From Understanding Telephone Scams to Implementing Authenticated Caller ID Transmission

by

Huahong Tu

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2017 by the
Graduate Supervisory Committee:

Adam Doupé, Co-chair
Gail-Joon Ahn, Co-chair
Dijiang Huang
Yanchao Zhang
Ziming Zhao

ARIZONA STATE UNIVERSITY

December 2017

ABSTRACT

The telephone network is used by almost every person in the modern world. With the rise of Internet access to the PSTN, the telephone network today is rife with telephone spam and scams. Spam calls are significant annoyances for telephone users, unlike email spam, spam calls demand immediate attention. They are not only significant annoyances but also result in significant financial losses in the economy. According to complaint data from the FTC, complaints on illegal calls have made record numbers in recent years. Americans lose billions to fraud due to malicious telephone communication, despite various efforts to subdue telephone spam, scam, and robocalls.

In this dissertation, a study of what causes the users to fall victim to telephone scams is presented, and it demonstrates that impersonation is at the heart of the problem. Most solutions today primarily rely on gathering offending caller IDs, however, they do not work effectively when the caller ID has been spoofed. Due to a lack of authentication in the PSTN caller ID transmission scheme, fraudsters can manipulate the caller ID to impersonate a trusted entity and further a variety of scams. To provide a solution to this fundamental problem, a novel architecture and method to authenticate the transmission of the caller ID is proposed. The solution enables the possibility of a security indicator which can provide an early warning to help users stay vigilant against telephone impersonation scams, as well as provide a foundation for existing and future defenses to stop unwanted telephone communication based on the caller ID information.

ACKNOWLEDGMENTS

Reviewing upon the last few years, I would never have been able to finish my dissertation without the guidance of my advisors and the support of my family.

First of all, I would like to express my sincere gratitude to my advisor Dr. Adam Doupé. Dr. Adam Doupé has been a mentor, a colleague, and a friend. His continuous support of my Ph.D. study has made this work possible. It has been an exciting and rewarding journey to finally be able to finish this dissertation under his guidance. I have learned so many things from him in research and teaching. Without his support and guidance, this dissertation would not have been possible.

I would like to thank Prof. Gail-Joon Ahn, who gave me the opportunity to work with the brilliant people in SEFCOM. I had a turbulent relationship with my ex-advisor and without his lab, I would not have the opportunity to continue my Ph.D. study. He has been monumental in attracting talents and acquiring sponsors at our lab. As a young faculty aspiring to establish my own lab, Prof. Ahn is someone I look to as a shining example.

I would like to thank Dr. Ziming Zhao, for his advice and for being a friend. He has always encouraged me with his best wishes. Our similar cultural backgrounds made it easy for us to share our thoughts and aspirations. We are both young and hungry to make an impact in our world. It would have been a lonely lab without him.

I would like to thank my fellow labmates in SEFCOM, for the stimulating discussions and for all the fun we have had in the last few years. It was not only fun to talk about our work, but also interesting to talk about things other than just our papers.

Finally, I would like to thank my beloved mother, for raising me and supporting me throughout my life. Despite the struggles that come with growing up with my mother, we eventually overcame the problems that life throws at us and became stronger than ever. I would not have been who I am today without my mother.

Thank you all.

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 INTRODUCTION 1

1.1 Dissertation Outline 6

2 BACKGROUND 8

2.1 Public Switched Telephone Network 8

2.1.1 PSTN Call Setup Signaling 10

2.2 How Telephone Spam Works 14

2.2.1 Spammer Operation 16

2.2.2 Telephone Spam Samples 21

2.3 Overview of the Caller ID 23

2.3.1 How Caller ID Spoofing Works 24

3 UNDERSTANDING WHY TELEPHONE SCAMS WORK 26

3.1 Background and Insights 27

3.2 Study Design 28

3.2.1 Attributes 29

3.2.2 Experiments 32

3.2.3 Population 33

3.2.4 Procedure 35

3.2.5 Ethical Compliance 44

3.2.6 Dissemination 45

3.3 Results and Analysis 46

CHAPTER Page

3.4 Survey Responses 57
3.5 Limitations 59
3.6 Discussion 61
3.7 Related Work 62
3.8 Conclusion 63
4 IDENTIFYING KEY CHALLENGES AND EXISTING COUNTERMEASURES 65
4.1 Key Challenges 66
4.1.1 Immediacy Constraint 66
4.1.2 Difficulty of Working with Audio Streams 67
4.1.3 Lack of Useful Header Data 67
4.1.4 Hard to Gain User Acceptance 67
4.1.5 Caller ID Spoofing 67
4.1.6 Difficulty of Tracing Spam Calls 68
4.1.7 Entrenched Legacy Systems 69
4.1.8 Lack of Effective Regulations 69
4.1.9 Lack of Globalized Enforcement 70
4.2 Basic Techniques and Countermeasures 71
4.2.1 Call Request Header Analysis 71
4.2.2 Voice Interactive Screening 77
4.2.3 Caller Compliance 79
4.3 Assessment Criteria 83
4.3.1 Usability Criteria 84
4.3.2 Deployability Criteria 85

CHAPTER Page

4.3.3 Robustness Criteria 86
4.4 Combining Techniques 87
4.5 Related Work 91
4.6 Conclusion 93
5 PROPOSING AUTHENTICATED CALLER ID TRANSMISSION 95
5.1 The Rise of Caller ID Spoofing 96
5.2 Solution: Security Indicators 99
5.3 The Underlying Caller ID Authentication Scheme 101
5.3.1 Caller ID Verification 106
5.3.2 Authenticated Call Request 108
5.3.3 Security Considerations 111
5.3.4 Local Deployment Considerations 113
5.4 Related Works 115
5.5 Conclusion 116
6 IMPLEMENTING PROTOTYPE WITH EVALUATIONS 117
6.1 Prototype Design 118
6.1.1 Caller ID Verification 119
6.1.2 Authenticated Call Request 121
6.2 Performance Analysis 126
6.3 User Study 129
6.3.1 Study design 130
6.3.2 Participant Recruitment 131
6.3.3 Evaluation 135
6.3.4 Interview Findings 140

CHAPTER Page

6.4 Discussion 143
6.5 Conclusion 144
7 CONCLUSION 146
REFERENCES 149

LIST OF TABLES

Table		Page
3.1	List of All Experiments and Their Attributes	34
3.2	Summary of Recipient Inputs from All Experiments	47
3.3	Estimating the Number of Recipients Tricked into Entering Their Real SSN Information	50
3.4	Summary of Statistical Hypothesis Testing Results	56
3.5	Summary of Recorded Survey Responses	58
4.1	MDMF Message Sample in the Existing POTS Protocol	75
4.2	Evaluation of Various Standalone Techniques Against the Criteria Described in Section 4.3	88
4.3	Summary of Various Anti-Spam Solutions Using a Combination of Standalone Techniques	90
5.1	List of Extended IAM Parameters	111
6.1	Performance Testing Results of the Prototype Implementation	126
6.2	Summary of Data Collected From Each Participant	137
6.3	Demographics of the Participants	140

LIST OF FIGURES

Figure		Page
1.1	Recent US Government Statistics on Phone Fraud and Call Complaints	2
2.1	Simplified Overview of the PSTN hierarchy	10
2.2	Overview of the Parties Involved in the Transmission of a Call Request	11
2.3	Sequence of Basic Call Control Signaling	12
2.4	Routing of a Spam Call	15
2.5	Flow of Money in the Telephone Spam Ecosystem	15
2.6	Specifications of the Calling Party Number (CPN) Parameter	24
3.1	Procedure of Each Experiment	37
3.2	Incoming Call Screen of Different Experiments	38
5.1	Examples of Security Indicators in HTTP and Email communication	99
5.2	Example of the Proposed Caller ID Security Indicator for an Incoming Call	100
5.3	Overview of the Parties Involved in the Transmission of a Call Request	101
5.4	Overview of the Existing Call Request Transmission Process	103
5.5	Overview of the Proposed Architecture	104
5.6	Sequence Diagram of the Steps to Obtain a Caller ID Certificate	108
5.7	Sequence Diagram of the Steps to Initiate an Authenticated Call Request	110
6.1	Registering a Phone Number and the Public Key With the CA	121
6.2	Receiving the Caller ID Certificate From the CA	122
6.3	Making a Call With an Authenticated Call Request	123
6.4	Tutorial of Security Indicators	124
6.5	Types of Security Indicators Displayed During an Incoming Call	125
6.6	Recruitment Poster	132
6.7	Procedure of Initial Set Up During First Run	133

Figure Page

6.8 Collecting User Feedback After the Display of a Security Indicator 135

6.9 Reviewing Past Security Indicators and Announcements 136

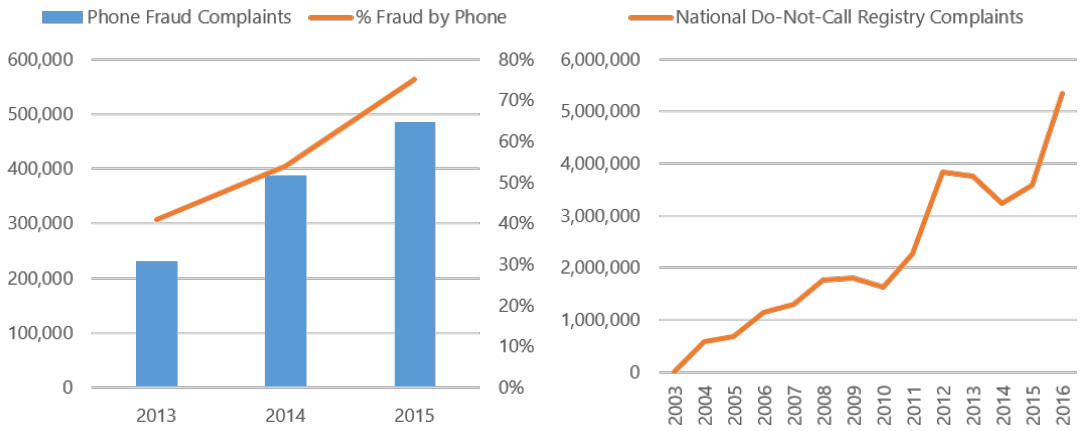
Chapter 1

INTRODUCTION

Since its introduction in 1876, telephone communication is an integral part of modern society and a critical component of our modern infrastructure and economy. The Public Switched Telephone Network (PSTN) ecosystem is at the core of today's telecommunication systems. The national and global telephony system is a critical component of our modern infrastructure and economy. The PSTN is an aggregate of various interconnected telephone networks that adhere to the core standards recommended by the International Telecommunication Union Telecommunication Standardization Sector (ITU-T), allowing most telephones around the world to intercommunicate. The current set of PSTN core standards have been around for several decades, allowing any phone to reach any other phone through a vast worldwide interconnection of switching centers. In the United States, almost every person in the US can be reached with a telephone number as mobile telephone subscribership penetration rate has already surpassed 100% of the country's population [1]. Using 2013 statistics, there are about 335 million mobile telephone subscribers [2], with 136 million fixed-telephone subscribers [3], and 34 million VoIP subscribers [4] in the US with population of 318 million. Each day more than 240 million hours are spent on telephone calls [5], equating to more than 88 trillion hours each year.

However, with the pervasiveness of telephone subscribership, and the introduction of IP access to the Public Switched Telephone Network (PSTN), telephone spam has also become an increasingly prevalent issue. The high reachability of telephone numbers has led to telephony being an attractive spam distribution channel. Spammers are leveraging recent advances in the telephony technology to distribute massive automated spam and scam calls, also known as robocalls.

Today, the PSTN is rife with telephone spam, namely voice, voicemail, and SMS spam. Voice phishing, vishing, or phone fraud has now become a significant and rapidly growing problem in many countries, including the US [6] and UK [7]. Despite various products and services aimed at stopping telephone spam, scam and robocalls, complaints about illegal calls have reached record high levels in recent years. According to recent US government reports, the number of phone fraud complaints in the US more than doubled in just a matter of two years from 2013 to 2015 [6]. During the 2016 fiscal year, the national Do-Not-Call Registry faced a near 50% surge in the number of consumer complaints about unwanted telemarketing calls, and the total number of complaints that year has grown to more than 5.3 million [8]. In the US, more than 75% of the reported fraud and identity theft attempts are now communicated over the phone [6].



(a) Phone Fraud Complaints Each Year Received by the FTC Consumer Sentinel Network

(b) Call Complaints Each Year Received by the National Do-Not-Call Registry

Figure 1.1: Recent US Government Statistics on Phone Fraud and Call Complaints

Spam calls are significant annoyances for telephone users unlike email spam, which can be ignored, spam calls demand immediate attention. When a phone rings, a call recipient generally must decide whether to accept the call and listen to the call. After realizing that

the call contains unwanted information and disconnects from the call, the recipient has already lost time, money (phone bill), and productivity. A study in 2014, it was found that 75% of people listened to over 19 seconds of a robocall message and the vast majority of people, 97%, listen to at least 6 seconds [9]. Even when the recipient ignores or declines the call, today spammers can send a prerecorded audio message directly into the recipient's voicemail inbox [10]. Deleting a spam voicemail wastes even more time, taking at least 6 steps to complete in a typical voicemail system.

Telephone spam is not only a significant annoyance, they also result in great financial loss in the economy, mostly due to scams and identity theft. According to a survey report in 2014, Americans lost more than $8.6 billion due to phone scams [11].

To deal with this issue, governments, including the US [12] and UK [13], have enacted laws to restrict most forms of unwanted telephone calls. Furthermore, some governments have established regulatory agencies and telephone number registries that allow consumers to explicitly opt out of unwanted calls [14, 15]. For decades, despite the legal actions prohibiting robocalling and telephone spamming, complaints on illegal calls have reached record numbers year in many recent years, which indicates that the laws have not deterred the spammers.

In addition to government efforts, there are also consumer and commercial products that are made to defend against unwanted calls. In the consumer market, there are many physical call-blocking devices for landline telephones, and various modern smartphone apps, that can block unwanted calls from offending caller IDs. According to a 2013 consumer poll, 22% of US smartphone users used a call-blocking app or a feature to block calls on their device [16]. Among business and network operators, there is also a supplementary network feature known as MCID (Malicious Call Identification) that allows the destination operator to request identification of the offending calling party [17].

Despite various efforts that reduce telephone spam, scam and robocalls, complaints on illegal calls has been making record numbers in recent years. Many people have reported that these countermeasures do not work and want relief [18]. Clearly, all these countermeasures have so far failed at reducing the growth of telephone spam. Illegal callers today have access to various technologies aimed at circumventing call blockers and evading identification. Among them, a practice known as *caller ID spoofing* is particularly effective at defeating call blockers, evading identification, and furthering a variety of scams.

The key technical component enabling telephone spam, is that, in the current caller ID scheme, the caller ID is trivially spoofed. Falsifying the caller ID enables illegal callers to make their victim believe that they are speaking to someone trusted. Although most modern mobile phones support the capability to block unwanted calls or SMS using caller ID. However, Caller ID spoofing also helps malicious callers to defeat anti-spam defenses that rely on caller ID blacklisting, a malicious caller can easily bypass caller ID blacklisting by spoofing any number not blacklisted. As most telephone spam defenses today (including law enforcement) rely on feedback from the users, caller ID spoofing has made identification and feedback completely irrelevant. At the root cause of this issue, not only has telephone spam become economically viable due to VoIP and autodialers, illegal callers today have access to caller ID spoofing great at circumventing call blockers and evading identification.

Caller ID spoofing is also pervasively used to assist malicious callers to further a wide variety of phishing scams. A malicious caller can spoof the caller ID to make it appear as if a call or SMS originated from a trusted entity, tricking the recipients into divulging sensitive information such as usernames, passwords, credit card details, and take harmful actions such as remitting money to the scammers. Many telephone phishing scams are directly attributable to caller ID spoofing, such as credit card verification scams, IRS tax

agent scams. Some of these phishers even use audio duplicated from the real institution, such as the bank, to trick the recipient into divulging their sensitive information.

Caller ID spoofing is often used in a variety of phone scams. To show an example of how caller ID spoofing is used in phone scams, one type of phone fraud that occurs frequently is the credit card verification scam, where the spammer spoofs the caller ID of a credit card issuing bank [19], and then mimicking or duplicating the audio from the credit card issuer's interactive voice response system to trick his recipients [20]. The audio recording tells the victims that their credit cards have been deactivated due to fraud, and is in urgent need of verifying their personal information to reactivate their account. The true motive of this scam is to steal the recipients' credit card and personal information.

Furthermore, caller ID spoofing can also frame true owners of spoofed caller IDs with illegal behavior. A malicious caller could spoof a known number to commit crimes, such as making phishing calls, making fake purchase orders, or sending police to a person's address for harassment [21]. As a result, true owners of spoofed caller IDs could also end up in trouble. For organizations, severe brand damage can come as a result of caller ID spoofing.

Because of the prevalence of caller ID spoofing, it has also led to many become overly suspicious of phone calls, even when it is for communicating legitimate and critical information. In a recent revelation about Russian cyberattacks on the Democratic National Committee (DNC), the DNC blew off FBI's repeated hack warnings because the worker could not differentiate a real FBI agent call from an impostor [22]. Because of caller ID spoofing, many users couldn't tell apart the real important calls from the malicious calls.

In the US and many other regions, the telephone number follows a numbering format that identifies the region code, central office code, and subscriber number [23]. If the telephone number is spoofed, the government and law enforcement agencies would typically

lose key information that could trace the source of the malicious caller. This has led to increasingly difficult legal enforcement against illegal telephone callers.

To make matters worse, with the increasing availability of VoIP-to-PSTN services, a spammer may distribute outbound calls from an overseas location, beyond the jurisdiction of law enforcement. Furthermore, the Internet provides plenty of opportunities for the malicious caller to hide his true location, such as with proxy, VPN or TOR. Due to the lack of legal repercussions, foreign-operated phone scams have become increasingly common.

This situation requires an effective solution, given the significant gains made in reducing email spam, this raises the question: *are there email spam solutions that could also be used to stop telephone spam?* Unfortunately, this issue is not easily solved, and, in fact, most of the simple and effective techniques against email spam cannot be applied to telephone systems. There are significant differences and unique challenges in the telephone ecosystem that require novel approaches. Many existing solutions have failed to overcome these challenges and, as a result, have yet to be widely adopted.

1.1 Dissertation Outline

The rest of this proposal is organized as the following: Chapter 2 will first describe the background information on the public switched telephone network, how the telephone spam operation works, and how the caller ID is transmitted; Chapter 3 will present a study on understanding how telephone scam works and why recipients fall victim to them; Chapter 4 will elaborate on the key challenges in dealing with telephone spam and present a survey of the existing countermeasures and techniques; Chapter 5 will propose a novel architecture and method to authenticate the caller ID which facilitates security indication of the caller ID transmission; Chapter 6 will present the prototype implementations of the proposed authenticated caller ID transmission scheme and present the results and evaluations

of its performance and the end-user experience; Chapter 7 will summarize the contributions and conclude this dissertation.

Chapter 2

BACKGROUND

2.1 Public Switched Telephone Network

The public switched telephone network (PSTN) [24] is an aggregation of the world's telecommunications networks that is designed for continuous real-time voice communications. The network consists of national, regional, and local telephony operators, which are connected together form the PSTN. At the core of the PSTN are switching centers, and switches can be physically interconnected through fiber optic cables, microwave transmission links, communications satellites, and submarine telephone cables, etc.

Historically, the network works by establishing a direct connection between any two points to carry analog signals modulated to voice frequencies, i.e. circuit-switching. Today, the core PSTN has evolved to carry almost entirely digital signals over fiber optic cables, sharing bandwidth with various services including Internet and TV, due to its much greater capacity. Despite the digital advancements, the core principles of the PSTN remain focused on being a circuit-based or connection-oriented system designed for the reliable delivery of voice. Compared to the Internet, the PSTN differentiates for its ability to provide highly reliable and low latency voice communications to its subscribers.

The interconnected public switched telephone network adheres to the core Signaling System No. 7 (SS7) standards [25] created by the International Telecommunication Union Telecommunication Standardization Sector (ITU-T), allowing most telephones around the world to intercommunicate. All telephone terminals in the PSTN use the E.163 and E.164 telephone numbering scheme created by the ITU-T to reference different end-point ad-

dresses. The combination of ITU-T core standards is what allows telephones around the world to communicate with each other seamlessly and reliably.

Due to the much greater capacity of IP infrastructure and the wide availability of IP-based equipment, some telephony service providers have shifted their network infrastructure to IP-based solutions, and the operation cost of the telephone network has dramatically decreased. While the core PSTN infrastructure has evolved to be almost entirely IP-based, the core signaling protocols have not changed. The entire ecosystem still relies on the three-decade-old Signaling System No. 7 (SS7) [26] suite of protocols, allowing any phone to reach any other phone through a worldwide interconnection of switching centers. Today, the cost of a telephone call via the internet is a tiny fraction of what it used to cost for traditional landlines.

The PSTN network topology is arranged in a hierarchical fashion, and this allows the network operators to efficiently interconnect and bill for the calls. Each step of the call path is routed and managed by the network operator and its switching systems. In the United States and many other regions, the telephony operators generally fall into two categories: *Interexchange Carrier* and *Termination Carrier*.

Interexchange Carrier (IXC), also known as a long distance carrier, is a cross-regional carrier that carries call traffic between telephone exchanges over long distances. They are points of high traffic aggregation and they cover larger geographical distances. High-speed transport links (such as fiber optic cables) are typically used between transit switches. IXCs typically own or share the various high-bandwidth, fiber-optic trunk lines that span across the country, and some also provide high-speed switched digital services for data and multimedia communication.

Termination Carrier, also known as local exchange carrier, is a carrier that provides call routing services within a local network that terminates at its end users. The termination carrier may be operating a landline, mobile, or IP-based telephone network. A termination

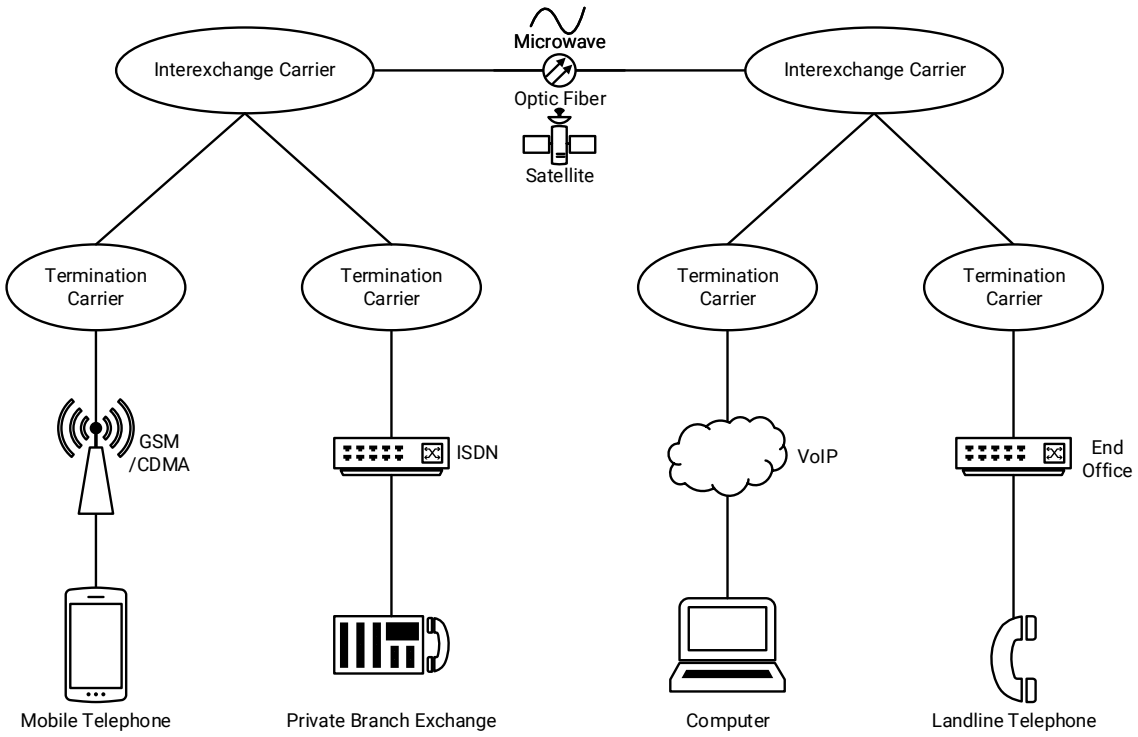


Figure 2.1: Simplified Overview of the PSTN hierarchy

carrier is also presubscribed to one or more IXCs to provide calls originating or terminating outside its network. Most consumers and businesses rely on termination carriers for their telephone subscriber services. The termination carrier usually has full control over their own local network, allowing them to define a vertical stack of equipment and protocols to be used within the local exchange network.

2.1.1 PSTN Call Setup Signaling

Before we discuss the technical detail of the underlying call setup process of establishing a telephone call, we first present an overview of the parties involved in the transmission of a call request.

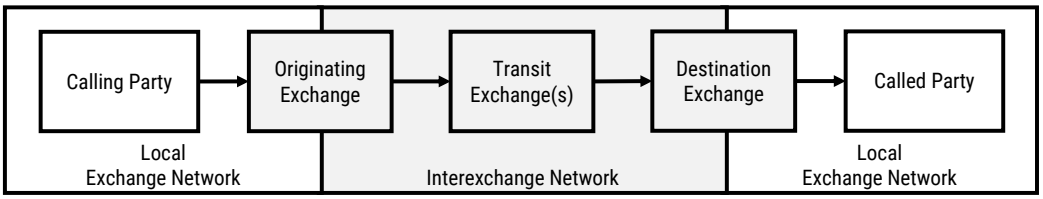


Figure 2.2: Overview of the Parties Involved in the Transmission of a Call Request

Calling Party is the party initiating the call request with a user equipment (UE), such as a mobile phone, desk phone, or software client that connects with the originating exchange.

Originating Exchange is a switch in the PSTN that generates and transmits the call request to the destination exchange pertaining to the call request from the calling party.

Transit Exchange is an interconnecting switch in the PSTN that helps to route the call request from the originating exchange to the destination exchange.

Destination Exchange is the terminating switch in the PSTN that receives the IAM and sets up the ring with the called party.

Called Party is the party with a user equipment or software client of the intended called party for the call request.

Also in Figure 5.3, we can see that the overall network can be divided into two categories: 1) *Local Exchange Network*, and 2) *Interexchange Network*.

In general, the sequences within a local exchange network define how user equipment interacts with the local exchange carrier during a call setup, and the sequences within the interexchange network define how SS7 switches interact with each other during a call setup. More details of basic call control and signaling procedures can be found in Q.764.2 [27].

The SS7 process of setting up a telephone call (i.e. dialing and connecting to phone number) is summarized in Figure 2.3.

As we can see from Figure 2.3, when requesting a telephone call, the calling party first initiates a dialing or call setup process with Originating Exchange. The specifics of this process varies between local exchange networks, there are currently more than 8 different

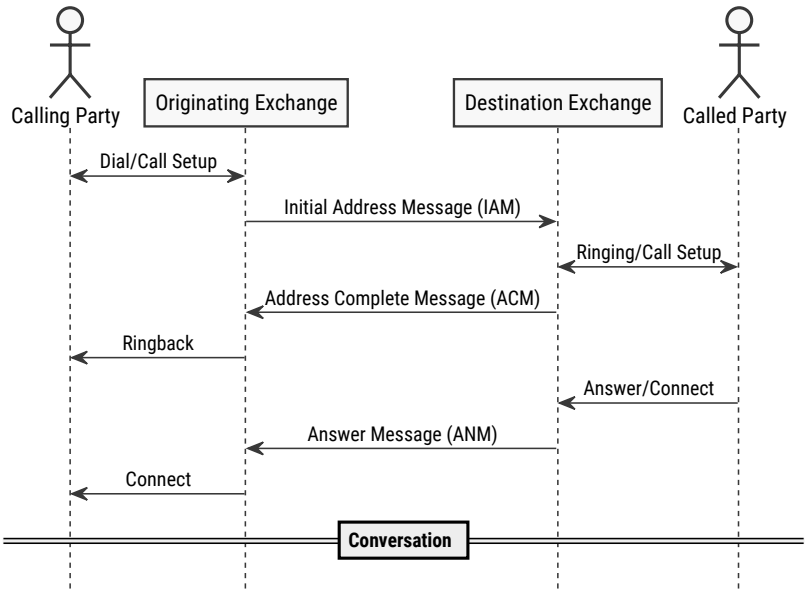


Figure 2.3: Sequence of Basic Call Control Signaling

standards used to set up a call within a local exchange network, with each standard defining its own call setup sequence: GSM [28], CDMA [29], 3G UTMS [30], SIP [31], H.323 [32], IMS [33], VoLTE [34], V5.2 [35], and ISUP [36]. Each type of local exchange network has it own setup sequence which can be extremely complicated, hence we will not explain the details of each. However, the general idea is that the local exchange network acts like a proxy to establish the PSTN call request to the destination exchange from the originating exchange on behalf of the calling party.

After completing the call setup process within the local exchange network, the originating exchange generates an initial address message (IAM) and transmits it to the destination exchange. The initial address message in principle contains all the information that is required to route the call to the destination exchange and connect the call to the called party. The IAM is a type of ISDN user part message which follows the ISUP (ISDN User Part) message format as defined in Q.763 [37]. Typically, the IAM must contain information such as, nature of connection indicators (Q.763.3.35), forward call indicators (Q.763.3.32), call-

ing party's category (Q.763.3.11), transmission medium requirement (Q.763.3.54), called party number (Q.763.3.9) [37]. In addition, to allow the transmission of the caller ID number (the calling party's number) to the called party, today almost all originating exchange also includes a calling party number (Q.763.3.10) parameter in the IAM.

After receiving the IAM from the originating exchange, assuming the IAM is valid, the destination exchange generates a call setup process with the called party. During this process, the called party's user equipment or software client would get a call setup request from the destination exchange and start ringing. If the calling party's number is included in the IAM, the destination exchange would also transmit this information to the called party which allows display of the caller ID number during the ring. If the called party number is valid, the destination exchange also sends back an Address Complete Message (ACM) to indicate the successful reception of IAM.

Typically, the ACM must include backward call indicators (Q.763.3.5) which contains information such as charge indicator (indicates whether a call should be charged), called party's status indicator (indicates whether the subscriber is free), called party's category indicator (indicates the general category of the called party, such as a payphone), etc.

After receiving the address complete message at the originating exchange. The calling party receives a ringback tone, which provides an audible indication that the called party is ringing.

When the called party answers the ringing phone, the destination exchange generates an Answer Message (ANM) and transmits it to the originating exchange. The ANM does not contain any mandatory fields other than the message type. After receiving the ANM, the originating exchange and destination exchange will set up a connection and conversation between the two parties is established.

2.2 How Telephone Spam Works

We define *telephone spam* as the mass distribution of unwanted content to modern telephones in the PSTN, which includes *voice spam* that distributes unwanted voice content to answered phones, and *voicemail spam* that distributes unwanted voice content into the recipient's voicemail inbox. While email spam is arguably the most well-known form of spam, telephone spam has now become more prevalent than ever.

A very common way of disseminating telephone spam is *robocalling*, which uses an autodialer that automatically dials and delivers voice or voicemail messages to a list of phone numbers. An *autodialer* is a generic term for any computer program or device that can automatically initiate calls to telephone recipients. Today, an autodialer is usually a computer program with Voice over Internet Protocol (VoIP) connectivity to a high volume VoIP-to-PSTN carrier, that may include features such as voicemail and SMS delivery, customizable caller ID, Call Progress Analysis, scheduled broadcast, text-to-speech, Interactive Voice Response, etc.

Key Players of Telephone Spam

To understand the telephone spam ecosystem, we will first identify and explain the roles of all players who take part in the routing of a telephone spam. Figure 2.4 shows a graphical depiction of the routing process: The spammer connects through the Internet to an *Internet Telephony Service Provider*, then the call is routed through an *Interexchange Carrier*, before finally being accepted by the *Termination Carrier*, who then routes the call to the victim.

Another way to understand the ecosystem is to show how money flows through the system, which we display in Figure 2.5: the money flows from the victim to the spammer, and the spammer uses this money to obtain leads (new phone numbers to spam) and to

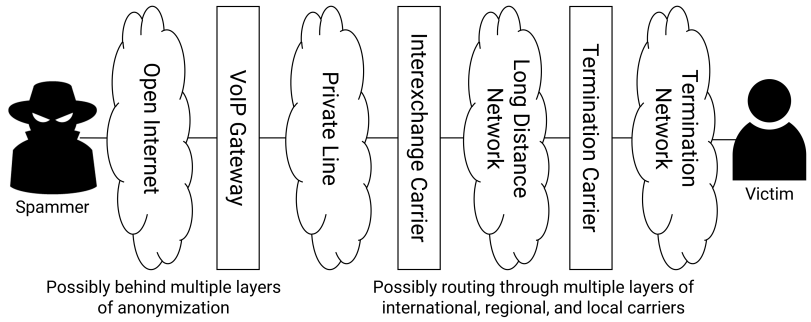


Figure 2.4: Routing of a Spam Call

pay for the spam calls, the *Internet Telephony Service Provider* receives the money from the spammer and pays the *Interexchange Carrier*, who then pays the *Termination Carrier*. Next, we examine each of these roles in turn.

