗†, Aurélien Francillon∗, Payas Gupta†, Mustaque Ahamad§

∗Eurecom, Sophia Antipolis, France
†Monaco Digital Security Agency
‡Pindrop, Atlanta, USA
§Georgia Institute of Technology, USA

Abstract—Telephone networks first appeared more than a hundred years ago, long before transistors were invented. They, therefore, form the oldest large scale network that has grown to touch over 7 billion people. Telephony is now merging many complex technologies and because numerous services enabled by these technologies can be monetized, telephony attracts a lot of fraud. In 2015, a telecom fraud association study estimated that the loss of revenue due to global telecom fraud was worth 38 billion US dollars per year. Because of the convergence of telephony with the Internet, fraud in telephony networks can also have a negative impact on security of online services. However, there is little academic work on this topic, in part because of the complexity of such networks and their closed nature. This paper aims to systematically explore fraud in telephony networks. Our taxonomy differentiates the root causes, the vulnerabilities, the exploitation techniques, the fraud types and finally the way fraud benefits fraudsters. We present an overview of each of these and use Caller Name (CNAM) revenue share fraud as a concrete example to illustrate how our taxonomy helps in better understanding this fraud and to mitigate it.

1. Introduction

Telephony, which used to be a closed system, has undergone fundamental changes in the past several decades. The introduction of new communications technologies and convergence of telephony with the Internet has added to its complexity. Despite (or because) of having been deployed for virtually hundreds of years, security challenges for telephony are neither well understood nor well addressed.

In this paper, we focus on the fraud and cybercrime ecosystem surrounding voice telephony (over all three networks - the Public Switched Telephone Network or PSTN, cellular and IP networks). We aim to provide a systematization of knowledge relevant to understanding telephony fraud. Our taxonomy allows to classify the techniques and fraud schemes without ambiguity. We believe that a good understanding of telephony fraud will provide insights for future research, increase cooperation between researchers and industry and finally help in fighting such fraud.

Although, we focus on telephony fraud, our work has broader implications. For example, a recent work shows how telephony fraud can negatively impact secure creation of online accounts [1]. Also, online account takeovers by making a phone call to a call center agent have been reported in the past [2], [3]. Telephony is considered as a trusted medium, but it is not always. A better understanding of telephony vulnerabilities and fraud will therefore help us understand potential Internet attacks as well.

1.1. Fraud in Telecommunication Networks

Existing definitions of telecommunications fraud usually focus on obtaining free telecommunications services and gaining financial benefits [4], [5]. In this work, we narrow our perspective to voice telephony but we do not limit frauds to financial benefits (a definition will be given in Section 3). Perpetrators of voice fraud may be any actor of the telephony ecosystem, such as operators, third party service providers, customers, employees and any other external party with the means and motivation to commit fraud. On the other hand, victims of voice fraud can be the operators, customers and enterprises that use telecom networks.

A survey of telecom service providers in 2015 estimates the losses due to fraud to 38.1 billion US dollars. This constitutes 1.69% of the estimated global revenue [6]. In addition to the financial losses, fraud aiming at service disruption or reputation damage may have devastating effects, because the telecommunications network is a critical infrastructure with millions of users relying on it. On the other hand, consumers are also victims of such fraud, the United States Federal Trade Commission (FTC) receives an average of 400,000 complaints per month [7].

Perpetrating fraud in telecom networks is relatively easy. Most of the attacks can be performed remotely and they do not require major equipment or high level of technical expertise. Moreover, it is often very easy to obtain a financial benefit from telephony fraud [8]. Often, fraud is buried
in massive volume of traffic and large variety of services. Therefore, it is difficult to identify, detect and prevent.

Having a comprehensive understanding of telephony fraud is a challenging task. For this, one needs to have a good understanding of the telephony ecosystem, its history, underlying technologies, regulations and international agreements.

Telecom industry embodies different communities such as operators, regulators and users. Every actor in this ecosystem experiences or approaches fraud in a different way. Moreover, each community has its own terminology, context and resources regarding fraud, which is a major obstacle in understanding fraud. We next explain the related work from each community and their limitations.

1.2. Related Work

Operators and service providers usually share fraud related information among their partners and various industry associations (e.g., TMForum, i3Forum, GSMA, FIINA, CFCA). Unfortunately, such groups are often restricted to vetted members and do not make their documents publicly available. The point we make in this paper is the opposite: we will only be able to fight fraud efficiently if it is well understood and openly discussed. A first attempt to create a fraud classification system that distinguishes enabling techniques and fraud types was proposed by the TM Forum [9], an approach that we extend in this work.

A lot of information about fraud schemes can be found in white papers by companies selling fraud detection systems [10], [11], but those often present an incomplete view because of the possible commercial interests.

In the academic literature, there is no previous systematic survey of telephony fraud. However, there are resources that handle part of the problem or try to reduce the problem into a single dimension such as actors (fraudster or victim) [12], underlying service and technology [13], [4], attack methodology or attack motivation [14]. However, the fraud ecosystem is too complex to be explained with a binary classification. [15] studies the telecom system security, covering many fraud related topics. It concludes that information on phone fraud is scattered and no single resource brings everything together. Another important work on telecommunications crime [16] presents historical information and some of the more recent fraud schemes. Existing surveys [14] address fraud detection, but do not try to systematize the fraud itself.

In [17], [18] authors analyze data from phone honeypots uncovering several fraud schemes affecting end users. [19] analyzes the voice spam ecosystem and evaluates existing solutions. There are also many books on telecom security related topics such as revenue assurance [20], fraud and quality of service management [13], UC (Unified Communications) [21] and VoIP network security. However, none of these resources provide a comprehensive view of the fraud ecosystem.

International bodies, such as ITU (an agency of the United Nations) and BERC (an agency of the European Union), and regulatory bodies, such as the Federal Communications Commission (FCC) and the Federal Trade Commission (FTC), are also concerned about some aspects of the fraud (e.g., number misuse [22], robocalling [23]) but they also do not aim at providing a comprehensive view on the telephony fraud.

1.3. Goals

In this paper, we aim to clarify telephony fraud and provide a holistic understanding of telephony fraud by considering its causes, the techniques, the fraud schemes and the reasons why fraud may be profitable. We hope that a better understanding of fraud mechanisms will foster research on this topic. We also believe that it is required to understand telephony fraud well to address it efficiently.

A fraud scheme often has multiple names, e.g., describing a variant, the technical aspect or the user visible part of the iceberg. In other cases, one name is used to describe several different schemes. We therefore also aim to clarify the inconsistencies in existing fraud terminology.

Moreover, this study may be beneficial to increase fraud awareness among users and operators that are not members of any industry group. In fact, a survey conducted among the wholesale operators in 2013 [24] shows that around 73% of the operators are not members of any of the industry groups.

1.4. Paper Organization

In the next section (2) we provide the necessary background information on the telephony ecosystem, summarizing the key concepts and money flows. Then we present an overview of our methodology in Section 3. In the four next sections we describe fraud in our taxonomy: root causes (4), weaknesses (5), techniques (6), the main fraud schemes (7), and the way fraud can benefit fraudsters (8). We then present a case study in Section 9. Finally, we conclude the paper in Section 10.

