

# Lawful Interception and Countermeasures

In the era of Internet Telephony

ROMANIDIS EVRIPIDIS



**KTH Information and  
Communication Technology**

Master of Science Thesis  
Stockholm, Sweden 2008

COS/CCS 2008-20

**LAWFUL INTERCEPTION AND  
COUNTERMEASURES:  
IN THE ERA OF INTERNET  
TELEPHONY**

**BY**

Romanidis Evripidis

Examiner: Gerald Q. Maguire Jr.

Supervisor: Gerald Q. Maguire Jr.

2008-09-22

Thesis submitted in partial fulfillment of the requirements for a Master's of Science Degree.

School of Information and Communication Technology  
Royal Institute of Technology  
Stockholm, Sweden

## **Abstract**

Lawful interception and the way it is performed have played a significant role in the effectiveness of this type of communication monitoring. Although the secrecy of interception and the related equipment are supposed to provide correct information to a law enforcement agency, there are some countermeasures that can be taken by the subject that can seriously undermine the collection of correct and accurate data.

This thesis project attempts to identify the problems that exist for interception of telephony (be it fixed, mobile, or via the Internet). Moreover, there are some suggestions for improvements how lawful interception should be performed in order to avoid possible attacks that could decrease the credibility of the intercepted data.

Numerous publications (in print or distributed on the Internet) have described weaknesses in the current state of the art lawful interception when using equipment that can be purchased in the market. This thesis presents improvements in how LI can be conducted in order to avoid these vulnerabilities. Additionally, there is a description of the key escrow systems and the possibility of avoiding one of their most significant vulnerabilities.

The main problem of the lawful interception is the rapid changes in telecommunications and the complicated architecture of the telecommunication networks, as both make monitoring vulnerable to specific countermeasures. An analysis of how lawful interception can take place and current countermeasures for lawful interception of Internet telephony are vital in order to identify the problems in carrying out such intercepts today and to make suggestions for improvements.

This topic is especially relevant given the current Swedish “FRA lagen” regarding interception of electronic communication going into, out of, and through Sweden. Not only is it important to understand how lawful interception can be performed or prevented, but it is also important to understand how information obtained from lawful interception could be purposely misleading or falsified.

## **Acknowledgements**

I would like to thank my supervisor, Professor Gerald Maguire, for his excellent guidance during the whole thesis. Even if we had to communicate through emails – because of our long distance – our communication did not confront any problems. His answers to my questions were very helpful and his proposals very illuminating. I really want to thank him for not having a problem of letting me working by distance.

## Table of Contents

Abstract .....	i
Table of Figures .....	v
List of Acronyms and Abbreviations.....	vi
1 Introduction .....	1
1.1 A brief introduction to lawful intercept .....	1
1.2 Methods to analyze LI and countermeasures against LI.....	1
1.3 Problem to be addressed by this thesis project.....	2
1.4 Importance of this problem.....	2
1.5 Thesis overview .....	3
2 Lawful Interception .....	4
2.1 What is lawful Interception? .....	4
2.2 Architecture of LI elements.....	5
2.3 Laws.....	8
2.3.1 <i>United States of America</i> .....	8
2.3.2 <i>Europe</i> .....	10
2.3.3 <i>Telecommunication Data Retention</i> .....	10
2.4 Greek tapping.....	11
3 LI in telephony (fixed and mobile) .....	13
3.1 How does LI work in a fixed telephony setting? .....	13
3.1.1 <i>Wiretapping methods</i> .....	14
3.2 Vulnerabilities in Wiretapping Systems .....	16
3.2.1 <i>Countermeasures against loop-extender taps</i> .....	17
3.2.2 <i>Signaling countermeasures against CALEA taps</i> .....	21
3.2.3 <i>Suggestions for reducing vulnerabilities</i> .....	22
4 LI in VoIP .....	24
4.1 How VoIP works .....	24
4.2 Similarities and differences in PSTN and Internet.....	26
4.3 SRTP in VoIP .....	27
4.4 Security Problems in interception of VoIP calls .....	28
4.5 Example of problematic VoIP interception.....	31
4.6 Solutions for VoIP interception.....	32
4.6.1 <i>A formal architecture for VoIP interception</i> .....	32
4.6.2 <i>Use of a Trojan</i> .....	33
4.6.3 <i>Watermark technique</i> .....	33
5 Key escrow .....	35
5.1 Government's key escrow goal.....	35
5.2 Clipper chip .....	36
5.3 Advantages of a key escrow system.....	37
5.4 Disadvantages of a key escrow system .....	38
6 SRTP/MIKEY and Key Escrow.....	43
6.1 Secure Real-time Transport Protocol .....	43
6.2 Multimedia Internet KEYing (MIKEY).....	46
6.3 Scenario: Minisip and Key Escrow .....	46
6.4 Problems to consider.....	48