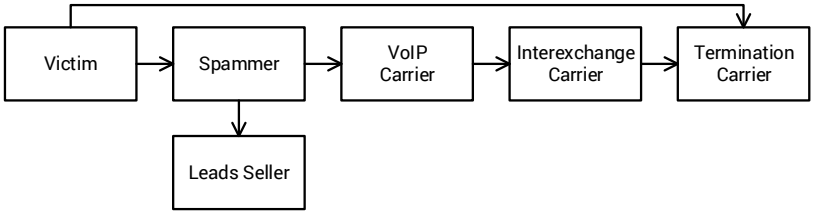


Figure 2.5: Flow of Money in the Telephone Spam Ecosystem

Spammer is the agent that carries out the spamming operation. The spammer could be part of an organization or an independent contractor that offers spamming-as-a-service. The goal of the spammer is usually to extract money from victims through sales and scams. For cost efficiency, spam calls are typically initiated using an autodialer connected to an *Internet Telephony Service Provider* to reach the PSTN victims. We will describe the spammer's operation in more detail in Section 2.2.1.

VoIP carrier, also known as an **Internet Telephony Service Provider** (ITSP), is a type of termination carrier that offers telecommunications service over the TCP/IP network, i.e. the Internet. The ITSP typically offers high volume calling at a lower cost compared to traditional carriers and generates revenue based on the minutes of calls hosted. Whenever

the spammer makes an outbound call to a PSTN number, the ITSP will convert the signaling protocol from VoIP to SS7 and route the converted signal through an interexchange carrier.

Interexchange Carrier (IXC), also known as a long distance carrier, is a cross-regional carrier that carries call traffic between telephone exchanges over long distances. The IXC charges its subscribers (mainly termination carrier such as the ITSPs and local mobile/land-line carriers) for handling long distance phone calls and compensates the next-hop carrier (such as the recipient's termination carrier) for access. Unlike the peering model between Internet service providers [38], the IXC negotiates access rates with other carriers, known as intercarrier compensation. In the US, intercarrier compensation [39] is a complex system in which the rates vary according to traffic origination, location, carrier, and traffic type, and the rates are governed by federal and state regulators. In general, when two carriers are directly connected, the originating carrier compensates the next-hop carrier for routing the call in the next-hop carrier's network.

Termination Carrier, also known as local exchange carrier, is a carrier that provides call routing services within a local network that terminates at its end users. The termination carrier may be operating a landline, mobile, or IP-based telephone network. Most consumers and businesses rely on termination carriers for their telecommunications services. The termination carrier typically bills the IXC for the amount of incoming traffic, known as the access charge. In the US and some other countries, the recipient subscriber may also be partially billed for incoming calls.

2.2.1 Spammer Operation

Spamming (regardless of the medium) requires three basic elements: a recipient list, content, and a cost-effective mass distribution channel. In addition, a more sophisticated spammer may employ circumvention measures to defeat spam countermeasures, and to avoid being stopped by law enforcement agencies.

Telephone Numbers

Spamming first requires a list of potential victims to contact, and in the case of telephone spam: a list of phone numbers. While there are many ways a spammer could gather phone numbers, the simplest method is to purchase the numbers from a leads seller. We did a simple Google search (keyword "leads for sale") and found hundreds of websites that offers access to millions of curated phone numbers for less than $100. There are also other ways to harvest phone numbers, such as crawling the web, collecting form submissions, downloading leak databases, covertly gathering through smartphone apps, or simply generating the numbers based on phone numbering plans. However, we do not know for sure the most popular means of obtaining a list of phone numbers for spamming, due to the lack of existing studies. Once the spammer gathers a list of phone numbers, the spammer can load it in an autodialer for mass distribution of the content.

Voice Content

The content of telephone spam is typically a prerecorded audio stream made by either recording human voice or by using a text-to-speech synthesizer program. Telephone spam can also deliver interactive voice content, with the use of an Interactive Voice Response (IVR) system. When the recipient answers a call from an autodialer with interactive content, the recipient can interact with the system through voice and keypad inputs, and an automated voice message is played back based on the interaction.

There are a wide variety of spam types, such as telemarketing, impersonation scam, debt collection, political campaigns, one-ring scam, and so on. To provide insight into the telephone spam content, we collected 100+ audio samples from various publicly available sources where audio recordings of voice or voicemail spam are uploaded. We perform this analysis to gain a general understanding of voice and voicemail spam, and we emphasize

that, due to the biased method of data collection, these results do not constitute measurements that reflect trends on the whole of voice and voicemail spam. However, these results provide needed background and insight into actual voice and voicemail spam. We will describe the following prevalent types of spam: credit card verification scam, fake tax agent scam, and political robocalls.

In the *credit card verification scam* samples, the called recipients are informed that their credit card account was deactivated, and they are asked to enter their credit card and social security number over the phone to verify their identity and get the account reactivated. While we were only able to listen to the audio of the call, based on comments from some of the uploaders, the scammers would spoof the caller ID to make it look as if the call originated from the credit card issuer. All of these scam calls used an Interactive Voice Response system to interact with the recipients and collect their credit card information. We found that the audio from the scammer's IVR system came from either a synthesized voice or *audio duplicated from the IVR system of the real credit card issuer*. From what we observed, the use of caller ID spoofing and sound duplicated from the real credit card issuer's IVR system made it almost indistinguishable from a real credit card verification call.

In the *fake tax agent scam* samples, the recipient receives a call from the scammer identifying himself as a tax agent of the Internal Revenue Service (IRS) and provides a fake badge number. The scammer proceeds to tell the recipient that he or she owes a specific amount of money to the IRS. Often, the scammers demand immediate payment and threaten jail, deportation, or loss of driver's license if the victim does not pay. Based on the comments from the uploaders, the scammers would spoof their caller ID to make it look as if the call originated from a government agency by showing an area code from 202 (Washington, DC). These scammers seem to target immigrants [40]. We found that the majority used a live person to interact with the victim, and the rest used a prerecorded

synthesized voice without an IVR system. One thing we noted was that all of the live person scammers had a South Asian accent, and in our opinion, the accent had made the call sound highly suspicious and easy to recognize as a scam (which might explain why it was posted online as a scam).

In the *political robocall* samples, the typical content is a prerecorded message making a political advertisement, or a poll asking the recipient about their political opinion. In the United States, political robocalls are exempt from regulation by the national Do-Not-Call Registry and the Telephone Consumer Protection Act of 1991. Before a national or state level election, they are distributed in high frequency using voice and voicemail broadcasting autodialers. All of the audio samples contained a prerecorded message, and most polls used an IVR system to interact with the recipient.

Mass Distribution

Mass distribution is the next critical step to a successful spam operation. The goal is to massively and cost-effectively deliver the spam content to a list of telephone numbers.

Using VoIP service to distribute calls to PSTN numbers, the content can be disseminated at a much higher volume, and at a fraction of the cost compared to traditional telephony. To understand the distribution cost of spamming, we researched the prices and found hundreds of VoIP service providers offering pay-by-the-minute calling service to US telephone numbers priced around $0.01 per minute. We also found some fixed monthly-fee pricing model with unlimited calling for about $150, however, these service providers tend to target small businesses, and these plans usually come with throttling, so high volume calling services are almost always offered with a pay-by-the-minute model.

Some VoIP service providers (such as CallFire [1] and Call-Em-All [2]) even cater specifically to telemarketers, providing features such as integrated autodialer and customizable caller ID in their service.

Circumvention

Spamming is an adversarial game, as spam defenses are widely introduced, the spammer has an incentive to defeat them. According to a poll conducted by Harris Poll on behalf of WhitePages in 2013, 22% of US smartphone users used a call-blocking app or a feature to block calls on their device [16]. Most mobile phones today has the basic capability to automatically block calls from a list of unwanted callers.

For the spammers today, two common ways to defeat them is to use *voicemail injection* and *caller ID spoofing*.

Voicemail injection is a recent extension of the autodialer which delivers prerecorded voice messages into the recipients' voice mailbox (voicemail). Typically, when a phone call is unanswered or declined, it gets forwarded to an answering machine that lets the caller leave a voice message. A voicemail broadcasting autodialer uses Answering Machine Detection (AMD) [41] technology to automatically complete the process of inserting a prerecorded voice message into the recipient's voicemail. A more recent type of voicemail broadcaster can even deliberately trigger the recipient's voicemail, a technique known as Forced Busy Channel [10], to directly inject a voice message into the recipient's voicemail without waiting for the call to be unanswered or declined.

Caller ID spoofing is the practice of deliberately falsifying the caller ID information sent to the recipient that identifies the caller of a phone call. It is particularly effective for defeating the call blockers and helps to further a variety of scams. The caller ID service

[1] https://www.callfire.com/

[2] https://www.call-em-all.com/

provides the caller's telephone number (and in some cases the caller's name) to the recipient before or during the ring of an incoming call. It allows the recipient to decide whether to answer a call based on the caller ID information, or to call back if the call could not be answered. The caller ID number is also widely used in other non-voice communication services, such as SMS, MMS, and many smartphone apps. The caller ID number is typically provided by the caller's switch, which can control what caller ID number is sent on a call-by-call basis. For general consumers, a legally mandated privacy feature allows them to hide the calling number [42]. However, malicious callers can also take advantage of the declarative nature of the caller ID mechanism to spoof or block the caller ID number, in order to defeat spam filters and further a variety of scams. The caller ID number can be easily spoofed because there is no built-in authentication mechanism, and it is not immediately verifiable by the recipient. The caller's service provider does not have any legal obligation to ensure that the caller ID number in the call request header is indeed owned by the caller before it is transmitted. In fact, some ITSPs today advertise customizable caller ID as a service feature.

2.2.2 Telephone Spam Samples

We collected 100+ audio samples from various publicly available Internet sources such as SoundCloud, YouTube, and LiveLeak, where audio recordings of voice or voicemail are voluntarily uploaded by individuals and corporations.

In the *credit card verification scam* samples, the called recipients are informed that their credit card account was deactivated, and they are asked to enter their credit card and social security number over the phone to verify their identity and get the account reactivated. While we only were able to listen to the audio of the call, based on comments from some of the uploaders, the scammers would spoof the caller ID to make it look as if the call originated from the credit card issuer. All of these scam calls used an Interactive Voice

Response system to interact with the recipients and collect their credit card information. We found that the audio from the scammer's IVR system came from either a synthesized voice or *audio duplicated from the IVR system of the real credit card issuer*. From what we observed, the use of caller ID spoofing and sound duplicated from the real credit card issuer's IVR system made it almost indistinguishable from a real credit card verification call.

In the *fake tax agent scam* samples, the recipient receives a call from the scammer identifying himself as a tax agent of the Internal Revenue Service (IRS) and provides a fake badge number. The scammer proceeds to tell the recipient that he or she owes a specific amount of money to the IRS. Often, the scammers demand immediate payment and threaten jail, deportation, or loss of driver's license if the victim does not pay. Based on the comments from the uploaders, the scammers would spoof their caller ID to make it look as if the call originated from a government agency by showing an area code from 202 (Washington, DC). These scammers seem to target immigrants [40]. We found that the majority used a live person to interact with the victim, and the rest used a prerecorded synthesized voice without an IVR system. One thing we noted was that all of the live person scammers had a South Asian accent, and in our opinion, the accent had made the call sound highly suspicious and easy to recognize as a scam (which might explain why it was posted online as a scam).

In the *political robocall* samples, the typical content is a prerecorded message making a political advertisement, or a poll asking the recipient about their political opinion. In the United States, political robocalls are exempt from regulation by the national Do-Not-Call Registry and the Telephone Consumer Protection Act of 1991. Before a national or state level election, they are distributed in high frequency using voice and voicemail broadcasting autodialers. All of the audio samples contained a prerecorded message, and most polls used an IVR system to interact with the recipient.

2.3 Overview of the Caller ID

Since its introduction in the 1990s, caller ID service has now become ubiquitous in almost every form of telephone service. The caller ID is a generic name for a supplementary service offered by the called party's telephone company that presents the calling party's telephone number to the called party's user equipment during an incoming call. It helps the called party to decide whether to answer a call based on the caller's phone number, and, to call back the caller if the call could not be answered. Today, the caller ID is also used in other telephone services, such as the SMS and MMS, and, with the prevalence of smartphones, many apps and services also rely on caller ID for identification.

The core process of providing the caller ID is known as Calling Line Identification Presentation (CLIP), which was first defined in ITU-T Recommendation Q.731.3 [43] for the Signaling System No. 7 (SS7) network in 1993. The SS7 network is the backbone infrastructure for most of the world's public switched telephone network (PSTN) telephone calls. Even as the telephone backbone moves towards being carried by an IP packet-based infrastructure, Q.731.3 still plays a major role in providing the caller ID for telecommunications and will continue to do so for many years to come.

In all major existing call signaling protocols (SS7, H.323, and SIP), caller ID is either provided by the originating exchange or by the calling party. In SS7 and SIGTRAN (IP version of SS7), caller ID is defined by the calling party number (CPN) parameter, where the parameter is an optional part of the Initial Address Message (IAM). The IAM is sent to the destination exchange as part of the basic call procedures according to Q.764 [27] to initiate a call. The IAM routes through transit exchange switches until it reaches the destination exchange of the called party, in which the called party's local exchange carrier would convert and retransmit the CPN to a specific caller ID format for the called party's user equipment during the incoming call setup process, e.g. mobile or landline.

Recommendation Q.731

STAGE 3 DESCRIPTION FOR NUMBER IDENTIFICATION SUPPLEMENTARY SERVICES USING SIGNALLING SYSTEM No. 7

	8	7	6	5	4	3	2	1
1	O/E	Nature of address indicator						
2	NI	Numbering plan indicator			Address presentation restricted indicator		Screening indicator	
3	2nd address signal				1st address signal			
:								
:								
m	Filler (if necessary)				*n*th address signal			

Figure 11/Q.763 – Calling party number parameter field

Figure 2.6: Specifications of the Calling Party Number (CPN) Parameter

2.3.1 How Caller ID Spoofing Works

In the process of providing the caller ID, the originating exchange can control what caller ID number is sent on a call-by-call basis. As the PSTN (public switched telephone network) is traditionally regarded as a closed network of SS7 exchange switches between trusted operators, usually only an SS7 switch operator or a private branch exchange (PBX) owner has the capability to customize the caller ID. Since it was prohibitively expensive for individuals and small businesses to gain switch level access to the SS7 network, in most telephone services, their caller IDs are typically managed by the caller's telephone carrier.

However, with growing access to the PSTN from the Internet, there are now many internet telephone service providers (ITSPs) that provide telephone services over an Internet connection. With ITSPs, individuals and businesses are no longer limited to telephone services from their local telephone service providers. With an Internet connection, a malicious caller now has access to a world of ITSPs that can provide features such as caller ID customization/spoofing.

To spoof the caller ID, the caller's originating exchange would declare the CPN parameter with false information. In the US and many other jurisdictions, the caller's telephone service provider does not have any legal obligation to ensure that the caller ID is verified before it is transmitted. Even in jurisdictions that forbid telephone service providers from providing falsely declared caller IDs, with Internet access to an untrustworthy telephone service provider, it is easy for a malicious caller to start the call request from a different origin, and transmit the fake caller ID.

At the heart of the issue, there is a lack of authenticity and accountability in the transmission of telephone identities. The PSTN has transformed from a closed trusted network to a diverse global ecosystem, mutual trust can no longer be relied upon to guard against the abuses of trust in the caller ID transmission. Addressing this issue requires the core protocol to provide a mechanism to ensure authenticity of the caller ID. This is why we advocate for a standardized caller ID authentication scheme. By providing authentication to the caller ID, accountability of the caller ID can be enforced. However, for viable deployment of authenticated caller ID transmission, it requires mutual interoperability. Therefore, design a scheme for the core PSTN protocol is the key to transforming the telephone ecosystem into a community that could finally rely on the authenticity and accountability of caller IDs.

Chapter 3

UNDERSTANDING WHY TELEPHONE SCAMS WORK

The rise of telephone spam, scam, fraud, phishing or vishing, is a significant and growing problem. According to FTC reports, the national Do-Not-Call Registry received more than 5.3 million unwanted call complaints in 2016, a near 50% surge from 2015 [8]. Phone fraud complaints have also more than doubled from 2013 to 2015, and more than 75% of the reported frauds are now attempted over the phone. Impersonation scams were the top fraud [6], these impostors called the recipients while pretending to be anyone from the IRS to a family member in trouble, from tech support to a business partner that turned out to be bogus.

With the growing dissatisfaction of telephone scams, however, little research has been done to study why people fall for telephone scams. Although telephone phishing has been in the news and has continued to exist for several decades, no past study has specifically measured what contributes to the effectiveness telephone scam. In this chapter, we present the results of an empirical telephone phishing study, designed to systematically measure the different attributes in relation to the success rate of telephone scams. Although the current understanding of telephone scams might be accepted as conventional wisdom, none had specifically validated such claims with hard empirical evidence. From this study, we hope to dispel some myths about what is "scammy" and what is not. The knowledge obtained from the study has scientific merits in learning why some telephone scams work, and why some would not. With the understanding of the key attributes that make a scam convincing, we can focus on developing prevention methods to challenge the fundamentals of telephone phishing attacks. By shedding light on what causes the recipients to fall for

or be suspicious of telephone scams, we can help to educate the public the key factors that make them vulnerable these types of attacks.

The main contributions of this chapter are the following:

- We taxonomize telephone phishing as a type of social engineering attack where it is formulated with a set of visual and voice attributes.
- We describe a systematic approach to test the significance of each attribute and conduct an empirical study using the approach.
- We describe a set of analysis criteria to evaluate the results and used the criteria to analyze the results of the empirical study.
- We present our evaluation of the study and provide our recommendations for solving the telephone phishing problem.

This chapter contains work currently under review in the IEEE Symposium on Security and Privacy 2018.

3.1 Background and Insights

With the emergence of distribution technology, economical cost, high reachability, and computerized automation, the telephone has become an attractive medium for disseminating unsolicited information. The use of the telephone for phishing has become increasingly popular. According to the FTC, the telephone is the leading channel (>75%) of all communication channels for reported fraud attempts [44]. As with any form of spam, it ultimately boils down to three key ingredients: (1) the recipient list, (2) the content, and (3) the distribution channel, as Tu et al. point out [45]. With modern technology and services, the telephone has become a cost-effective channel to massively deliver content. Telephone scam relies on distributing *deceitful* voice content, whereas telephone spam or telemarketing primarily distributes marketing and advertising content. In telephone scam, fraud,

phishing or vishing, the goal of the voice content is trick the human victim into performing harmful actions for the benefit of the attacker. Hence, telephone phishing falls into the category of social engineering attack where the goal is to exploit human vulnerabilities.

Compared to other forms of phishing, such as email and website phishing, it differentiates by having the potential to make the scam more convincing by falsifying both visual and auditory perceptions to induce the victims into falling for the scam. Visually, the scam can be made more convincing by altering the caller ID, such as by spoofing the caller ID, manipulating the area code, and impersonating a familiar contact name. With the characteristic of demanding immediate decision upon an incoming call, a scammer can apply visual trickeries to make the recipient answer the phone. Once the scammer has successfully tricked the recipient into answering the call, the attacker then moves to using deceitful voice content to exploit the human recipient. Within the voice content, an attacker can spoof, duplicate or make the speech sound like it is from a known organization or a familiar personal contact. To provide a motivation for the recipient to divulge confidential or personal information, the scammer presents a scenario that satisfies the human need for a justification to perform a (harmful) action.

Hence, by looking at telephone phishing from a perspective that can be characterized by the visual and voice attributes which it embodies, a systematic approach can be used to study and understand why some scams work better than others. As many are now undertaking the crusade of curbing the telephone spam problem, understanding why telephone phishing works can help us design solutions that challenge the core foundations of telephone scam.

3.2 Study Design

The goal of the study is to design a systematic approach that can reveal the factors in telephone scams which make them effective. Our approach to designing the study is to first

identify the attributes that could lead to an effective telephone phishing scam. After that, design a set of experiments and procedures that allow comparisons of different variations of an attribute. The approach will then conducted with an empirical study. Each experiment will follow a standardized procedure and the procedure will be conducted on each group simultaneously. We will tabulate the results and perform analysis on the results. Finally, we provide a discussion on what could be learned from the analysis and provide our recommended solutions for solving the telephone phishing problem. The study was conducted with IRB approval.

3.2.1 Attributes

To identify the telephone scam attributes, we gathered and reviewed more than 100 existing real-world telephone scam samples from various Internet sources, including the FTC website, IRS website, news websites, YouTube, SoundCloud, user comments, and industry surveys. While reviewing the scams, we looked for properties that can be grouped into separate classes and then systematically categorized them into attributes that distinct from each other.

The attributes identified in our study are as follows:

Area Code

In North America, the area code is the first three digits on the caller ID. The area code specifies the geographic location associated with the caller's phone number, e.g. 202 is Washington, DC. In addition, a toll-free phone number is also identified by the three-digit prefix similar to a geographic area code, e.g. 800, 888, 877, etc. According to reports of real-world IRS impersonation scams [46, 47], many scammers appeared to have either spoofed or obtained a 202 area code or toll-free area code on their caller IDs to make it appear as if the IRS is calling. It seems that the area codes could be one of the attributes

that make the scam more convincing. To test this hypothesis, in our experiments varied the caller ID area code between—202 (Washington, DC), 800 (Toll-free), and 480 (the area code of the University Location).

Caller Name

Today, most telephone terminals have the capability of associating a name with a telephone number. With a stored contact, an incoming call from the stored contact would show the name associated with the caller ID, helping the recipient to identify the caller. To perform a spear phishing attack, a malicious caller could spoof the caller ID of a known stored contact, such as impersonating the HR department of a company, to scam other victims in the organization. This is definitely an attribute that could have important implications. For legal and IRB approval reasons, we could not actually spoof a known caller name. Instead, we asked our telephone service department to temporarily create a new contact in the university's internal phone directory and associated a legitimate sounding name with the telephone number. We used that telephone number in our scam experiments to produce a similar effect to caller name spoofing.

Voice Production

According to reports of real-world telephone scams, some used a robotic (synthesized) voice, while others used a human voice to communicate with the recipient [47, 48]. To test the effect of synthesized voices vs. human voices, we downloaded recordings of existing telephone scams, listened to them, and recreated the scams using a text-to-speech synthesizer to generate a speech similar to the real-world scams. To mimic the human voice version of the scams, we also recorded human voices speaking the exact same announcement message.

Gender

From listening to recordings of actual telephone scams, some used a male voice and some used a female voice. Perhaps the voice gender is an attribute that could make the scam more convincing. To test if the vocal gender of the voice could have an effect on the telephone scam, we varied the voice gender between male and female in the text-to-speech synthesizer.

Accent

From the reports of telephone scams, some spoke with an Indian accent and some others spoke with an American accent. It seems that recipients would be more wary of scams that spoke in a foreign accent, and would be less suspicious of scams that spoke in an American accent. To test if this could have an effect on the telephone scam, we varied the voice accent between Indian and American in our experiments.

Entity

From the gathering of real-world telephone scams, two types of scams stood out in term of the number of reports: IRS impersonation scams [49] and HR impersonation scams [50]. In these scams, the scammer claimed to be from the IRS or the company's HR department. While the IRS scams can affect any taxpayer in the US, the HR scams are usually targeted toward people in a specific company. It seems that a more targeted attack would have more success. To test if this attribute works, we varied the impersonated entity of our scams between the IRS and ASU's HR department. To simulate the real-world HR scams as closely as possible, we initially wanted to impersonate our university's HR department, however, our HR department had strong objections about using their name to conduct scam

experiments. As a compromise, our experiments claimed to be from a bogus but legitimate-sounding HR-like department called the ASU "W-2 Administration".

Scenario

Every scam has to provide a motivation for the targeted individuals to perform a certain (harmful) action. Real-world telephone scams created various scenarios to motivate their victims, such as tax lawsuits, payroll issues, credit card verification, etc. The type of motivation are generally either fear-based (such as lawsuits, audits, jail time, losses, etc.) and reward-based (prizes, money, gifts, etc.). It seems that the recipients might be more motivated by fear than reward due to loss aversion. In decision theory, loss aversion refers to people's tendency to prefer avoiding losses to acquiring equivalent gains. Whichever the scenario, it has to be related to the entity that the scammer is impersonating. In our empirical study, we crafted a fear-based and a reward-based scenario related to each entity. These scenarios were inspired by real-world IRS scams and HR scams. To test each type of scenario, our message announcements varied between Tax Lawsuit (IRS), Unclaimed Tax Return (IRS), Payroll Withheld (HR), and Bonus Issued (HR).

3.2.2 Experiments

To test these attributes, we designed the experiments such that variations of each attribute can be compared under similar environmental conditions. To study the effectiveness of two attribute variations, e.g. A and B, we would split the testing comparing two variations of the attribute to see which one performs better. Both variations of the experiment will be conducted simultaneously under the same controlled environment. At its core, the experiments are designed similar to the concept of A/B testing.

When performing experiments under same environmental conditions, one of the design issues is to decide whether to counterbalance the environmental conditions such that all

variations of background attributes are tested. This would theoretically avoid the of possible interference due to a specific set of background conditions. However, performing a counterbalanced measures design does not come without costs and absurdities. Counterbalancing the conditions is performed by splitting the experiments into groups of every possible order of attribute conditions. Given the large number of attributes that we have identified, and of each attribute with 2-4 variations that we have identified, it would require us to create 384 separate groups of experiments! Every new variation of an attribute increases the numbers of groups by a multiple, and every new attribute increases the numbers of groups by an order of magnitude. This is unfeasible for an empirical study when there are real-world time and resource constraints. Furthermore, when testing the scam attributes, counterbalancing the environmental conditions would be nearly impossible, as the possible attributes and variations are potentially infinite. There are still many other possible attributes and variations that have yet been identified. Therefore, we are never able to test all background conditions. As a realistic solution to this problem, instead of experimenting with an absurd number of background conditions, we compare variations of each attribute under a specific set of background conditions that seem to be the most popular in the real world. With what we gathered, we decide on a standard background condition: a phishing scam with area code 202, with no caller name, speaking in a synthesized, male voice, in American accent, impersonating the IRS, motivating the recipient with a tax lawsuit.

The set of 10 experiments and the variations of each attribute are listed in Table 3.1.

3.2.3 Population

To make our experiments more relatable to the real world, we initially sought to run our experiments on the general population as this is what we thought most scammers would target in the real world. However, the IRB only allowed us to conduct the experiments on our university's internal population. Under the Telephone Consumer Protection Act,

No.	Caller ID	Area Code Location	Caller Name	Voice Production	Gender	Accent	Entity	Scenario
E1	202-869-4555	Washington, DC	N/A	Synthesizer	Male	American	IRS	Tax Lawsuit
E2	800-614-1339	Toll-free	N/A	Synthesizer	Male	American	IRS	Tax Lawsuit
E3	480-939-5666	University Location	N/A	Synthesizer	Male	American	IRS	Tax Lawsuit
E4	202-869-2440	Washington, DC	N/A	Synthesizer	Female	American	IRS	Tax Lawsuit
E5	202-869-2442	Washington, DC	N/A	Synthesizer	Male	American	IRS	Unclaimed Tax Return
E6	202-849-5707	Washington, DC	N/A	Human	Male	American	IRS	Tax Lawsuit
E7	202-869-4024	Washington, DC	N/A	Human	Male	Indian	IRS	Tax Lawsuit
E8	480-462-2513	University Location	N/A	Synthesizer	Male	American	ASU	Payroll Withheld
E9	480-462-2515	University Location	W-2 Administration	Synthesizer	Male	American	ASU	Payroll Withheld
E10	480-462-2517	University Location	N/A	Synthesizer	Male	American	ASU	Bonus Issued

Table 3.1: List of All Experiments and Their Attributes

it is illegal for us to make unsolicited robocalls to residences, or we could be sued for $1,500 in damages for each violation [12]. Complying with this law would require us to obtain explicit permissions before sending the scam calls. However, doing so would nullify the authenticity of our scam experiments. Because if the subjects already knew that they were going to be scammed, it would affect their vigilance and the study results would be unrealistic. We have never heard of any reports of a scammer complying with the law asking for explicit permission in the real world. As law-abiding researchers, hence, the best alternative is to conduct the experiments without obtaining explicit permissions is to conduct the experiments on our university's internal numbers. We managed to obtained IRB permission to conduct the telephone phishing experiments on our university's internal numbers with a waiver for informed consent.

The population of the study were work telephone numbers that were associated with university staffs and faculties. The work phones were Cisco IP phones, model UC Phone CP-8961. Due to the huge population size of our university, we could only realistically contact a subset of the population as there are cost and time constraints associated with any study. Once the study has begun, the marginal benefit associated with gathering more information diminishes over time. In addition, in an internal setting where the awareness of an

ongoing telephone scam could spread out quickly, scam awareness could start influencing our data if a longer time is needed to disseminate the calls. With all these factors in mind, we decided that a population of 3,000 recipients (300 per experiment) is a reasonable size for the empirical study.

To compile the list of telephone numbers, we first wrote a custom tool to bulk download ASU's internal phone directory. For the real-world scammer, our ASU phone directory is also publicly available on the Internet [1] for crawling.

To minimize selection bias, the telephone numbers were randomly chosen from the university telephone directory, then the chosen contacts were randomly put into one of the 10 experiment groups. The sample selection procedure was as follows:

1. Compile the list of work telephone numbers associated with university staffs and faculties.
2. Remove telephone numbers of people already aware of the study.
3. Associate a computer-generated random ID to each telephone number.
4. Sort the random IDs.
5. Select the first 3,000 sorted random IDs.
6. For each 3,000 sorted random ID, incrementally assign an experiment number (1-10) with wraparound.

3.2.4 Procedure

There were several considerations that went into the design of the procedure. First, we need to make sure that the procedure is standardized across all experiments, such that

[1] https://isearch.asu.edu/asu-people/

the results are directly comparable to each other. Second, we need to ensure that process minimizes false positives and false negatives, otherwise, the study results could be unreliable. Last but not least, the procedures also need to be carried out ethically and minimize potential harm to the participant.

To ensure that the procedure is standardized, we decided to use an autodialer to automate the process of sending out the telephone calls and collecting the recipients' responses. We avoided using a live person to conduct the procedure as we believe using a live person could introduce inconsistency. No person can speak in the exact same way for every call, therefore, having a "Johnny" speaking to the recipients could potentially create disturbances to the attributes tested. Although using a live person to interact with the recipient is what many scammers have done, in the real world, telephone scammers have also been using autodialers to distribute scam calls for decades. Therefore, our procedure is applicable to many real-world scams.

Every experiment followed a standard procedure that is summarized in Figure 3.1. The procedure has several parts that require inputs from the recipient. It should be noted that a recipient could break off from the procedure at any point by simply disconnecting the phone, hence not every recipient would follow the procedure until the end. Similar to a real-world scam, a recipient may quickly realize that this is a scam or already know the scam exists, and therefore hang up on the phone call.

The procedure first begins with a ring on the recipient's work phone. When the phone is ringing, the incoming call screen shows the caller ID and the caller name (in experiment E9). The incoming call screen incorporates the visual attribute properties of each particular experiment. An example of the incoming call screen is shown in Figure 3.2a. In all of our experiments, the caller ID showed up as 91xxxxxxxxxx, where xxxxxxxxxx is the caller ID used in the respective experiment. Our university's work phone adds a 91 prefix to every incoming phone call from an external source as all of the calls were distributed from an

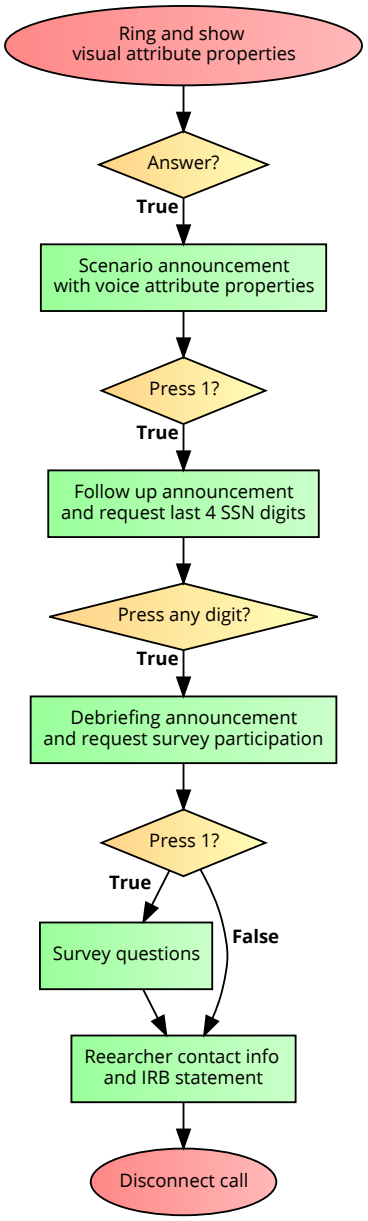
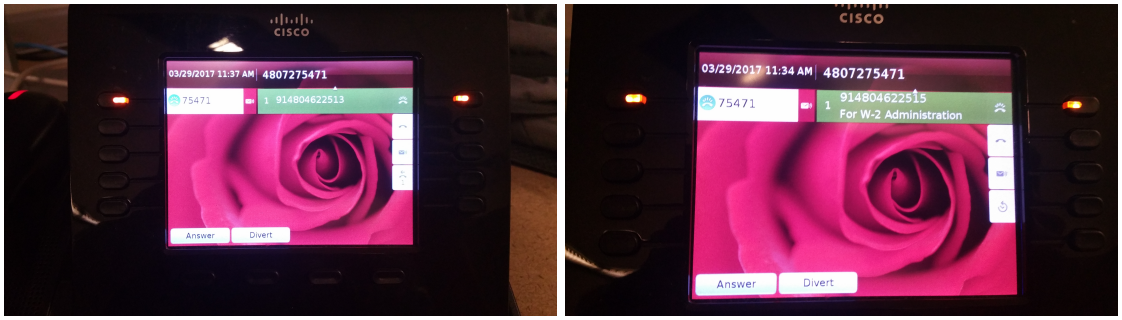


Figure 3.1: Procedure of Each Experiment

external telephone service provider, similar to what a real-world scammer would do. The 91 prefix is present in almost every phone call at our university and therefore we believe its effect is negligible. For Experiment 9, the incoming call screen also shows a caller name as shown in Figure 3.2b. This experiment was designed to simulate a scammer spoofing a known caller name. For legal and ethical reasons, we could not actually spoof a phone number. Instead, we asked our telephone service department to temporarily create a new contact in the university's internal phone directory and associated a legitimate sounding HR department name "W-2 Administration" with the telephone number. We used that temporary phone number to forward phone calls from our autodialer, which added the "For" prefix in the experiments. The end result of the caller name was "For W-2 Administration", which was shown as the caller name in the experiment. As far as we know, no one reported suspicion on the "For" prefix.