2. Overview of the Telephony Ecosystem: Networks and Money Flows

In this section, we provide a high level overview of voice telephony related networks and components that are required to understand fraud in voice telephony.

2.1. Telephony Networks and Components

Public Switched Telephone Network (PSTN). The historic core of telephony networks is formed of copper telephone wires that use circuit-switching technology to transmit analog voice signals (also called Plain Old Telephone Service (POTS)). Switches in operators’ Central Offices (PSTN CO) control call establishment by creating a dedicated physical circuit from the caller’s phone to the
calling system. Initially, the same circuit was used for the in-band signaling between the callee and the operator (e.g., dial tone and ringing) but also between operators (e.g., billing, call routing).

**Integrated Services Digital Network (ISDN).** ISDN allows digital transmission over the copper lines. Up to 30 lines can be multiplexed on a physical phone line (T1 or E1 Primary Rate Interface (PRI)) for transmitting data or voice. ISDN dedicates a separate channel for signaling (out-of-band signaling), which constitutes the user part of the Signaling System 7 (SS7) protocol [25]. Digital networks and out-of-band signaling solved some security problems (see Section 7.1) and introduced new features to telephony, e.g., voice mail, call forwarding and caller ID display.

**Mobile Networks.** Most of the mobile networks are still using GSM protocols and equipment (2G) but also support more recent protocols (3G and 4G/LTE). Each generation of mobile communication uses some form of encryption (over the wireless channel) and specific equipment to handle the communications and customer identification (e.g., Mobile Switching Center (MSC) and Home Location Register (HLR) in 2G). Mobile phones (except CDMA phones) use a SIM (Subscriber Identity Module) card with an International Mobile Subscriber Identity (IMSI) that uniquely identifies the user on the network. The SIM card contains a cryptographic key which is assigned by the operator and associated with the IMSI.

**Voice over IP (VoIP).** With the rise of the Internet, transmission of Voice over IP (VoIP) emerged as an alternative to traditional PSTN. Currently, telephone networks consist of various gateways between PSTN, cellular networks and VoIP telephony. Over-The-Top services (OTT) are services which work on top of data links and, in general, out of operators’ control. Such voice services (e.g., Skype, Viber) are attracting more users and are seen as a threat by the operators [26].

**Private Branch Exchange (PBX).** Enterprise customers usually use a PBX to manage their internal and external communication needs. A traditional PBX provides extensions, i.e., an internal phone number to reach each user within the enterprise. A PBX also has a connection (called a trunk) with an operator to reach the PSTN or mobile networks. The trunk usually supports a certain number of simultaneous communications, which may be different from the total number of phone numbers used by the company. A traditional PBX uses phone cables for all internal lines which is expensive to deploy and manage. On the other hand, an IP-PBX can connect IP phones or soft phones over IP (see Figure 1). IP-PBXs can use PRI trunks, SIP trunks or SIM cards for external communications.

## 2.2. Telephony Actors

Apart from the end-users, the main actors of the telephone networks are the operators (carrier, telecom service provider) and third party service providers.

**Operators.** The deregulation of the telecommunication markets has resulted in wide variety of service providers and operators. Some of these operators invest in, or own, the network infrastructure and equipment, whereas others only resell the service they buy from other operators (e.g., Mobile Virtual Network Operators (MVNO)).

**Third Parties.** Third party service providers and VoIP resellers [27] are important actors of the telephony ecosystem. Value added service (e.g., premium rate service) providers deliver content to end-users via phone calls, messaging or data network (e.g., gaming, chat lines or news). VoIP resellers buy communication services from carriers, and resell through VoIP gateways. They provide geographical numbers (numbers with country and area codes), mobile numbers, toll free numbers and premium rate numbers in every country. In recent years, cloud based communication services have appeared (e.g., Twilio [28]) and provide access to cheap bulk phone numbers (that are usually recycled), cloud PBX, SIP trunks or scripted Interactive Voice Response (IVR) systems.

## 2.3. Billing Systems and Call Routing

Understanding billing mechanisms is key to understand telephony fraud, as most of the fraud schemes aim at financial benefits. Operators keep Call Detail Records (CDR) for each call routed (originated, terminated or transited) over their networks. CDRs are created at the network switches and include various information, such as originating and destination phone numbers, inbound and outbound routes, date, call duration and call type. All CDRs generated at different switches are collected and processed in a central location and sent to the billing system to be charged. Operators deal with two types of billing: retail billing and wholesale billing.

**Retail Billing.** Most services (international or domestic landline, mobile, or data services) are billed to customers at the end of the billing period (post-paid). However,
mobile services are also often available as pre-paid. The post-paid billing process involves the collection of usage reports, validating them, applying the tariff plan and sending the final bill to the customers. To be sure to be paid, operators verify the personal and financial information of their post-pay customers. In the pre-paid billing, customer information check is often less strict, because the customer will only be able to use the service he already paid for.

**Wholesale Billing.** The wholesale (interconnect) market is mainly for international and long distance calls, as operators need to make interconnection agreements and to rely on transit operators to be able to provide worldwide coverage. Such interconnect agreements describe the prices for interconnect communications, but also policies and dispute resolution. There are also stock-exchange like platforms where operators can buy and sell minutes directly and even anonymously [29].

An example of money flow, in an international call, is shown in Figure 2. The call is initiated from the originating operator (Op-A) and goes through two transit operators, to finally reach the customer through Op-B. For this call, Op-A will bill his end customer for a collection charge of $1. However, the operator will pay 80c for routing the call, and keep 20c, similarly the transit operators each keep a 20c and finally the terminating operator Op-B will keep the termination fee of 40c. In other words, each upstream (originator) network pays to its downstream (terminator) network the cost of terminating the call [30] until the call reaches its final destination. Operators may have multiple routing choices to route a call. They choose the best route depending on the prices and quality of alternative routes. The process of checking the quality and reliability of a transit operator before the partnership agreement is called due diligence. Unlike in IP networks, the routing of a call is very often opaque. Each operator only knows the next hop of the upstream and downstream routes as well as the originating and the destination number.

![Figure 2: Overview of money flows in a call.](image)

3. Fraud Taxonomy: Overview and Systematization Methodology

One common fallacy in previous classifications is that the fraud descriptions are often bundling different problems together. For example, a fraud will be described by the technique it uses. However, techniques used by a given fraud often change, e.g., in reaction to the implementation of new countermeasures. The intricate combination of those concepts makes previous descriptions confusing or narrow.

2. The originating number may be absent or incorrect.

We propose to analyze the problem in several layers to clarify the cause and effect relations surrounding telephony fraud and explore a part of the problem at each layer. For this purpose, we base our classification on the following definition of fraud:

A **fraud scheme** is a way to obtain an illegitimate **benefit** using a **technique**. Such techniques are possible because of **weaknesses** in the system, which are themselves due to **root causes**.

Based on this general definition we further refine these concepts as follows:

- A **root cause** is an inherent characteristic of the telephony networks, standards and ecosystems which can result in weaknesses.
- A **weakness** is a vulnerability or a feature of the system, that can be manipulated in unintended ways.
- A **technique** is a mechanism, or service, which is used to abuse a weakness in a telephony system to commit a fraud. Such techniques may be illegitimate (e.g., compromising a PBX) or may have legitimate uses (e.g., conference calling) that are abused to commit a fraud.
- A **fraud scheme** is a method which is intentionally and knowingly used by a fraudster, relying on one or more techniques, to abuse a user, an entity, or a system with the goal of obtaining an illegitimate benefit.
- **Benefit:** The goal of a fraud scheme is to obtain a benefit, this can be a monetary benefit or not (e.g., competitive advantage, reputation, bypassing regulation).