7	Conclusions.....	49
8	Future work.....	50
9	REFERENCES.....	51
10	Appendix A.....	55

## Table of Figures

Figure 1: General Network Arrangements for Interception as proposed by ETSI (Adapted from [18]) .....	5
Figure 2: PacketCable Surveillance Model (Adapted from [3]) .....	6
Figure 3: Basic Elements in LI (Adapted from [3]) .....	8
Figure 4: Loop-extender architecture (Adapted from [12]) .....	15
Figure 5: CALEA wiretap architecture (Adapted from [12]).....	16
Figure 6: Dual-tone multi-frequency (DTMF) keypad and waveforms of generated tone (Adapted from [12]) .....	18
Figure 7: VoIP service with the use of analog telephone adapter (ATA).....	25
Figure 8: Problematic VoIP interception (Adapted from [22]).....	31
Figure 9: Experimental setup for real-time tracking of anonymous VoIP calls (Adapted from [31]) .....	34
Figure 10: Format of SRTP packet (Adapted from [44]) .....	43
Figure 11: SRTP Key Splitting (Adapted from [44]).....	44
Figure 12: SRTP encoding/decoding process (Adapted from [46]).....	45
Figure 13: Digital signatures for SRTP packets .....	47

## **List of Acronyms and Abbreviations**

ADMF	Administrator Function
ANSI	American National Standards Institute
ATA	Analog Telephone Adaptor
CALEA	Communication Assistance for Law Enforcement Act
CC	Contents of Communications
CCC	Call Contact Channels
CD	Call Data
CDC	Call Data Channel
CDR	Call Detail Records
CNID	Calling-number ID
CO	Central Office
DHCP	Dynamic Host Configuration Protocol
DNR	Dialed Number Recorder
DOJ	(U.S.) Department of Justice
DTMF	Dual-tone Multi Frequency
ETSI	European Telecommunication Standard Institute
E.U.	European Union
FBI	(U.S.) Federal Bureau of Investigation
FISA	(U.S.) Foreign Intelligence Surveillance Act
GUI	Graphical User Interface
HI	Handover Interfaces
IETF	Internet Engineering Task Force
IIF	Internal Intercept Function
IMS	Interception Management System

INI	Internal Network Interface
IP	Internet Protocol
IRI	Intercept Related Information
ISP	Internet Service Provider
KRC	Key Recovery Center
LEA	Law Enforcement Agencies
LEAF	Law Enforcement Access Field
LEMF	Law Enforcement Monitoring Facility
LI	Lawful Interception
MAC	Media Access Control
MF	Mediation Function
MIKEY	Multimedia Internet KEYing
MKI	Master Key Identifier
NSA	(U.S.) National Security Agency
PAS	Priority Access Service
PKI	Public Key Infrastructure
POTS	Plain Old Telephone Service
PSK	Pre-shared key
PSTN	Public Switched Telephone Network
PTN	Public Telecommunication Network
RES	Remote-control Equipment Subsystem
RFC	Request For Command
RTP	Real-time Transport Protocol
SMS	Short Message Service
SRTP	Secure Real-time Transport Protocol