(a) All Experiments Except Experiment E9

(b) Experiment E9 With Caller Name Displayed

Figure 3.2: Incoming Call Screen of Different Experiments

After displaying the incoming call screen, if the call is answered, it starts by playing a prerecorded scenario announcement message. The prerecorded scenario announcement message incorporates the voice attribute properties of each particular experiment. We crafted the announcement messages to mimic what a real-world scammer would say. Many

of the words and sentences in the announcement messages were borrowed from the 100+ scam samples that we collected.

In experiments using the IRS Tax Lawsuit scenario, the transcript of the announcement message is as follows:

> This is an urgent notice! This is a final notice from the IRS. The reason for this call is to inform you that the IRS is filing a lawsuit against you. Your action is required immediately, or a penalty will be assessed. To speak to an IRS agent and get more information about this case, please press 1 on your phone now.

In experiments using the IRS Unclaimed Tax Return scenario, the transcript of the announcement message is as follows:

> This is an urgent notice! This is a final notice from the IRS. The reason for this call is to inform you that the IRS has an unclaimed tax return for you that is due to expire within three days. Your action is required immediately. To speak to an IRS agent and get more information about claiming your tax refund, please press 1 on your phone now.

In experiments using the ASU Payroll Withheld scenario, the transcript of the announcement message is as follows:

> Dear ASU employee. This is an urgent notice! This is a final notice from the ASU W-2 administration office. The reason for this call is to inform you that to process your next Friday payroll, you are required to update your 2016 tax information immediately. To speak to a staff agent and get more information, please press 1 on your phone now.

Our university has a publicly available payroll calendar on the HR department's website [2] , hence a real-world scammer could also use this information to craft an announcement message based on the payroll information.

In experiments using the ASU Bonus Issued scenario, the transcript of the announcement message is as follows:

> Dear ASU employee. This is an urgent notice! This is a final notice from the ASU W-2 administration office. The reason for this call is to inform you a performance bonus has been issued to your account. Your action is required immediately. To speak to a staff agent and get more information, please press 1 on your phone now.

Our university has a publicly available webpage listing the types performance-based bonuses for faculties and staffs [3] , hence a real-world scammer could also use this information to craft an announcement message based on the performance bonus information.

Every scenario announcement message requests the recipient to enter `1` to continue to the next step. The purpose of this action is to reduce the likelihood of recipients making some random input actions without first finish hearing the scenario announcement with the voice attribute properties of each particular experiment. The action also helps to filter out answers from answering machines.

If the recipient pressed `1` shortly after the scenario announcement message, a follow-up announcement message will be played. The follow-up announcement message is as follows:

[2] https://www.asu.edu/fs/documents/2017-BiWeekly-Calendar.pdf

[3] https://cfo.asu.edu/compensation

> Please wait for the next available agent. Thank you for holding. Your call will be connected shortly. Please enter the last four digits of your social security number on your phone now.

This follow-up message was the same for every experiment. The goal of the follow-up message is to trick the recipients into divulging the last four digits of their social security number. in the real world, the last four digits of the social security number can be all a criminal need to perpetuate financial and identity fraud [51]. Other parts of the social security number can also be inferred from the recipient's phone number [52]. To minimize potential risk to the recipient, we did not record which digits were pressed, we instead recorded if any particular (0–9) digit was pressed.

Pressing any digit shortly after the follow-up message will immediately lead to a debriefing announcement and a request to participate in our phone survey. The debriefing message is as follows:

> Hi, I am a graduate student under the direction of Professor Adam Doupé in the department of computer science at Arizona State University. I am conducting a research study to measure the effectiveness of telephone phishing. The reason you are receiving this message is because I would like to inform you that what you just did could potentially lead you becoming exploited in a real telephone scam. However, I would like to assure you that this is not an actual scam, none of your social security information was actually collected.
>
> We would like to invite you to participate in our phone survey, to help us better understand your thoughts about the scam. You will be able to listen to the survey questions right after this message. Your participation in this survey is voluntary. There are no foreseeable risks for your participation. If you choose not to participate or to withdraw from the survey at any time, there will be

> no penalty. Your responses will be anonymous. The results of this study may be used in reports, presentations, or publications but your identity will not be used. Please press 1 to listen to the survey questions or participate in the phone survey.

The debriefing announcement and survey questions were recorded with the researcher's real voice emphasize the fact that whatever they listened to was not a real scam.

The survey consisted of two questions. If the recipient pressed `1` shortly after the debriefing message, the first survey question is as follows:

> Thank you. Could you please help us understand if the scam was able to convince you to enter your social security number? Please use the number on your keypad to answer this question. If "yes", please press 1. If "no" please press 0.

We recorded the input digit shortly after the first survey question. If `1` was pressed, the second question is as follows:

> Thank you. Could you please help us understand what was the most important factor that made the scam convincing? We will record your voice response for this question. At the tone, please state briefly what you thought was the most important factor. When you are finished, please press the pound key to end recording.

If `0` was pressed shortly after the first survey question, the second question is as follows:

> Thank you. Could you please help us understand what was the most important reason you did not believe in the scam? We will record your voice response for this question. At the tone, please state briefly what you thought was the most important reason. When you are finished, please press the pound key to end recording.

We recorded the participant's voice recording for the second question. After the second survey question, the autodialer system plays the following ending message:

> Thank you. This is the end of the research experiment. If you have any questions concerning the research study, please contact the research team at 480-420-8250. If you have any questions about your rights as a participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Institutional Review Board, through the ASU, at 480-965-6788. Thank you for your participation. Goodbye.

During each procedure, the autodialer was configured to collect the following inputs from the recipient:

Continued

Whether the recipient pressed 1 during or shortly after the scenario announcement message is played.

Entered SSN

Whether the recipient pressed any digit during or shortly after the follow-up announcement message is played.

Convinced

Whether the recipient pressed 1 stating that they were convinced to enter the last four digits of their SSN during or shortly after the first survey question is played.

Unconvinced

Whether the recipient pressed 0 stating that they were not convinced to enter the last four digits of their SSN during or shortly after the first survey question is played.

Recording

The recipient's recorded voice response for the second survey question.

3.2.5 Ethical Compliance

To address the ethical issues our experiments, we worked with our university's IRB to obtain approval for this experiment. We strove to design the experiments ethically to protect the subjects as much as possible. To do so, we implemented several safeguards in the experimental design to minimize the harm to participants.

The nature of this experiment, studying telephone phishing attacks, involves deception as well as involuntary participation. Both aspects are critical to receiving scientifically valid results—informing the participants of the study would significantly bias the results. However, the use of deception could result in complaints and negative reactions from our recipients. Before proceeding with the experiment, we also worked with our university's IT security group to provide them with the information that would help to alleviate the concerns our participants. This IT security group at ASU is responsible for the security of all aspects of the university. We shared with the security group the experiment contact list, the experimental design, and the incoming phone numbers so that the help desk personnel could be prepared to handle the requests and reports.

In recording the results, we also strive to do so ethically and in accordance with established IRB protocols. One of the major safeguards is that we did not record the social security number. While a spammer would typically want the social security number (or en-

tire credit card number in different scam scenarios), all that is recorded is the fact that they pressed any digit. To further reduce potential harm, we recorded very little meta information from each call (each record is only the call time, call duration, answer state, and input responses to the interactive system). This further reduces potential harm to the participants. Although these measures may diminish the strength of our data, we believe ethics is a more important aspect of designing a telephone phishing study.

3.2.6 Dissemination

We ran the previously described procedure using the 10 described experiments during the work week of April 27th–31st of 2017, during core working hours of 10:00am–5:00pm each day. We used an Internet-hosted autodialer to automate the process of sending out the telephone calls to the 3,000 recipients. Each experiment's calls were simultaneously distributed during the experiment period at a rate of no more than 3 live calls per experiment.

We associated each experiment with a unique caller ID. In all experiments, vast majority of the outbound calls did not reach a live recipient and was answered by a voicemail answering machine. Every ASU work phone had voicemail enabled by default and our scenario announcement message would be recorded by the voicemail answering machine. If a recipient could not answer the phone, the recipient could use the caller ID in their call history to call us back. The recipient may also listen to the voicemail message left by the scenario announcement message prior to calling back.

As each experiment had a unique caller ID, the return call would be directed to that particular experiment's procedure. When a recipient called back, the same procedure was administered where a prerecorded scenario announcement message is first played. The follow-up processes are also same as the outbound calling process with the same data collected.

While disseminating the phone calls, several unexpected events might have impacted our study.

The ASU school of journalism and mass communication identified the scam call incidents only 2:45 hours from the launch of the experiments on the first day. Instead of reporting it to the university help desk (who were prepared and aware of our study), the school sent out a mass email warning all journalism staffs and faculties at 4:28 hours from launch.

At 4:22 hours from the launch of the experiments, our university's telephone service office also started blocking our phone calls as they were receiving system alerts of too many incoming phone calls exhausting the telephone trunk routes. We managed to work with the telephone service office to get our calls unblocked within the next 4 hours as we dialed down the simultaneous call rate of our phone calls to 1 per experiment.

Meanwhile, the IRB office also received some complaints regarding the scam call experiments, which resulted in the experiments being paused for roughly 12 hours on the third day since launch, as we waited for the IRB committee to review and allow us to proceed with the experiments.

In the end, despite the unexpected events, we finished sending out the telephone calls to the 3,000 recipients as planned before the end of the work week.

3.3 Results and Analysis

The goal of this section is to present the results and provide a methodology of analyzing the results. We apply the methodology to the results from the 3,000 phishing calls and provide a discussion on the outcomes of the study.

After running the experiment for a work week, the results were collected from the 3,000 phishing calls. The input data collected from the recipients are presented in Table 3.2.

No.	Continued		Entered SSN		Convinced		Recordings		Unconvinced		Recordings	
E1	12	4.00%	6	2.00%	0	0.00%	0	0.00%	4	1.33%	2	0.67%
E2	19	6.33%	15	5.00%	3	1.00%	0	0.00%	3	1.00%	3	1.00%
E3	13	4.33%	8	2.67%	1	0.33%	1	0.33%	2	0.67%	1	0.33%
E4	23	7.67%	13	4.33%	2	0.67%	0	0.00%	3	1.00%	2	0.67%
E5	9	3.00%	2	0.67%	1	0.33%	0	0.00%	1	0.33%	1	0.33%
E6	9	3.00%	8	2.67%	2	0.67%	2	0.67%	2	0.67%	1	0.33%
E7	13	4.33%	9	3.00%	3	1.00%	1	0.33%	5	1.67%	4	1.33%
E8	53	17.67%	30	10.00%	8	2.67%	3	1.00%	9	3.00%	8	2.67%
E9	60	20.00%	35	11.67%	7	2.33%	3	1.00%	4	1.33%	3	1.00%
E10	45	15.00%	22	7.33%	8	2.67%	7	2.33%	4	1.33%	2	0.67%
Total	256	8.53%	148	4.93%	35	1.17%	17	0.57%	37	1.23%	27	0.90%

Table 3.2: Summary of Recipient Inputs from All Experiments

Across all 10 experiments of 3,000 total recipients, we had 8.53% (256/3000) of all recipients continued after listening to the scam scenario announcement, 4.93% (148/3000) of all recipients entered at a digit when requested to enter the last four digits of their social security number, 1.17% (35/3000) of all recipients explicitly stated that they were convinced by the scam, and 1.23% (27/3000) of all recipients explicitly stated that they were not convinced by the scam.

Before presenting our analysis of the experiments, we provide a discussion on the development of our methodology to systematically analyze their relative effectiveness.

The first step of performing the analysis is to decide on a set of good metric(s) that will be used for as the standard of measurement. To chose an ideal metric, we believe a good metric should not only be quantifiable but also be a proxy for what ultimately matters. From the telephone scammers' perspective, we understand what matters to them is to collect as many social security numbers as possible for the purpose of conducting identity fraud. Due to the ethical considerations discussed in the previous section, we could not to directly

collect and verify the recipients' social security numbers. Therefore, we need to derive a metric that could provide us with a reasonable estimate of the actual number of real SSNs given to us in each experiment. This is how we arrived at that metric:

Continued

This is the total number of instances where digit `1` was pressed during or shortly after the scenario announcement is played. After listening to the scenario announcement, the recipient would be interested in getting more information about the scenario by pressing `1`. This metric could help to infer the effectiveness of the experiment. However, there is still a possibility that the recipient may continue after the scenario announcement and not fall for the scam.

Entered SSN

Because there is still a possibility that the recipients continue after the scenario announcement and not provided the last 4 digits of their SSNs, we could use the total number of instances where a digit was pressed shortly after the follow-up announcement. At this point, the recipients would have already listened to the scenario announcement, pressed `1`, listened to the follow-up announcement, and then would have probably tried to enter the last four digits of their social security number. Although this seems like an idea metric to estimate the number of SSNs collected, however, there is still a possibility that the recipient may have tried to enter a fake social security number. In some of the recordings, a few recipients stated that they did not enter their real social security number information.

Convinced

This is the number of recipients that explicitly stated that they were convinced by the scam after the first survey question. At this point, the recipients would have already listened

to the scenario announcement, pressed `1`, listened to the follow-up announcement, and then would have probably tried to enter the last four digits of their social security number, then listened to the debriefing announcement, pressed `1`, then listened to the first survey question, and then pressed `1`. This metric would be most conservative for estimating attack success. However, looking at the low number of responses, participants rarely made it to that step. People are generally not inclined to stay on a robocall for too long. Using this metric would exclude a large number of recipients that fell for the scam but declined to participate in the phone survey after the debriefing announcement.

Possibly Tricked

Since we cannot assume that all SSNs entered were real, to reduce these types of false positives, we could remove the participants that entered their SSNs and then subsequently stated that they were unconvinced by then scam during the survey process. This number, which we call *Possibly Tricked*, provides a more reasonable estimate of the actual number of recipients that fell for the scam by entering the last four digits of their social security number. Compared to the previous metrics, this metric provides a good balance of conservativeness and sample size, therefore, we decided on using this metric for our analysis.

Table 3.3 presents a table of the number of possibly tricked recipients for each experiment, ranked from most successful to least successful. Comparing the tricked result between experiments, experiment E9 had the highest tricked rate among all experiments, with an estimate of 10.33% (31/300) of recipients tricked into entering the last four digits of their social security number. On the other hand, experiment 5 had the lowest success rate among all experiments, with an estimate of only 0.33% (1/300) of recipients tricked into entering the last four digits of their social security number. At a glance, the ASU scams performed distinctly better than the IRS scams as they took the top 3 spots on our list.

No.	Entered SSN	Unconvinced	Possibly Tricked	
E9	35	4	31	10.33%
E8	30	9	21	7.00%
E10	22	4	18	6.00%
E2	15	3	12	4.00%
E4	13	3	10	3.33%
E3	8	2	6	2.00%
E6	8	2	6	2.00%
E7	9	6	3	1.00%
E1	6	4	2	0.67%
E5	2	1	1	0.33%
Total	148	37	111	3.70%

Table 3.3: Estimating the Number of Recipients Tricked into Entering Their Real SSN Information

The next step for us is to decide on an appropriate method of data analysis on the chosen metric.

With a myriad of possible data analysis methods, model-based analysis, such as regression, Bayesian, and machine-learning models, is one category of the methods we considered using. Model-based analysis can produce a model that describes an optimal mapping of attribute properties to the results. However, such methods often achieve optimality through an ideal combination of parameters, which tend to overfit the spurious correlations that occur in our training data. Looking at the training data, it would be hard to construct a trick success rate model based on the attribute properties without overfitting as it is a Small Data problem [53].

Therefore, we ultimately decided on using statistical hypothesis testing approaches for analysis, as it is a more suitable analysis technique in this situation. Statistical hypothesis testing also has the added benefit of providing more comprehensible answers to the questions that matter. The goal of our study to perform analysis such that the knowledge

extracted from the data can be discussed in relevance to the real world, not our training data.

To begin using statistical hypothesis testing approaches, we asked, "what are the hypothesis questions that our data can provide an answer for?" We will provide a discussion on the hypothesis questions we decided to ask and how we applied a data analysis method to provide a contextual answer to the hypothesis questions.

Can manipulating the area code have a significant effect on the attack success of a telephone scam?

In the real world, we observed telephone scammers used area code manipulation in many instances. To provide an answer to this question, we can compare the number of possibly tricked between similar experiments that used different area codes, i.e. E1, E2, and E3. We see that E1 had 0.67% possibly tricked, E2 had 4% possibly tricked, and E3 had 2% possibly tricked. In our question concerning the significance of area code, since E1 and E2 have the greatest difference in the number of possibly tricked precipitants, we test if using toll-free area code is significantly more effective than Washington, DC area code in the context of the IRS scam example. So we perform a right-tailed p-value statistical hypothesis testing approach on the chosen experiment groups.

The use of right-tailed p-value approach in statistical hypothesis testing is a way to answer if assuming the null hypothesis is true whether the improved alternative hypothesis is "likely" in the direction of the alternative hypothesis. With regards to the choice of using Bayesian vs Frequentist methods, since there no past knowledge of similar experiments conducted, we can only use Frequentist methods to calculate the statistical significance on the underlying truths using only data from the current experiment. In addition, not only do we want to know if the improvement to attack success is significant, it is also important to know the magnitude of improvement. To avoid making statements like "E2 is 5 times more

effective than E1", instead of measuring the relative difference, we calculated Cohen's d to measure the effect size for comparison between the two groups.

Using the right-tailed p-value approach, we have a Z-score of 2.721 and a p-value of 0.0033. Using an arbitrary confidence level of 95% (p-value < 0.05), it is very likely that using toll-free area code can result in a more successful attack than using than Washington, DC area code in the context of the IRS scam example. The two groups also have a Cohen's d of 0.222, which suggests it has a small effect according to Cohen [54] and has a somewhat educationally significant effect according to Wolf [55]. Therefore, we could say that the area code can have a significant effect on the attack success of telephone phishing scam.

Can manipulating the type of voice production have a significant effect on the attack success of a telephone scam?

In the scam samples collected, we heard many scammers used a synthesized voice and some others used recorded human voice. To provide an answer to this question, we can compare the number of possibly tricked between similar experiments that used different types of voice production, i.e. E1 and E6. In our question concerning the significance of voice production, we test if using recorded human voice is significantly more effective than using synthesized voice in the context of the IRS scam example.

Using the same right-tailed p-value approach, we have a Z-score of 1.426 and a p-value of 0.0769. Using an arbitrary confidence level of 95% (p-value < 0.05), it is unlikely that using recorded human voice can result in a more successful attack than using than using synthesized voice in the context of the IRS scam example. The two groups have a Cohen's d of 0.117, which also suggests the effect size is very small and not educationally significant. Therefore, we are not able to conclude at this time if the type of voice production had a significant effect on the attack success of a telephone phishing scam.

Can manipulating the voice gender have a significant effect on the attack success of a telephone scam?

For the telephone scammer, the voice gender of the voice synthesizer can be easily changed with a simple option click in the autodialer. To provide an answer to this question, we compare the number of possibly tricked between similar experiments that used different voice genders, i.e. E1 and E4. In our question concerning the significance of voice gender, we test if using a female synthesized voice is significantly more effective than using male synthesized voice in the context of the IRS scam example.

Using the same right-tailed p-value approach, we have a Z-score of 2.343 and a p-value of 0.00955. Using an arbitrary confidence level of 95% (p-value < 0.05), it is likely that using a female synthesized voice can result in a more successful attack than using a male synthesized voice in the context of the IRS scam example. The two groups have a Cohen's d of 0.192, which suggests the effect size is small and not educationally significant. Therefore, it is hard for us to conclude at this time if the voice gender has a significant effect on the attack success of a telephone phishing scam.

Can manipulating the voice accent have a significant effect on the attack success of a telephone scam?

In the IRS scam samples, many spoke with an Indian accent while some spoke with an American accent. To provide an answer to this question, we compare the number of possibly tricked between similar experiments that used different accents, i.e. E6 and E7. In our question concerning the significance of voice accent, we test if speaking with an American accent is significantly more effective than speaking with an Indian accent in the context of the IRS scam example.

Using the same right-tailed p-value approach, we have a Z-score of 1.008 and a p-value of 0.157. Using an arbitrary confidence level of 95% (p-value < 0.05), it is unlikely that speaking with an American accent can result in a more successful attack than speaking

with an Indian accent in the context of the IRS scam example. The two groups also have a Cohen's d of 0.082, which suggests the effect size is very small and not educationally significant. Therefore, we are not able to conclude at this time if the voice accent had a significant effect on the attack success of a telephone phishing scam.

Can spoofing a known caller name have a significant effect on the attack success of a telephone scam?

In the real world, we observed telephone scammers spoofing a known caller name in some instances, especially in targeted attacks. To provide an answer to this question, we compare the number of possibly tricked between similar experiments that has a difference in the display of a caller name, i.e. E8 and E9. In our question concerning the significance of spoofing caller name, we test if displaying a HR-department like caller name "W-2 Administration" is more effective than not displaying a caller name in the context of the HR scam example.

Using the same right-tailed p-value approach, we have a Z-score of 1.454 and a p-value of 0.0730. Using an arbitrary confidence level of 95% (p-value < 0.05), it is unlikely that displaying a HR-department like caller name can result in a more successful attack than displaying a caller name in the context of the HR scam example. The two groups also have a Cohen's d of 0.119, which suggests the effect size is very small and not educationally significant. Therefore, we are not able to conclude at this time if spoofing a known caller name had a significant effect on the attack success of a telephone phishing scam.

Can impersonating an internal entity have a significant effect on the attack success of a telephone scam?

With any form of spear phishing, it involves impersonating an internal entity that the recipient is familiar with. The scammer then has to create a scenario that is tailored to the entity, as the "Entity" cannot be set independently from "Scenario". To provide an answer

to the hypothesis question, we compare the number of possibly tricked between similar experiments that used different entity-scenarios, i.e. comparing E1 and E5 with E8 and E10. In our question concerning the significance of impersonating an internal entity, we test if impersonating the HR department using the name "W-2 Administration" is more effective than impersonating the IRS with the context of the scenarios tested.

Using the same right-tailed p-value approach, we have a Z-score of 5.732 and a p-value of 4.97E-9. Using an arbitrary confidence level of 95% (p-value < 0.05), it is likely that impersonating the HR department using the name "W-2 Administration" can result in a more successful attack than impersonating the IRS with the context of the scenarios tested. The two groups also have a Cohen's d of 0.331, which suggests the effect size is small and educationally significant. Therefore, we could say that impersonating an internal entity had a significant effect on the attack success of a telephone phishing scam.

Can manipulating the type of motivation have a significant effect on the attack success of a telephone scam?

To motivate recipient into taking some harmful action, the scammer could either use fear or reward. To provide an answer to the hypothesis question, we compare the number of possibly tricked between similar experiments that used different types of motivation, i.e. comparing E1 and E8 with E5 and E10. In our question concerning the significance of the type of motivation, we test if fear-based scenarios are more effective than reward-based scenarios the context of the entities tested.

Using the same right-tailed p-value approach, we have a Z-score of 0.628 and a p-value of 0.265. Using an arbitrary confidence level of 95% (p-value < 0.05), it is unlikely that fear-based scenarios can result in a more successful attack than reward-based scenarios with the context of the entities tested. The two groups also have a Cohen's d of 0.036, which suggests the effect size is very small and not educationally significant. Therefore, we are

Hypothesis	Group A	Possibly Tricked	Group B	Possibly Tricked	p-value	Significant ($p < 0.05$)	Cohen's d	Effect Size	Conclusive
Can manipulating the area code have a significant effect on the attack success of a telephone scam?	E1	2/300	E2	12/300	0.0033	Yes	0.222	Small & somewhat educationally significant	Somewhat
Can manipulating the type of voice production have a significant effect on the attack success of a telephone scam?	E1	2/300	E6	6/300	0.0769	No	0.117	Very small & not educationally significant	No
Can manipulating the voice gender have a significant effect on the attack success of a telephone scam?	E1	2/300	E4	10/300	0.00955	Yes	0.192	Small & not educationally significant	Hardly
Can manipulating the voice accent have a significant effect on the attack success of a telephone scam?	E7	3/300	E6	6/300	0.157	No	0.082	Very small & not educationally significant	No
Can spoofing a known caller name have a significant effect on the attack success of a telephone scam?	E8	21/300	E9	31/300	0.073	No	0.119	Very small & not educationally significant	No
Can impersonating an internal entity have a significant effect on the attack success of a telephone scam?	E1 + E5	3/600	E8 + E9	39/600	4.97E-9	Yes	0.331	Small & educationally significant	Yes
Can manipulating the type of motivation have a significant effect on the attack success of a telephone scam?	E5 + E10	19/600	E1 + E8	23/600	0.265	No	0.036	Very small & not educationally significant	No

Table 3.4: Summary of Statistical Hypothesis Testing Results

not able to conclude at this time if manipulating the type of motivation had a significant effect on the attack success of a telephone phishing scam.

Summary

The summary of our statistical hypothesis testing results are shown in Table 3.4. Based on the statistical hypothesis results, we found that impersonating an internal entity had the most significant effect on the attack success of a telephone phishing scam. We also found that manipulating the area code (using toll-free area) can have a somewhat significant effect. On the contrary, manipulating the type of motivation, voice production, voice accent, and caller name, individually had an insignificant effect on the attack success. It is also hard for us to conclude whether manipulating the voice gender have a significant effect even though the result was statistically significant.

3.4 Survey Responses

In this section, we discuss the recorded survey responses that asked the participants for the reasons they were convinced or unconvinced to enter the last four digits of their social security number. We listened to all 44 recorded voice responses and summarized the responses in Table 3.5.

Based on the limited number of voice responses from the survey respondents (1.4%, or 44/3,000), no one provided an explicit voice response on why they were convinced by the IRS scams. The four recordings we received were either silent or contained no useful information. In general, participants were less willing to report the reasons why they were convinced by the scam after they were explicitly told that they had fallen victim to an attack.

On why the IRS scams were unconvincing, most of the survey respondents stated that they already knew that the IRS would not make a call like this or that they were already vigilant about IRS scam calls. This is understandable because there are numerous media reports about the IRS scams, and the IRS posted many public warnings on not to trust these

No.	Reasons Convinced	Reasons Unconvinced
E1		Would never enter SSN on incoming call; No name mentioned for the charge
E2		IRS won't make a call like this (x2); Already aware of scams like this
E3		IRS won't make a call like this
E4		IRS won't make a call like this; Didn't sound legitimate
E5		IRS won't make a call like this
E6		IRS won't make a call like this; Already aware of scams like this
E7		IRS won't make a call like this (x4); Indian accent (x2)
E8	To get paid (x2); Sounded legitimate; Trusted work phone; Only asked for last 4 SSN; Caller ID showed local ASU number	ASU won't make a call like this (x5); Not from ASU number (x2); Synthetic voice;
E9	Sounded legitimate; Only asked for last 4 SSN; Caller ID showed ASU W-2	Should have asked for complete SSN (x2); Would never enter SSN on incoming call
E10	To get bonus (x2); Trusted work phone; From ASU number; Asked to do so	ASU won't make a call like this; Not ASU number

Table 3.5: Summary of Recorded Survey Responses

types of scams. This further supports the hypothesis that the impersonated identity and the corresponding scenario was the most significant factor. Interestingly, in experiment E7, two respondents mentioned that the Indian accent was also one of the reasons they were unconvinced.

The ASU phishing scams had more survey responses. On why the ASU scams convinced them, most of the survey respondents described something related to the scam scenario, which means that the impersonated entity and the scenario were the key factors. Three respondents also believe that the caller ID was from ASU and stated caller ID was one of the reasons they believed in the scam, even though none of the caller IDs were actually from ASU.

On why the ASU scams did not convince them, most of the survey respondents described that they were quite certain that ASU would not make a call like this or they were already vigilant about giving their SSN information over an incoming call. Interestingly, two respondents mentioned that the scenario only asked for the last four digits of their SSN, and should have asked for their complete SSN if it was really payroll related, which quite possibly meant that those two might have given out their complete SSNs if the phishing scam had asked for it. The external caller ID and synthetic voice were also mentioned for being one of the factors that made the survey respondents suspicious.

3.5 Limitations

The experiments were conducted in a university setting where the recipients are university staffs and faculties. The demographics of the recipients in our study is not a representative sample of the general population of telephone users in the US. It is possible that general population of telephone users in the US have different tricked rates than university staffs and faculties. We suspect that the general population would be more susceptible to telephone phishing scams than the population tested in our experiments, as all ASU staffs

and faculties were required to take information security awareness training within the first month of employment and annually thereafter. However, our study would not pass IRB approval or would have required explicit permissions if it was conducted on the general population.

The experiments only sent out calls to a specific brand of work phones. The type of phone in our study is not a representative sample of the entire population of telephones in the US. The vast majority of the telephones are mobile phones and it is possible that these have a different tricked rates than work phones. We suspect that the mobile phones have a higher answering rates and therefore would have a higher rate of attack success compared to work phones. However, personal cell phone numbers are considered residential property and our study would not pass IRB approval or would have required explicit permissions if it was conducted on mobile telephones.

The experiments only collected SSN information based on estimates. The experiments had several safeguards and the process was tightly regulated to ensure risks to the human research subjects were minimized. This prevented us from collecting any actual social security numbers from the recipients. We suspect that collecting actual social security numbers would have strengthened the results of our study. However, our study would not pass IRB approval or would have required explicit permissions if it was designed to collect social security numbers.

In the real world, there are infinite ways a telephone scam could be constructed. Our study, at best, shines a slither of light on the universe of all possibilities. We designed our scam experiments based on some popular telephone scams being reported to government and user websites, but there are some other possible scams that we have not yet covered. At this time, our study focused on attributes and scams that we presented in this chapter. The study of other possible attributes and scams is what we could do in a future work.

3.6 Discussion

Our experiments have shown that automated telephone phishing attacks can be effective. One experiment, E9, which simulated a targeted phishing attack with caller name spoofing, achieved a 10.33% tricked rate, where recipients possibly divulged the last four digits of their social security numbers.

We have also validated some potential key attributes that can have a significant effect on the scam effectiveness, i.e., impersonating an internal entity and announcing a relevant scenario. Manipulating the caller ID to a toll-free area code may also somewhat improve the scam effectiveness. Other attribute properties such as human voice, female voice, American accent, caller name spoofing, fear-based scenario also improved the scam effectiveness in our empirical study, however, at this time we are not able to conclusively prove that they have a significant effect. Nonetheless, given how easy it is for a scammer to manipulate all these attributes, a scammer would seek to incorporate all attribute properties that made an improvement to the attack success, i.e. a phishing scam with toll-free area code, spoofing known a caller name, speaking in a recorded human, female voice, in American accent, impersonating an internal entity, motivating the recipient with a relevant fear-based scenario.

Our process of designing the phishing scams and disseminating the scam calls have also shown that this type of massively scalable phishing attack can be carried out with little barriers to entry: A scammer can compile a recipient list from a company's publicly available telephone directory; The voice announcement can be recorded easily with a text-to-speech synthesizer or by recording a human voice; A convincing scam scenario can be crafted by researching the company's departments; The caller IDs are easily bought or spoofed with an Internet-hosted service provider; The scam calls can be distributed cheaply and automatically with an autodialer. Furthermore, the Internet provides plenty of opportunities to

avoid legal repercussions. Hence, this user study is a reminder that telephone scam is a technically and economically viable operation, which provides evidence as to their rise in popularity.

To prevent falling victim to these types of phishing scams, we believe that the key is to target and prevent *impersonation*. Our statistical results have shown that impersonating an internal entity had a significant effect on the scam effectiveness. To address the impersonation issue, feedback from our survey participants suggest that vigilance was an important reason for not falling for a scam. Many surveyed subjects expressed distrust towards our scam calls when they were already vigilant about the scam scenario. It is possible that the information from existing scam reports and the security awareness training could have helped the subjects stay vigilant against the phishing calls. Therefore, we recommend education and awareness of telephone phishing as a countermeasure. On technical solutions, we recommend a similar approach to help the subjects stay vigilant against the phishing calls. There are solutions that can provide early warnings against impersonated calls, such as, caller ID authentication [56–58], which has strong safeguards against caller ID impersonation and could help to warn the users against malicious calls with a reputation system.

3.7 Related Work

We have not found prior empirical user studies on telephone phishing. The only similar work we found was by Aburrous et al., who performed a phone phishing experiment on a group of 50 employees contacted by female colleagues assigned to lure them into giving away their personal e-banking usernames and passwords. They were able to deceive 32% of the employees to give out their e-banking credentials [59]. However, in the experiment, the 50 employees already knew the female colleagues that contacted them, which suggests that this is more of an experiment studying an insider attack rather than an impersonation attack.

Other related works studied phishing using different channels. Dhamija et al. performed a website phishing study on 22 university participants and their best phishing site was able to fool more than 90% of participants [60]. Egelman et al. performed an email and website phishing experiment on 60 in-person participants to test the effectiveness of various web browser phishing warnings at that time, and it was found that 79% of Internet Explorer 7.0 participants heeded the active phishing warnings and only 13% of them obeyed the passive warnings [61]. Jagatic et al. performed a social media spear phishing study on 481 targeted Indiana University student emails obtained by crawling social network websites and it had a 72% success rate of recipients authenticating themselves on a redirected website [62]. Vidas et al. performed a QR code phishing study where the experiment distributed 139 posters containing QR codes at various locations at Carnegie Mellon University and the city of Pittsburgh, the experiment was able to trick 225 individuals to visit the associated website in four weeks [63].