### 3.1. Applying the Taxonomy on Wangiri Fraud

To present our taxonomy in a concrete way, we analyze Wangiri fraud, which is a well-known voice scam, within this context. Figure 3 summarizes this example.

**Wangiri** (‘one ring and cut’ in Japanese) fraud is also called **callback scam,** **ping call** or **one-ring scam.** In this scam, the fraudster leaves missing calls on a huge number of (usually randomly chosen) victims’ phone numbers. The call only rings once, so that the victim does not have the opportunity to answer. As a result, the curious victim calls back the phone number, which usually turns out to be a premium rate number (PRN) owned by the fraudster. The **fraud benefit** in this scheme is financial: the premium rate service provider pays the fraudster a certain share of the call.

![Figure 3: Overview of our fraud taxonomy with an example of Wangiri fraud.](image)
revenue for each minute of call received by this premium rate number.

To generate the large number of calls, the fraudster can use multiple techniques, e.g., autodialers or compromised PBX systems and spoof the originating phone number (caller ID) as the premium rate number. The fraudster can easily set up this scheme using online premium rate service providers or resellers. Such online services even include ready to use IVR systems to keep victims on the phone for a longer duration.

These techniques abuse several weaknesses in the telephony ecosystem, which could be related to the underlying technologies (e.g., lack of caller ID authentication), third party services (e.g., abusive PRN resellers) or end users (e.g., users’ lack of fraud and security awareness).

Finally, we can identify the root causes that result in these weaknesses, such as the presence of legacy protocols, convergence of multiple technologies and variety of service providers. Analyzing the problem at these different layers can help us see the overall picture and anticipate the outcome of possible actions to fight this fraud.

3.2. Methodology

Our goal is not to provide an exhaustive list of frauds, but to provide a comprehensive survey of the topic. For this, our first source of information was the literature, books, publications but also white papers from industry groups and fraud management companies. However, this is not enough as information on the topic is scattered and often incomplete.

To make sure we had a good understanding of the ecosystem, we interviewed several experts in the field and participated to industry forums. We also sent a questionnaire to a selected list of experts and well identified mailing lists to obtain feedback on the first version of our taxonomy. We only had 15 answers to this questionnaire (so we don’t present statistics) but most were from experts in fraud management or those working in the field. Their feedback and some discussions with the respondents allowed us to refine our taxonomy, to better understand fraud, and to discover new fraud schemes.

Figure 4 shows a detailed view of the taxonomy. However, it is not feasible to draw all the relations between each component of the figure in one page. Therefore, we created a dynamic picture showing all the links between the components which is available, with a copy of the questionnaire, at: https://telephony-fraud.github.io/taxonomy/.

Finally, our goal with this classification is to help explain each component of telephony fraud without ambiguity. In the next sections, we describe the taxonomy: root causes and weaknesses, techniques, fraud schemes and benefits in more detail.

4. Root Causes of Telephony Fraud

Root causes are inherent to the telephony ecosystem and are unlikely to be solved in the near term.

The legacy systems that lie in the core of telephony network were not designed with security in mind. This was not an issue when telecom networks were a closed and controlled environment where all the entities were trusted (monopolistic operators). However, this can cause various weaknesses in today’s environment. Unfortunately, upgrading these legacy systems on a global scale is not feasible in near future, due to high costs.

Telecommunication networks comprise of different, interconnected technologies, services and products, which are usually obscure and poorly understood [13]. This turns telephony networks into a large attack surface. All actors in the ecosystem have to adopt themselves to new technologies, while remaining vigilant against possible attacks.

As the telecom market became more liberalized, a large number and variety of operators have gotten involved in the market. As a result, it is not possible to make sure that all parties are carrying good intentions. It is also not possible to reduce the number of operators, as this would damage the competition and liberalization, and prevent the growth of new technologies and diversity of services.

5. Weaknesses of Telephony Networks

Weaknesses are consequences of the root causes, but they can be addressed or mitigated, if they are properly identified. We classify weaknesses in 4 categories related to protocol and network, regulation, billing and human factors.

5.1. Protocol and Network Weaknesses

Telecom networks are an interconnection of PSTN, cellular and IP networks, all of which have different weaknesses and vulnerabilities. In particular, the lack of security mechanisms in SS7 signaling leads to many problems, SS7 itself does not have any encryption or authentication mechanisms. Therefore, operators using SS7 (or anyone with access to signaling links) can tamper with SS7 messages or interact with SS7 systems [25]. The SIGTRAN protocol suite was introduced as a transport layer for SS7 messaging over IP, which can use TLS or IPsec [31]. However, there is no end-to-end security and each transit operator can modify the SS7 messages.

With deregulation and Internet convergence, it became easy to access SS7 networks, i.e., access is no more restricted to a small number of trusted operators. Nowadays, operators employ traffic screening mechanisms and filtering rules to discard unwanted incoming signaling messages [25]. Indeed, it became easier for external parties to have partial or complete access to signaling through femtocells, SIP/PRI trunks, operator partnerships (e.g., value added services) or by attacking telecom equipment [32]. Legal interception gateways, which operators often have to install to comply with laws, also have direct access to SS7, and have been sources of vulnerabilities [33], [34].

The SS7 protocol also does not support a mechanism to trace the route of a call. Each switch has its own routing
table and select the appropriate outbound link based on the destination phone number, pricing and commercial agreements. Thus, they only have a partial view of the call route which leads to a lack of route transparency. VoIP interconnections make it even harder to trace the calls. Similarly, during a phone call, caller ID (identification) information is transmitted between operators through the signaling system of the underlying telecommunication service. However, this information cannot be trusted as SS7 [25] or most of the IP based signaling protocols lack caller ID authentication. Wireless and VoIP networks also often lack proper authentication or encryption, e.g., between a mobile device and a base station, leading to the possibility to use IMSI Catchers [35]. Most of the problems in mobile network protocols are addressed starting with third generation networks. However, legacy technologies are still widely deployed, opening the possibility of downgrade attacks [36]. In addition, cellular and VoIP networks inherit some vulnerabilities from PSTN, as calls still traverse PSTN networks [37]. LTE networks involve both VoIP and cellular network related issues, and can be vulnerable to billing, DoS and caller ID spoofing attacks [38].

5.2. Regulatory, Contractual and Legal Weaknesses

Arbitrage, as a concept in economics, is the manipulation of price discrepancies in different markets. In telecommunications, price discrepancies can occur between mobile/PSTN/VoIP originated calls or domestic/international calls. Fraudsters can circumvent the high cost route or terminate a high cost call in a low cost market to profit from the price difference. Countries with high international call termination rates (usually developing countries with heavy regulations [39]) are frequently manipulated by fraudsters.