TTP Trusted Third Party

U.S.A. United States of America

VoIP Voice over IP

## 1 Introduction

### 1.1 A brief introduction to lawful intercept

Governments always desire the ability to monitor people (inside or outside their own nation) in order to control these people and to formulate their policies. Even though there were several illegal means to achieve this, within the scope of democracy, the respect of human rights and the existence of different types of telecommunications many different legal ways of monitoring have been employed. The result of this trend was the introduction of the term of **lawful intercept (LI)**, which is used to describe both the **means** and **mechanisms** for law enforcement agencies (LEAs) and other government agencies to have *the technical ability and the authorization to perform* a lawful interception of the communication of persons that are believed to have committed illegal actions (or in some case to have the *intention* of committing serious crimes). However, as in most facet of human life, even though lawful intercept may often times provide useful information concerning illicit actions (i.e., terrorism or kidnapping) there have been several instances where interception has also been used for *illegal monitoring* of persons who were simply political adversaries or persons that had different opinion from the current government. As a consequence, many people mistrust lawful interception (even if it is legal under certain circumstances); as they believe that it violates one of their fundamental human rights – the right of privacy in communication and that the statutes governing lawful interception may potentially be misused to conduct *illegal* interception.

### 1.2 Methods to analyze LI and countermeasures against LI

Lawful intercept is not stable, as it must always adapt to new ways and patterns of communication. When LI was first used only the traditional telephone was used (although prior to this telegraph messages could also be intercepted). However, with the advent of mobile communication system and the Internet and their rapid & widespread adoption, governments and their agencies have begun to consider how these methods of communication can be included in their lawful monitoring (which includes how to modify and adapt earlier laws designed for telephony, telegraphy, and radio communication to address these new communication technologies).

In fixed and mobile telephony, interception can be considered quite easy as the communication passes through a small number of networks that belong to public or private companies. It is easy for these companies to control their networks and to define where agencies can easily connect their monitoring devices. In most cases this is particularly easy as these networks have very centralized control and there are relatively speaking few connections from one network to another. Interception is especially easy for fixed and mobile telephony operators as the agencies do not have to setup any devices as the operators have generally already installed the appropriate software for interception (in some cases this was a legal requirement for them as an operator in a given country). However, for Internet and particularly in Voice over IP (VoIP) interception there appears to be very weak control by government agencies. In part this happens due to the nature of the Internet. As the Internet is a global network that has many different routes to send information to the same destination and lacks centralized control. There has also been no

economic reason for Internet service providers (ISPs) to maintain detailed records of packet traffic, but rather to simply keep statistics (for example, total number of packets per month or total number of gigabytes sent or received in a billing cycle). Additionally, keeping more detailed traces of traffic would only increase the operator's cost and would neither improve service for the customer nor generate additional revenue. This means that it is not so easy to track the route of a conversation or to intercept all of a conversation of a possible criminal activity.

However, monitoring of traditional telephony & telegraph communications is not always successful. Target users can use several countermeasures in order to avoid interception, depending on the medium of communication they use (i.e. Internet, fixed, or mobile telephony). Additionally, these countermeasures may also result in false or misleading information being recorded.

### **1.3 Problem to be addressed by this thesis project**

As countermeasures often spoil lawful interception, governments have tried (and still try) to circumvent these countermeasures. One of the ways to surpass the threat (to LI) of increasing use of encryption was the use key escrow systems. With these systems, trusted organizations maintain a depository of the master keys of all the conversations in order for a LEA to be able to intercept later of someone who is a "target" of an LEA. These key escrow systems were thought to give the government absolute control for their monitoring as key escrow eliminates the problems of the LEA being unable to access the plain text of either the signaling or the communication - when faced with encryption and no-cooperation of companies that provide a communication service or in some cases even the end users themselves. In order to be certain success, the use of a key escrow system must include the majority – if not all – of the products and services that provide encrypted communication.

However, many doubts exist as to the potential for success of a large scale key escrow system. Many people including individual cryptographers and individuals point to severe security flaws and violation of privacy (respectively) as reason not to pursue key escrow system. The analysis of the opponents of the key escrow systems successfully raised serious doubts as to the possibility of a successful implementation of a key escrow system.

### **1.4 Importance of this problem**

There is no doubt that today many governments conduct LI. The exact number of intercepts which are conducted is not always clear, but a clear statistic is that the number of LIs in recent years is substantial [48]. Additionally the growing use of voice over IP is introducing new problems both for the regulations for LI of this means of communication and the technical abilities of the LEAs [49].