3.8 Conclusion

This chapter presented the methodology, design, execution, results, analysis, and evaluation in the quest of answering why telephone phishing works. The methodology involved first taxonomizing the visual and voice attributes of telephone phishing scams, and then developing a systematized approach designed to test if each attribute has a significant effect on the success of the scam. The study was executed using 10 experiments simulating telephone phishing attacks, administered to 3,000 works phones of university staffs and faculties over the course of a work week. The results were collected from the inputs and survey responses of the phone recipients. We analyzed the results by performing statistical hypothesis testing methods on a chosen metric derived from the inputs and we were able to identify at least one attribute that had a significant effect. We provided a discussion on

how to effectively prevent such types of telephone phishing scams, and we believe that the best countermeasures should target impersonation and instill vigilance.

Chapter 4

IDENTIFYING KEY CHALLENGES AND EXISTING COUNTERMEASURES

With an understanding of why telephone scams work from the previous chapter, the question we ask is: what are the existing defenses against spam and scam calls? Given the significant gains made in reducing email spam, this raises the question: are there any simple and effective solutions that could stop telephone spam? Unfortunately, we found that this issue is not easily solved as there are unique challenges in the telephone ecosystem that require novel approaches. Many existing solutions have failed to overcome these challenges and, as a result, have yet to be widely implemented. In this chapter, we will present a survey of the key challenges in dealing with telephone spam and the existing telephone spam countermeasures. With such understandings, we derive a set of evaluation criteria that we use to analyze the failings of the current techniques and present the metrics that are critical to an acceptable solution. We believe that this work will help guide the development of effective telephone spam defenses, as well as provide a framework to evaluate future defenses.

The main contributions of this chapter are the following:

- We describe the challenges in reducing telephone spam, describing the technical and regulatory challenges that make telephone spam distinct from email spam.
- We develop a taxonomy that classifies the existing anti-spam techniques into three categories, providing a high-level view of the benefits and drawbacks of each type of technique.

- We provide a systematization of assessment criteria for evaluating telephone spam countermeasures, and we evaluate existing techniques using these assessment criteria.

- We provide a discussion on why we believe in stopping caller ID spoofing to be the best direction of solving the telephone spam problem.

This chapter contains work published in IEEE Symposium on Security and Privacy 2016 [45].

4.1 Key Challenges

Before diving into the various anti-spam techniques and countermeasures to prevent telephone spam, it is important to first discuss the challenges and constraints in combating telephone spam to understand how it affect the design of a countermeasure.

We identify several challenges in combating telephone spam—that are significantly different from email spam—some of which are technical and some of which are regulatory.

4.1.1 Immediacy Constraint

Unlike email, which can be queued for later analysis, a voice call has an immediacy constraint. A telephone call request is immediate and therefore must be analyzed as soon as it appears, and the telephone anti-spam system must complete analysis and take action within a short window of time to reduce the delay. If a solution adds too much delay to a call request, the legitimate caller may assume that the recipient could not answer the phone and hang up.

4.1.2 Difficulty of Working with Audio Streams

The content of a voice call is difficult to parse and analyze: the content is an audio stream as opposed to the text of an email. To make matters worse, the content of a voice call is only revealed when the call is answered, and both the caller and the recipient will be affected if an anti-spam system answers the call. Whereas an email anti-spam system can easily analyze the content of an email, and neither the sender nor the receiver is affected.

4.1.3 Lack of Useful Header Data

Voice calls lack the rich header data of email. When a call arrives at the recipient, it contains little useful header information. When a phone call arrives at the recipient, usually the only useful information available to the recipient is the caller ID. An email header, however, has well-defined and information-rich SMTP headers—before the content of the email. It is also difficult to omit the sender's IP address and the domain name of the email. This is in stark contrast to a call request header, where the header data is easily omittable by a spammer.

4.1.4 Hard to Gain User Acceptance

The bar for user acceptance of a telephone spam countermeasure is much higher compared to email. Consumers, rightly, have a very low tolerance for false positives of blocked calls. Phone call tends to be more urgent and important compared the email, and once a phone call is wrongfully blocked it could have severe consequences.

4.1.5 Caller ID Spoofing

The *Caller ID* service is an information service that provides the recipient with information of the caller before answering the phone, which could be useful for blocking spam

calls. However, caller ID fundamentally has no authentication mechanism and is easily spoofed. The only security mechanism comes from having the TSP send the caller ID on behalf of the caller. This security mechanism is eroded when the spammer subscribes to a telephone service that allows customization of caller IDs. With the rise of VoIP services that provide features such as caller ID customization over the Internet, it is trivial for any caller to cheaply and effectively spoof the caller ID. Thus, any telephone spam defense technique that relies on the caller ID is now vulnerable to caller ID spoofing.

4.1.6 Difficulty of Tracing Spam Calls

One way to combat spam is to make it illegal and enforce those laws. In the history of email spam, a small number of players were responsible for the majority of the spam, hence taking action against these big targets resulted in significant drops in spam volume. For instance, shutting down the Rustock botnet reduced global spam levels by around 40% [64]. It is reasonable to assume a similar distribution of telephone spammers. Unfortunately, identifying the actual distribution of telephone spammers is difficult due to the technical and regulatory challenges of monitoring PSTN traffic and the prevalence of caller ID spoofing.

It is difficult to locate the true origin of a call after it has been initiated. PSTN calls are designed to work on the principle of forwarding tables and circuit switching. Each time a call is placed, only the destination number is used for routing. It works by establishing individual circuits down a sequence of neighboring switches until it ends up at the recipient's terminal. The outbound switch(es) do not necessarily need to know whether the optional caller ID number in the call request header would route back to the caller's terminal. If the outbound switch also serves as the caller's inbound switch, then the TSP could perhaps verify the true owner of the caller ID number from its own records. However, the TSPs do not have a legal obligation to perform any verification, or to share that information with the

recipient, thus, without the cooperation of the caller's TSP, tracing a spam call is almost impossible.

To make matters worse, as spam calls can now be initiated over the Internet, a spammer can further hide behind proxies, VPNs, or Tor networks, or even distribute outbound calls using a botnet, adding even more difficulty in tracing the exact whereabouts of a spammer.

4.1.7 Entrenched Legacy Systems

The PSTN ecosystem has been around for several decades, allowing any phone to reach any other phone through a vast interconnection of switching centers. While the core networks have evolved to be almost entirely carried by an IP-based infrastructure, the signaling protocols have not changed (to ensure legacy compatibility). Even though VoIP is touted as a major revolution of voice communication, the legacy of PSTN protocols will remain for many years to come. Change is difficult when the entire ecosystem must ensure that the majority of legacy systems will work, and therefore wholesale replacement of the core telephony system is a nonstarter. As a result, telephone spammers can exploit the weaknesses in the legacy technology (such as the lack of caller ID verification) to run a successful spamming operation.

4.1.8 Lack of Effective Regulations

Unfortunately, there is also a lack of incentive for the industry to participate in the anti-spam effort. Unlike email and Internet traffic where the peering model [38] incentivizes the Internet service providers to reduce the load of spam traffic on their systems, telephony service providers profit from the spam-generated traffic and intercarrier compensation fees. Most players (phone number collectors, lead sellers, telephony service providers, and backbone carriers) in the PSTN ecosystem profit from telephone spam, except the consumer. Although TSPs may benefit in other ways by reducing telephone spam (for instance, in

better public relations or charging spam-filtering service as a fee), there exists, at least, a minor monetary disincentive.

Further complicating matters, the current United States law ensure that TSPs are immune from liability for servicing spam calls [65] under the Telephone Consumer Protection Act of 1991, which means that they cannot be held liable for servicing spam calls. Classified as common carriers, *TSPs have an obligation to move all phone traffic with no exceptions* [66]. Therefore, it is difficult to implement anti-spam solutions at the most natural place: the TSP who has a direct view of the telephony network.

4.1.9 Lack of Globalized Enforcement

In the United States, a number of laws and regulation exist at both the federal and state levels, such as making robocalling illegal (with some exemptions) [12], making caller ID spoofing illegal (with some exemptions) [67], and the establishment of a national Do-Not-Call Registry [68]. The FTC is also interested in stopping telephone spam, and they have held numerous competitions to combat robocalling [69]. Despite resolute efforts by the US government, robocalling and caller ID spoofing is still an unsolved problem. Technology and globalization have resulted in *telephony networks shifting from a national ecosystem to a global ecosystem*. With the use of VoIP service, a telephone spammer can cheaply distribute outbound calls from an overseas location. Because the spammers lie beyond the jurisdiction of US law enforcement authorities, it is hard for law enforcement to prosecute those spammers for breaking the law. Effective control of telephone spam would require cross-border enforcement. However, cross-border jurisdiction of telephone spam has yet to catch up with the present technology, and many countries would have no incentive to cooperate with US regulatory and enforcement agencies.

4.2 Basic Techniques and Countermeasures

There are various anti-spam techniques are used to prevent telephone spam. To identify the state-of-the-art in preventing telephone spam, we gathered existing techniques from academic, industry, SPam over Internet Telephony (SPIT), and Internet domain. The techniques can be systematically grouped into the following classes: (1) Call Request Header Analysis, (2) Voice Interactive Screening, and (3) Caller Compliance.

4.2.1 Call Request Header Analysis

Call Request Header Analysis is a category of techniques that filters calls based on the header information associated with the call request. For instance, the caller ID is a popular type of request header information that can be used to analyze a call. The effectiveness of Call Request Header Analysis depends on the accuracy of the information collected, which could be severely impacted when spoofing or omission is possible.

Caller ID Blacklisting rejects a call if the caller's phone number (captured from caller ID or Automatic Number Identification service) appears on a blacklist, otherwise, calls from all other phone numbers are accepted. This can be used to block spam calls by blacklisting phone numbers that are known to be spamming, and the recipient's terminal would silently block all phone calls from those phone numbers without disturbing the recipient. Caller ID Blacklisting only blocks phone numbers that are explicitly added to a blacklist, hence it tends to be permissive to all other callers. As caller ID service has become ubiquitous in all telephone services, Caller ID Blacklisting does not face compatibility issues. Caller ID Blacklisting is easy to implement and requires very little computational resources, and it is a common feature in modern smartphones[70, 71]. However, a blacklist must be well populated to be effective against spam, therefore compiling a comprehensive list would not

be scalable for the recipient. A spammer could defeat Caller ID Blacklisting by spoofing any number not known to be blacklisted, hence it is not effective against most forms of call request header manipulation.

Caller ID Whitelisting only accepts calls from phone numbers that appear on a whitelist, otherwise, calls from all other phone numbers are rejected. This can be used to block spam calls by whitelisting phone numbers that are known to be trusted, and the recipient's terminal would silently block phone calls from all other phone numbers without disturbing the recipient. Caller ID Whitelisting is easy to implement and requires very little resources, and it is easy to find implementations on modern smartphones[72, 73]. Caller ID Whitelisting blocks all calls that are not added to a whitelist, and does not need to be well populated to be effective against spam, hence it is quite scalable for the recipient when defending against spam. It is usually quite easy to populate a whitelist, as the numbers could be derived from the recipient's contacts list. However, unknown legitimate callers would always get blocked in Caller ID Whitelisting. A spammer could defeat Caller ID Whitelisting by spoofing the caller ID of a number known to be trusted by the recipient, however this is more difficult without prior knowledge about the recipient's whitelist.

Caller Reputation System uses reputation or trust associated with a caller's phone number to determine if the caller is a spammer. A Caller Reputation System maintains and publishes reputation scores associated with individual callers, in which the reputation scores are computed based on various caller-related information such as recipient black/whitelists [74–77], caller behavior [75, 78, 79], recipient behavior [74, 80, 81], caller's domain reputation [76, 82], social connections [80, 83–86], and recipient feedbacks [74, 75, 77, 82, 87, 88]. There are also many opportunities to improve a Caller Reputation System by developing better scoring algorithms. The Caller Reputation System can be used to filter spam

calls by configuring the recipient's terminal to block calls from callers associated with poor reputation. A Caller Reputation System generally requires a large amount of data, which are usually crowdsourced from many recipients, and the data would need to be curated by an administrative third party. It would also require frequent maintenance to ensure quality and freshness of data in order to be effective. However, large scale collection of personal information could be at risk of violating privacy. Caller Reputation System could be vulnerable to Sybil attacks, where a malicious caller obtains multiple identities to gain a large influence over its own (or other caller's) reputation. Because the reputation of a caller is associated with the caller's phone number, a spammer could defeat the Caller Reputation System by spoofing the caller ID to a number with a good reputation. A malicious caller could also sabotage someone by deliberately making junk calls while spoofing the caller ID number, such that the victim gets a poor reputation.

Caller Behavior Analysis uses the call behavioral features associated with a caller's phone number to determine if the caller is a spammer, using behavioral features such as call count/velocity [75, 79, 85, 89–95], call duration sum/mean/variance [75, 85, 90–92, 94–96], call rejection count/ratio [81, 85, 90, 92, 93, 95, 97, 98], recipient diversity count/ratio [90, 91, 95, 98], invalid recipient count/ratio [85, 93–95, 97], repeated call count/ratio [91, 98], outbound-to-inbound ratio [79, 94, 97, 99, 100], simultaneous calls [92], and caller's domain behavior [78, 97]. There are also many opportunities to improve the technique by developing better classification algorithms. Acquiring the caller's behavioral information usually requires participation from the caller's telephony service provider or a honeypot of telephones [79, 81]. If not required by regulation, it is usually not in the TSP's business interest to report on or impose a call behavior restriction on their callers. The callers' behavioral information would need to be updated frequently to ensure accuracy and freshness in order to be effective. Large scale collection of callers' call behavior could

also face privacy issues and numerous obstacles from legal regulations. Because the call behavior of a caller is associated with the caller's phone number, a spammer could defeat the Caller Reputation System by spoofing the caller ID to a number with good calling behavior. Furthermore, a spammer could hide its illegitimate call behaviors by using multiple caller identities.

Device Fingerprinting collects a variety of metadata from the call request header for the purpose of creating a device fingerprint of a caller's terminal. Device fingerprinting improves the accuracy of determining the caller's identity by using only a set of information that meets the properties of diversity and stability. Device Fingerprinting has been proposed for SPIT prevention by blacklisting or whitelisting the device fingerprints of SIP-based terminals [101]. However, in PSTN, device fingerprint information is a scarce resource. This is due to the little amount of header information in PSTN call requests compared to SIP or email, resulting in having too little workable information for device fingerprinting to work effectively.

Caller ID Anomaly Detection searches for anomalous patterns in the caller ID, such as invalid format, invalid number, unavailable number, toll-free number, area codes, regular expression, to determine if the caller is a spammer. Caller ID Anomaly Detection is quite easy to implement and requires very little computational resources and, therefore, is easy to find in several call blocking apps [102, 103]. Caller ID Anomaly Detection does not track information associated with any individual caller, instead, it looks for general patterns in the caller ID that can be used to differentiate spammers and legitimate callers. As Caller ID Anomaly Detection tend to find matches more broadly, it tends to be easier to manage and maintain. However, some patterns may be potentially prone to false negatives, and therefore may restrict some legitimate callers, such as VoIP users or privacy enabled callers. A

Description	Decimal	ASCII	Hex
Message Type (MDMF)	128		80
Message Length	33		21
Parameter Code (Date & Time)	1		01
Parameter Length	8		08
Month (November)	49	1	31
	49	1	31
Day (28)	50	2	32
	56	8	38
Hour (3pm)	49	1	31
	53	5	35
Minutes (43)	52	4	34
	51	3	33
Parameter Code (CPN)	2		02
Parameter Length (10)	10		0A
From (6062241359)	54	6	36
	48	0	30
	54	6	36
	50	2	32
	50	2	32
	52	4	34
	49	1	31
	51	3	33
	53	5	35
	57	9	39
Parameter Code (Name)	7		07
Parameter Length (9)	9		09
Name (Joe Smith)	74	J	4A
	111	o	6F
	101	e	65
	32		20
	83	S	53
	109	m	6D
	105	i	69
	116	t	74
	104	h	68
Checksum	88		58

Table 4.1: MDMF Message Sample in the Existing POTS Protocol

spammer could defeat Caller ID Anomaly Detection by carefully crafting the caller ID to not trigger any known anomalous patterns.

ANI-CPN Matching checks whether the Calling Party Number (CPN) captured by the caller ID service matches with the Automatic Number Identification (ANI) number captured by the ANI service [104]. Automatic Number Identification service [105] is a separate type of calling line identification service that can capture the calling number information even when the caller ID is not presented. It was originally designed to obtain the calling party's billing number from a local exchange carrier to any interconnecting carrier for billing of long distance calls. In most cases, the billing number is the same as the CPN, and usually when a mismatch happens it is likely due to caller ID spoofing, or the caller is calling from a private branch exchange (PBX). ANI-CPN Matching assumes that a legitimate caller's CPN matches the ANI number whereas a malicious caller would spoof the CPN which results in a mismatch. However, ANI service are usually not made available to regular consumers (usually only offered to 800 toll-free, 900 premium-rate, or 911 emergency service lines), therefore, only some businesses would benefit from this technique. ANI service is also not always reliable at capturing the caller's ANI number. Placing a legitimate call using an outbound VoIP service or a calling card service would result in a non-working or a generic ANI number being captured. As a result, false positives may frequently occur which hinders user acceptance.

ANI-II Filtering can be used to filter spam calls by blocking certain types of origin service captured by the ANI-II service. ANI-II [106] is an extension of the ANI service that identifies the type of service associated with the originating switch. Each type of service is represented by a two-digit code. ANI-II Filtering assumes that legitimate callers would have a valid (00 or 61) ANI-II code, whereas, malicious callers would be making VoIP

calls that would have an invalid ANI Failure (02) code, and therefore should be blocked. However, with the growing use of VoIP service by regular consumers, this technique could potentially result in too many false positives if all calls with ANI Failure codes are blocked. Only some businesses would benefit from an implementation of this technique, as ANI-II service is usually offered only to premium-rate, toll-free, or emergency lines. Therefore, this technique would not be accessible or cost effective for the regular consumers.

4.2.2 Voice Interactive Screening

Voice Interactive Screening is a category of techniques that forces the caller to interact with a voice input-based interactive system and decide if the call is spam after analyzing the caller's interaction. The system either requires active or passive interaction from the caller. An active interaction system relies on the caller providing a response to a specific task which requires some effort from the caller, whereas a passive interaction system silently gathers the caller's response without explicitly informing the caller. Voice Interactive Screening techniques do not need to rely on the caller ID or any other call request header information, hence they are generally not vulnerable to caller ID spoofing. However, Voice Interactive Screening techniques generally require processing of audio signals, which tends to be more complex to implement. Because these techniques can only work *after* recording a length of the caller's voice, all Voice Interactive Screening techniques have a screening period, therefore, would introduce additional delay to the caller. Due to the recording of the caller's voice during the screening, in the US, some states require explicit consent of recording the conversation, which could hinder the screening process or invoke privacy fears from some legitimate callers. As telephone audio can be manipulated, and tends to contain artifacts such as background noise, network dropouts, or compression

losses, Voice Interactive Screening techniques are generally more prone to errors.

Audio Fingerprinting uses the voice recording of the caller, or audio features extracted from the voice recording of the caller, to analyze for similarity to a set of known spam call profiles. If the voice recording is similar to an audio stream of a known spam profile, then the call is classified as spam. Audio Fingerprinting has been proposed to combat replayed voice spam in several works [107–113]. However, the performance of Audio Fingerprinting depends on the completeness of spam profiles, which is usually not feasible for a recipient to collect. Audio Fingerprinting would usually require a thirty-party to continuously collect and maintain the known-spam audio profiles to ensure effectiveness. However, a spammer could potentially defeat the mechanism by dynamically creating variations of the spam audio message (such as adding audio artifacts or using personalized messages) to avoid identification.

Speech Content Analysis first records the caller's voice, then makes use of speech recognition technology to transcribe the voice into text. The text is then analyzed with text profiles of known spam calls to classify if the call is spam. As opposed to managing audio recordings, a corpus of text data is usually much easier to manage. As many spam calls are simply variations of a call script, a keywords-based classification model could be used against variations of a same type of spam [114]. However, the effectiveness of this technique depends on the accuracy of speech recognition, and of course the effectiveness of the classification model. In practice, automatic speech recognition of telephone voice is an ongoing research problem [115], which tends to be prone to errors, and still has several years to go to reach human-level performance [116].

Acoustic Pattern Analysis extracts distinguishing acoustic patterns from the caller's audio stream, such as signal losses [117], peak uniformity [117], noise uniformity [117], voice activity [118, 119], and double talks [118–120], to determine if the call is spam. Audio Fingerprinting looks for general patterns in the audio signal that can broadly distinguish spam calls from legitimate calls. Unlike Audio Fingerprinting and Speech Content Analysis, Acoustic Pattern Analysis does not require a large collection of known-spam profiles, which could be difficult to gather and maintain. However, some patterns may be prone to false positives and could be easily defeated with manipulation of the audio stream.

CAPTCHA/Turing Test is an interactive challenge-response technique that requires the caller to complete a reverse Turing test to determine whether the caller is a human or robo-caller. The tests are designed to be difficult for a computer but easy for a human to complete. For instance, the test could ask the caller to key in what they hear from a distorted audio stream of random numbers [121–123]. However, CAPTCHA/Turing Test would need to be careful not to discriminate against certain groups of people, such as people with poor English or disabilities, while not giving too much leeway for abuse by "decaptcha" systems [124]. On the other hand, CAPTCHA/Turing Test would also need to be careful not to be illegible even for users with no handicaps, as the legitimate caller may become irritated by the obstacles of initiating a call with the recipient. Because CAPTCHA/Turing Test is highly interactive, it tends to require a high degree of effort, and cause significant delays to the caller.

4.2.3 Caller Compliance

Caller Compliance is a category of techniques that require the caller to first satisfy a compliance requirement prior to or during a call request. If the caller is able to satisfy the

compliance requirement, then the caller is allowed to communicate with the recipient. Satisfying the requirements should be easy for a legitimate caller but difficult (or costly) for a spammer. Some compliance measures require special changes made to the call setup process or to the communicating terminals. Some techniques require prior instructions given to the caller.

Do Not Call Registry simply provides a registry of phone numbers that spammers are legally prohibited from calling in most circumstances. The spammer may be subject to substantial fines if they fail to comply. The registry is usually maintained by the national government, in the US [68], the list is maintained by the Federal Trade Commission. However, the recipients would need to actively provide feedbacks for the government to legally act on spammers violating the law. The Do Not Call Registry can act as a good deterrence for domestic law-abiding telemarketers, however, it would have little effectiveness on spoofed numbers and overseas spammers.

Graylisting [125] first rejects the initial call request from a caller and then accepts the next call request from the same caller made within a short period of time. This technique defends against autodialers that simply call a list of phone numbers and do not make repeated call attempts. The technique also assumes that if an uninformed (about the defense) caller is calling about legitimate business, the caller will try again. The implementation is simple and does not require changes to the infrastructure. However, the legitimate caller must make two calls for every call request, which introduces additional delay and calling cost. A spammer could easily defeat the Graylisting mechanism by configuring the autodialer to automatically call again if a call goes unanswered, but at the cost of higher phone bills and reduced efficiency.

Consent-based Communication first requires the caller to send a consent request to the recipient before initiating a call. For instance, the request could be a forwarded greeting message where an answering machine first records the name spoken with the caller's voice and then plays it to the recipient [126–128]. The recipient then decides whether to accept the caller's request to communicate. If the call is spam, the recipient is only limited to being exposed to an abridged recording (or the request message) of the spam call. However, the recipient is still disturbed for every unconsented caller, therefore it is not scalable, and the recipient is not spared from the disturbance of a spam call. It also adds delay to each call, as legitimate callers are forced to wait for consent before each call.

Call Back Verification first rejects an initial call from a caller, then forces the caller to wait for the recipient to call back the caller. Call Back Verification is a good defense against caller ID spoofing, as it forces the caller to provide a genuine caller ID. The basic mechanism is simple, and some implementations try to automate this process [129, 130]. However, it requires the caller to first own a reachable inbound number, which could restrict communication from legitimate VoIP users and telephone extension terminals. Call Back Verification also add delays to each communication, as the legitimate caller must wait for the recipient to call back. Calling back could also add calling cost on both the caller and recipient in PSTN, which can be especially significant for premium or international numbers.

Weakly Secret Information requires the caller to demonstrate knowledge of a weakly secret information before allowing communication with the recipient. Weakly secret information could be in various forms such as a passcode, an extension code, a limited-use phone number, or a message identifier [131]. However, the recipient would first need to share the weakly secret information to all trusted callers, hence it may not be scalable for

a recipient with a large contact list. Legitimate calls from unknown callers would also be restricted from communicating with the recipient.

Payment at Risk is a micropayment, cost-based, technique where the caller is required to deposit a small amount of money before making a call. If the recipient reports that the call is spam, then the deposit is confiscated or kept by the recipient, otherwise, the money is refunded to the caller. This was proposed as a method for SIP spam prevention [84]. This technique prevents spamming by making it prohibitively expensive to send out a large amount of spam calls, while costing very little for legitimate callers. However, the solution requires a universal micropayment system that collects payment on every call, which may require significant resources to create and administer. There also are many questions regarding the legality of this approach, for instance on the lawful confiscation of payments and abuse of spam reporting. The value amount of the deposit would also affect the number of recipients needed to report on the spam caller to effectively make spamming unprofitable.

Proof of Work is a computational, cost-based, technique where the caller's terminal is required to produce a proof-of-work, such as hashcash [132], that is moderately hard to compute (being computational or memory-bound) but easy for the recipient to verify, before allowing communication with the recipient. As the amount of work increases, it would be prohibitively inefficient to distribute large amounts of spam calls. A legitimate caller would not be significantly affected when making a few phone calls. On one hand, Proof of Work has an advantage over Payment at Risk by not requiring a micropayment system, therefore avoiding the administrative and legality issues. On the other hand, Proof of Work faces a trade-off problem between permissiveness and anti-spam effectiveness. In PSTN, due to the significant share of low-end telephone terminals, the difficulty of the work would need

to be low enough to ensure permissiveness. However, this may allow a spammer using moderately powerful computerized terminals to easily generate as much work as needed for spamming. Legitimate callers with high outbound calls, such as a bank, may also be obstructed from doing legitimate business if it is prohibitively costly to generate the proof-of-works to contact a large number of customers.

Proof of Identity requires the caller to send a verifiable identity token that would authenticate the credentials of the caller whenever making a call. This technique has been proposed for SIP domain users [129, 133–135], due to the availability of SSL/TLS certificates and maturity of the underlying public key infrastructure. This technique prevents spamming by ensuring that the caller could be held responsible for making illegal calls, and prevents scams by ensuring that the caller cannot impersonate as someone else. Proof of Identity could also prevent a spammer from using multiple identities when identity verification is required. Proof of Identity has an advantage over Proof of Work by not having the issue of deciding the right difficulty level of proof-of-work which could either obstruct calls from low-end telephone terminals or give too much leeway for spamming. However, the scheme could be hard to deploy in PSTN, as it would require establishment of a certificate authority for issuing and verifying caller identities, and may require significant changes to the call request protocols in PSTN.

4.3 Assessment Criteria

It is clear that there is no shortage of techniques to combat telephone spam, but what would an ideal telephone spam defense entail? Therefore, we propose a set of assessment criteria.

We separate the assessment criteria into three categories: (1) Usability, which evaluates the ease-of-use from either the caller or recipient's perspective, (2) Deployability, which evaluates the ease of installation, deployment, and operation, and (3) Robustness, which evaluates the technique's resilience against errors and effectiveness against a spammer actively evading the defense. We define each of the identified criteria and give a mnemonic name.

4.3.1 Usability Criteria

No-Disturbance-to-Recipient When a known-spam call arrives, the technique does not disturb the recipient, such as prompting for additional action from the recipient.

Scalable-for-Recipient The technique does not increase the burden of work on the recipient with an increasing number of spam calls. The technique can handle a large variety of spam calls with minimal input from the recipient.

Effortless-for-Caller When initiating a call, the technique requires minimal or zero effort from the caller.

Negligible-Changes-to-Call-Setups The technique requires negligible changes to the existing call setups or configurations in the callers' terminals.

Negligible-Delays When initiating a call, the technique adds negligible or unperceivable delay to the caller, other than the typical time to connect and time waiting for the recipient to answer the phone.

Permissive-for-VoIP-Callers The technique would not restrict any legitimate calls that use VoIP service. For instance, some outbound-only VoIP users (such as Skype) tend to have a generic (or unavailable) caller ID number and cannot receive incoming PSTN calls.

Permissive-for-Unknown-Callers The technique would not restrict calls from a legitimate caller not known by the recipient.

4.3.2 Deployability Criteria

Negligible-Changes-to-Infrastructure The technique requires zero or negligible changes to existing PSTN protocols, terminals, or infrastructure.

No-Third-Party-Involvement The technique does not require a third-party. A compromise of the third-party would not result in mishandled calls or in a breach of privacy.

Low-Resource-Requirement The technique is lightweight and does not require a significant amount of resources (e.g., people, equipment, engineering, or funding) to initiate and deploy.

Low-Maintenance The technique requires low maintenance, in terms of administrative cost, time, or resources, to maintain good working order.

Negligible-Cost-per-Call The technique adds negligible cost to each call, taxed on the legitimate caller, recipient, third-party, or carriers. The cost could also be indirect, such as reduced efficiency or capacity.

4.3.3 Robustness Criteria

Effective-Against-Dynamic-Caller-ID-Spoofing The technique is robust even when the spammer spoofs different caller IDs nondeterministically.

Effective-Against-Targeted-Caller-ID-Spoofing The technique is robust even when the spammer spoofs a specific caller ID known to be trusted by the recipient.

Effective-Against-Unavailable-Caller-ID The technique is robust even when the spammer makes the caller ID unavailable or sends a faulty caller ID to cause errors.

Effective-Against-Multiple-Identities The technique is robust even when the spammer initiate calls from multiple sources, such as using multiple subscriber accounts or a telephone botnet, to disseminate spam calls. This is different from caller ID spoofing where the caller IDs are not necessarily spoofed but are instead initiated from different sources.

Effective-Against-Answering-Machine-Detection The technique is robust even when the spammer uses Answering Machine Detection technology, which is a feature in autodialers that can distinguish human pick-ups from answering machines. With AMD, an autodialer can be configured to call again later if the call was not answered by a human, or to deliver the audio message into the recipient's voicemail.

Effective-Against-Dynamic-Audio-Content The technique is robust even when the spammer uses an autodialer capable of personalizing or altering the audio messages for different recipients. This is usually featured in autodialers that are able to synthesize text to speech.

We evaluate each technique using the criteria proposed in Section 4.3 and Table 4.2 visually summarizes this evaluation. Each technique is evaluated as either satisfying the criteria (denoted as ●), may satisfy the criteria (denoted as ◐), or not satisfying the criteria (denoted as ○). "May satisfy the criteria" means that the technique can be made to satisfy the criteria depending on the implementation or configuration, while some implementations do not fully satisfy the criteria.

Of course, this analysis requires some opinion, and in each case, we evaluated each technique and criteria to the best of our abilities. While others may disagree with the exact assessment of each technique, we believe that the criteria outlined in Section 4.3 will help to guide future telephone spam defenses and to provide a framework to evaluate these defenses.

4.4 Combining Techniques

From analyzing all the standalone techniques, it is clear that there is no single technique that can satisfy all the criteria. No technique is a complete solution to the spam problem, and each has trade-offs between usability, deployability, and robustness. Therefore, an improved anti-spam system would look to combine different techniques, to leverage the positives and compensate the negatives. We outline the different ways in which a solution could use a combination of standalone techniques.