Numbering plans allow to decode phone numbers and find the operator or type of service for a given number. The E.164 standard describes a globally routable phone numbering structure and assigns number ranges (country codes) to countries [40]. Each country has its own regulatory body to further assign and control its national number range, but number portability blurs the lines. There is no global numbering plan listing all valid number ranges that are in use, although some databases allow partial lookups3. Therefore, an operator may not know for sure, if a phone number in another country is currently in use [41]. VoIP protocols use the notion of contacts instead of phone numbers. However, if the call traverses a VoIP/PSTN gateway, a phone number should be associated with the contact [42]. Many OTT providers use phone numbers to identify and authenticate their users (e.g., Viber, WhatsApp).

Telephony ecosystem embodies a large variety of regulations and laws, and the notion of legality can significantly vary depending on the country and the communication medium. For example, some countries ban VoIP usage, e.g.,

to protect their revenue from international call termination [43]. Some countries try to bound OTT providers by the same regulations that operators are subjected to [44], [45]. In general, the need for regulation may not be perceived before the system is manipulated. Therefore, it can be difficult for regulators to anticipate regulation needs.

The lack of cooperation is another weakness of telephony ecosystem. Law enforcement authorities have difficulties in international law enforcement, which makes identification of fraudsters difficult, even when the fraud is detected [24]. Moreover, despite the presence of international organizations, there is a lack of joint industry initiative to fight fraud. Due to the privacy issues and competition, operators are usually not willing to share their pricing terms, routing options or fraud related findings [46]. In addition, not all the operators have the same incentives to fight fraud. Indeed, sometimes the losses due to fraud at one operator can benefit another, innocent, operator in addition to the fraudster. In other cases, fighting small scale fraud can be more expensive than the losses due to the fraud itself.

Having a large number of operators brings the inevitable need for partnerships between them. Lack of due diligence in these partnership agreements make call traffic vulnerable to fraud, if one party has fraudulent intentions. Especially the competitive transit operators may ignore route quality and make use of cheap routes to grow their business.

5.3. Billing Related Weaknesses

Complexity of billing mechanisms have increased with the introduction of new technologies and services. Any mistake in the billing process (e.g., inaccurate or late billing, errors in pre-paid credit tracking) can be manipulated by fraudsters [13], [20]. Most of the time, operators are reluctant to change the legacy billing systems, due to the high cost and backward compatibility problems. Errors in the complicated tariff plans can also be manipulated.

Billing of value added services is another weakness, because it adds a third party to the system. Because of their high fees, they can result in significant losses. Operators should be careful in identification of value added service numbers and registration of entities who use these numbers. Unfortunately, fraudsters often abuse complex networks of resellers and service providers and are therefore difficult to identify.

Mobile roaming services also complicate billing. Roaming CDRs are not immediately available to the home operator, so detecting and stopping fraud quickly is difficult. To address this issue, Near Real Time Roaming Data Exchange (NRTRDE) systems have been developed. Nevertheless, using NRTRDE, the transmission of CDRs from the visited network to the home network still takes about four hours [47], which is a long enough time window for the fraudsters to make profit.

5.4. Human Negligence

Humans interact a lot with telecom networks. This leads to various weaknesses due to their negligence or naivety. Lack of security and fraud awareness is frequently manipulated by fraudsters [13]. On the enterprise level, lack of internal control systems (such as access control), poor deployment practices (weak passwords, neglecting updates) and lack of vulnerability management in software and hardware systems are some other sources of weaknesses [21].

6. Techniques Used in Fraud Schemes

In this section, we describe the techniques which enable various fraud schemes. Some of these techniques may have legitimate uses as well. We group them by the kind of access they require (e.g., operator level) or their purpose (e.g., increasing profit).

6.1. Operator Level Techniques

Number range hijacking occurs when a fraudulent operator advertises very cheap rates for a destination number range and attracts traffic from other operators [48]. For example, in Figure 1, there are several possible routes to the terminating operator. Assume that routes 1 and 3 are the usual routes. In the event that the transit operator 3 suddenly advertise a very cheap rate (possibly for a very small range of numbers), the originating operator may select route 2 for delivering the calls. In this case, the calls to the victim number range will be hijacked and routed/terminated fraudulently [49], [50]. Lack of due diligence in operator partnership agreements facilitates this technique.

A parallel can be made between phone number range hijack and BGP hijacks [51]. In both cases, a part of the traffic is redirected by a malicious entity that advertises false (or misleading) information. For phone number ranges, this is the price for a destination, while in BGP, this is the prefix advertisements. However, as opposed to the IP networks, call routing is opaque (Section 5.1), which makes detection more difficult. Furthermore, there is no mechanism in telephony networks to directly authenticate the owner of a number range or check if an operator really has the connectivity to route the call to that number range. Like with BGP, deploying security mechanisms would face significant practical difficulties [52].

Manipulation of call routing is possible as the operators have full control over the calls that transit through their networks (either legitimately or because of a hijack). A fraudulent transit operator can divert a call or send it over illegitimate routes to perform different fraud schemes. In case of call short-stopping, the transit operator directly terminates the call (e.g., to an IVR) instead of sending the call to the legitimate destination. It can also selectively short-stop only some of the calls. Due to the lack of route transparency, the originating operator cannot know if the call was routed normally and has reached the correct destination.
Manipulation of call signaling messages is also easy for the operators. For example, the caller ID can be changed to fake the origin of the call (which may affect billing). Call setup signals can be tampered to answer the call before it is actually answered by the customer (early answer) or to not disconnect the call immediately (late disconnect) [53]. The call will be longer than it should, which will affect the revenue (False Answer Supervision, see Section 7.3).

Pricing confusion is the use of multiple and varying pricing plans to confuse customers about the real market price of a service. Such operators constantly provide new offerings and special introductory discounts, to be competitive [15], but quickly change the prices once customers are registered.

6.2. Techniques For Increasing Profit

Here we present techniques which can be used to make a fraud scheme more efficient, however, many of them have a legitimate use.

Traffic pumping, or artificial inflation of traffic, is the act of generating a high level of call traffic to some phone numbers deliberately. This can be achieved by creating and advertising 3rd party services such as conference calling [54], free radio broadcast over phone [55] or adult entertainment. By providing such services for free (or at a very low cost) many users are attracted, which, in turn, generates a high volume of calls. Value added services or arbitrage opportunities can make traffic pumping advantageous in certain fraud schemes (Sections 7.4 and 7.3).

Initiating multiple simultaneous calls allows the fraudster to increase the profit of a fraud scheme in a certain time window. Multiple outgoing calls, or conference calls, can be generated on compromised PBXs [10], VoIP accounts [56], or SIM cards. Up to 6 simultaneous calls can be generated from a single SIM card [57]. Finally, call forwarding can be used to forward all incoming calls to a certain fraudulent phone number.

6.3. Value Added Services

Premium Rate Numbers (PRN) are used to provide wide range of services such as gambling, live chat, adult services; through voice call or SMS. To cover the cost of services provided, the cost of calling a premium rate number is much higher than a regular call. In most countries, a fixed number range is allocated for PRNs which allows users to easily distinguish them, however, it is not true everywhere. Users sometimes tend to confuse the number ranges and call PRNs unwittingly. Such premium rate numbers may be abused, e.g., when the promised service is not delivered, the cost of the service is not clearly stated or artificial traffic is created to these numbers [48]. The abusive premium rate services usually manipulate the lack of due diligence between number resellers and numbering plans in which the premium rate number range is not clearly identifiable by users [58]. Many online sites offer premium rate number services, which gives a cash back on calls reaching this premium number.