Many people in government agencies still believe that a key escrow system can be implemented and would definitely help them in intercept the communication of terrorists and other criminals. Even if it is considered difficult to implement a key recovery system which would include all the communication products that provide encryption, there are thoughts that it could be used at least in some of the products that are used by millions of

users worldwide (i.e. Skype or other VoIP hardware/software) [39]. This belief becomes more apropos as the pressure from governments to control communications increases every year, highlighted by global terrorism and the perceived need to eliminate it. Could key escrow systems yet be the answer for the interception of communications that were previously difficult to be intercepted? Still many people (cryptographers, human rights organizations, and individuals) do not agree that key escrow is the solution. More details of this are presented in Chapter 5.

## **1.5 Thesis overview**

This thesis project's main focus is to examine if a key escrow system could be viable when utilized with VoIP software. As many different types of VoIP software are used by millions of users around the world and many provide encryption between the parties in a conversation it is considered important by many government agencies to find a way that will provide them with easy and continuous access to these encrypted communications sessions. One of the solutions has been proposed several times are key escrow systems. However, there are strong indications that these kinds of systems cannot be considered a wise solution, as they suffer from security flaws and human rights violation. This thesis will examine key escrow in the context of a software VoIP system from a technical point of view in order to determine if there are also technical reasons why such a solution will not be feasible or successful in achieving the aims of many governments.

The objective of the overall thesis is to provide the reader with a general understanding of what lawful interception is and some of the problems that law enforcement agencies are facing. Chapter 2 gives a description of what lawful interception is and why is it desired so much by many governments. Chapter 3 describes how lawful interception can be performed in fixed and mobile telephony and enumerates some of the problems that LEAs can face due to specific countermeasures that targeted users can currently employ. Chapter 4 introduced VoIP communication and the problems that arise when attempting to monitor this communication. Chapter 5 presents an analysis of key escrow systems and the advantages and disadvantages of using them in communication products and software that encrypts the signaling and/or session content. Chapter 6 describes a way to avoid the problem of insider misuse of a key escrow system through the use of digital signatures. Finally, chapters 7 and 8 presents conclusions and future work (respectively).

## 2 Lawful Interception

### 2.1 What is lawful Interception?

Lawful Interception (LI) is not something new in our lives. During the last 50-60 years governments around the world have evaluated and used systems that are able to trace and intercept telecommunications in an attempt to prevent possible social hazards that might arise or to gather evidence for criminal prosecution of individuals for serious crimes. In the beginning the interception of the public telecommunications were done without any legal authorization. However, this situation has changed; over the years various legislatures have created and introduced laws defining the legal framework and authority for the government to perform interception.

One definition of Lawful Interception (LI) is “the acquisition of call identifying information and the interception of communications contents” by law enforcement agencies (LEAs) after receiving proper authorization from competent authorities [1]. The competent authorities are the mediator organizations that stand between the agency that wants to perform the interception and the network operator(s). Lawful interception can be performed in all modern systems -- including the Public Switched Telephone Network (PSTN), wide area wireless networks (e.g., mobile telephony), cable television systems, and the Internet. Although the monitoring of the fixed lines is a well-known procedure of law enforcement and intelligence services, in recent years challenging new obstacles have arisen. One such problem is tracking an end-user that uses a mobile phone or Voice over IP (VoIP); as both of these technologies allow the user to easily change their physical location (hence both offer *mobile communications*). Such mobile communications is significantly more difficult to intercept than traditional fixed telephony. Hence, users of such mobile communications are becoming much harder to track.

Although the details of LI differ between the different types of interceptions, there are some basic requirements that every interception has. A LI system must provide transparent interception of only the specified traffic, and the subject must **not** be aware of the interception. Moreover, during an interception other telephone users must **not** be affected in any way by degrading their provisioned service. Furthermore, in every interception there are some minimum data that must be collected and recorded in order for the intercept to be used later as evidences in a legal proceeding. So, in every type of interception there is a need to determine the presence, identity, and location of the parties of the specified communication.

Due to the increasing number of the terrorist attacks all over the world, governments and their intelligence services have re-examined the importance of legally monitoring telecommunications. Furthermore, in most countries these authorities assert that LI can be used in all cases of *wireless emergency calling* and *priority access calling* [2]. As far as the wireless emergency service is concerned, LI can be used for tracking and prosecution of the persons who intentionally make **false** emergency calls (i.e., calls to ‘911’ in Unites States and ‘112’ in Europe). Making false emergency calls is an illegal act, because it ties up resources that might be needed by someone who actually has an emergency. Moreover, LI can be used to enable a Priority Access Service (PAS). PAS can provide the necessary access to the authorities and governmental officials to make

priority calls *despite* the telecommunication networks being congested or otherwise lacking resources. Again there is a need to ensure that PAS is only used for proper purposes and that abuses can be prosecuted [7].