Phased Decisions combine several techniques into a linear sequence (i.e., a pipeline process) of decision stages. If an earlier technique determines the call is spam, then it may not be necessary to run the evaluation techniques at later stages. This is suitable for combining techniques that use information that are obtained chronologically, such as first using Call Request Header Analysis, followed by Voice Interactive Screening. We found the use of Phased Decisions approach in related works by Niccolini and Quitek et al. [142, 143],

		References	Usability						Deployability						Robustness					
			No-Disturbance-to-Recipient	*Scalable-for-Recipient*	*Effortless-for-Caller*	*Negligible-Delays*	*Permissive-for-VoIP-Callers*	*Permissive-for-Unknown-Callers*	*Negligible-Changes-to-Infrastructure*	*Negligible-Changes-to-Call-Setups*	*No-Third-Party-Involvement*	*Low-Maintenance*	*Low-Resource-Requirement*	*Negligible-Cost-per-Call*	*Effective-Against-Dynamic-Caller-ID-Spoofing*	*Effective-Against-Targeted-Caller-ID-Spoofing*	*Effective-Against-Unavailable-Caller-ID*	*Effective-Against-Multiple-Identities*	*Effective-Against-Answering-Machine-Detection*	*Effective-Against-Dynamic-Audio-Content*
Call Request Header Analysis	Caller ID Blacklisting	[70, 71]	●	◐	●	●	●	●	●	●	●	◐	●	●	○	○	○	◐	●	●
	Caller ID Whitelisting	[72, 73]	●	●	●	●	●	○	●	●	●	◐	●	●	●	○	○	●	●	●
	Caller Reputation System	[74–88, 136]	●	◐	●	●	●	◐	●	●	◐	○	◐	●	◐	○	○	◐	●	●
	Caller Behavior Analysis	[75, 78, 79, 81, 85, 87, 89–100, 137, 138]	●	●	●	●	●	◐	●	●	◐	○	◐	●	◐	○	○	◐	●	●
	Device Fingerprinting	[101]	●	●	●	●	●	◐	●	●	◐	○	◐	●	◐	○	◐	◐	●	●
	Caller ID Anomaly Detection	[102, 103]	●	●	●	●	●	●	●	●	◐	◐	●	●	◐	○	◐	○	●	●
	ANI-CPN Matching	[104]	●	●	●	●	◐	●	●	●	●	●	●	◐	◐	○	◐	○	●	●
	ANI-II Filtering	[104]	●	●	●	●	◐	●	●	●	●	●	●	◐	◐	○	◐	○	●	●
Voice Interactive Screening	Audio Fingerprinting	[107–113]	●	●	◐	○	●	●	●	●	◐	○	○	◐	●	●	●	●	◐	◐
	Speech Content Analysis	[108, 114]	●	●	◐	○	●	●	●	●	◐	◐	○	◐	●	●	●	●	◐	◐
	Acoustic Pattern Analysis	[117–120]	●	●	◐	○	●	●	●	●	◐	◐	◐	◐	●	●	●	●	◐	◐
	CAPTCHA/Turing Test	[121–123]	●	●	○	○	●	●	●	●	●	◐	◐	◐	●	●	●	●	●	●
Caller Compliance	Do Not Call Registry	[68]	◐	◐	◐	●	●	●	●	●	○	◐	◐	●	○	○	○	◐	●	●
	Graylisting	[120, 125]	●	●	◐	○	●	◐	●	◐	●	●	●	○	●	●	●	●	●	●
	Consent-based Communication	[126–128]	○	○	◐	○	●	◐	●	◐	●	●	●	◐	●	◐	●	●	●	●
	Call Back Verification	[129, 130]	◐	◐	◐	○	◐	◐	●	●	●	●	●	○	●	●	●	●	●	●
	Weakly Secret Information	[131]	●	◐	◐	◐	●	○	◐	◐	●	◐	●	●	●	●	●	●	●	●
	Payment at Risk	[84]	●	○	◐	●	●	○	○	○	○	○	○	○	●	●	●	●	●	●
	Proof of Work	[132, 139–141]	●	●	●	●	●	○	◐	○	●	●	◐	●	●	●	●	●	●	●
	Proof of Identity	[129, 133–135]	●	●	●	●	●	○	◐	○	◐	●	◐	●	●	●	●	●	●	●

●= satisfy the criteria ◐= may satisfy the criteria ○= does not satisfy the criteria

Table 4.2: Evaluation of Various Standalone Techniques Against the Criteria Described in Section 4.3

Schlegel et al. [144], Gritzalis and Mallios [145, 146], and Azad and Morla [85].

Weighted Scoring combines several techniques by running each technique individually and then combining the outputs to produce a final score by applying a weighted scoring method. The classification of whether the call is spam is based on the final score. As Weighted Scoring need to collect outputs from all standalone techniques, it is suitable for combining techniques that can be performed simultaneously, such as the various standalone Call Request Header Analysis techniques. We found the use of Weighted Scoring approach in related works by Dantu and Kolan [147], Niccolini and Quitek et al. [142, 143], Schlegel et al. [144], Hansen et al. [148], and Mathieu et al. [149].

Conditional Procedures combine several techniques based on a predefined set of rules (i.e., a policy or an algorithm). This allows for higher flexibility of combining the techniques, such as using a different sequence of standalone techniques based on the preference of each recipient or the reputation of each caller. We found the use of Conditional Procedures approach in related works by d'Heureuse et al. [150], Dritsas et al. [151], Scata and La Corte [152], and Soupionis and Gritzalis [93].

We evaluate existing solutions using a combined approach and summarized which standalone techniques those solutions incorporated in Table 4.3. All of these works are mainly focused on defense against SPIT, and some of these may include SPIT-specific techniques that do not appear in our table. Again, this analysis requires some opinion, and we evaluated each solution to the best of our abilities. We believe that the various strategies of combining techniques outlined in Section 4.4 will help to improve future telephone spam defenses.

	[142, 143]	[144]	[145]	[146]	[147]	[148]	[149]	[150]	[151]	[152]
Phased Decisions	✓	✓	✓	✓						✓
Weighted Scoring	✓	✓			✓	✓	✓	✓		
Conditional Procedures								✓	✓	✓
Caller ID Blacklisting	✓	✓	✓	✓	✓	✓	✓		✓	✓
Caller ID Whitelisting	✓	✓		✓	✓	✓	✓		✓	
Caller Reputation System	✓		✓	✓	✓	✓			✓	✓
Caller Behavior Analysis	✓	✓	✓		✓	✓	✓		✓	✓
Device Fingerprinting										✓
Caller ID Anomaly Detection										
ANI-CPN Matching										
ANI-II Filtering										
Audio Fingerprinting				✓						
Speech Content Analysis	✓									✓
Acoustic Pattern Analysis	✓									
CAPTCHA/Turing Test	✓	✓	✓	✓		✓		✓		✓
Do Not Call Registry										
Graylisting	✓					✓	✓			✓
Consent-based Communication	✓		✓							✓
Call Back Verification										
Weakly Secret Information	✓									
Payment at Risk										
Proof of Work	✓							✓		
Proof of Identity			✓	✓					✓	

Table 4.3: Summary of Various Anti-Spam Solutions Using a Combination of Standalone Techniques

4.5 Related Work

While in this chapter we have compared and analyzed the state-of-the-art research in telephone spam defense, we will now discuss related survey papers. Most of the papers focus on spam in the Voice over IP (VoIP) domain, so-called SPam over Internet Telephony (SPIT), rather than the larger PSTN telephony network.

Keromytis [153, 154] presented two comprehensive surveys of VoIP security, which summarized previous works related to VoIP security and organized them according to an extended version of the VoIP Security Alliance (VoIPSA) Threat Taxonomy. The papers reviewed many previous works addressing every type VoIP threat in the VoIPSA taxonomy, with the social threats of spamming as one of the categories.

Baumann et al. [155] presented a survey of potential solutions to SPIT. The paper provided an overview and classification of SPIT prevention methods based on detection using Signaling versus Voice and order-based Before Call versus After/While Call. The paper also proposed a Biometric Framework for SPIT Prevention as a way to bind identities to each caller.

Phithakkitnukoon et al. [156] presented a survey focused on five primary types of VoIP attacks, SPIT being one of them. The authorized provided an introduction to the basic knowledge of VoIP systems and its available security tools, and summarized a list of proposed solutions for SPIT from previous literature.

Quinten et al. [157] presented a survey evaluating the techniques to prevent and reduce SPIT. The authors evaluated the effectiveness of techniques by dividing them into four categories: unsuitable techniques, techniques with potential, suitable techniques, and combinations of techniques.

Dantu et al. [158] presented a survey discussing the attacks and solutions in VoIP, with VoIP Spam and Phishing being one of the attacks. The authors reviewed previous work

addressing all types of VoIP attacks and proposed a high-level security architecture to make the VoIP infrastructure more secure and robust.

Dritsas et al. [159] presented a survey reviewing a list of SPIT identification criteria that can be used by anti-SPIT mechanisms and identified the different detection stages. The authors propose two generic categories of SPIT identification criteria: SIP Message criteria and SIP User Agent criteria. They also proposed a two-fold evaluation framework for discovering possible SPIT messages.

Marias et al. [160] presented a survey assessing the threats and vulnerabilities that the SIP protocol introduces. The authors also reviewed existing anti-SPIT mechanisms and classified them into three classes: Prevent, Detect, and Handle. The paper also proposes a list of qualitative and quantitative criteria to assess the effectiveness of the anti-SPIT countermeasures.

Khan et al. [161] presented a survey reviewing various existing methods for preventing spam in IP telephony. The paper also presented a discussion on the implementation costs of different types of techniques and commented that no single technique is sufficient and therefore a framework of multiple techniques is recommended.

Rosenberg et al. [162] presented an open memo reviewing various solutions that might be possible to deal with SIP spam. The author also presented some borrowed techniques that have been employed to deal with email spam. In conclusion, the author recommends using identity related techniques, while also commented that identity techniques may be vulnerable when a SIP request without an authenticated identity cannot know whether the request lacked such an identity because the originating domain didn't support it, or because a man-in-the-middle removed it.

In general, most existing survey papers focus on techniques against SPIT or more specifically spam in the SIP protocol. This chapter is focused on techniques to address spamming in the PSTN telephony network. Some techniques for SPIT are not applicable

to PSTN due to protocol differences. As far as we are aware, this is the first survey paper specifically addressing spam calls directed to the PSTN telephony network. In terms of evaluation differences, we are the first to propose a taxonomy to classify the existing standalone techniques into three categories, the first to evaluate the standalone techniques based on three sets of assessment criteria, and the first to outline the three strategies of combining standalone techniques.

4.6 Conclusion

From analyzing and evaluating the existing solutions that attempt to address telephone spam, we reach the conclusion that there is no universally acceptable solution to telephone spam. Every approach thus far has different tradeoffs, specifically between usability, deployability, and robustness.

From our analysis of the telephone spam ecosystem and defensive techniques, we believe that usability is the most important criteria for evaluating a defense. Unlike email, which can be delayed or possibly lost due to a false positive, telephony solutions have a high bar for user acceptance. Telephone users will not adopt techniques that impose an excessive burden on both the caller and recipient. Therefore, future research into this area must consider the usability of the defense from both the caller and the recipient perspective.

We believe that one promising avenue of research is using a combination of techniques, which should improve the robustness of standalone techniques, and potentially each technique could address the weaknesses of the others. However, as the telephony system has real-time immediacy constraints, care must be taken so that the combination of techniques will not degrade the user experience due to higher complexity. Our intuition leads us to recommend combining no more than two standalone techniques, as we observed that a good balance of usability, deployability, and robustness could be achieved by using two standalone techniques.

One glaring issue that continually reoccurs when analyzing the telephone spam ecosystem is caller ID spoofing. We believe that the key to combating telephone spam is to make the caller ID trusted and verifiable, while making minimal changes to existing infrastructure. For instance, from our evaluation of Call Request Header Analysis techniques, they provide the best overall usability and deployability, however, they suffer from robustness due to the spammer's ability to spoof the caller ID. If caller ID spoofing can be effectively prevented, then we believe that Call Request Header Analysis would satisfy all of our evaluation criteria.

Telephone spam is poised to increase significantly, defrauding consumers of billions of dollars. Therefore, an effective telephone spam defense is critical. However, the techniques and approaches in combating email spam are inappropriate when applied to telephone spam. We attribute this to differences not only in the technology used but more fundamentally to the key challenges in dealing with telephone spam. This is why a survey of the telephone spam defenses is necessary: to highlight these differences and to define an ideal criteria. This chapter provided a framework to help guide and shape future telephone spam defenses.

Chapter 5

PROPOSING AUTHENTICATED CALLER ID TRANSMISSION

With a broad understanding of the current state of the art from the survey research, the next phase of the research plan is to propose a countermeasure toward preventing telephone spam. As we have discussed in the previous chapters, one glaring issue in telephone spam is caller ID spoofing. In the current calling line identification presentation scheme, the caller ID is trivially spoofed. Telephone spammers are able to use spoofed caller IDs to trick their victims into answering unwanted calls and defeat a variety of countermeasures. To provide a solution to this problem, this chapter will analyze the fundamental causes of caller ID spoofing and, by analyzing the root cause, design an authentication scheme that addresses the aforementioned fundamental security flaws for the current caller ID scheme. The key idea of this proposal is to help prevent users from falling victim to phone impersonation scams by using a security indicator, as well as provide a foundation for future defenses to stop unwanted calls based on the real caller ID information.

The main contributions of this chapter are the following:

- We describe how caller ID spoofing works and the reasons that lead to its prevalence.
- We propose a caller ID authentication solution that results in the display of a security indicator during the incoming phone call and describe why it matters for the user.
- We describe the design of the underlying authentication and verification mechanism behind the security indicator.
- We provide a discussion on the security considerations for the deployment of the authenticated caller ID infrastructure.

This chapter contains prior work from the ITU Kaleidoscope 2016 [56], the IEEE Communications Standards Magazine 2017 September Issue [163], a USPTO non-provisional patent application [164], and a pending technical standards contribution at the study group 11 of the ITU Telecommunication Standardization Sector.

5.1 The Rise of Caller ID Spoofing

The caller ID is a generic name for a supplementary service offered by the called party's telephone company that provides the calling party's telephone number to the called party's user equipment during an incoming call. It helps the called party to decide whether to answer a call based on the caller's phone number, and, to call back the caller if the call could not be answered. Since its introduction in the 1990s, the caller ID service has now become ubiquitous in almost every telephone service. Today, the caller ID number is also used in other telephony services, such as SMS and MMS, and, with the prevalence of smartphones, many smartphone apps and services also rely on the caller ID for identification.

The core process of providing the caller ID is known as Calling Line Identification Presentation (CLIP), which was first defined in ITU-T Recommendation Q.731.3 [43] for the Signaling System No. 7 (SS7) network in 1993. The SS7 network is the backbone infrastructure for most of the world's public switched telephone network (PSTN) telephone calls. Even as the telephone backbone moves towards being carried by an IP packet-based infrastructure, Q.731.3 still plays a major role in providing the caller ID for telecommunications and will continue to do so for many years to come.

In all major existing call signaling protocols (SS7, H.323, and SIP), caller ID is either provided by the originating exchange or by the calling party. In SS7 and SIGTRAN (IP version of SS7), caller ID is defined by the calling party number (CPN) parameter, where the parameter is an optional part of the Initial Address Message (IAM). The IAM is sent to the destination exchange as part of the basic call procedures according to Q.764 [27]

to initiate a call. The IAM routes through transit exchange switches until it reaches the destination exchange of the called party, in which the called party's local exchange carrier would convert and retransmit the CPN to a specific caller ID format for the called party's user equipment during the incoming call setup process, e.g. mobile or landline.

However, because the PSTN was traditionally regarded as a closed trusted network, the caller ID scheme was designed with little security in mind. The telephone network relied upon trust in switch operators to play by the rules. In the process of providing the caller's telephone number, the originating exchange can arbitrarily declare what caller ID number is sent on a call-by-call basis.

Traditionally, a caller would need to gain control of an SS7 switch to have the capability to customize the caller ID. In consumer telephony services, their caller IDs are typically managed by the caller's telephone carrier, preventing general users from spoofing the caller ID. It was also prohibitively expensive for individuals and small businesses to gain switch level access to the SS7 network, which kept the number of people with caller ID spoofing capability small.

However, with the recent rise of IP access to the PSTN, cheap IP-based client protocols (such as SIP [165]) are replacing the expensive traditional bulk telephone services (such as ISDN). Cheap and accessible Voice-over-IP (VoIP) bulk telephony services are now becoming the norm. Today, the SS7 network is no longer exclusive to traditional telephone carriers, there are many internet telephony service providers (ITSPs) that provide telephony services over an Internet connection. With the popularity of the cloud services business model, access to SS7 switch level capability can be sold as a service and is becoming more available to untrusted parties. Some ITSPs *sell customizable caller ID as a service feature*, along with mass distribution technologies such as voice broadcasting, voicemail broadcasting, and SMS broadcasting, all provided over an Internet connection. With ITSPs, individuals and businesses are no longer limited to telephone services from their local

telephone service providers. With an Internet connection, a malicious caller now has access to a world of ITSPs that can provide features such as caller ID customization/spoofing.

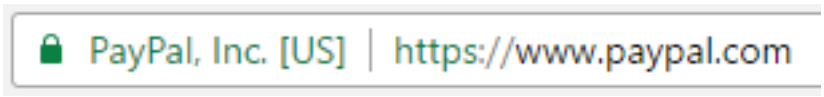
The caller ID is defined by the calling party number (CPN) parameter. The parameter value of the CPN is placed within the optional part of the initial address message. The IAM follows the ISUP (ISDN User Part) message format as defined in Q.763 [37]. The CPN parameter follows a structured binary coding format as defined in Q.763.3.10. The calling party number parameter contains several codes, and the specifics of each code can be found in Q.763 subclause 3.10 and 3.9.

To spoof the caller ID, the caller's originating exchange or the calling party would declare the CPN parameter with false information. In the US and many other jurisdictions, the caller's telephone service provider does not have any legal obligation to ensure that the caller ID number is genuine before it is transmitted. Even in jurisdictions that forbid telephone service providers from providing falsely declared caller ID information, with Internet access to an untrustworthy telephone service provider, it is easy for a malicious caller to start the call request from a different origin, and transmit the false caller ID to the destination exchange of the called party.

At the heart of the issue, there is a lack of authenticity and accountability in the transmission of telephone identities. The PSTN has transformed from a closed trusted ecosystem to a diverse global ecosystem, mutual trust can no longer be relied upon to guard against the abuses of trust in caller ID transmission. Addressing this issue requires the core protocol to provide a mechanism to ensure authenticity and accountability. This is why we advocate for a standardized caller ID authentication scheme. By providing authentication to the caller ID, authenticity and accountability of the caller ID can be assured. However, for viable deployment of authenticated caller ID transmission, it requires mutual interoperability. Therefore, standardization is the key to building a telephone ecosystem that could rely on the assurance of caller IDs.

5.2 Solution: Security Indicators

To provide a solution to this problem, we drew inspiration from the Internet. The Internet is widely known for its exposure to intrusion and man-in-the-middle attacks from untrusted parties around the world. In such a relatively untrusted environment, solutions were developed to combat the identity spoofing.



(a) An Example of HTTPS Security Icon in Google Chrome

(b) An Example of Authentication Icon in Gmail

Figure 5.1: Examples of Security Indicators in HTTP and Email communication

In the Internet ecosystem, the HTTP and email communication are arguably the most popular types of communication used today. In HTTP communication, the universally recognized padlock indicator with the name of the company displayed in the address bar of modern web browsers (such as the one shown in Figure 5.1a) provides users with immediate trust in the website's domain and entity name identity.

In email communication, the key-shaped security indicator of the email sender (such as the one shown in Figure 5.1b) in some email clients provides the users with immediate trust in the identity of the email sender.

An example of a possible caller ID security indicator for an incoming call is shown in Figure 5.2. The security indicator can be similarly attached to other forms of telecommunication such as SMS and MMS.

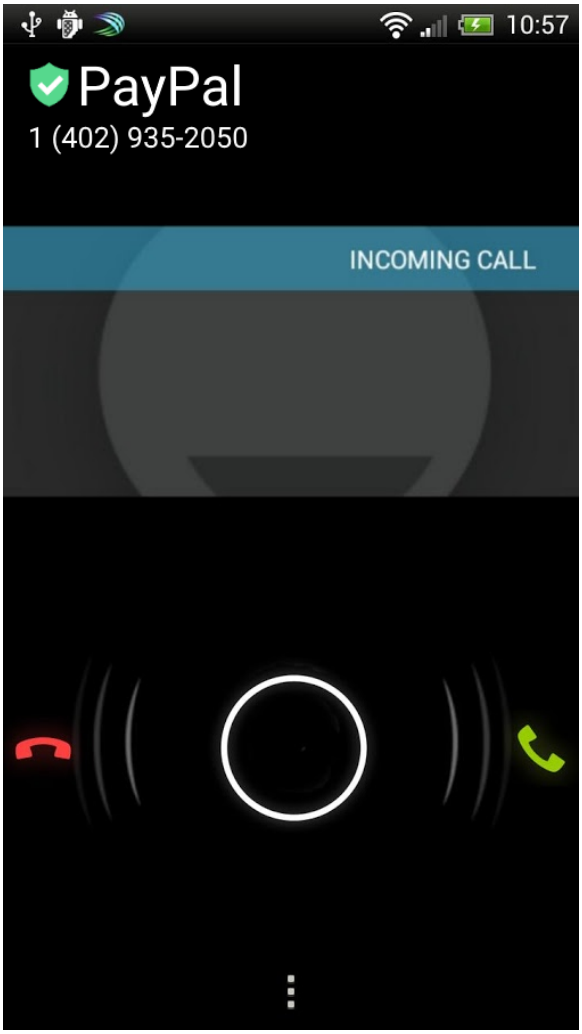


Figure 5.2: Example of the Proposed Caller ID Security Indicator for an Incoming Call

These security indicators are crucial to informing the user that the information is from a verified source. The availability of the security indicator provides an immediate indication of the authenticity of the sender's identity. The recognizability of the security indicator icon provides an immediate understanding of the functionality of the indication. By simply recognizing an icon, users are able to quickly determine if the sender is authentic to protect themselves from phishing and impersonation scams. The prevalence of security indicators promotes awareness that the user should only trust senders that are verified, which would inspire users to be more vigilant of calls and messages from unverified sources.

Having a security indicator for telecommunication would also be an effective solution against telephone spam. Apps and services can be built on top of the security indicator to analyze if a call comes from an untrusted source to more effectively block unwanted callers. With the growing prevalence of phone fraud, calls from billing, banking government, law enforcement organizations would also benefit from providing authenticity of their caller IDs, as their recipients would be certain that the caller is real and not an impostor, therefore feel better assured receiving communication over the phone. Authenticated caller IDs may also be useful for immediate customer identity verification, without relying on (possibly stolen or guessable answers of) security questions to verify the identity of customers. As there are also scam calls that spoof the caller IDs of existing customers, which the malicious callers then trick the institution into emptying their customers' bank account [166].

5.3 The Underlying Caller ID Authentication Scheme

Before we discuss the technical detail of designing the underlying caller ID authentication scheme behind the security indicator, we first present a quick overview of the parties involved in the transmission of a call request.

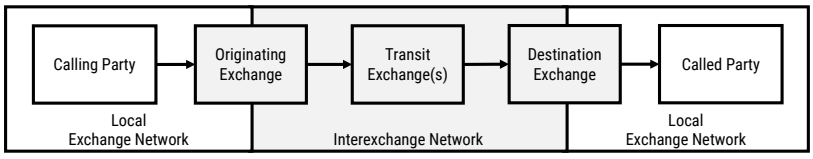


Figure 5.3: Overview of the Parties Involved in the Transmission of a Call Request

Calling Party is the party initiating the call request with an user equipment (UE) or software client that connects with the originating exchange.

Originating Exchange is a switch in the PSTN that generates and transmits the IAM to the destination exchange pertaining to the call request from the calling party.

Transit Exchange is an interconnecting switch in the PSTN that helps to route the messages from the originating exchange to the destination exchange.

Destination Exchange is the terminating switch in the PSTN that receives the IAM and sets up the call with the called party.

Called Party is the party with an user equipment or software client of the intended called party for the call request.

In general, the sequences within a local exchange network define how user equipment interacts with the local exchange carrier during a call setup, and the sequences within the PSTN define how SS7 switches interact with each other during a call setup.

The current caller ID transmission scheme has two fundamental insecurities: (1) a lack of verification and authentication of the declared caller ID and (2) a lack of integrity protection of the transmitted caller ID. The current calling line identification presentation scheme allows the CPN to be declared arbitrarily. There are currently no mechanisms to protect the CPN from unwanted modification during transmission. Even if the caller has proven that she indeed owns that phone number, an actor (perhaps in association with the caller) along the transit link may still intercept and alter the caller ID number.

Therefore, the design principles of a prospective caller ID authentication scheme must address the aforementioned fundamental security flaws: (1) ensuring the caller ID is verified and authenticated (can only be produced by the calling party or the originating exchange) before transmission, and (2) ensuring the caller ID is guarded against unwanted modification during transit. Furthermore, it is crucial that the users of caller ID authentication enjoy the same user experience as before, hence (3) it must also be able to coexist with the existing call control signaling protocols.

When designing an authenticated caller ID scheme, an immediate idea is to model it after the SSL/TLS protocol of the Internet. However, this design, although can be used to secure the caller ID, is ill-suited for the PSTN. The PSTN primarily uses the SS7 protocol

stack to service telephone calls, whereas SSL/TLS was mainly designed to encrypt data communication, which has a significant transport and latency overhead. After establishing an initial end-to-end connection with a TCP 3-way handshake, the SSL/TLS process requires two additional round-trips (4-way handshake) to establish a secure connection. Whereas, in the SS7 call request, this "handshake" is a one-way forward transmission (as shown in Figure 5.4), where the originating exchange sends an initial address message to the destination exchange, to reduce the delays of initiating a call. Implementing the SSL/TLS scheme for SS7 would require all exchange switches to support the multi-way handshake process, which not only require critical changes, it could potentially add significant delays to the call request process. In addition, SSL/TLS is designed for a client-server web environment, which requires the server ("called party") to first acquire a certificate from a certificate authority (CA), whereas, in the PSTN scenario, we are mainly concerned with authenticating the client ("calling party").

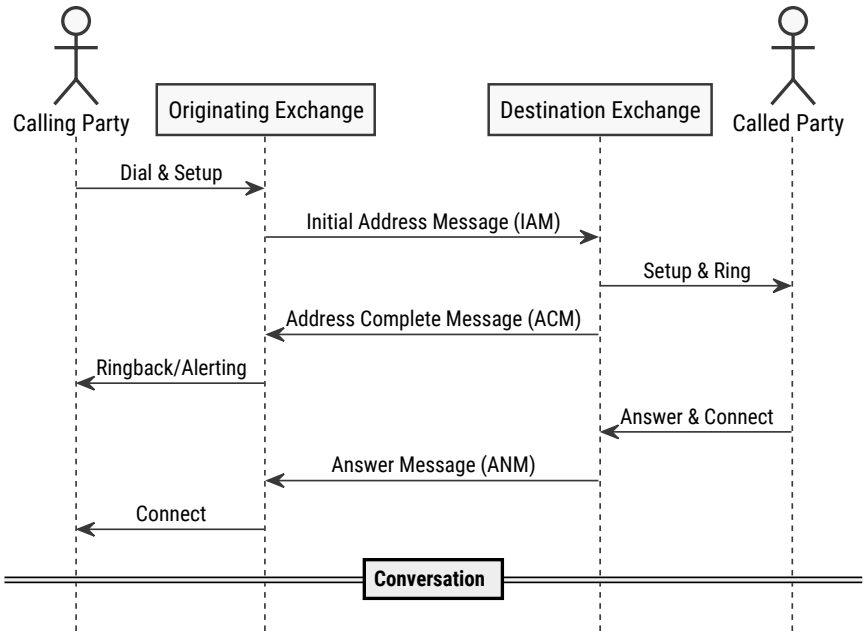


Figure 5.4: Overview of the Existing Call Request Transmission Process

Therefore, we need to design an authentication scheme better suited for the PSTN. Designed as an initial reference, we propose a caller ID authentication scheme, which will guide and shape an authenticated calling line identification presentation process for the SS7 ecosystem.

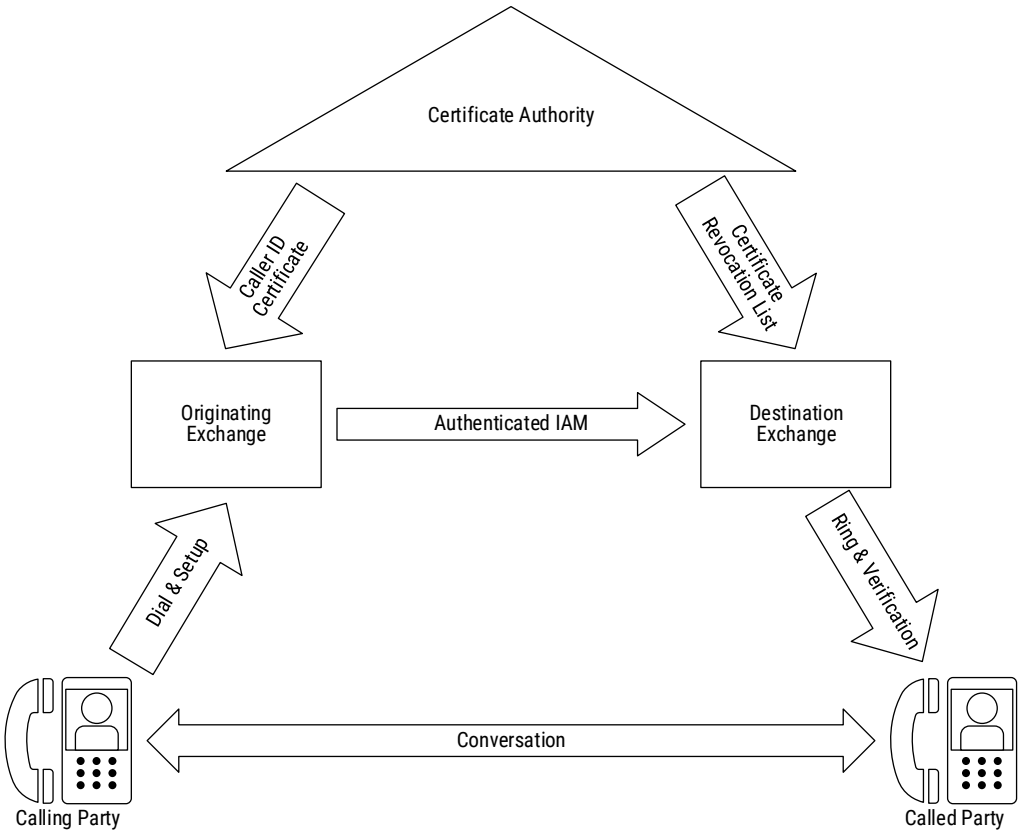


Figure 5.5: Overview of the Proposed Architecture

The high-level idea of the scheme is to introduce a public key infrastructure (PKI) scheme for the PSTN. The high-level architecture of the proposed scheme is shown in Figure 5.5. The scheme will have Certificate Authorities (CAs) certify, issue, and revoke caller ID certificates (CICs) for the calling parties that have proven ownership of their respective telephone numbers. After successfully obtaining the CIC, the calling party's originating exchange can then use the caller ID certificate to generate an authenticated call request, by extending the existing initial address message. Upon receiving an IAM call

request, the destination exchange then checks for the presence and validity of authenticated call request parameters and presents the validated caller ID using a security indicator during the call setup to the called party.

The role of each actor with regard to the caller ID authentication scheme is as follows:

Certificate Authority is an entity in the PSTN that verifies phone number ownership and issues caller ID certificates to a requester that successfully provided proof of phone number ownership. The CA is a trusted third party, trusted both by the calling party and by the called party relying upon the certificate. The CA is also responsible for revoking caller ID certificates if needed.

Calling Party sets up a call request with the originating exchange for the called party. Under the caller ID authentication scheme, the calling party or the originating exchange may initiate a request to obtain a caller ID certificate from the CA.

Originating Exchange obtains and stores the caller ID certificate from the CA for the calling party's phone number. Upon a call request, the originating exchange generates an authenticated IAM on behalf of the calling party and transmits it to the destination exchange.

Destination Exchange receives the authenticated IAM and checks the validity and authenticity of the call request, and it sets up the call with the called party with a security indicator showing the caller ID verification status. The destination exchange may also forward the authenticated IAM to the called party to allow verification to be performed at the terminal for better security.

Called Party receives the call/ring request with a verification status or authenticated IAM. The terminal displays an incoming call with a security indicator.

The processes of the authentication scheme can be logically divided into 2 parts: Caller ID Verification and Authenticated Call Request.

5.3.1 Caller ID Verification

In the Caller ID Verification process, the goal is for a CA to verify a calling party's ownership of a phone number, i.e. the phone number actually routes to the calling party, and then issue a certificate. The verification process can take advantage of the fact that *receiving* a call or message is proof of phone number ownership in the PSTN. In the actual process, more steps are involved to ensure the authenticity the CA's identity and integrity of the certificate request. The calling party/originating exchange will thus need to generate a public-private key pair and store the private key securely. After proving to the CA that the calling party/originating exchange is really the owner of the phone number and the public key, the public key of the calling party is signed by the CA with attributes indicating phone number ownership information, turning it into a caller ID certificate.

In Caller ID Verification, the core process is sequenced as follows:

Prerequisites to the process: (1) the CA's public key P_S is publicly known, and (2) the CA has his private key Q_S.

1. Originating exchange or calling party generates a public-private key pair for the calling party's phone number, P_A and Q_A.
2. Originating exchange sends calling party's phone number From_A and public key P_A to the CA.
3. CA creates an encrypted nonce ENonce_S by first generating a random nonce Nonce_S and then encrypting it with the calling party's public key. ENonce_S = Encrypt(P_A){Nonce_S}. This ensures that only someone with the calling party's private key can decrypt ENonce_S.
4. CA signs the ENonce_S to create a signature $\mathit{ENonce\text{-}Sig}_S$. This is to safeguard the authenticity of the nonce during transmission.
5. CA sends ENonce_S and $\mathit{ENonce\text{-}Sig}_S$ back.

6. Originating exchange verifies the signature *ENonce-Sig*$_S$ to ensure CA's identity.
7. If *ENonce-Sig*$_S$ is verified, the originating exchange decrypts *ENonce*$_S$ with private key Q_A to obtain *Nonce*$_S$.
8. Originating exchange sends decrypted *Nonce*$_S$ to CA, proving that the originating exchange/calling party is really the owner of the public key.
9. CA verifies *Nonce*$_S$ and, if valid, sets a short expiration time *Expiry*$_A$ and generates a caller ID certificate (CIC) for the calling party *CIC*$_A$ by signing the calling party's phone number *From*$_A$, public key P_A, and *Expiry*$_A$ using the CA's private key.
10. CA sends *CIC*$_A$ to the calling party's telephone number *From*$_A$. The phone number should route to the originating exchange or calling party.

A sequence diagram of the Caller ID Verification process is shown in Figure 5.6.

In actual deployment, there can be several CAs, allowing different users, such as in different networks or regions, to verify with an appropriate CA.

With regards to the caller ID certificate format, the certificate could be based on ITU-T X.509 format [167], and the telephone number in the certificate could be based on international E.164 format [168]. The required critical extension field for the X.509 certificate could be as follows (in RFC5280 style [169]):

```
Extensions ::= SEQUENCE {intlPhoneNumber  E.164}
E.164 ::== PrintableString (SIZE (3..15))
```

Although verification of a caller ID can also be done directly by the called party, where the called party can check for the authenticity of a caller ID by simply calling/messaging back the calling party's phone number, which had been proposed in previous works [129, 130], however, this type of scheme add delays to each communication, and is repetitive for each call request. With a PKI certification model, it eliminates the need for a connection-oriented, repetitive callback verification on every call request.