CNAM (Caller Name) lookup service provides a 15-character long caller name string (associated with caller’s phone number), to help users easily identify a caller [59]. In the USA, operators are responsible for making the CNAM lookup (dip) for the calls received by their customers. A CNAM service usually comes as part of the landline package and it is enabled by default. However, there is no centralized CNAM database in North America. Instead, multiple independent CNAM providers allow operators to lookup the CNAM information for a fee [60]. Fraudsters can use a CNAM service to register a false caller name for their phone number, or to abuse the payment mechanism (Section 9).

Toll free numbers are phone numbers which do not incur any charges to the caller. Instead, the call is charged to the toll free customer (call recipient), which are usually call center services. Toll free numbers use a prefix allocated by the regulator. For toll free numbers, charge collection is reversed: The toll free customer pays the toll free number provider (usually the terminating operator) for all incoming calls. Toll free providers keep a part of the profit and passes a share to the originating operator, as the caller does not pay for the call [61].

6.4. Protocol Related Attacks

SS7 tampering by external parties became possible with easy access to SS7 networks. This can lead to attacks such as locating the phone users, intercepting the calls or denial of service [62], [63].

VoIP protocol attacks can manipulate the implementation flaws, underlying network platform or the voice application layer [64]. Various attacks, such as SIP scanning, registration hijacking, redirection attacks, session tear down, SIP phone reboot and audio insertion are demonstrated in [21], [37]. Billing systems can be manipulated through VoIP attacks as well [65].

IMSI catchers (or stingray) [35], [36] are fake GSM base stations that are used to identify phones in proximity (catch their IMSI), intercept calls and communications, or even to send out spam and fraud messages [66]. IMSI catchers manipulate the lack of authentication from the network to the device in GSM. Such a fake base station can be built using operator grade equipment or open-source software and cheap hardware [67], [68], [69]. The phone is deceived to connect to the false base station and usually the mobile device is forced to not use encryption or downgrade to an insecure mode (e.g., 3G to 2G) [36]. More recent mobile protocols (like LTE) are using authentication but are not immune to such attacks; first, the authentication keys could be leaked (or seized); second, the IMSI catcher may abuse vulnerabilities in the protocol stacks [70]. Some discrepancies in the perceived network features can be used to detect IMSI catchers [71], [72], [73], [74].

Caller ID spoofing requires transmission of fake caller IDs in the signaling system. Even though there are certain
legitimate ways of using caller ID spoofing [75], it is definitely a catalyst for voice fraud. PRI and SIP trunks enable transmission of fake caller ID to the SS7 network, as a result of the lack of caller ID authentication [41], [25]. Various online services and mobile applications provide caller ID spoofing, via the service provider’s IP-PSTN gateway connections. Spoofing caller ID between two VoIP applications is even easier, because the caller ID can be inserted in SIP requests.

Apart from enabling various fraud schemes, caller ID spoofing can be damaging when used against services (such as banking systems, voice mails or emergency services) where caller ID information is used to authenticate or locate the users [76].

STIR (Secure Telephony Identity Revisited) working group [77] attempts to provide a SIP header authentication mechanism to authenticate the caller ID [78], [79]. However, operators may be reluctant to deploy this solution due to the implementation overhead [23]. Moreover, with cloud VoIP services, phone numbers became an extremely cheap commodity and fraudsters could easily obtain bulk phone numbers and change authenticated phone numbers faster than they can be blacklisted.

Researchers also propose caller ID authentication mechanisms for telephony networks. AuthLoop [80] is a TLS-inspired protocol that uses the voice channel to provide end-to-end authentication irrespective of the underlying technology. Tu et al. [81] proposes a caller ID authentication scheme via the transmission of a pre-computed security indicator over the SS7 network. However, these solutions have their own limitations (e.g., difficulty of deployment, scalability or overhead) and have not yet found extensive, practical use. Another related work includes the use of call audio features to determine the source of the call and the types of traversed networks [82].

6.5. Other Techniques

PBX hacking is a significant threat for enterprises. PBX systems (hardware or software) are often not properly administrated and secured. Attackers can find PBX systems by calling a large range of numbers (e.g., via wardialing, in Section 7.5) or through the enterprise’s publicly known phone numbers. IP-PBXs can also be identified using SIP scanners [83]. Once a PBX is identified, attackers will typically gain access to maintenance interfaces or voice mail systems (e.g., by abusing weak passwords), or use social engineering [84], [85] to compromise and reconfigure the PBX.

Cellular phones and SIM cards can be abused by techniques like cloning and theft. In CDMA networks, phone cloning is done by reprogramming phone’s electronic serial number and mobile identification number. In GSM networks, the phones are identified by their IMEI number. Tampering IMEI can be useful in some countries to circumvent state control on phones, or to avoid blacklisting of stolen phones. SIM swap is a service provided by operators to the customers to register an existing phone number on a new SIM card. This service can be manipulated by fraudsters to obtain the ownership of the phone line [86]. To this end, the fraudster contacts the operator claiming, e.g., that the SIM card was stolen, and uses social engineering techniques to impersonate the SIM card owner. If the fraudster can convince the operator to register a new SIM card for a particular phone number, he can generate calls that will be billed on someone else’s account [87]. SIM swap can also affect two factor authentication mechanisms, including banking [88].

SIM boxes, or GSM Gateways, are devices that can act as a gateway to the mobile network (e.g., GSM) and provide a VoIP or PRI trunk. Those devices can be used to provide mobile connectivity to a PBX. The device is essentially composed of one, or more, mobile modems to which SIM cards can be attached, the modems are then controlled by a computer which converts the calls to VoIP or ISDN. SIM boxes have legitimate uses (such as providing GSM gateways to enterprise PBX systems) that are permitted by operators and regulators [89], [90]. However, there are many frauds that rely on SIM boxes, in particular interconnect bypass (Section 7.3) and IRSF (Section 7.4) [91], [92].

Autodialers are systems that automatically dial telephone numbers, randomly or given a predefined list [93]. Once the call has been answered by the other party, a sophisticated autodialer can analyze the incoming audio stream to predict whether it has been picked by a real human being or an answering machine. Autodialers can either play a recorded message or connect the call to a live person upon the determination of human pickups.

Telephony Denial of Service (TDoS) attacks are performed by sending a very large volume of call traffic to a target number, to deprive the system resources (such as the trunk capacity) and disrupt the telephony service of the targeted customer. There are different ways to initiate a TDoS attack, such as organizing people on social media to call a specific number, or using autodialers [21]. TDoS attacks became easier with the convergence of telephony and internet. The attacker can use a VoIP-PSTN gateway to generate cheap calls, a call generation software and some audio content to stalk the target with a realistic scenario [21]. The VoIP-PSTN gateway can be a free IP PBX software (e.g., Asterisk [94]) with an access to a SIP trunk. Other methods can be using a compromised PBX, a botnet or an online TDoS service [95].