Nevertheless, there are many organizations and people that do not believe that the usefulness of interception in preventing, detecting, deterring, and prosecuting criminal or terrorist acts is worth the risk of the violation of individual's rights, privacy, and personal integrity. Thus in most jurisdictions there are a need for balance between interception and privacy.

## 2.2 Architecture of LI elements

Even though there are some differences in the details of how interception is performed, the main architecture of LI all around the world is basically the same. The primary LI requirements and standards have been developed by the European Telecommunications Standard Institute (ETSI) in Europe and by ANSI in the United States of America (U.S.A.). In the USA, the regulations concerning LI are spelled out in the Communications Assistance for Law Enforcement Act (CALEA) [3]. In both Europe and the U.S.A., a dominant theme was to design LI systems such that information concerning an interception is communicated *only* to those persons in the telecommunication operator's network that *must* be involved in a given intercept, in order to reduce the possibility that someone might compromise an investigation; while simultaneously ensuring that *only* the legally authorized type of interception is applied and that it is *only* applied *to the specified targets of the investigation* (i.e., to avoid intercepting traffic which is not the target of a legal intercept).

The organization that primarily defined the architecture of LI systems (not only in Europe but also worldwide) is ETSI. Figure 1 gives a generalized overview of the proposed ETSI LI systems architecture.

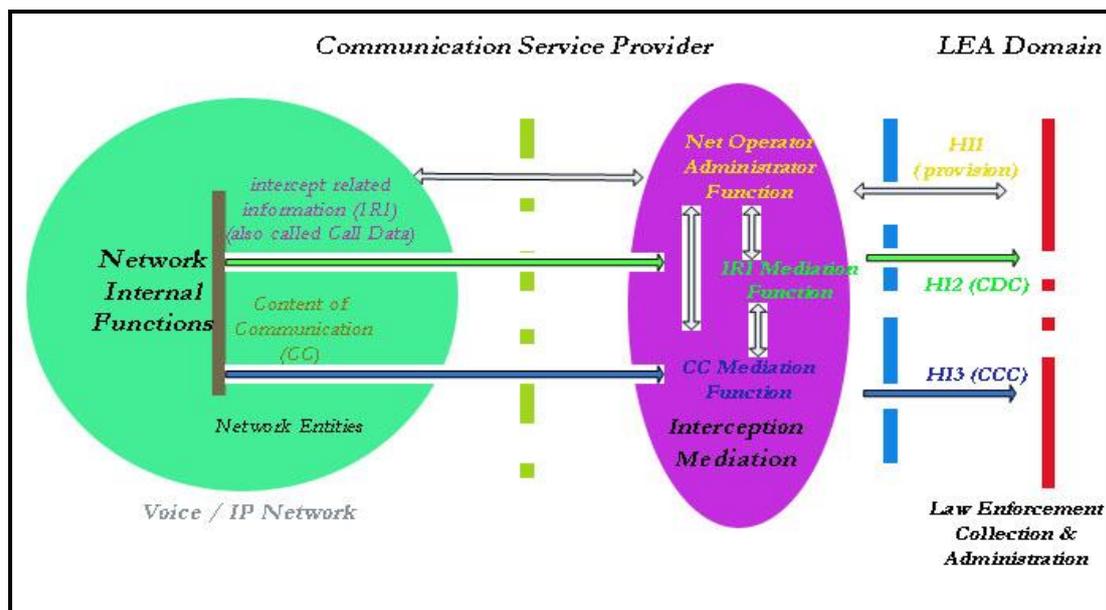


Figure 1: General Network Arrangements for Interception as proposed by ETSI (Adapted from [18])

This architecture describes how network operators and law enforcement agencies (LEAs) interact; even as networks expand and as these networks provide new services. This general architecture can be applied to all types of intercepts, e.g., to fixed lines, mobile calls, instant messaging, email, and VoIP. In the U.S.A., after the enactment of CALEA, the PacketCable surveillance model [3] is being used (See Figure 2). This approach has many similarities with the ETSI LI model and primarily differs in the terminology used.