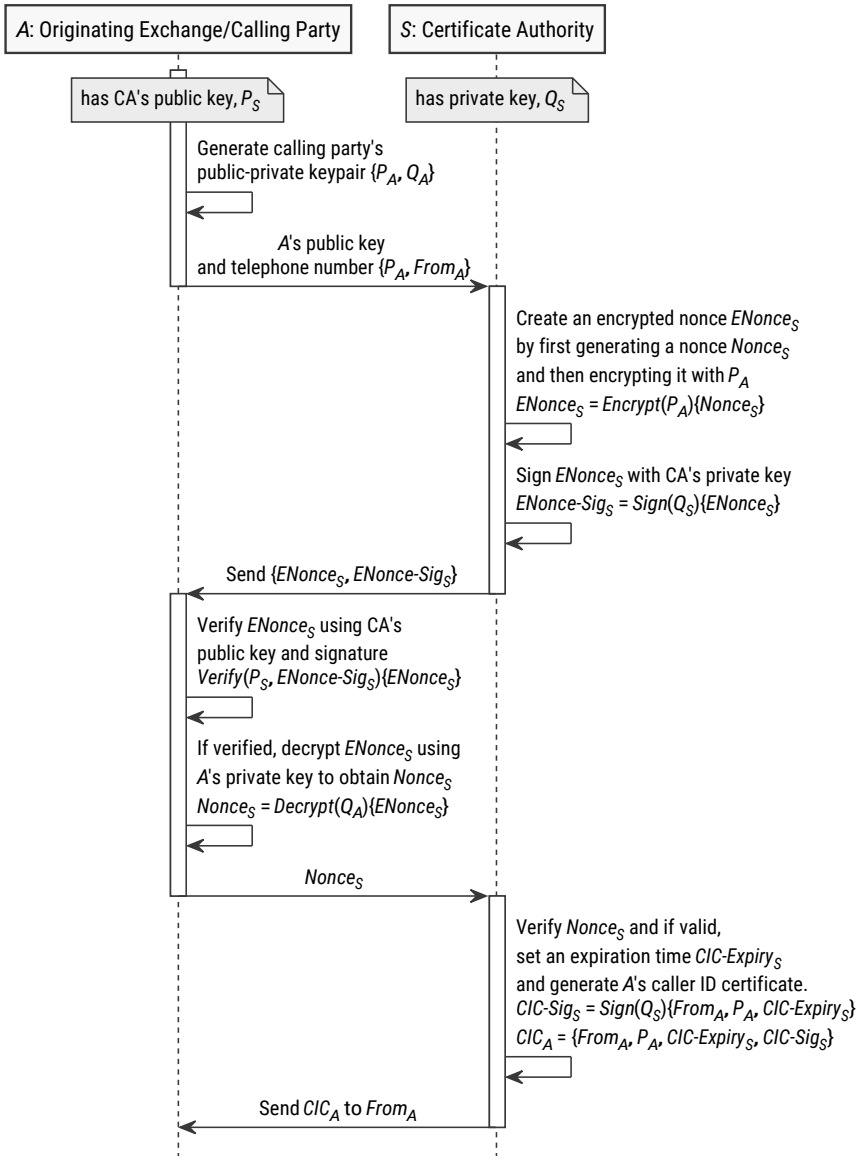


Figure 5.6: Sequence Diagram of the Steps to Obtain a Caller ID Certificate

5.3.2 Authenticated Call Request

In the Authenticated Call Request process, the goal is for a certified calling party to generate an authenticated call request that the called party trusts that the CA has guaranteed that the caller ID really belongs to the calling party. When initiating a call request, the calling party's originating exchange will generate an extended IAM that includes some

additional parameters that authenticate the call request. These additional parameters are designed to prove that the caller ID is authentic, and the request is transient and unique (non-repeatable) to guard against "cut and paste" or replay attacks by a man-in-the-middle or malicious called party. Upon receiving the extended IAM, the destination exchange checks the authenticity and validity of the call request and sets up the call with the called party with a security indicator showing the caller ID verification status. The destination exchange may also forward the extended IAM to the called party to allow verification to be performed at the terminal for better protection against man-in-the-middle attacks.

In Authenticated Call Request, the core process is sequenced as follows:

Prerequisites: (1) the originating exchange has CA's public key P_S, and (2) the originating exchange has caller ID certificate CIC_A and his private key Q_A.

1. Originating exchange generates an IAM for the call request as usual.
2. Originating exchange generates an IAM Signature $IAM\text{-}Sig_A$ by signing all enclosed fields in the IAM along with current the current UTC timestamp $Time_A$. The inclusion of a UTC timestamp ensures that the call request is transient and unique with regards to time and destination, in order to guard against "cut and paste" and replay attacks.
3. Originating exchange attaches the UTC timestamp $Time_A$, IAM Signature $IAM\text{-}Sig_A$, and Caller ID Certificate CIC_A in the optional part of the IAM and sends the extended IAM to the destination exchange.
4. Destination exchange obtains the extended IAM and checks if CIC_A is valid, expired or revoked.
5. If the CIC_A is valid, verify IAM signature against all the enclosed fields.
6. If the IAM signature is valid, check if the UTC timestamp is valid (within a reasonable delay and clock drift), and check if the called party number is correct.

7. Setup the call request with the called party and present a security indicator for the verification result.
8. Destination exchange sends address complete message (ANM) with verification result back to the originating exchange.

A sequence diagram of the Authenticated Call Request process is shown in Figure 5.7.

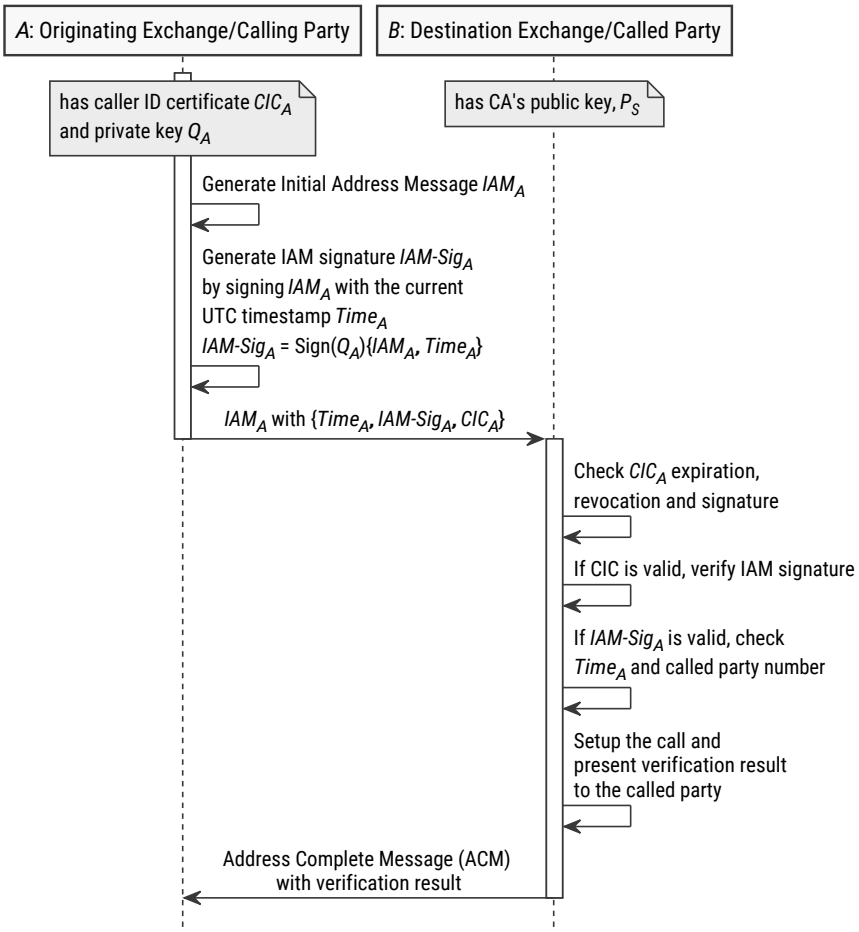


Figure 5.7: Sequence Diagram of the Steps to Initiate an Authenticated Call Request

The existing parameters of the IAM is listed in Q.763 [37] Table 32. The proposed extended IAM parameters could be as follows in Table 5.1.

To ensure transit compatibility, the extended IAM would include a *Parameter Compatibility Information* parameter to instruct the existing transit exchanges to transparently

Parameter	Type	Length (octets)
UTC Timestamp	Optional Part	4-?
Signature Algorithm	Optional Part	1-?
Signature	Optional Part	16-?
Caller Identity Certificate	Optional Part	32-?

Table 5.1: List of Extended IAM Parameters

forward the extended parameters to the destination exchange. The specifics of Parameter Compatibility Information parameter can be found in Q.764 (12/99) section 2.9.5.3.2 [27].

The Authenticated Call Request process does not change the existing one-way process of transmitting the call request using the IAM, which is what enables a call request to be delivered quickly and thus can be implemented without adding perceivable delay to the existing user experience of initiating a call.

To inform the originating exchange that the authenticated call request has successfully pass verification at the destination exchange, we also recommend including a *Request Verification Status* parameter in the optional part of the address complete message to provide a feedback on the verification result. This would be useful for the originating exchange to determine if the extended IAM has been successfully verified by the destination exchange and make corrections if needed.

After the last step, the called party decides whether to answer the call request based on the caller ID and the verification result.

5.3.3 Security Considerations

Even as we outlined the scheme to authenticate the caller ID, we also need to assume that there is a constant threat of malicious actors stealing the caller's identity, such as by mobile phone theft, or using a malware to steal the private key. Furthermore, having a valid

caller ID certificate does not imply that the caller should *always* be trusted. As a critical security measure, the certificate authority must also be able to deal with revocations of a previously issued certificate.

Learning from the pains of revoking certificates on the Internet, where using Certificate Revocation List (CRL) [170] has the disadvantage of distributing bulky lists for large number of revocations, and the alternative Online Certificate Status Protocol (OCSP) [171] has the disadvantage of requiring the receiving party to open a real-time connection with the CA, potentially stalling the communication, therefore, we need to explore a more suitable approach for handling certificate revocations in the PSTN.

With that in mind, first, we recommend using CRL over OCSP when verifying revoked certificates. A phone call is more urgent compared to email and web communication, if a phone call is stalled by the certificate verification process, it will severely affect the user experience. It is important that the authentication scheme does not cause perceivable delays, otherwise, some users may even choose to abandon security verification. CRL has an advantage over OCSP in this regard, because the revocation list can be cached at the destination exchange for immediate verification. The downside of CRL is that it does not receive real-time revocation updates, however, the risks can be mitigated by having the originating exchange or calling party choose to use shorter-term certificates, and by having the destination exchange choose to update the revocation list more frequently.

Second, unlike domain certificates which are typically valid for years at a time, in PSTN, we recommend the CA issue short-term caller ID certificates to limit the expiration period. There are two reasons for having short-term certificates. First, it reduces the risk from a successful theft of the certificate private key or phone number by containing the impersonation threat within a bounded period. Second, it significantly reduces the size of revocation lists as the CA would only need to revoke unexpired certificates within the bounded period. Of course, the downside of having short-term certificates is that caller ID

certificates must be renewed frequently. However, unlike domain certificates which can hours due to DNS propagation delay, caller ID certificate renewals could be completed within seconds as the process of verifying a telephone number can be fully automated. Furthermore, because the number of future certificate renewals is largely predictable, the CAs would be able to pre-adjust the quality of service to meet future demands, and perhaps even pre-generate some caller ID certificates to further improve service efficiency.

Finally, we recommend the CA issue caller ID certificates for conditional usage, limiting the usage to a specific method of contact or capability of the calling terminal, such as by whitelisting/blacklisting features such as SMS, MMS, call forwarding, etc. This further reduces the risk from a caller identity theft by containing the threat to limited methods of contact. For instance, it is unlikely that a customer support phone would need to contact individuals using SMS or MMS, hence, a successful theft of the company's caller identity would force the attacker to use voice when contacting the victims, which could make the scam sound suspicious.

5.3.4 Local Deployment Considerations

As we outlined the process to verify the calling party number at the destination exchange, we also need to consider how the security indicator for the caller ID verification status would be transmitted and presented to the called party.

At the destination exchange, the local exchange carrier would present the caller ID verification status in a local exchange call setup format (e.g., POTS, GSM, SIP, etc.). Hence, for a local exchange network to support the caller ID verification scheme, some type of modification/extension to the local call setup format is required, since each SS7 call request will need to be converted to a local call setup format. An immediate thought is to implement the caller ID verification status as a simple indicator flag/parameter added to the local exchange call setup format. However, there are some risks in such implementation.

We would like to provide some discussion on how a conversion of an authenticated call request should be implemented.

In mobile telephone services, caller ID is typically a parameter within a SETUP message transmitted to the called party's user equipment via an encrypted wireless signal. Assuming that the wireless transmission is well encrypted, a key consideration here is whether the identity of the base station is authenticated. In technologies that provide mutual authentication between the mobile phone and the base station, the presentation can be implemented as a flag indicator parameter, after performing the call verification at the destination exchange. However, in technologies where base station authentication is missing or flawed (such as in GSM), the local exchange network should not use the flag indicator approach, because the verification status flag would be vulnerable to being spoofed by an attacker that could spoof a base station. Instead, the presentation of caller ID verification status should be implemented as a forwarding of the extended IAM parameters, transmitted to the called party, to allow the called party's user equipment to perform verification of the authenticated call request.

In landline telephone services, namely the POTS (Plain Old Telephone Service), caller ID is a parameter within the header message encoded in SDMF (Single Data Message Format) or MDMF (Multiple Data Message Format), transmitted to the called party in FSK (Frequency Shift Keying) signal. Assuming that the connection to the central office exchange is secure (such as from physical protection), a key consideration here is whether the call request header is integrity protected. In POTS, the call request header is potentially vulnerable to "Orange box" attacks, where a malicious caller is able to alter the SDMF/MDMF header with spoofed FSK signals, as a result, the verification status flag would be vulnerable to being spoofed by the malicious caller. Hence, in such cases, the caller ID verification status should also be implemented as a forwarding of the extended IAM parameters to protect the integrity of the authenticated call request.

Therefore, in summary, when implementing the presentation of caller ID verification status to the called party, only in scenarios where (1) the local exchange network connection is secured, (2) the identity of the local exchange carrier is authenticated, and (3) the call request header is integrity protected, should the local exchange carrier implement the presentation of verified caller ID as an indicator flag, otherwise, the caller ID verification status should be implemented as a forwarding of the extended IAM parameters to allow the called party's user equipment to perform verification of the call request.

5.4 Related Works

Peterson et al. [57] recently proposed an identity authentication mechanism for end users that originate SIP (Session Initiation Protocol) requests. The scheme proposes having the SIP proxies generating and inserting a PASSporT object [172] (a type of identity token) in the Identity header of every SIP request. Other than transport protocol and data format differences, the scheme uses a similar identity-token based mechanism in authenticating and verifying the caller identity. However, Peterson et al's proposal requires TLS connection for every communication, for reasons mentioned before, is difficult to adapt to the PSTN.

Reaves et al. [173] recently proposed an in-band modem for executing a TLS-inspired authentication protocol over the voice channel of the conversation. The modem is designed to overcome the challenges of low transmission bitrate due to voice codec and transmission losses. After the in-band modem established a data channel between the two parties over the voice channel, the scheme uses a cryptographic challenge-response based scheme to verify the caller's identity. The scheme can provide strong security guarantees comparable to the TLS. However, the verification process require both parties' telephone terminals to support read-write access and live processing of the voice signals, which would require significant computation power on both parties' telephone terminals. It could also invoke

privacy fears due to voice recording capability, and potentially add significant delay prior to the voice conversation.

5.5 Conclusion

With increasing abuse of PSTN's insecurities from untrusted parties, telephone spam, phone fraud and caller ID spoofing is poised to increase significantly. To ensure a sustainable future for the PSTN, the SS7 is in critical need of an upgrade of its core robustness. This chapter proposes a standardized authentication scheme for the caller ID that enables the possibility of a security indicator for SS7 telecommunication. The goal of this proposal is to help prevent users from falling victim to telephone spam and scams, as well as provide a foundation for future and existing defenses to stop unwanted telephone communication based on the caller ID information.

Chapter 6

IMPLEMENTING PROTOTYPE WITH EVALUATIONS

With a description of the architecture, protocols, and security mechanisms of the caller ID authentication scheme, the final phase of the study is to implement a prototype and evaluate the implementation in terms of performance and user experience. The evaluation result can help us validate the assumptions made in the proposed scheme and inspire future improvements and refinements. In this chapter, we describe the design and implementation of an end-user prototype based on the proposed caller ID authentication scheme. The prototype was implemented using the Android framework for Android smartphones. After implementing the prototype, we conducted a performance analysis and an end user study based on the participants' behavior data and feedback. With the user study, we were able to better understand how to design telephone security indicators that meet the users' expectations and learn what improvements can be made for the future iterations of end-user implementation.

The main contributions of this chapter are the following:

- We design an end-user prototype and implement the proposed security indicators and caller ID authentication scheme.
- We analyze the performance of different design variants and present our analysis for future performance improvements.
- We conduct a user study by collecting user behavior and feedback and present our findings from interacting with the users.
- We provide a list of recommendations for the future implementations of authenticated caller ID transmission.

6.1 Prototype Design

In the previous chapter, we described the proposed security indicators and caller ID authentication scheme based the adding new components to the existing SS7 protocols and infrastructure. We initially wanted to develop a prototype implementation of the SS7 infrastructure itself as that would provide the best demonstration of our idea. However, after carefully studying what is needed to develop a prototype implementation for the SS7 infrastructure, we eventually realized that such implementation would unfeasible for a Ph.D. student project. As every project has time and resource constraints, implementing anything directly on the SS7 infrastructure would require tremendous time and money. The core customers of the SS7 infrastructure are telephone service providers, with only a small circle of huge multinational telecommunications equipment vendors supplying the entire global market of telecommunications providers. If we were to implement a prototype on the SS7 infrastructure, we would need to acquire the equipment and set up a wired or wireless telecommunications network similar to a telephone service provider. The telecom is a capital-intensive industry, a telecom company typically invests millions of dollars to set up equipment on their network. Although we did not specially ask for how much a telecom equipment costs, e.g. for just an SS7 gateway, we estimated that it would be at least tens of thousands of dollars. Not only that, even after acquiring the equipment, it does not guarantee that we will be able to make any changes to their proprietary firmware. Hence, we need to look for other feasible ways to implement the proposed scheme.

As a viable alternative, we implemented a prototype to demonstrate the caller ID authentication scheme that requires no modification to existing protocol standards or infrastructure. While an ideal deployment would be to embed the authenticated call request data in the SS7 IAM itself, we implemented a version of the authenticated call request that relies on out-of-band delivery using SMS instead of the IAM. Both the IAM and SMS

use the CPN as the destination address to deliver information. As long as we can extract the authenticated call request from the SMS before or during the call request, we can still implement the caller ID authentication scheme with the same features.

In the vast ecosystem of telephone users, the vast majority are mobile phone users. In many countries, mobile subscribers already surpassed the national population. Today, Android smartphone users represent the vast majority of new mobile subscribers [174]. Globally, Google's Android operating system runs on about 80% to almost 90% of the smartphone marketshare [175].

Based on the described caller ID authentication scheme, we developed an Android app, named "Hamout", for the caller to verify the caller ID to the CA and transmit the authenticated call requests upon an outbound phone call. We implemented another Android app, named "Caller ID Verifier", for the recipient to perform verification of the authenticated call request and display a security indicator upon an inbound phone call. For the caller, the app first communicates with the CA to obtain the caller ID certificate. Upon a call request, the app generates an Authenticated Call Request and delivers it to the recipient using SMS. For the recipient, the app automatically extracts and verifies the Authenticated Call Request from the SMS and displays a security indicator during the ring on an incoming call. To implement the CA, we also developed a web and SMS server to provide the Caller ID Verification service for the caller terminals. Both the caller's and recipient's Android app are preinstalled with the CA's public key certificates. Each Android app is signed, which prevents modifications to the app code and the embedded public key certificates.

We will go into detail how each process is performed in the prototype implementation:

6.1.1 Caller ID Verification

Caller ID verification allows the CA to verify the caller's phone number and provides the caller with a caller ID certificate (CIC) for making future authenticated calls upon

successful verification. The Caller ID verification process is only needed if the caller does not already have a valid CIC, such as a first time user. In our prototype implementation, when running the Caller's app, it detects whether a valid CIC is present in the app storage, and, if not, it will initiate the registration process by automatically reading the device's phone number and, in case the user has a phone number that is not directly associated with the device (e.g., a Google Voice number), optionally prompting the user to enter their phone number. Figure 6.1 shows the phone number submission activity. When the user selects OK, it generates an elliptic curve (NIST curve P-256 aka prime256v1) keypair, and then securely sends the phone number and the public key to the CA over HTTPS. These actions implement steps 1 and 2 of the proposed Caller ID Verification process as described in the previous chapter.

When the CA server receives the request from caller's app, it first generates a 24-bit nonce and pins it to the caller's phone number and public key for an expiration time of 300 seconds. The nonce, caller's phone number, and public key are stored in a self-expiring database. The CA then encrypts the nonce with the caller's public key. The encrypted nonce is replied back to caller's app as a HTTPS response which ensures the response message is secured and signed. When the caller's app receives the HTTPS response from the CA, it then decrypts the nonce. After that, it sends a new HTTPS request to the CA with the decrypted nonce and waits for the caller ID certificated to arrive in SMS. These actions implement steps 3 to 8 of the proposed Caller ID Verification process.

When the CA server receives the decrypted nonce, it looks up the caller's public key and phone number using the nonce. The CA generates the caller ID certificate (CIC) by signing the phone number, public key, and Expiry time using the CA's private key. The CA then sends the CIC to the caller's phone number using SMS. Figure 6.2 shows the activity that waits for the Caller ID Certificate to arrive in SMS. The app tries to obtain the Caller ID Certificate from the SMS automatically, if that failed, the user can manually copy

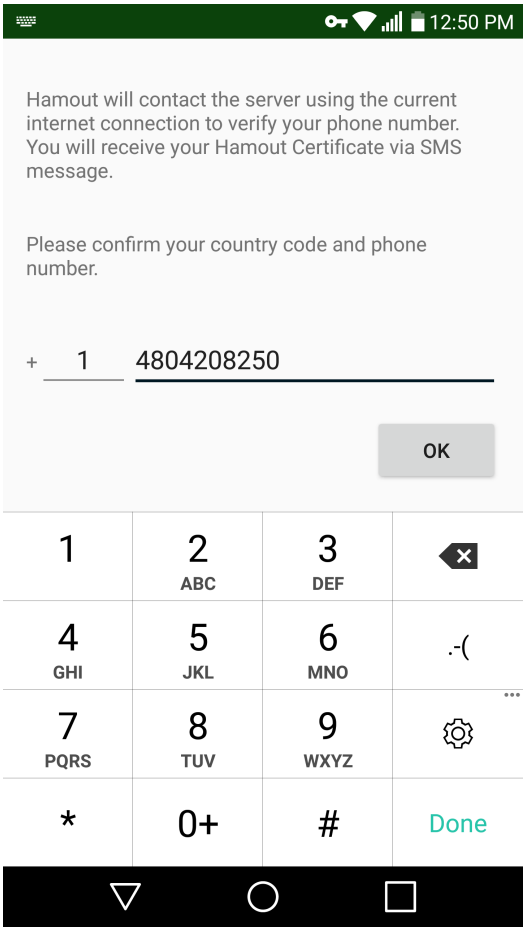


Figure 6.1: Registering a Phone Number and the Public Key With the CA

and paste the Caller ID Certificate into the app. When the user selects OK, it verifies the Caller ID Certificate signature, and if successful, the app is ready to make Authenticated Call Requests in the next activity. These actions implement steps 9 and 10 of the proposed Caller ID Verification process.

6.1.2 Authenticated Call Request

Authenticated Call Request allows a certified calling party generate a transient and unique call request designed to prove that the caller ID is authentic. In our prototype implementation, after the caller has successfully obtained a caller ID certificate in the Hamout

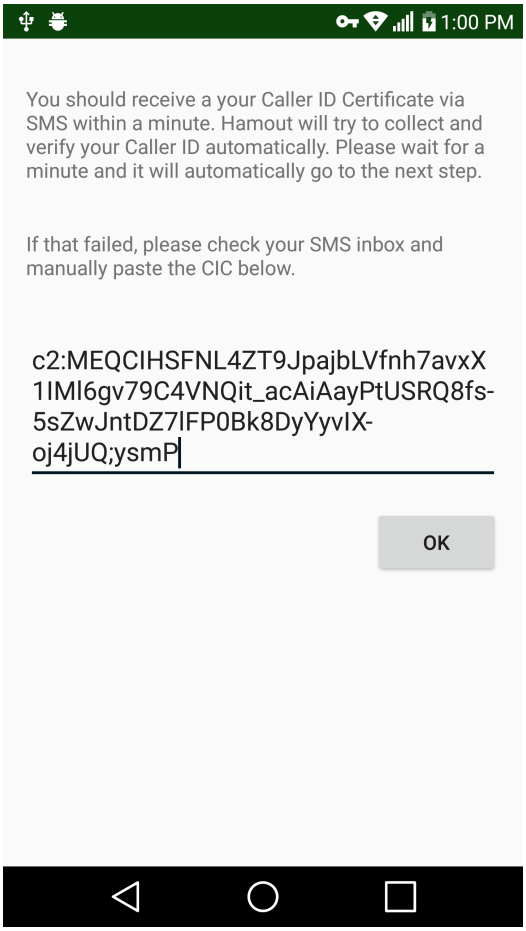


Figure 6.2: Receiving the Caller ID Certificate From the CA

app, the user can make a call with Authenticated Call Request. Figure 6.3 shows the activity that the user can make a phone call with Authenticated Call Request. The user just needs to enter a phone number that he/she would like to call. When the "CALL" button is selected, the app automatically generates an authenticated call request and sends it to the recipient's phone number via SMS. The SMS message is encoded in Base64, and sent in multi-part SMS. After sending out the SMS, it automatically initiates a phone call to the recipient's phone number after a brief delay.

For the recipient, we implemented a separate app called "Caller ID Verifier". The app is designed to process and verify the Authenticated Call Request from the caller, and display

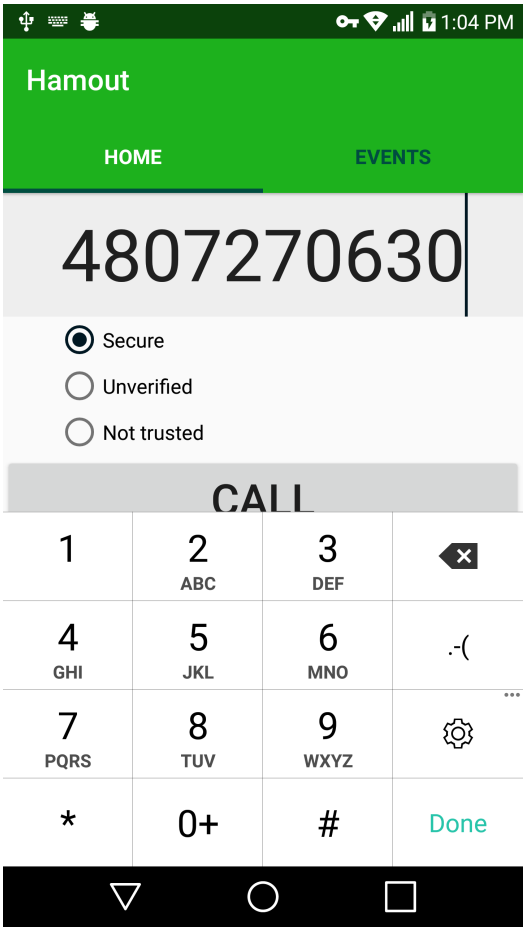


Figure 6.3: Making a Call With an Authenticated Call Request

a security indicator upon an incoming call. During first run, the app presents a tutorial to help the users get familiar with the security indicators. Figure 6.4 shows the tutorial activity.

After the user is familiar with security indicators, the Recipient's app runs in the background which has an event hook for incoming SMSes. When the Recipient's app receives an SMS message with the Authenticated Call Request, it automatically extracts and stores it in a self-expiring cache. The Authenticated Call Request of the caller will be retrieved from the cache when the incoming call arrives.

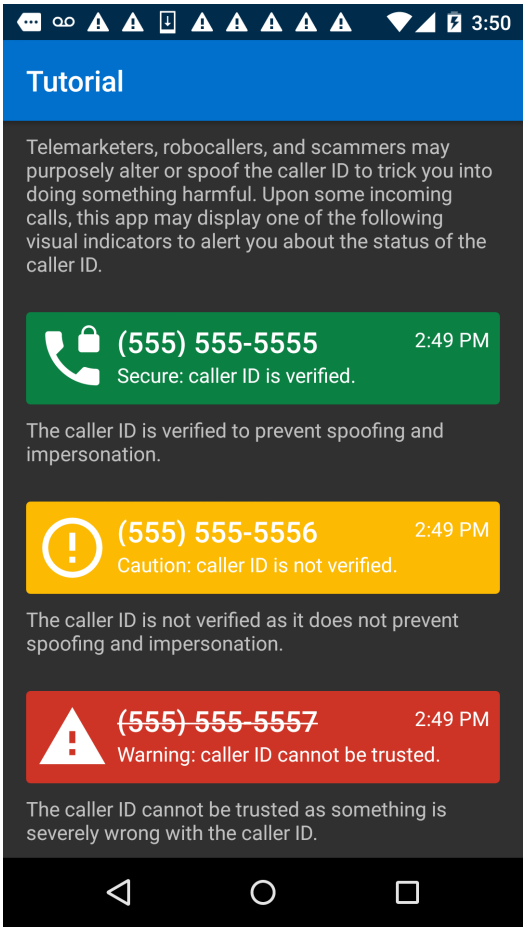
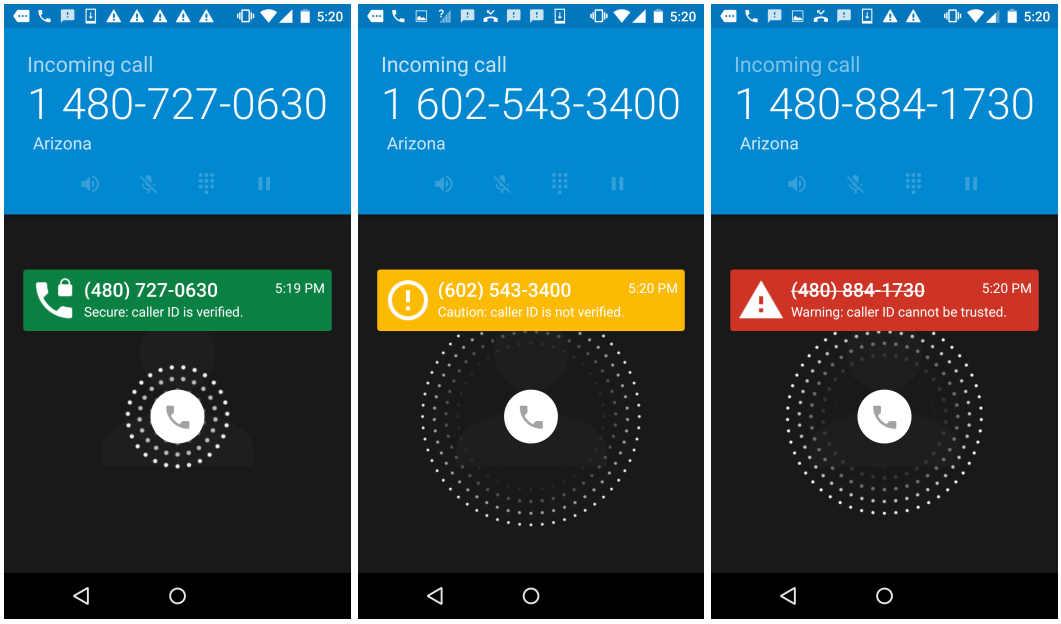


Figure 6.4: Tutorial of Security Indicators

Upon the incoming call, the app first retrieves the Authenticated Call Request of the caller from the self-expiring cache. After completing the verification process, one of three types of security indicators would be shown: Secure, Caution, and Warning. The security indicator display an overlay window over the incoming call screen. Figure 6.5 shows the security indicators overlayed on the incoming call screen. If the ACR is in the valid format and is verified for the caller ID certificate owner, then a secure indicator is displayed upon the incoming call, as shown in Figure 6.5a. The secure indicator informs the recipient that the caller ID is verified. If the ACR is missing (not received from the caller), then a caution indicator is displayed upon the incoming call, as shown in Figure 6.5b. The caution

indicator informs the recipient that the caller ID is not verified. If the ACR is invalid for any reason (such as invalid signature), then a warning indicator is displayed upon the incoming call, as shown in Figure 6.5c. The warning indicator informs the recipient that the caller ID cannot be trusted.



(a) Secure indicator (b) Caution indicator (c) Warning indicator

Figure 6.5: Types of Security Indicators Displayed During an Incoming Call

The designs of the security indicators were kept consistent, which consists of an icon, the caller ID/caller name, a text description of the caller ID status, time stamp, and a colored background. To ensure design consistency, the colors and icons were based on Android material design. When designing the text descriptions, the caller ID status messages were inspired from Google Chrome's HTTPS text description of the domain certificate status.