Social engineering is the process of manipulating people to act in a certain way or to give up confidential information [21]. Social engineering attacks exploit individuals’ lack of security and fraud awareness. Tricking company employees to give up their passwords or persuading people to call a certain phone number are some examples of social engineering attacks. Telephony has been a preferred channel for social engineering because of the trust people put in the telephone and because impersonation is easier over the phone [15]. Malware infecting smartphones and VoIP phones can steal personal data, e.g., helping in social engineering.
engineering, but can also initiate calls or send SMS [96], [97].

7. Fraud Schemes

Abuse of telecommunication infrastructure started in late 50’s. Driven by technical curiosity, the phreakers (from ‘phone freaks’) explored the telephone network by reverse engineering the tone patterns in in-band signaling, and by social engineering telecom company’s personnel [98]. Detailed information on the history of telephony fraud can be found in [99], [98], [15], [100].

In this section, we provide an overview and categorization of the modern fraud schemes. We focus on the current fraud ecosystem, but some of the schemes may be applicable or profitable only in some parts of the world. Also, we do not provide legal opinion on the fraud types as it depends on country specific laws.

7.1. Toll Evasion Fraud

Toll evasion fraud aims at making calls without the obligation of paying the call charges. Toll evasion is the oldest type of fraud in telephone networks [16] and can be categorized in four main categories.

Subscription fraud is the act of using stolen identity credentials, or providing false information while subscribing for a service, in order to avoid service charges [13], e.g., SIM card subscription for a post-paid account using false information. Researchers have proposed solutions to detect subscription fraud using data mining [101] and other classification techniques [102].

Similarly, superimposed fraud aims to take control over a legitimate customer’s account via cellular cloning or theft, and burden call charges on his account [103], [104], [105].

Enterprises are also frequent targets of toll evasion fraud. In PBX dial-through fraud, compromised PBXs can be used to make free calls, while the call charges are ascribed to the PBX owner. [106] demonstrates various attacks aiming at toll evasion fraud, using a IP-PBX honeypot with PSTN connections. Internal fraud is usually committed by employees of a telecom company who have access to user accounts, tariff plans and billing system [13]. Fraudulent employees can, for example, deactivate billing for certain accounts, tamper the call records or manipulate the tariff plans to avoid or reduce the call charges [21].

7.2. Retail Billing Related Fraud

In this section, we analyze the fraud schemes related to retail billing (see Section 2.3). The fraudulent actor may be the customer, the operator, or a 3rd party service.

Due to the complexity of billing systems or confusing pricing policies, tariff plans can be inconsistent or can contain errors. Customers or other operators can exploit the mistakes in tariff plan or campaigns. One example of tariff plan abuse occurs when customers gain credits as they receive calls. In this case, customers inflate traffic to their own phone numbers to gain free credits. Another example is the abuse of unlimited, or flat rate calling plans by customers or businesses to call specific destinations, or by wholesale operators to terminate large volumes of traffic.

Over-billing fraud is performed by operators against their customers or partners, to increase operator’s revenue illegitimately. Cramming is one example of over-billing fraud, where telecom operators or other service providers intentionally place unauthorized charges on client’s bill, for providing services such as voice mail [107]. Customers may be deceived into accepting these charges while signing promotional materials, or through social engineering techniques like negative option marketing [108]. Another scheme called Slammering occurs when a fraudulent telecom operator switches the local or international service provider of the customer to itself, without the customer’s consent and explicit notice [15]. The operator may additionally charge the customer for high call termination rates.

Unauthorized call reselling scheme involves reselling of fraudulently obtained calls (e.g., obtained via toll evasion fraud) for prices lower than the market rates. Reselling calls over a compromised PBX is also called call transfer fraud [10].

7.3. Wholesale Billing Related Fraud

This section covers the fraud schemes related to the inter-carrier billing processes in the wholesale market.

Call routing abuses are possible due to the lack of route transparency in the telecom networks. This fraud usually aims at profiting from the arbitrage opportunities. Route blending occurs when an operator illegitimately sends part of its transit traffic over a low quality and low cost network, violating its inter-carrier contracts and service level agreements [24].

Re-originating (also known as re-file, hubbing, refiling) uses an intermediate country to decrease the termination charges. The international traffic is first sent to a hub country, which has a more competitive market and more favorable interconnect agreements [109], [110]. The hub country changes the origin of the call, by modifying the Calling Line Identity (CLI) in SS7 signaling, and sends the call to the destination country. In this way, a cheaper call termination fee is paid to the destination country. Both the originator and intermediate operator profit from this operation; whereas, the operator in destination country loses revenue.

These fraud schemes became less profitable as the pricing anomalies reduced with the increased competition in telephony markets.

Interconnect bypass fraud, or gray routing, can be defined as the use of illegitimate gateway exchanges to avoid the legitimate gateways and international termination fees5. These routes are sometimes used to coordinate

---

5. Because SIM Boxes and PBXs are frequently used to enable these illegitimate routes, this fraud is also named as Leaky PBX, SIM Box fraud and GSM Gateway fraud.
criminal activities, because the calls are more difficult to intercept [111]. Gray route fraud exploits, e.g., the price differences between wholesale and retail markets [112] to reduce the cost of an international call. This is achieved by fraudsters establishing a “bypass” in either the destination or an intermediate country.

The bypass mechanism mostly involves the use of a SIM Box or a PBX [11]. A basic form of interconnect bypass occurs when a fraudulent transit operator routes an international call over the IP network and terminates it as a domestic mobile or landline call in the destination country. However, interconnect bypass can take many forms, depending on the arbitrage opportunity. A more recent type of bypass is performed by terminating the regular international calls on the Over-The-Top (OTT) applications installed on recipients’ smartphones [113], [114].

Interconnect bypass fraud leads to financial losses for the destination operator as well as the bypassed transit operators who were supposed to get a share of the call revenue.

Commercial approaches to detect interconnect bypass fraud include test call generators and statistical fraud management systems [11]. In addition, there exists machine learning [92], [91] and call audio analysis [90] techniques in the literature focusing on the detection of SIM Boxes and blocking of bypassed calls.

**False Answer Supervision (FAS)** fraud enables (transit) operators to fraudulently increase their revenue from each call, by performing one of the following [24]:
- **False answer**: (also called short-stopping fraud) the operator diverts a call (short-stops it) to a recorded message and starts charging, instead of transmitting the call to the real network.
- **Early answer**: the operator increases the duration of the call fraudulently by, e.g., answering the call and playing a fake ringing tone until the callee actually answers [115].
- **Late disconnect**: the operator delays the transmission of call disconnection message to the calling party and therefore bills for a longer call.

According to a survey conducted in 2013, FAS was the top fraud reported by the wholesale carriers [24]. It may also damage the reputation of retail carriers, as they may receive customer complaints about incorrect billing issues.

### 7.4. Revenue Share Fraud

Revenue share fraud occurs when an operator (or third party service provider) makes an agreement with another party which will generate calls to predefined numbers (the revenue share numbers). The operator who owns the revenue share numbers, usually advertises these numbers through an online premium rate service reseller (see Section 6.3). A fraudster can easily obtain a revenue share number and start generating calls to this number.

Revenue share fraud often involves a combination of multiple fraud schemes. In this section, we will first examine the schemes used for traffic generation to the revenue share numbers. Then we will analyze common fraud agreement schemes and in particular fraudulent termination.