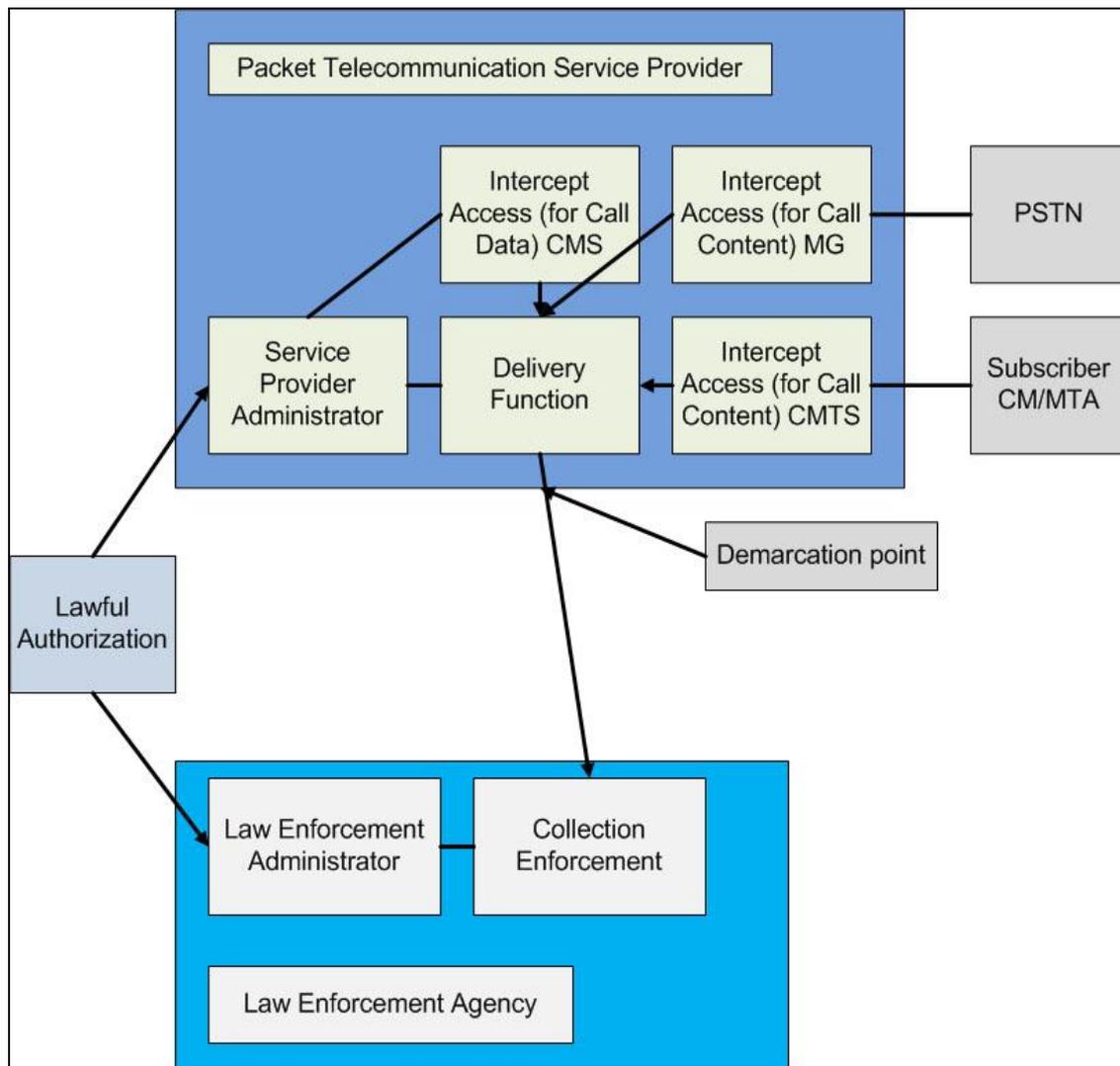


Figure 2: PacketCable Surveillance Model (Adapted from [3])

The call data can be separated into two broad categories: Intercept Related Information (IRI) {Europe} or Call Data (CD) {U.S.A.} and the Contents of Communications (CC). IRI/CD includes specific signaling information about the calling parties, such as the destination of a call (the called party), the source of a call (the caller), the time of the call,

call duration, etc. The CC includes the actual content of the communication, such as the message, video, voice, or other contents.