Security Provider	AndroidKeyStore 5.0				SpongyCastle 1.52	
Signature Algorithm	SHA256withECDSA		SHA256withRSA		SHA256withECDSA	
ECC Curve	prime192v1	prime256v1	-		prime192v1	prime256v1
Key Size	192 bit	256 bit	1024 bit	3072 bit	192 bit	256 bit
Avg. Key Pair Generation Time	999.724 ms	1002.18 ms	1719.42 ms	23247.54 ms	424.88 ms	608.09 ms
Avg. ACR Sig Generation Time	14.51 ms	20.14 ms	35.85 ms	621.02 ms	402.39 ms	602.54 ms
Avg. ACR Sig Verification Time	7.2 ms	9.55 ms	5.81 ms	6.16 ms	525.16 ms	646.58 ms
ACR Size	71 bytes	87 bytes	143 bytes	399 bytes	71 bytes	87 bytes
ACR + CIC Size	217 bytes	265 bytes	484 bytes	1240 bytes	217 bytes	265 bytes
Median SMS Delivery Time	8579 ms	11480 ms	18610 ms	39762 ms	-	

Table 6.1: Performance Testing Results of the Prototype Implementation

6.2 Performance Analysis

To understand the runtime performance and suggest future performance improvements to the implementation of the caller ID authentication service, we conducted a systematic test of various cryptographic algorithms for the authenticated call request and their respective SMS delivery times.

The Android framework and its underlying Java ecosystem provide a multitude of cryptographic providers. Therefore, we measured the empirical run-time performance of different cryptographic algorithms that are potential candidates for future app implementation. We also measured the data sizes of authenticated call request generated by different signature algorithms, to help us evaluate which type of Authenticated Call Request is small enough to be delivered before the incoming call. We used an LG G3 which uses Snapdragon 801 processor on T-Mobile USA network as our benchmarking setup, and we performed the experiments by varying the security providers and signature algorithms. The results of our performance analysis are shown in Table 6.1.

From the benchmark table, we can see that using ECDSA brings in tremendous performance advantage over RSA. In comparison with RSA, Android's ECDSA performed better in almost every operation, except for signature verification where the difference is unnoticeable at less than 4 milliseconds.

Comparing different security providers of ECDSA, AndroidKeyStore [176] (Android's stock security provider) performed much better than SpongyCastle [177] (an open source security provider based on BouncyCastle), only requiring about 20 milliseconds for signature generation and 10 milliseconds for signature verification. Key pair generation using AndroidKeyStore's ECDSA implementation took about 1 seconds compared to about 600 milliseconds in SpongyCastle's implementation, which is perceivably longer, however, we consider it as acceptable in real-world usage because in the Hamout app the key pair generation is only performed once during caller ID certificate registration. Because signature generation and verification are the most common operations in our scheme, based on the benchmarked results, using AndroidKeyStore's ECDSA implementation would provide the best overall performance. The operation times are averages of 5 tests for each case, with very little variances observed.

Furthermore, we are able to show that the ECDSA ACR fields would indeed fit within the current call request headers, including POTS where the header message has the lowest size limit of 256 bytes. From the test results, using SHA256withECDSA with prime192v1 curve with safely fit the fields within an MDMF header message in all cases. This shows that the proposed ACR fields could be made compatible with existing call request specifications for universal PSTN adoption.

We also tested the SMS delivery delays using two Android phones running on two separate GSM networks in the United States (AT&T and T-Mobile). Using the two separate networks avoids in-network routing, and provides a better representation of the real world performance. We have also tuned the caller's SMS send rate such that we are reasonably

sure that network throttling did not affect the delay duration. However, there are still many other uncontrollable variables, such as traffic volume, equipment performance, wireless interference, etc., that could affect the SMS delay. For each experiment, we sent out no less than 250 SMSes containing both the ACR and CIC, and recorded their delivery times. After running the tests, we found that the SMS delivery delay durations were not as consistent as we thought. The SMS delivery times tend to have long-tail outliers. This finding was surprising as we thought the SMS would be consistently delivered in near instantaneous time. In reality, only about 94% falls within a narrow range of delay, the outliers have delays that can be measured in minutes. Hence, we have decided to show the median instead of average delay, in order to provide a much more representative number for our tests.

From the SMS performance test results, we can see that the SMS delivery delays makes the biggest impact on the user experience, adding about 8-12 seconds delay to the authenticated call request process. Because of this long consistent delay, the app will need to add a long wait time after sending the ACR to the recipient. This long delay is generally unacceptable in the real-world, as this meant the caller may have to wait at least 8 seconds to ensure the authenticated call request is delivered. Because of this issue, we felt that it is important to provide our recommendations to solve this problem in the future production versions of the authenticated caller ID transmission scheme.

We propose three solutions to solve the SMS delivery delay issue. The first solution is to use MMS instead of SMS. MMS was introduced in the third-generation (3G) mobile phones and has a theoretical data transfer rate up to 750X the rate of SMS transmission. The transmission rate is much faster due to the introduction of faster network technology, such as 3G HSPA and 4G LTE. However, the main issue with MMS is that it currently costs 1-3x more to send a MMS message compared to SMS in the US. Furthermore, MMS also does not solve the fundamental issue of insistent delays from the carriers.

The other alternative solution is to use the Cloud Messaging. Cloud Messaging a service that uses the Internet infrastructure to facilitate messaging between app servers and client apps, either over WiFi or cellular data network. Google's Firebase Cloud Messaging (FCM) is an example of a cross-platform cloud messaging provider [178]. The cost of cloud messaging is much cheaper compared to MMS, with Google's FCM, the cost is free for low usage and scales cheaply with high volume. Theoretically, a cloud message could be delivered as fast as WiFi or cellular data speeds. However, we do not yet know if this solution would significantly reduce the ACR delivery time. In the worldwide production version of the caller ID verifier app, we believe cloud messaging would be the most scalable approach for delivering the ACR.

Ultimately, the goal of caller ID authentication scheme is to integrate the ACR in the fields within the call setup header itself, making it almost indistinguishable from the existing calling procedures. The long-term vision is to make authenticated caller ID transmission an essential part of the call request process. In the future, every important call should be authenticated. From the table, we are able to show that the ACR fields can indeed fit within all existing call request headers using SHA256withECDSA with prime192v1 curve. The ACR requires as low as 71 bytes and 217 bytes in the most conservative implementation, therefore, can safely fit within call request headers such as SS7/ISUP, UMTS, LTE, and including POTS where the header message has the lowest size limit of 256 bytes. However, this solution requires standardization and modification to the SS7 infrastructure.

6.3 User Study

To understand the user behavior when they see the security indicators upon an incoming call and learn about the potential use cases of caller ID authentication, we conducted a systematic user study of user behavior by using the caller ID verifier app to collect the

users' call actions and feedback responses. Furthermore, the responses collected from the participants can be used to inspire future improvements.

6.3.1 Study design

Our approach to designing the study is to first identify the user behavior and feedback data that we will collect. After that, modify the caller ID verifier app such that the user data could be collected and transmitted to our database. We then design a set of experiments and procedures that allow comparison of different variations of the app. During the execution, we recruited a population of participants and split the population equally among the experiments. Each experiment has a standardized procedure and the procedure will be conducted on each group simultaneously. Finally, tabulate the results and perform analysis on the results. The study was conducted with IRB approval.

In the quest to understand what the user thinks about having a security indicator upon an incoming call, we modified the caller ID verifier app to automatically collect and send the incoming call action to us upon some incoming calls using the caller ID verifier app. We collected the user's action and the type of security indicator displayed upon an incoming call. More specifically, we collect what kind of action was taken, i.e., answered, or declined/ignored. This will give us an idea of which security indicator tend to lead to what kind of action.

To reduce the privacy concerns, we only collected action data of incoming calls called from our phone numbers. To also ensure that the participants are aware that the app collects data from them, on the first run, the app will show a prompt screen to ask the participants for their consent to participate in the research experiments and to collect information from them. We managed to obtain IRB permission to conduct the study to collect data from the participants.

6.3.2 Participant Recruitment

Selecting participants size has several considerations. Unlike the experiments where we conducted the telephone phishing research, this research requires greater participation from each member. Recruiting a large group of participants may be regarded as statistically rigorous, however, their characteristics in relation to the real-world may still be questionable at this stage. Having a smaller group allows us to focus our attention on a few participants, whereas a larger group can lead to "social loafing" [179]. In fact, research shows that collecting data in large sample sizes can lead to analysis in the topic of interests lacking sufficient depth [180].

As a proof-of-concept prototype, the design of the prototype can heavily influence the outcome of the statistical result. A production-ready real-world implementation of such caller ID verifier can be designed very differently from the research prototype. Especially for software design, making design iterations is a constant process. Therefore, any statistical analysis of the research prototype would easily lose touch with a real-world implementation. The recruitment objective, in this case, is not collect data for statistical analysis but rather to evaluate of the user behaviors and feedback and inspire refinements for future implementation.

During the recruitment process, we disseminated recruitment posters through various communication channels, including emails, bulletin boards, facebook, and local craigslist ads. An example of the recruitment poster is shown in Figure 6.6. As an incentive to participate in our experiments, we advertised that the first 50 participants will receive a $10 Amazon gift card. To make it easy for the participant to download and install our app, we posted our app on the Google play store and posted a link (with QR code) to it with a shortened URL. We used the university's URL shortener service to help inform the participant the app is from the university and not a potential scam.

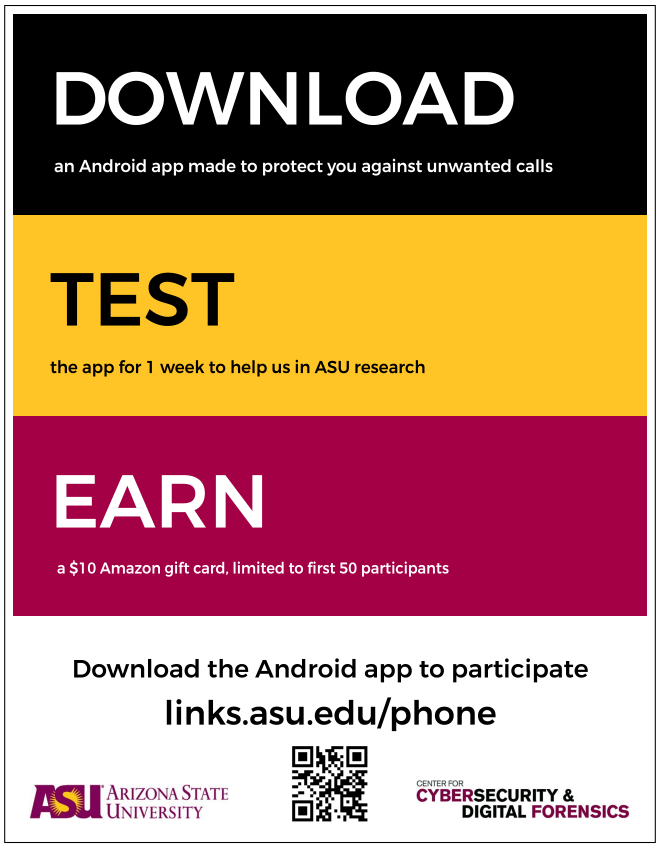
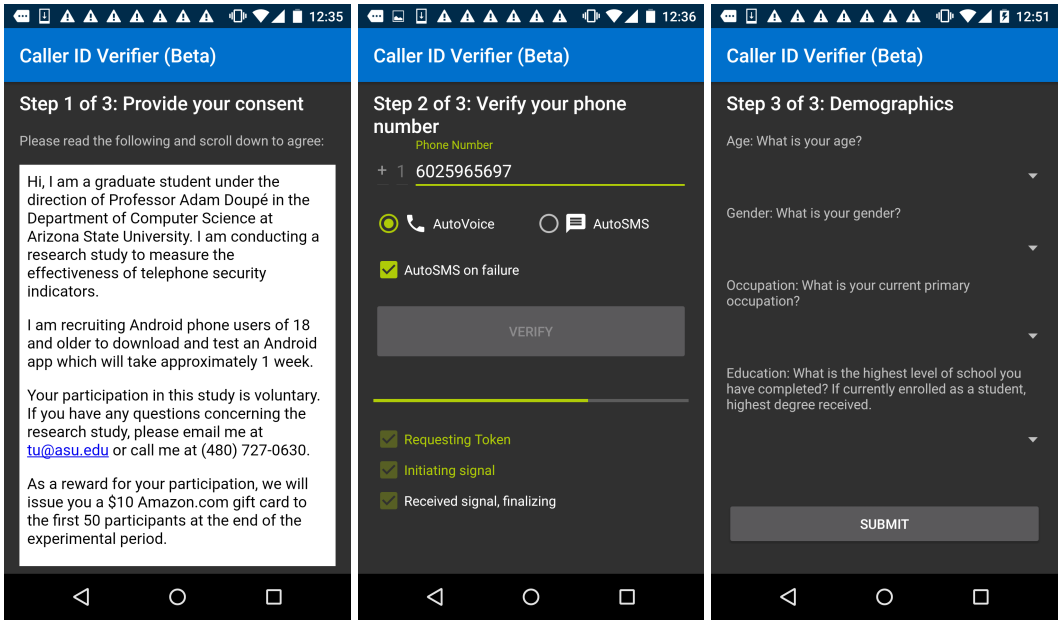


Figure 6.6: Recruitment Poster

At the end of the recruitment effort, our app was installed 70 times and we recruited a total of 57 participants for the experiments.

After the user installed our app, on the first run, the app shows a prompt screen asking the participant give explicit consent to participate in the research experiments and to collect information from them.

The user will need to read and scroll to the end of the consent information to find the agree button. After the user agrees, the app will show an activity for the participant to provide their phone number. The participant's phone number will be automatically verified by receiving a phone call or SMS message with a confirmation code through the app. During the process, we used a third-party service for the auto verify service to check if the phone number works. Finally, the app asks for the user's demographic information (age, gender,



(a) Consent Form (b) Phone Number Verification (c) Demographics Information

Figure 6.7: Procedure of Initial Set Up During First Run

education, occupation). Providing the demographic information is optional the user. After the user selects "SUBMIT", the user becomes a participant and the app sends the participant's, phone number, and demographic information to our server via an encrypted internet connection. The server randomly assigns the participant a group number (0-2) which determines the participant's experiment group. After that, the app is all set and is ready to display the security indicators from our calls.

Experiment Procedure

The app prototypes are designed to display the telephone security indicator and understand if the security indicator can influence the participants' behavior upon receiving a call. To

subject the participants to different types of influence, first, we evenly divided the participants in our study into 3 groups: group 0, group 1, and group 2.

Group 0 is the control group, where the app in this group would not display any security indicators upon any incoming call. The participants in this group do not see any security indicators. Hence, the participants in this group can be considered as similar to existing phone users, having no influence of the security indicators.

Group 1 is the simulated group, where the app in this group would immediately display a security indicator upon our incoming call. The participants in this group would see a security indicator when the call from one of our caller IDs. We preselected three different caller IDs for each of the three types of security indicators. This group is designed to show the influence of the security indicators to call recipient behavior.

Group 2 is the prototype group, where the app in this group would first receive an ACR via SMS and then display a security indicator upon our incoming call. We called the participants in this group using our Hamout app. Both the secure and warning security indicators require a valid or invalid ACR to be first delivered to the recipient prior to the call. Caution security indicators do not require an ACR from the caller. This group is designed to give us an idea of the SMS delivery mechanism on the effect of the security indicator, as our performance analysis has shown that the ACR delivery is not always consistent.

During the experiment period, we encouraged the participant to communicate openly with the researcher. In the app, we also added a menu shortcut to compose an email to write to the researcher. After receiving a call from us, the app displays a notification asking the participant to review the security indicator and provide a feedback. We collected the user feedback in the form of questionnaires and free-form comments. The activity to collect the feedback is shown in Figure 6.8.

In the app, we also provided a dashboard where the users could review the past security indicator alerts, where selecting on one would lead us to the same activity to provide us

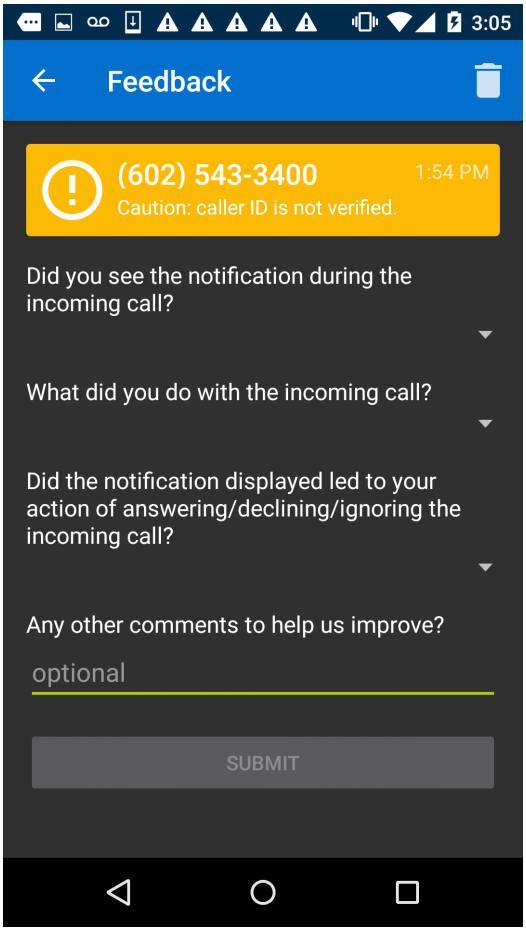


Figure 6.8: Collecting User Feedback After the Display of a Security Indicator

with a feedback. The apps also incorporated a feature to receive announcements from the researcher and the users could review the announcements in the dashboard.

The data in which the app collected from each participant is summarized in Figure 6.2.

6.3.3 Evaluation

During the experiment period of about 1 week, we sent out 168 phone calls in total, about 3 phone calls to each participant. To ensure that the procedure is standardized, we wrote a program to automate the process of sending out the telephone calls for group 0 and

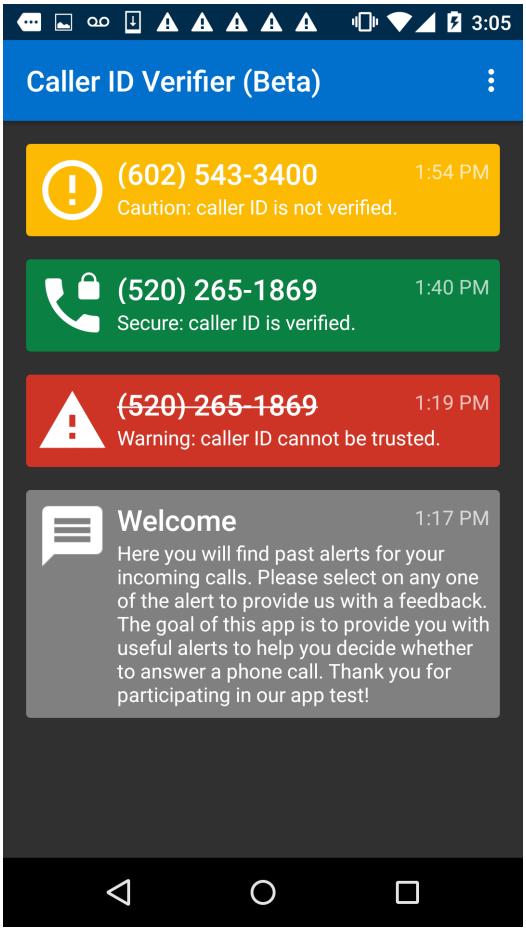


Figure 6.9: Reviewing Past Security Indicators and Announcements

group 1 participants. The content of the phone calls were all empty. During the experimental period, we called during the daytime (10am-5pm).

After disseminating the phone calls, we received a total of 136 records. Each record consists of a collection of action and feedback data described in Table 6.2. Our evaluation will present a discussion of the users' data in each group.

Group 0 Participants

We received a total of 44 records from the participants in group 0. Looking at their call action data, they have a 11:17 ratio of answering vs declining/ignoring (39% answer rate)

Data Collected	Description
Phone number	The verified phone number of the participant.
Gender	Optional response to the demographic question asking the gender of the participant.
Occupation	Optional response to the demographic question asking the occupation of the participant.
Education	Optional response to the demographic question asking the education of the participant.
Incoming call action	The action taken upon an incoming call from us.
Incoming call action delay	The amount of time from the start of incoming call to the action taken.
Incoming call count	The number of incoming calls received from us.
Security Indicator Type	The type of security indicator shown during the incoming call.
Has seen notification	Participant's feedback on whether the he/she saw the security indicator.
Action feedback	Participant's feedback on the action taken upon the incoming call.
Did notification led to action	Participant's feedback on whether the security indicator led to the action taken.
Other comments	Participant's feedback on other comments.

Table 6.2: Summary of Data Collected From Each Participant

upon an incoming call from us. The participants in this group are more likely to decline call than answering, and we can use answering ratio as a basis to compare the influence of security indicators to user behavior in the other groups.

We also received 3 feedback comments, and the 3 participants commented that they did not receive any notification while they got the call. A participant also emailed me asking if the app was working. This was deliberate so we did not reveal any information on why their app does not seem to be working.

Looking that the call count information, only 5 of the 17 participants remained to submit their data to us on the 3rd call. The other groups have almost twice as much participants remained to submit their data to us on the 3rd call. The shows that the other participants in group 0 must have uninstalled the app because the app did not provide any useful function for the participants. This is an example of how low app engagement can lead to low retention. As shown by many industry research, smartphone apps are generally subject to anywhere from a 30 to 80% attrition rate of users within the first month. If an app does

not provide enough engagement, the users may see the app as a useless tool and therefore choose to uninstall it. With consideration of these results, we can learn that the design of future implementations should think about user engagement.

Group 1 Participants

When disseminating the phone calls to group 1 participants, we called each phone number with 3 different security indicators, as we wanted to compare if an individual would behave differently when seeing different security indicators. To avoid the order effect of showing the security indicators, we separated participants into 3 groups, each group receiving the calls in a different order.

We received a total of 38 records and 9 comments from the participants in group 1. Looking at their call action data, they have a 21:17 ratio of answering vs declining/ignoring (55% answer rate) upon an incoming call from us. Looking at the break down of ratios of different security indicators: secure (caller ID is verified) indicators has the highest ratio of 9:4 (69% answer rate), caution (caller ID is not verified) indicators has a ratio of 6:5 (55% answer rate), warning (caller ID cannot be trusted) indicators has the lowest ratio of 6:8 (43% answer rate) upon an incoming call. In the comments, one participant stated that "The notification was very beneficial as it told me the ID could not be trusted. Thanks for the warning! :)". Another participant stated that "this made me cautious of answering. " We also received 2 emails from the participants near the end of the study in which they commented that they thought the app was helpful. This is good preliminary evidence showing that the respective security indicator can influence the user's incoming call action in the right direction.

One question that rose up in our minds is that: why are group 1 participants generally more likely to answer than group 0? We only expected participants to more likely answer when being shown a secure indicator, but why the caution and warning indicators also

made them more likely to answer? To answer this question, we looked at the participant's comments. One participant stated that "I only answered because I was curious about the efficacy of this app. If I trusted this app and saw the red warning I wouldn't have answered". Another participant said that "It was spam like it said, I just answered to see." The participants stated that even when being shown a warning indicator, and they still answered because they curious with the content of the call. They were curious to find out what is the content like in a phone call that displays a warning indicator. With consideration of these results, a user is probably not going to fall for a scam call that after answering a call with a warning or caution indicator. This is because the user would be already be expecting to hear a scam. As we have demonstrated in the telephone phishing study in chapter 3, vigilance is the best defense against telephone scams. If an individual is already expecting to hear a scam, we can say that the individual is already vigilant against the scam.

Group 2 Participants

When disseminating the phone calls to group 2 participants, we called each phone number with 3 different security indicators, using our Hamout app. To avoid the order effect of showing the security indicators, we separated participants into 3 groups, each group receiving the calls in a different order.

We received a total of 54 records and 6 comments from the participants in group 3. Looking at the data, the participants in this group has a 23:31 ratio of answering vs declining/ignoring (43% answer rate) upon an incoming call from us. Looking at the break down of ratios of different security indicators: secure (caller ID is verified) indicators has the highest ratio of 6:2 (75% answer rate), caution (caller ID is not verified) indicators has a ratio of 10:17 (37% answer rate), warning (caller ID cannot be trusted) indicators has the lowest ratio of 6:11 (35% answer rate) upon an incoming call. In the comments, we also see that the participants stated the notification was useful and some stated that they were

curious and just answered to see what the content of the call might be. These results are not surprising as it can be explained by the same discussion in group 1.

One thing that we notice is that we have a larger than expected number of records on caution indicators. This is unexpected because we disseminated the same number of phone calls for each of the three types of security indicators. The larger than expected caution indicators mean that either some of the ACRs sent were not delivered or some of the ACRs were not delivered before the phone call arrived. This further validates the performance analysis results that the SMS delivery mechanism is not always consistent. In future design iterations, we will need to incorporate a more consistent and reliable ACR delivery mechanism. As mentioned before, we believe that cloud message can be a good solution but this will need further testing to confirm. We also believe that the best future should be the integration of the ACR within the call request header itself.

Age		Gender		Occupation		Education	
18-35	35	Male	41	Student	49	Bachelor's	19
25-34	17	Female	12	Faculty	1	Master's	7
Unspecified	5	Unspecified	4	Employed	1	Some college credit	7
				Self-Employed	1	High school	6
				Unspecified	5	Doctorate	2
						Associate	2
						Unspecified	14

Table 6.3: Demographics of the Participants

6.3.4 Interview Findings

After the experiments, we followed up with some of the participants to further understand what they thought about the user experience of the app. We made email correspon-

dence with 7 participants. The interviews targeted active users of the mobile app based on our feedback data. Our interviews were conducted in an open response feedback style, we asked the interviewees for feedback on a set of leading questions and then recorded their responses. Most of the questions were open response questions, and many were generated on the spot to follow up on the interviewee's answer. With the interviews, we discovered many areas of interest that were not gathered from the app, and we were able to go deep into discussing the area of interest. We did not set a time limit on our interview process, we only stop the interviewing process when we felt that a state of saturation is reached. In total, we carried out 3 interview sessions where each participant comes from each different group.

When being asked about the app's user experience, all of our interviewees at some point mentioned that they hoped the app could help them identify unknown and unwanted calls. All interviewees stated that they received robocalls and scams calls at some point in their life and they wish that this app could help to prevent these types of calls for them. Although the prototype app was designed to only identify calls that are from our phone calls, they wanted to see a security indicator on phone calls that are not from us. Our group 0 participant also mentioned that she suspected if the app was working when she received phone calls and no security indicators showed up. When being asked about whether showing security indicators frequently would annoy them, all of them stated that they would not mind having frequent displays. If the app displays the security indicators too frequently, the apps could be designed such that there is an option to removed security indicator for certain types of phone calls, such as for some existing contacts or for a specific phone number. Therefore, in future designs, we should design the app to consistently or more frequently provide security indicator on telephone calls. This could be achieved by collecting more information from the users during the initial setup process. Also, we could allow other recipients to submit information about certain phone numbers.

All of our interviewees did not realize that the caller ID is spoofable and could be used to impersonate a known contact. The interviewees did not think that our app was mainly designed to prevent caller ID spoofing. They were more interested in the informational aspect of the security indicator. Interestingly, one of the interviewees revealed that he uses a similar call notification app, Truecaller. Truecaller is a caller name directory app and his experience with Truecaller is that it tends to be inaccurate because it allowed users to arbitrarily submit a name associated with their phone number without verification. Truecaller often shows a wrong name associated with the phone number. When being asked about what differentiates our app from Truecaller, he thinks that our app seems to be more security focused because of the phone number verification and therefore has the potential to be more accurate. He hopes that our app can be more accurate than Truecaller while being more capable of providing warnings.

On the design of the security indicator, one of interviewee stated that the warning indicator should be bigger and more visible than other indicators. He also suggested adding animation to the warning indicator might help to make it more visible. Because the goal of the app is to prevent the recipient from falling victim to phone scams, he was worried that some people might ignore the indicator or forget that it is a scam call. He stated that for example, his grandparent might be the kind of user that would ignore the indicator or forget that it is a scam call. We can draw a similarity to the way Google Chrome shows its warnings, where the web browser shows a red color warning in full page for a malicious site. Google Chrome shows a much more visible indicator for malicious sites than secure sites. This is the reason why we should design the telephone security indicator to give more emphasis on malicious phone calls.

Two interviewees also recommended that the security indicator should give more information, such that information about the type of caller and the location of the caller. These are information can help to determine whether to answer the call. For example, they wanted

the yellow caution to show if the caller is a telemarketer. It would also be interesting to know the real GPS location of the caller because the area code location of the caller is often misleading.

On the benefits of green/secure indicators, the interviewees generally believe that it would be most useful for urgent and important calls from known organizations. One interviewee also suggested that it would useful for calls from package delivery services or store pickup notifications, where it would be helpful to know when a package is waiting to be picked up. The interviewee also stated that the green/secure indicators would also help to determine whether to call back after missing the incoming call.

On whether the app performance was satisfactory, the interviewees generally think that the app was responsive enough to display the security indicators during the incoming call. None of the interviewees felt that the app is was causing performance issues on their smartphones. The interviewees generally did not think the app was trading off performance for the feature.

6.4 Discussion

With this empirical study, the findings showed that the display of security indicators for authenticated caller ID transmission can help users to determine whether to answer an incoming phone call. The findings of this study also suggest several recommendations for the implementation of authenticated caller ID transmission to make it more reliable, useful, and user-friendly.

The first recommendation is that the delivery mechanism for the Authenticated Call Request should be made more reliable by using an alternative out-of-band channel such as cloud messaging, or integrating the ACR as an essential part of the call request procedure. This is important, because if the ACR could not be consistently delivered, the users may miss out on important secure or warning indicators.

The second recommendation is that the end-user implementation should be more engaging for the user. It needs to provide frequent displays of security indicators to help the users realize that the app is working. While doing so, it would also be useful provide customization options for the users to disable the indicators for certain known callers that do not require high security. The ability to provide feedback for certain callers is another way that we can engage the user. With more user engagement, users are more likely to feel satisfied and less likely to abandon the feature.

The third recommendation is that the implementation should focus more on the design of warning indicators. The warning indicator should be made bigger and more visible than other indicators. The warning indicators should provide more notice to discourage the user to answer the call or help to instill vigilance against the caller. We also think that it might be helpful to add more information on why the call might be harmful.

Finally, we recommend adding more information to the security indicators, especially for unknown callers. For instance, it would be helpful to show if the caller is a telemarketer and the real GPS location of the caller. While doing so, it would important the ensure that the information is accurate. Existing caller directory apps suffer from inaccuracies, which is what our implementation can differentiate from existing products. Our app provides phone number verification, therefore, the information associated with the caller is more accurate. In future implementation, the app should also provide more information about the caller while keeping the information accurate.

6.5 Conclusion

In this study, we have explored how a mobile app might serve as an implementation of security indicators for authenticated caller ID transmission to help telephone users distinguish between important and potentially unwanted telephone calls. Our approach has been to design and evaluate the Hamout and Caller ID verifier apps, which implements

authenticated caller ID transmission and security indicators from caller ID verification. To evaluate the encryption performance, we provided a performance analysis of different encryption implementations of Hamout and was able show to that the encryption scheme would not introduce a perceivable delay to the call request. We also evaluated the delivery mechanism for the Authenticated Call Request and found that SMS delivery makes a significant impact on the user experience due to its inconsistent delay from carriers. This inspired us to recommend using cloud messaging or the call header itself to deliver the ACR in future designs. We also performed an empirical user study of using the Caller ID verifier app to evaluate the user experience and behavior upon seeing telephone security indicators. By collecting the participants' incoming call behavior, we learned that the telephone security indicators overall did help to influence the participants to answer important calls and avoid unwanted calls. Through feedback and interviews from participants, we identified several strategies to further improve future designs. In future work, we plan to expand this work from an proof-of-concept prototype conducted on a small focus group to a production-quality implementation conducted on a larger community to explore the benefits of authenticated caller ID transmission, and work toward a self-sustainable community where telephone spam, scam, fraud, phishing or vishing can be effectively prevented.

Chapter 7

CONCLUSION

Telephone spam, including scam, phishing, vishing, robocall, and telemarketing, has become an increasingly prevalent issue worldwide. Today, solutions primarily rely on blacklists of known scam numbers do not work effectively on spam calls when the caller ID has been spoofed. In this dissertation, we have presented motivation, background, survey, design, architecture, implementation and evaluation of the proposed caller ID authentication scheme for the telephone network that provides the possibility of a security indicator that will help prevent users from falling victim to telephone spam and scams. In this chapter, we summarize the major achievements of this dissertation that contributed to solving the telephone spam and scam problem.

With telephone scams becoming increasingly prevalent, it is crucial to understand what causes the recipients to fall for these scams. Armed with this knowledge, effective countermeasures can be developed to challenge the key foundations of successful telephone phishing attacks. In chapter 3, we presented the methodology, design, execution, results, analysis, and evaluation in the quest of answering why telephone scam works. The study performed 10 telephone phishing experiments on 3,000 university participants without prior awareness over the course of a work week. Overall, we were able to identify at least one key factor that had a significant effect in tricking the victims into revealing their social security information. Our evaluation recommended solutions designed to target impersonation and instill vigilance to prevent users from falling victim to telephone phishing. The work in this chapter is currently under review in the proceedings of IEEE Symposium on Security and Privacy 2018.

As many are now undertaking the crusade of curbing the telephone spam problem, surveying the existing solutions in combating telephone spam and, by analyzing the failings of the current techniques, we can derive the requirements that are critical to an acceptable solution. In chapter 4, we described the key challenges in solving the telephone spam problem, specifically focusing on the differences between email and telephone spam. Then, we survey the existing telephone spam solutions and, by analyzing the failings of the current techniques, derive evaluation criteria that are critical to an acceptable solution. This work helped to guide the development of effective telephone spam defenses, as well as provide a framework to evaluate future defenses. The work in this chapter was published in the proceedings of IEEE Symposium on Security and Privacy 2016 [45].