#### 7.4.1. Traffic Generation Schemes

In general, fraudsters will be attracted to generate calls as long as the revenue share they receive is higher than the cost of generating the calls.

Toll Evasion Fraud (see Section 7.1) can be used to create traffic without bearing any charge to the fraudster. Common techniques involve exploiting PBXs, roaming SIM cards, or using dialer malware (e.g., smartphone malware which dials revenue share numbers without user’s permission [96], [116]). Fraudsters may also abuse unlimited (or low cost) international tariff plans.

Call generation schemes are often combined with other techniques to maximize profit (Section 6.2) of revenue share fraud in a limited time, before the fraud is detected. For instance, PBX dial through fraud can be combined with conference calling, multiple call and call forwarding techniques [10]. Moreover, social engineering and scams (Section 7.6) can be used to deceive people into calling the revenue share numbers. For instance, fraudster can send a phishing SMS or leave voice mail to many phone lines, promising them a profit or service if they call a specific number. Social engineering is also used to keep the caller on the phone for a long duration.

#### 7.4.2. Fraud Agreement Schemes

Various different fraud agreement schemes are possible. For example, in *domestic revenue share fraud*, parties making the fraud agreement operate in the same country, whereas the *international revenue share fraud* involves international traffic. CNAM revenue share fraud is also possible and will be described in more details in Section 9.

**Access stimulation fraud** is a form of domestic revenue share fraud seen in the USA, which abuses the high termination rates in rural areas. In this fraud, an operator in a rural area makes an agreement with a company that bears high volume of inbound calls [117]. Such companies use various techniques (see Section 6.2) to inflate incoming traffic to the number range of the operator in rural area. Finally, the operator shares some of its revenue with the company.

**In toll free number fraud**, the fraudster makes an agreement with an operator (often a competitive local carrier) and uses this operator to initiate a large volume of calls to a toll free number [10]. Because of the reverse payment flow used for toll free numbers (Section 6.3), the originating operator earns revenue from this call volume, and shares the profit with the fraudster who generated the calls.

**In International Revenue Share Fraud** (IRSF), a fraudulent operator, or third party service provider, advertises a range of phone numbers as *International Premium Rate Numbers* (IPRN) in various parts of the world [22], [58]. This victim number range often belongs to a small, developing country, or to a satellite operator with a high interconnect termination fee. In general, IPRNs are not actually part of a real premium range in the target country.
A fraudster can obtain one of these ‘pseudo’ international premium rate numbers from various websites\(^6\) that offer revenue in exchange for traffic generation to these numbers [118]. The number of such reseller websites increased by 400\% between 2009 and 2013 [119]. Depending on the revenue share mechanism and fraud agreement, the owner of the victim number range may or may not be aware that its numbers are used for IRSF [58]. E.g., a transit operator may short-stop the calls to the victim number range, keeps the termination fee, and shares it with the fraudster who generates the calls, as depicted in Figure 5. The owner of the number range may only become aware of the fraud because he may become unreachable from some originations [120].

IRSF was initially performed by using stolen SIM cards in roaming to generate calls to IPRNs [57], [121]. In an attempt to stop such fraud, NRTRDE (Section 5.3) was introduced [122], [123]. However, looking at IRSF within our taxonomy (Figure 6), we can see that this only partially mitigates one form of toll evasion, but does not deal with the real causes behind IRSF. The estimated loss due to IRSF was $10.76 Billion in 2015 [6].


7.5. Targeted Fraud

Fraud schemes targeting a certain company or individual may not be as common as other types of fraud, but they may have significant consequences, affecting many users or resulting in huge losses. Impostering is the act of stealing someone else’s identity, and performing operations on his/her behalf. Techniques like caller ID spoofing, fake base stations, mobile malware or social engineering can be used to spoof identity, steal one-time passwords and trick companies (such as banks) to perform unauthorized operations. Another fraud targeting individuals is call interception and eavesdropping operations. SS7 access, mobile technologies (IMSI catcher), insecure VoIP systems (compromised PBXs) and legal interception gateways enable eavesdropping on phone calls, SMS messages as well as location tracking [21], [62], [124]. Although, it is possible to have end-to-end encrypted VoIP calls and secure VoIP protocols (like SIPS, SRTP) exist, majority of voice communications are still vulnerable to eavesdropping, and subject to legal interception. Another form of targeted fraud can be performed by tracking an individual’s call records (via phone bills or operator’s CDR database) and extracting information from the call metadata (e.g., source and destination numbers, duration) [21], [125]. To prevent this kind of attacks, researchers have proposed a call forwarding infrastructure that will obfuscate the metadata by routing the call over a series of telephony relay nodes [125].

Wardialing is the use of autodialers to scan a phone number range, e.g., to identify modems, fax machines or voice mail accounts. This can be used for reconnaissance purposes targeting a company, and further attacks can be initiated [126]. In CNAM datamining, a fraudster calls himself multiple times by spoofing the caller ID of other numbers, to obtain the caller name information for those numbers. The calls are never answered but the attacker’s carrier is forced to make a CNAM lookup for the spoofed number [127].

Disruption of service, e.g., using TDoS (see Section 6.5), can have devastating effects for an operator, an enterprise or service (such as health or police emergency lines), and is often used in blackmailing schemes [128], [129].

7.6. Voice Spam and Scams

Voice spam is one of the most visible types of voice fraud targeting customers. It includes all types of unsolicited and illegitimate calls. Scammers can obtain phone number lists from leaked databases, form submissions, or simply by purchasing them online [19]. They mostly use autodialers to generate large number of calls and use prerecorded messages (robocalling) which may be later forwarded to live call center agents to interact with the victims. Caller ID spoofing and social engineering techniques are frequently used to deceive people to do certain actions or to reveal sensitive information. Due to the low cost and scalability of VoIP based calling systems, scammers can make millions of calls and easily expand the scam ecosystem. A recent work [19]
describes the challenges that arise in fighting voice spam by comparing it to email spam. It then analyzes existing anti-spam techniques and provides an assessment criteria. Voice spam can take many forms, but we will explain some of the most common schemes (see [19] for more examples).

Telemarketing is a method of direct marketing in which a salesperson entices customers to buy products or services over the phone [17]. Telemarketing can be illegitimate in certain jurisdiction, e.g., if the telemarketer did not take prior consent from the call recipient [130].

In voice phishing (also known as vishing), the caller imitates a legitimate organization, person or entity and tries to gain access to private, personal and financial information using social engineering [131], [21]. Caller ID spoofing is often used by scammers to hide their real identity and makes it difficult to block the spam calls or to take legal actions against fraudsters [19].

Many other types of scams can make use of telephony. For example, in the tech support scam, fraudsters try to convince people that their computer is infected with malware (mostly by tricking them into installing remote access tools) and request a payment to solve the so-called problem [132], [133]. In advance fee fraud (419 scam), the victim is being tricked into making some up-front payment to be able to receive a larger sum of money, such as a bogus lottery prize [134]. A similar scam is the free cruise scam, where fraudsters advertise a free cruise opportunity, but later on require additional payments [135].

8. Benefits

Benefits are central to the fraud, and without the benefits, there would be no fraud. Most of the fraud schemes on telephony networks target financial benefits, but financial aspects are not the only motivation. For example, a goal may be to learn individuals’ opinions about a political election, through voice spam [19].