Figure 3 illustrates the distinct separation between the Public Telecommunication Network (PTN) and the networks which are used for the processing and distribution of the intercepted information. Generally, there are three basic elements in this architectural model [3]:

- An Internal Intercept Function (IIF) which is located in the PSTN operator's network nodes is responsible for collecting the targeted data – both the Intercept Related Information (IRI) and the Contents of Communication (CC) according to the specific LI request [3].
- The Administration Function (ADMF) is located inside the PTN and communicates with the IIFs and Mediation Function through an Internal Network Interface (INI). The Administrative Function (ADMF) is responsible for managing the orders that specify interceptions [3]. The ADMF is accessed through a web based graphical user interface (GUI) that is only accessible to authorized users. ADMF organizes all the tasks necessary for the interception. For each instance of an intercept order, a Warrant ID and Case ID are assigned by the LEA. For each of these intercept orders a starting and ending date is specified – this defines the exact duration of the interception. The intercept order also specifies the kind of the intercepted data that is to be collected (IRI, CC, or both).
- A Mediation Function (MF) is located inside the PTN which communicates with IIFs using INIs and the Handover Interfaces (HI2 and HI3) to communicate with the Law Enforcement Monitoring Facility (LEMF). Before transmitting data of these interfaces, the intercepted data (IRI and/or CC data) – which comes from the IIFs – are formatted properly. After validating that this data is to be provided to a specific LEMF, based upon the ADMF target details, the specified data is sent to the specified LEMF through HI2 (IRI) and/or HI3 (CC) [3].