Caller ID is at the heart of stopping telephone spam—a variety of apps and services, including law enforcement, rely on the caller ID information to defend against unwanted callers. However, spammers are using spoofed caller IDs to defeat call blockers, to evade identification, and to further a variety of scams. To provide a solution to this problem, in chapter 5, we proposed a standardized authentication scheme for the caller ID that enables the possibility of a security indicator for telecommunication. The goal of this proposal is to help prevent users from falling victim to telephone spam and scams, as well as provide a foundation for future and existing defenses to stop unwanted telephone communication based on the caller ID information. Calls from legitimate callers, such as billing, delivery, banking, government, and law enforcement organizations, would also benefit from providing authenticity of their caller IDs, as their recipients would be certain that the caller is real and not an impostor, therefore feel better assured receiving communication over the phone. The work in this chapter was published in the proceedings of the ITU Kaleidoscope 2016 [56], the IEEE Communications Standards Magazine 2017 September Issue [163], a USPTO non-provisional patent application [164], and a pending technical standards contribution at the study group 11 of the ITU Telecommunication Standardization Sector.

With a description of the architecture, protocols, and security mechanisms of the caller ID authentication scheme, the next step is to follow up the proposed authentication scheme with a user study to understand the user behavior when they see the security indicators, learn about the potential use cases of caller ID authentication, and make future improvements to caller ID authentication implementation. In chapter 6, we described an end-user prototype design and implementation of the proposed caller ID authentication scheme. After implementing the prototype, we conducted a performance analysis and an end user study and collected the participants' feedback and behavior data. We presented the results of the user study and provide a performance analysis and evaluation of the feedback and behavior data from the study. Finally, we provided recommendations for the future end-user implementations of security indicators for the authenticated caller ID transmission.

REFERENCES

[1] C. T. W. Association, "Annual wireless industry survey." http://www.ctia.org/your-wireless-life/how-wireless-works/annual-wireless-industry-survey, 1 2016. (Visited on 01/29/2016).

[2] Statista, "Annual Wireless Industry Survey." http://www.ctia.org/your-wireless-life/how-wireless-works/annual-wireless-industry-survey.

[3] International Telecommunication Union, "Fixed-telephone subscriptions." http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2014/Fixed_tel_2000-2013.xls.

[4] Statista, "Countries by number of Voice over Internet Protocol (VoIP) subscribers in 1Q 2013." http://www.statista.com/statistics/236824/number-of-voip-subscribers-by-leading-countries/.

[5] U.S. Bureau of Labor Statistics, "American time use survey fact sheet." http://www.bls.gov/tus/atussummary.pdf, June 2015.

[6] Federal Trade Commission, "Consumer Sentinel Network Data Book for January - December 2015," 2016.

[7] Financial Fraud Action UK, "FFA - Fraud the Facts 2016," 2016.

[8] Federal Trade Commission, "National Do Not Call Registry Data Book for Fiscal Year 2016," 2016.

[9] S. H. Kimball, T. Levy, H. Venturelli, and S. Miller, "Interactive Voice Recognition Communication in Electoral Politics: Exploratory Metadata Analysis," *American Behavioral Scientist*, 2014.

[10] T. Mobarak and A. Han, "Method and apparatus for forcing a call to a carrier provided voice mail facility," 2013. US Patent 8,605,869.

[11] TrueCaller, "Americans lost $8.6 billion in phone scams: Learn to protect yourself." http://www.truecaller.com/blog/americans-lost-86-billion-in-phone-scams-learn-to-protect-yourself, 2014.

[12] Federal Communications Commission, "Telephone Consumer Protection Act 47 U.S.C. 227." https://transition.fcc.gov/cgb/policy/TCPA-Rules.pdf, 1991.

[13] OFCOM: The Office of Communications, "Nuisance calls and messages," 2012.

[14] Federal Trade Commission, "National Do Not Call Registry." https://www.donotcall.gov/, 2016.

[15] TPS: Telephone Preference Service, "Register." http://www.tpsonline.org.uk/tps/number_type.html, 2016.

[16] Marketwired, "From Stalkers to Spam, WhitePages Study Breaks Down Reasons Americans Block Calls." http://www.marketwired.com/press-release/from-stalkers-to-spam-whitepages-study-breaks-down-reasons-americans-block-calls-1900134.htm.

[17] International Telecommunication Union, "Q.731.7 : Stage 3 description for number identificationsupplementary services using SignallingSystem No. 7: Malicious call identification(MCID)," 1997.

[18] L. Leung, "Fed up with rising robocalls, millions say 'do not call' list doesn't work and want relief," *The Orange Country Register*, Oct 2016.

[19] A. Johnson, "Scammers can fake caller id info | consumer information." https://www.consumer.ftc.gov/blog/scammers-can-fake-caller-id-info, 5 2016. (Accessed on 04/03/2017).

[20] Numbercop, "Smishing & vishing news ? weekly summary 3/2-3/8." https://numbercop.tumblr.com/post/115252732893/weekly-summary-32-38, 4 2015. (Accessed on 04/03/2017).

[21] J. Pepitone, "'swatting' celebrities is far too simple." http://money.cnn.com/2013/04/14/technology/security/swatting-caller-id/, 4 2013. (Accessed on 04/03/2017).

[22] Eric Lipton, David E. Sanger and Scott Shane, "The perfect weapon: How russian cyberpower invaded the u.s. - the new york times." https://www.nytimes.com/2016/12/13/us/politics/russia-hack-election-dnc.html, 12 2016. (Accessed on 04/06/2017).

[23] NANPA : North American Numbering Plan Administration, "About the North American Numbering Plan." https://www.nationalnanpa.com/about_us/abt_nanp.html.

[24] Wikipedia, "Public switched telephone network — wikipedia, the free encyclopedia," 2016. [Online; accessed 23-November-2016].

[25] I. T. Union, "Q.700: Specifications of Signalling System No. 7," 1993.

[26] A. R. Modarressi and R. Skoog, "Signaling System No. 7: A Tutorial," *IEEE Communications Magazine*, 1990.

[27] International Telecommunication Union, "Q.764 : Signalling System No. 7 - ISDN User Part signalling procedures," 12 1999.

[28] E. Inc, "Gsm originating call." http://www.eventhelix.com/RealtimeMantra/Telecom/GSM_Originating_Call_Flow.pdf. (Visited on 02/08/2016).

[29] M. Grayson, K. Shatzkamer, and K. Wierenga, *Building the Mobile Internet*. Pearson Education, 2011.

[30] E. Inc, "3g umts originating call flow." http://www.eventhelix.com/umts/originating-call/. (Visited on 02/08/2016).

[31] E. Inc, "Sip tutorial: Sip to pstn call flow (detailed)." http://www.eventhelix.com/RealtimeMantra/Telecom/SIP_PSTN_Call_Flow.pdf. (Visited on 02/08/2016).

[32] E. Inc, "H.323 call setup involving h.224, q.931, h.245, rtp and rtcp p? message sequence chart." http://www.eventhelix.com/RealtimeMantra/Telecom/h323_call_flow.pdf. (Visited on 02/08/2016).

[33] E. Inc, "Ims originating to pstn isup call." http://www.eventhelix.com/ims/ims-to-pstn-call/ims-to-pstn-callflow.pdf. (Visited on 02/08/2016).

[34] E. Inc, "Voice over lte (volte) originating call." http://www.eventhelix.com/lte/volte/volte-originating-call.pdf. (Visited on 02/08/2016).

[35] E. Inc, "V5.2 to v5.2 call – dtmf, called subscriber goes onhook." http://www.eventhelix.com/RealtimeMantra/Telecom/V52_Call_Processing.pdf. (Visited on 02/08/2016).

[36] E. Inc, "Isup to isup successful call." http://www.eventhelix.com/RealtimeMantra/Telecom/ISUP_ISUP_Call.pdf. (Visited on 02/08/2016).

[37] International Telecommunication Union, "Q.763 : Signalling System No. 7 - ISDN User Part formats and codes," 1999.

[38] B. Woodcock and V. Adhikari, "Survey of Characteristics of Internet Carrier Interconnection Agreements," tech. rep., Packet Clearing House, 2011.

[39] Federal Communications Commission, "Intercarrier compensation." https://www.fcc.gov/encyclopedia/intercarrier-compensation, 2015.

[40] P. Casanova, R. Bandyopadhyay, and V. Balasubramaniyan, "Largest IRS Phone Scam Likely Exceeded 450,000 Potential Victims in March." http://www.pindropsecurity.com/irs-phone-scam-live-call_analysis/, 2015.

[41] C. A. Hamilton, "Machine answer detection," Dec. 6 1994. US Patent 5,371,787.

[42] Federal Communications Commission, "Calling Number Identification Service–Caller ID." http://www.gpo.gov/fdsys/pkg/FR-1995-06-05/pdf/95-13760.pdf, 2015.

[43] International Telecommunication Union, "Q.731.3 : Stage 3 description for number identification supplementary services using Signalling System No. 7 : Calling line identification presentation (CLIP)," 1993.

[44] Federal Trade Commission, "Consumer Sentinel Network Data Book for January - December 2016," 2017.

[45] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn, "SoK: Everyone Hates Robocalls: A Survey of Techniques against Telephony Spam," in *Proceedings of the 37th IEEE Symposium on Security and Privacy*, IEEE, 2016.

[46] I. R. Service, "Irs repeats warning about phone scams." https://www.irs.gov/uac/newsroom/irs-repeats-warning-about-phone-scams, 8 2014. (Accessed on 04/20/2017).

[47] Andrew Johnson, Division of Consumer and Business Education, FTC, "Voicemail from an irs imposter? | consumer information." https://www.consumer.ftc.gov/blog/voicemail-irs-imposter, 9 2016. (Accessed on 04/20/2017).

[48] J. Pavia, "Sadly, irs phone scams are very successful 'businesses'." http://www.cnbc.com/2016/10/18/sadly-irs-phone-scams-are-very-successful-businesses.html, 10 2016. (Accessed on 04/20/2017).

[49] I. R. Service, "Irs warns of pervasive telephone scam." https://www.irs.gov/uac/newsroom/irs-warns-of-pervasive-telephone-scam, 10 2013. (Accessed on 04/17/2017).

[50] I. R. Service, "Irs alerts payroll and hr professionals to phishing scheme involving w-2s." https://www.irs.gov/uac/newsroom/irs-alerts-payroll-and-hr-professionals-to-phishing-scheme-involving-w2s, 3 2016. (Accessed on 04/17/2017).

[51] K. Queen, "Guard the last 4 digits of your social security number - theyâĂŹre all id thieves need." http://blogs.creditcards.com/2015/11/social-security-last-4-digits.php, 19 2015. (Accessed on 08/30/2017).

[52] A. Acquisti and R. Gross, "Predicting social security numbers from public data," *Proceedings of the National academy of sciences*, vol. 106, no. 27, pp. 10975–10980, 2009.

[53] A. E. Deeb, "What to do with small data." https://medium.com/rants-on-machine-learning/what-to-do-with-small-data-d253254d1a89, 5 2015. (Accessed on 09/25/2017).

[54] J. Cohen, "Statistical power analysis for the behavioral sciences (revised ed.)," 1977.

[55] F. M. Wolf, *Meta-analysis: Quantitative methods for research synthesis*, vol. 59. Sage, 1986.

[56] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn, "Toward Authenticated Caller ID Transmission: The Need for a Standardized Authentication Scheme in Q.731.3 Calling Line Identification Presentation," in *Proceedings of ITU Kaleidoscope 2016 - ICTs for a Sustainable World*, Nov. 2016.

[57] C. Jennings, J. Peterson, and E. Rescorla, "Authenticated Identity Management in the Session Initiation Protocol (SIP) draft-ietf-stir-rfc4474bis-09," *IETF*, 2016.

[58] B. Reaves, L. Blue, and P. Traynor, "AuthLoop: End-to-End Cryptographic Authentication for Telephony over Voice Channels," in *Proceedings of the USENIX Security Symposium (USENIX)*, 2016.

[59] M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, "Experimental case studies for investigating e-banking phishing techniques and attack strategies," *Cognitive Computation*, vol. 2, no. 3, pp. 242–253, 2010.

[60] R. Dhamija, J. D. Tygar, and M. Hearst, "Why Phishing Works," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 581–590, ACM, 2006.

[61] S. Egelman, L. F. Cranor, and J. Hong, "You've been warned: an empirical study of the effectiveness of web browser phishing warnings," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1065–1074, ACM, 2008.

[62] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Communications of the ACM*, vol. 50, no. 10, pp. 94–100, 2007.

[63] T. Vidas, E. Owusu, S. Wang, C. Zeng, L. F. Cranor, and N. Christin, "Qrishing: The susceptibility of smartphone users to qr code phishing attacks," in *International Conference on Financial Cryptography and Data Security*, pp. 52–69, Springer, 2013.

[64] E. Park, "Rustock Takedown's Effect on Global Spam Volume." `http://www.symantec.com/connect/blogs/rustock-takedown-s-effect-global-spam-volume`, 2011.

[65] M. Carney, "Courts deem CallFire a common carrier, setting a major precedent at intersection of telecom and tech law." `http://pando.com/2015/02/27/courts-deem-callfire-a-common-carrier-setting-a-major-precedent-at-intersection-of-telecom-and-tech-law/`, Feb. 27, 2015.

[66] K. Cox, "FTC: Totally Fine By Us If Phone Companies Block Robocalling Numbers." `http://consumerist.com/2015/01/27/ftc-totally-fine-by-us-if-phone-companies-block-robocalling-numbers/`, Jan. 27, 2015.

[67] Public Law 111-331 111th Congress, "Truth in caller id act of 2009." `https://www.congress.gov/111/plaws/publ331/PLAW-111publ331.pdf`.

[68] Federal Trade Commission, "National Do Not Call Registry." `https://www.donotcall.gov/`.

[69] Federal Trade Commission, "Robocalls | consumer information." `https://www.consumer.ftc.gov/features/feature-0025-robocalls`, 2015.

[70] E. Montejo, "How to block phone calls on your Android phone." http://www.androidauthority.com/how-to-block-phone-calls-numbers-android-phone-246484/.

[71] Apple Inc., "Block calls and block or filter messages on your iPhone, iPad, or iPod touch." https://support.apple.com/en-us/HT201229.

[72] T. Nimmerjahn, "Whitelist Call Blocker." https://play.google.com/store/apps/details?id=de.tn_software.callblocker.

[73] NQ Mobile Security, "NQ Mobile Call Blocker." http://en.nq.com/callblocker.

[74] P. Kolan and R. Dantu, "Socio-technical defense against voice spamming," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 2, no. 1, p. 2, 2007.

[75] F. Wang, Y. Mo, and B. Huang, "P2p-avs: P2p based cooperative voip spam filtering," in *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, pp. 3547–3552, IEEE, 2007.

[76] P. Patankar, G. Nam, G. Kesidis, and C. R. Das, "Exploring anti-spam models in large scale voip systems," in *Distributed Computing Systems, 2008. ICDCS'08. The 28th International Conference on*, pp. 85–92, IEEE, 2008.

[77] R. Zhang and A. Gurtov, "Collaborative reputation-based voice spam filtering," in *Database and Expert Systems Application, 2009. DEXA'09. 20th International Workshop on*, pp. 33–37, IEEE, 2009.

[78] C. Sorge and J. Seedorf, "A provider-level reputation system for assessing the quality of spit mitigation algorithms," in *Communications, 2009. ICC'09. IEEE International Conference on*, pp. 1–6, IEEE, 2009.

[79] V. B. Payas Gupta, Bharat Srinivasan and M. Ahamad, "Phoneypot: Data-driven Understanding of Telephony Threats," in *Proceedings of the Symposium on Network and Distributed System Security (NDSS)*, 2015.

[80] P. Kolan, R. Dantu, and J. W. Cangussu, "Nuisance level of a voice call," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 5, no. 1, p. 6, 2008.

[81] T. S. Corporation, "Nomorobo." https://www.nomorobo.com.

[82] K. Srivastava and H. G. Schulzrinne, "Preventing spam for sip-based instant messages and sessions," 2004.

[83] Y. Rebahi and D. Sisalem, "Sip service providers and the spam problem," in *Proceedings of the 2nd VoIP Security Workshop*, 2005.

[84] Y. Rebahi, D. Sisalem, and T. Magedanz, "Sip spam detection," in *Digital Telecommunications,, 2006. ICDT'06. International Conference on*, pp. 68–68, IEEE, 2006.

[85] M. A. Azad and R. Morla, "Multistage spit detection in transit voip," in *Software, Telecommunications and Computer Networks (SoftCOM), 2011 19th International Conference on*, pp. 1–9, IEEE, 2011.

[86] M. A. Azad, R. Morla, "Caller-rep: Detecting unwanted calls with caller social strength," *Computers & Security*, vol. 39, pp. 219–236, 2013.

[87] Y.-S. Wu, S. Bagchi, N. Singh, and R. Wita, "Spam detection in voice-over-ip calls through semi-supervised clustering," in *Dependable Systems & Networks, 2009. DSN'09. IEEE/IFIP International Conference on*, pp. 307–316, IEEE, 2009.

[88] F. Wang, F. R. Wang, B. Huang, and L. T. Yang, "Advs: a reputation-based model on filtering spit over p2p-voip networks," *The Journal of Supercomputing*, vol. 64, no. 3, pp. 744–761, 2013.

[89] D. Shin, J. Ahn, and C. Shim, "Progressive Multi Gray-Leveling: A Voice Spam Protection Algorithm," *IEEE Network*, 2006.

[90] H.-J. Kim, M. J. Kim, Y. Kim, and H. C. Jeong, "Devs-based modeling of voip spam callersâĂŹ behavior for spit level calculation," *Simulation Modelling Practice and Theory*, vol. 17, no. 4, pp. 569–584, 2009.

[91] M.-Y. Su and C.-H. Tsai, "A prevention system for spam over internet telephony," *Appl. Math*, vol. 6, no. 2S, pp. 579S–585S, 2012.

[92] M. Amanian, M. H. Y. Moghaddam, and H. K. Roshkhari, "New method for evaluating anti-spit in voip networks," in *Computer and Knowledge Engineering (ICCKE), 2013 3th International eConference on*, pp. 374–379, IEEE, 2013.

[93] Y. Soupionis and D. Gritzalis, "Aspf: Adaptive anti-spit policy-based framework," in *Availability, Reliability and Security (ARES), 2011 Sixth International Conference on*, pp. 153–160, IEEE, 2011. http://dx.doi.org/10.1109/ARES.2011.29.

[94] N. Chaisamran, T. Okuda, and S. Yamaguchi, "Trust-based voip spam detection based on calling behaviors and human relationships," *Information and Media Technologies*, vol. 8, no. 2, pp. 528–537, 2013.

[95] R. J. B. Chikha, T. Abbes, W. B. Chikha, and A. Bouhoula, "Behavior-based approach to detect spam over IP telephony attacks," *International Journal of Information Security*, 2015.

[96] H. Sengar, X. Wang, and A. Nichols, "Call Behavioral analysis to Thwart SPIT attacks on VoIP networks," in *Security and Privacy in Communication Networks*, pp. 501–510, Springer, 2012.

[97] H. J. Kang, Z.-L. Zhang, S. Ranjan, and A. Nucci, "Sip-based voip traffic behavior profiling and its applications," in *Proceedings of the 3rd annual ACM workshop on Mining network data*, pp. 39–44, ACM, 2007.

[98] Y. Bai, X. Su, and B. Bhargava, "Adaptive Voice Spam Control with User Behavior Analysis," in *Proceedings of the IEEE International Conference on High Performance Computing and Communications (HPCC)*, 2009.

[99] R. MacIntosh and D. Vinokurov, "Detection and mitigation of spam in IP telephony networks using signaling protocol analysis," in *Advances in Wired and Wireless Communication, 2005 IEEE/Sarnoff Symposium on*, pp. 49–52, IEEE, 2005.

[100] S. Phithakkitnukoon, R. Dantu, R. Claxton, and N. Eagle, "Behavior-based adaptive call predictor," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 6, no. 3, p. 21, 2011.

[101] H. Yan, K. Sripanidkulchai, H. Zhang, Z.-Y. Shae, and D. Saha, "Incorporating active fingerprinting into spit prevention systems," in *Third annual security workshop (VSWâĂŹ06)*, Citeseer, 2006.

[102] EveryCaller, "Call Control." `http://www.everycaller.com`.

[103] Budaloop, "Regex Call Blocker." `https://play.google.com/store/apps/details?id=com.budaloop.regexblocker`.

[104] Pindrop Security, "Fraud Detection System." `http://www.pindropsecurity.com/fraud-detection-system`.

[105] S. J. Brolin and S. Colodner, "Automatic number identification in subscriber loop carrier systems," Nov. 1 1977. US Patent 4,056,690.

[106] I. Neustar, "Nanpa : Ani ii digits - view assignments." `https://www.nationalnanpa.com/number_resource_info/ani_ii_assignments.html`, 2015.

[107] S. Horvath and T. Kasvand, "Voice identification pre-screening and redirection system," Sept. 6 2002. US Patent App. 10/236,810.

[108] D. Reich and R. Szabo, "Method and system of determining unsolicited callers," Apr. 28 2004. US Patent App. 10/833,515.

[109] C. Pörschmann and H. Knospe, "Analysis of Spectral Parameters of Audio Signals for the Identification of Spam Over IP Telephony.," in *CEAS*, 2008.

[110] C. Pörschmann and H. Knospe, "Spectral Analysis of Audio Signals for the Identification of Spam Over IP Telephony," in *Proceedings of the NAG/DAGA International Conference on Acoustics*, 2009.

[111] D. Lentzen, G. Grutzek, H. Knospe, and C. Porschmann, "Content-based Detection and Prevention of Spam over IP Telephony-System Design, Prototype and First Results," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2011.

[112] J. Strobl, B. Mainka, G. Grutzek, and H. Knospe, "An Efficient Search Method for the Content-Based Identification of Telephone-SPAM," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2012.

[113] S. A. Iranmanesh, H. Sengar, and H. Wang, "A Voice Spam Filter to Clean Subscribers' Mailbox," *Security and Privacy in Communication Networks*, 2013.

[114] F. Maggi, "Are the Con Artists Back? A Preliminary Analysis of Modern Phone Frauds," in *Proceedings of the IEEE International Conference on Computer and Information Technology (CIT)*, 2010.

[115] L. R. Rabiner, "Applications of speech recognition in the area of telecommunications," in *Automatic Speech Recognition and Understanding, 1997. Proceedings., 1997 IEEE Workshop on*, pp. 501–510, IEEE, 1997.

[116] David R. Wheeler, "Voice recognition will always be stupid." `http://www.cnn.com/2013/08/20/opinion/wheeler-voice-recognition/`.

[117] 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*, CCS '10, (New York, NY, USA), pp. 109–120, ACM, 2010.

[118] H. Hai, Y. Hong-Tao, and F. Xiao-Lei, "A SPIT Detection Method Using Voice Activity Analysis," in *Proceedings of the International Conference on Multimedia Information Networking and Security (MINES)*, 2009.

[119] J. Quittek, S. Niccolini, S. Tartarelli, M. Stiemerling, M. Brunner, and T. Ewald, "Detecting SPIT calls by checking human communication patterns," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2007.

[120] J. Quittek, S. Niccolini, S. Tartarelli, and R. Schlegel, "Prevention of Spam over IP Telephony (SPIT)," tech. rep., NEC, 2006.

[121] J. Lindqvist and M. Komu, "Cure for spam over internet telephony," in *4TH IEEE CONSUMER COMMUNICATIONS AND NETWORKING CONFERENCE (CCNC 2007). ProceedingsâĂę vol., n*, pp. 896–900, 2007.

[122] A. Markkola and J. Lindqvist, "Accessible Voice CAPTCHAs for Internet Telephony," in *Proceedings of the Symposium on Accessible Privacy and Security (SOAPS)*, 2008.

[123] Y. Soupionis, G. Tountas, and D. Gritzalis, "Audio CAPTCHA for SIP-based VoIP," in *Proceedings of International Information Security Conference*, 2009.

[124] E. Bursztein and S. Bethard, "Decaptcha: breaking 75% of ebay audio captchas," in *Proceedings of the 3rd USENIX conference on Offensive technologies*, p. 8, USENIX Association, 2009.

[125] E. Harris, "The Next Step in the Spam Control War: Greylisting." http://projects.puremagic.com/greylisting/whitepaper.html, 2003.

[126] Google Voice, "Screen calls." https://support.google.com/voice/answer/115083.

[127] Verizon, "Call Intercept." https://www.verizon.com/support/residential/phone/homephone/calling+features/call+intercept/130058.htm.

[128] Phone.com, "Phone.com university - screening calls for your business line | phone.com." https://www.phone.com/blog/tips-tricks/2014/02/24/phone-com-university-screening-calls-business-line/, February 2014.

[129] N. Croft and M. Olivier, "A model for spam prevention in ip telephony networks using anonymous verifying authorities," 2005.

[130] H. Mustafa, W. Xu, A. R. Sadeghi, and S. Schulz, "You Can Call but You Can't Hide: Detecting Caller ID Spoofing Attacks," in *Proceedings of the Conference on Dependable Systems and Networks (DSN)*, 2014.

[131] K. Ono and H. Schulzrinne, "Have i met you before?: using cross-media relations to reduce spit," in *Proceedings of the 3rd International Conference on Principles, Systems and Applications of IP Telecommunications*, p. 3, ACM, 2009.

[132] A. Back, "Hashcash - A Denial of Service Counter-Measure." http://www.hashcash.org/hashcash.pdf, 2002.

[133] H. Tschofenig, R. Falk, J. Peterson, J. Hodges, D. Sicker, J. Polk, and A. Siemens, "Using saml to protect the session initiation protocol (sip)," *IEEE Network*, vol. 20, no. 5, pp. 14–17, 2006.

[134] S. Saklikar and S. Saha, "Identity federation for voip-based services," in *Proceedings of the 2007 ACM workshop on Digital identity management*, pp. 62–71, ACM, 2007.

[135] L. Kong, V. A. Balasubramaniyan, and M. Ahamad, "A lightweight scheme for securely and reliably locating sip users," in *VoIP Management and Security, 2006. 1st IEEE Workshop on*, pp. 9–17, IEEE, 2006.

[136] V. Balasubramaniyan, M. Ahamad, and H. Park, "Callrank: Combating spit using call duration, social networks and global reputation.," in *CEAS*, 2007.

[137] R. Dantu and P. Kolan, "Preventing voice spamming," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), Workshop on VoIP Security Challenges and Solutions*, 2004.

[138] B. Mathieu, Y. Gourhant, and Q. Loudier, "Spit mitigation by a network level anti-spit entity," in *Proc. of the 3rd Annual VoIP Security Workshop*, 2006.

[139] C. Dwork and M. Naor, "Pricing via Processing or Combatting Junk Mail," in *Advances in Cryptology (CRYPTO)*, 1992.

[140] N. Banerjee, S. Saklikar, and S. Saha, "Anti-vamming trust enforcement in peer-to-peer voip networks," in *Proceedings of the 2006 international conference on Wireless communications and mobile computing*, pp. 201–206, ACM, 2006.

[141] C. Jennings, "Computational puzzles for spam reduction in sip," 2007.

[142] S. Niccolini, "Spit prevention: state of the art and research challenges," in *Proceedings of the 3rd Workshop on Securing Voice over IP*, 2006.

[143] J. Quittek, S. Niccolini, S. Tartarelli, and R. Schlegel, "On Spam over Internet Telephony (SPIT) Prevention," *IEEE Communications Magazine*, 2008.

[144] R. Schlegel, S. Niccolini, S. Tartarelli, and M. Brunner, "Ise03-2: Spam over internet telephony (spit) prevention framework," in *Global Telecommunications Conference, 2006. GLOBECOM'06. IEEE*, pp. 1–6, IEEE, 2006.

[145] D. Gritzalis and Y. Mallios, "A sip-oriented spit management framework," *Computers & Security*, vol. 27, no. 5, pp. 136–153, 2008.

[146] D. Gritzalis, G. Marias, Y. Rebahi, Y. Soupionis, and S. Ehlert, "Spider: A platform for managing sip-based spam over internet telephony spit," *Journal of Computer Security*, vol. 19, no. 5, pp. 835–867, 2011.

[147] R. Dantu and P. Kolan, "Detecting spam in voip networks," *Proc. SRUTI*, vol. 5, 2005.

[148] M. Hansen, M. Hansen, J. Möller, T. Rohwer, C. Tolkmit, and H. Waack, "Developing a legally compliant reachability management system as a countermeasure against spit," in *Proceedings of Third Annual VoIP Security Workshop, Berlin, Germany*, 2006.

[149] B. Mathieu, S. Niccolini, and D. Sisalem, "SDRS: a voice-over-IP spam detection and reaction system," *Security & Privacy, IEEE*, vol. 6, no. 6, pp. 52–59, 2008.

[150] N. d'Heureuse, J. Seedorf, and S. Niccolini, "A policy framework for personalized and role-based spit prevention," in *Proceedings of the 3rd International Conference on Principles, Systems and Applications of IP Telecommunications*, p. 12, ACM, 2009.

[151] S. Dritsas, V. Dritsou, B. Tsoumas, P. Constantopoulos, and D. Gritzalis, "Ontospit: Spit management through ontologies," *Computer Communications*, vol. 32, no. 1, pp. 203–212, 2009.

[152] M. Scatá and A. L. Corte, "Security analysis and countermeasures assessment against spit attacks on voip systems," in *Internet Security (WorldCIS), 2011 World Congress on*, pp. 177–183, IEEE, 2011.

[153] A. D. Keromytis, "A survey of voice over ip security research," in *Information Systems Security*, pp. 1–17, Springer, 2009.

[154] Keromytis, Angelos D, "A comprehensive survey of voice over ip security research," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 2, pp. 514–537, 2012.

[155] R. Baumann, S. Cavin, and S. Schmid, "Voice over ip-security and spit," *Swiss Army, FU Br*, vol. 41, pp. 1–34, 2006.

[156] S. Phithakkitnukoon, R. Dantu, and E.-A. Baatarjav, "Voip security - attacks and solutions," *Information Security Journal: A Global Perspective*, vol. 17, no. 3, pp. 114–123, 2008.

[157] V. M. Quinten, R. Van De Meent, and A. Pras, "Analysis of techniques for protection against spam over internet telephony," in *Dependable and Adaptable Networks and Services*, pp. 70–77, Springer, 2007.

[158] R. Dantu, S. Fahmy, H. Schulzrinne, and J. Cangussu, "Issues and challenges in securing voip," *computers & security*, vol. 28, no. 8, pp. 743–753, 2009.

[159] S. Dritsas, Y. Soupionis, M. Theoharidou, Y. Mallios, and D. Gritzalis, "Spit identification criteria implementation: Effectiveness and lessons learned," in *Proceedings of The Ifip Tc 11 23rd International Information Security Conference*, pp. 381–395, Springer, 2008.

[160] G. F. Marias, S. Dritsas, M. Theoharidou, J. Mallios, and D. Gritzalis, "SIP vulnerabilities and anti-SPIT mechanisms assessment," in *Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on*, pp. 597–604, IEEE, 2007. http://dx.doi.org/10.1109/ICCCN.2007.4317883.

[161] S. F. Khan, M. Portmann, and N. W. Bergmann, "A Review of Methods for Preventing Spam in IP Telephony," *Modern Applied Science*, vol. 7, no. 7, p. p48, 2013.

[162] J. Rosenberg, C. Jennings, and J. Peterson, "The session initiation protocol (SIP) and spam," tech. rep., RFC 5039, January, 2008.

[163] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn, "Toward Standardization of Authenticated Caller ID Transmission," *IEEE Communications Standards Magazine*, 9 2017.

[164] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn, "Systems and methods for authenticating caller identity and call request header information for outbound telephony communications," 03 2016.

[165] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "Sip: session initiation protocol," tech. rep., 2002.

[166] D. Lawrence, "An Identity Thief Explains the Art of Emptying Your Bank Account," *Bloomberg Businessweek*, 2015.

[167] International Telecommunication Union, "Information technology - Open Systems Interconnection - The Directory: Public-key and attribute certificate frameworks," 2012.

[168] International Telecommunication Union, "E.164 : The international public telecommunication numbering plan," 2010.

[169] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and W. Polk, "Internet X. 509 Public Key Infrastructure Certificate and CRL Profile," *RFC5280*, 2008.

[170] R. Housley, W. Polk, W. Ford, and D. Solo, "Internet X. 509 public key infrastructure certificate and certificate revocation list (CRL) profile," tech. rep., 2002.

[171] M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams, "X. 509 Internet public key infrastructure online certificate status protocol-OCSP," tech. rep., 1999.

[172] C. Wendt and J. Peterson, "Persona Assertion Token draft-ietf-stir-passport-03," *IETF*, 2016.

[173] B. Reaves, L. Blue, and P. Traynor, "Authloop: End-to-end cryptographic authentication for telephony over voice channels," in *25th USENIX Security Symposium*, USENIX Association, 2016.

[174] Wikipedia, "List of countries by smartphone penetration — wikipedia, the free encyclopedia," 2017. [Online; accessed 7-September-2017].

[175] L. Sui, "Strategy analytics: Android captures record 88 percent share of global smartphone shipments in q3 2016," 11 2016. (Accessed on 09/07/2017).

[176] Google, "Android Keystore System." `https://developer.android.com/training/articles/keystore.html`.

[177] rtyley, "Spongy Castle." `https://rtyley.github.io/spongycastle/`.

[178] Google, "Firebase cloud messaging | send notifications across platforms for free Âă|Âă firebase." `https://firebase.google.com/products/cloud-messaging/`. (Accessed on 09/11/2017).

[179] R. C. Liden, S. J. Wayne, R. A. Jaworski, and N. Bennett, "Social loafing: A field investigation," *Journal of Management*, vol. 30, no. 2, pp. 285–304, 2004.

[180] M. Bloor, *Focus groups in social research*. Sage, 2001.