It is important to make a separation between fraud schemes and benefits, as different fraud schemes can target the same benefit. One example of problem when failing to separate fraud scheme from the benefit is the use of toll fraud term for various fraud schemes such as subscription fraud, International Revenue Share Fraud (IRSF) and PBX dial through fraud [21], [136], [137]. Here, “toll fraud” indicates that the fraud will have financial benefits. However, these fraud schemes are very different from each other and lead to different types of financial benefits. A fraud scheme may also generate multiple benefits. For example, telemarketing can involve increase of company revenue through sales, influencing people through advertisement, and getting a share of revenue through CNAM revenue share [138], [139]. Hence, separating fraud schemes and benefits provides a clear view of fraud, and solves common terminology problems.

9. Case Study: CNAM Revenue Share Fraud

CNAM revenue share is a not very well known fraud mechanism, this scheme was probably made possible because of the obscure and deregulated nature of the CNAM service. As we mentioned in Section 6.3, many CNAM (Caller Name) lookup (dip) service providers exist. Operators rely on them to provide caller name information to their customers. When a call is received by the terminating phone company, it performs a dip for a fee (CNAM dip fee). This compensation happens for every call where the calling party name is displayed to the called party, even if the call is not answered.

Figure 7: CNAM revenue share fraud.

CNAM revenue share fraud is similar to other revenue share fraud schemes where a fraudster generates call traffic to a partner service provider, who will share part of his revenue with the fraudster. However, the fraudster needs to ‘originate’ calls from the revenue share numbers (instead of ‘terminating’). This is because the CNAM lookup service uses the looked up caller ID to identify the revenue share partner. Moreover, call generation is cheap for the fraudster because the calls are generally not answered.

On the other hand, the CNAM dip fee is usually very small and to make a significant profit, the fraudster has to generate a large number of calls from the revenue share number. As it is not practical to generate thousands of calls from a single phone, the fraudster usually spoofs his caller ID and uses multiple autodialers to generate the calls (Figure 7).

For telecom customers, the visible effect of this fraud will be a large number of missed calls on their landline phones. These calls are most likely to be treated as voice spam (or ping calls) by customers. In fact, CNAM services are also often used by telemarketers to get some additional revenue from their calls [140], which could explain why it is uncommon for telemarketers to randomly spoof their caller IDs [141].

CNAM fraud advertisement in India. In [138], a CNAM provider offers a revenue share deal to a call center in India. In this example, 34 million calls were generated over a period of 10 months with 39% of the calls leading to a CNAM dip reaching the CNAM provider. The average revenue share was of $1.15 per thousand CNAM lookups (dips), generating a total revenue of $15,372. While it is difficult to fully trust this advertisement, it nevertheless provides an interesting view on the scale of this fraud. Another source confirms the order of magnitude by advertising a revenue of 50c to $1 per 1000 calls [139].

Honeypot findings. We also study the data from the telephony honeypot presented by Gupta et al. in [17]. We collected data during the course of 2.5 months from 17th July 2014 to 30th September 2014. The 39K phone
Figure 8: Taxonomy for CNAM revenue share fraud.

numbers linked to the honeypot received 2.3 million calls from 285K sources on 275K honeypot numbers. In addition to the numerous fraudulent calling patterns, e.g., telemarketing and robocalls [17], we also noticed traces of CNAM revenue share fraud. Table 1 shows the call volume from TABLE 1: CNAM revenue share fraud case study on the telephony honeypot.

<table>
<thead>
<tr>
<th>Source Number</th>
<th># calls</th>
<th>Source Number</th>
<th># calls</th>
</tr>
</thead>
<tbody>
<tr>
<td>141XXXX2353</td>
<td>13040</td>
<td>141XXXX2807</td>
<td>1684</td>
</tr>
<tr>
<td>141XXXX2328</td>
<td>7118</td>
<td>141XXXX2801</td>
<td>1389</td>
</tr>
<tr>
<td>141XXXX2538</td>
<td>6078</td>
<td>141XXXX2335</td>
<td>1334</td>
</tr>
<tr>
<td>141XXXX2368</td>
<td>3503</td>
<td>141XXXX2322</td>
<td>684</td>
</tr>
<tr>
<td>141XXXX2362</td>
<td>2918</td>
<td>141XXXX2710</td>
<td>597</td>
</tr>
</tbody>
</table>
the number block which was used by the attacker. A total of 39K calls hit the honeypot from this pool of 10 source numbers. The intuition that this was CNAM revenue share fraud was confirmed by a fraud management expert from a Tier 1 operator, who recognized the pattern and numbers used.

Fighting CNAM fraud. Figure 8 shows CNAM fraud in our taxonomy, we see that there are several ways CNAM fraud could be mitigated: by addressing the call generation schemes (regulating telemarketing, or autodialers), the protocol weaknesses (Caller ID authentication), or the financial benefits (for example by regulating cash back schemes). None of them are likely to completely solve the problem but the taxonomy provides a clear view of the CNAM ecosystem and the points at which the CNAM revenue share fraud can be addressed.

10. Conclusion

Telephony fraud has constantly evolved throughout history. Starting with curiosity of people and simple tricks to make free calls, it became a huge fraud ecosystem involving various actors, technologies, applications and networks. In this paper, we systematically studied the fraud in voice telephony. For this, we proposed a taxonomy that provides a holistic framework to better understand and analyze such frauds. Our work provides a comprehensive view of the causes, weaknesses, techniques, schemes and benefits obtained by fraudsters. Based on a clear fraud definition, new fraud schemes can be easily inserted in our taxonomy to extend it. This will help to analyze their relations and interactions with other fraud schemes and techniques, and finally pinpoint their root causes. Without such a holistic view of fraud schemes, they may be misunderstood. E.g., CNAM fraud can be easily misinterpreted as ping calls but it may not be addressed the same way because the revenue share mechanism is not the same. Despite the advances in fighting telephony fraud, it will likely continue to be an important topic in the foreseeable future. A good understanding of the problem is required to continue the fight against fraud. We hope that our work will foster more academic research on this topic and in particular help to understand effectiveness and implications of new countermeasures.

11. Acknowledgments

This research was partially funded by the Principality of Monaco. It was completed while Payas Gupta was employed at New York University Abu Dhabi. In addition, Mustaque Ahamad’s participation was funded by National Science Foundation award 1318167. We also thank all the experts who answered the questionnaire and the anonymous reviewers for their invaluable feedback.

References


F. Joachim and B. Rainer, “Method for identifying a mobile phone.”


“Snoopsnitch project,” opensource.xlabs.de/projects/snoopsnitch.


H. Mustafa, AhmadSadeghi, S. Schulz, and W. Xu, “You can call but you can’t hide: Detecting caller ID spoofing attacks,” in DSN’14.


B. Reaves, L. Blue, and P. Traynor, “AuthLoop: End-to-end cryptographic authentication for telephony over voice channels,” in

Authorized licensed use limited to: University of Virginia Libraries. Downloaded on April 10,2021 at 18:52:09 UTC from IEEE Xplore. Restrictions apply.


[115] “Details about FAS (False Answer Supervision),” Nexmo, Dec’16.


[137] “5 tips for businesses to prevent telecom toll fraud,” Transnexus Whitepaper, April 2015.