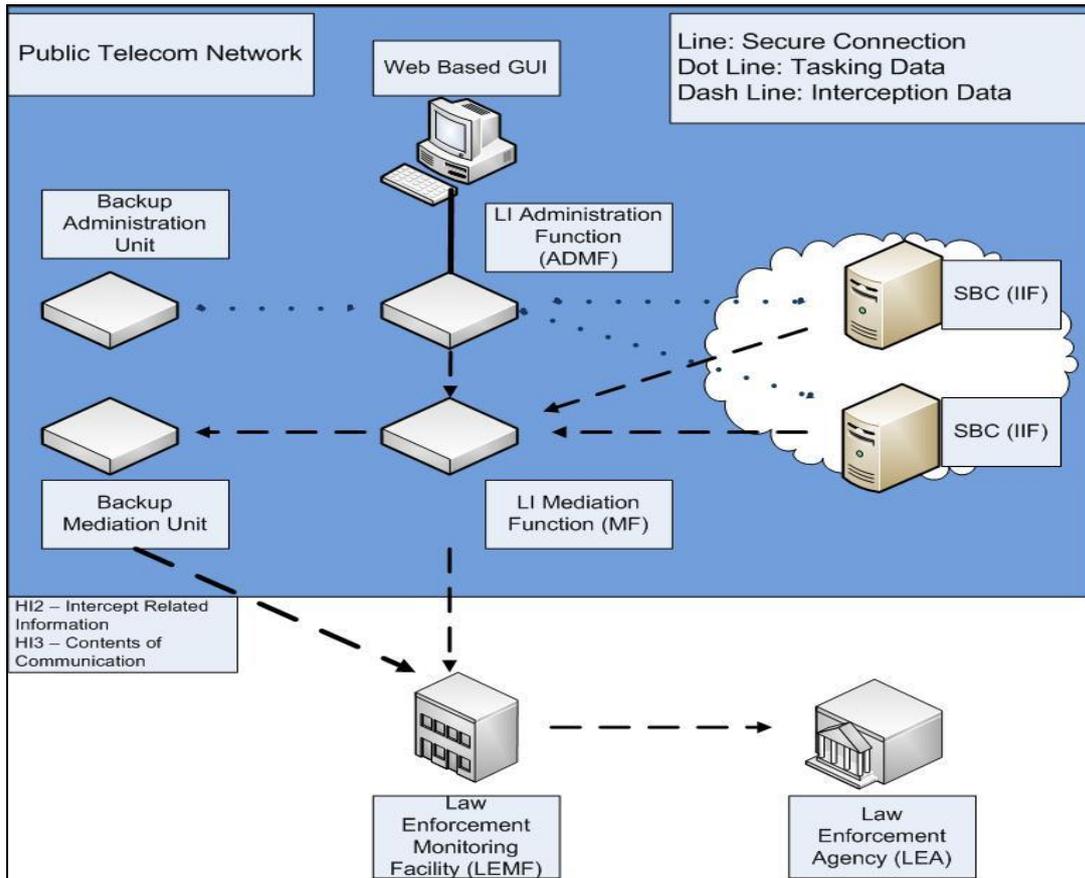


Figure 3: Basic Elements in LI (Adapted from [3])

## 2.3 Laws

A majority of the countries around the world have enacted laws that enable government agencies (including their intelligence services) to monitor telecommunications networks. This legislation attempts to protect the private lives of citizens against possible abuse of power, based upon unwarranted interception. The pioneer among the countries that established a legal framework for lawful interception was U.S.A. Other countries have imitated the United States in this field, but there are some important differences in the details of LI according to each country's perception of their own needs.

### 2.3.1 United States of America

As U.S.A. faced the problem of LI very early, their government passed a wiretapping law under Title III of the Omnibus Safe Streets and Crime Act of 1968 [19]. This law defined the wiretapping procedures for criminal investigations. Later when "national security" was viewed as being at stake, another federal law was enacted to allow the lawful electronic monitoring of communications regarded as "foreign intelligence information" [4]. This legislation specified how these intercepts were to be conducted

within territory that was controlled by the United States. This law is widely known as the Foreign Intelligence Surveillance Act (FISA) [4] and it permitted surveillance with warrants in three basic cases [4]:

- Any person in the United States that communicates through wire
- A U.S. person <sup>1</sup> in the United States that communicates through wire or radio
- Anyone inside the U.S.A. that communicates through radio with people - all of whom are in the United States

However, FISA included an important clear exception: There was no need for a warrant to intercept radio communications between people outside of and people inside the U.S.A. *unless the intelligence services were monitored a specific U.S. person that was inside the United States*. Although this was initially considered a temporary exception, there has been no amendment of this point of the law; hence this type of warrantless interception exception continues to exist. In February 2008, the U.S. Congress tried to pass a bill that would amend the FISA law [15], by granting new authorities for conducting electronic surveillance against foreign people. But the promised retroactive immunity for the telecommunication companies that helped the national agencies to perform such intercepts is still an open issue, as the House of the Representatives passed the bill *without* this retroactive immunity.

The increasing use of fiber optics has changed how international communications are performed; specifically the percentage of communications that is carried by radio communications links has decreased very significantly. As a direct consequence, the warrantless interception exception that existed under FISA became less applicable. There was a major need for an update to FISA; in response to this demand, the U.S. Congress enacted the Communications Assistance for Law Enforcement Act (CALEA) in 1994. The main intention of this law was to require “the telecommunications carriers ensure that their equipment, facilities, or services provided the necessary capabilities” [1] in order to assist the government in its electronic surveillance. Therefore, every telephone company in the U.S.A. had to create the necessary infrastructure that would allow the law enforcement agencies to monitor a phone call as transmitted over its network, and also make available Call Detail Records (CDR). In addition, the law requires that telephone companies provide a secure method of telephone surveillance that is not traceable by the target person. Finally, the data intercepted by a CALEA device is to be delivered to the premises of the LEA simultaneously with its capture.

After the terrorist attacks in the City of New York on the 11<sup>th</sup> of September 2001, the U.S. Congress enacted another law that enhanced both CALEA and FISA. This law is known as the USA Patriot Act. The main goal of this law is to extend the interception to include the publicly available broadband networks and VoIP services. Moreover, there was “no requirement of any warrant (in any medium) for communications of U.S. persons located in the United States with persons ‘reasonably believed to be located outside the United States’ ” [5]. However, since modern telecommunication technology (e.g., Internet, mobile phones, VoIP, etc.) often supports mobility, it is not very easy to

---

<sup>1</sup> U.S. citizens, permanent residents, and U.S. corporations

determine the real location of a targeted person. So, surveillance may include traffic of people that are **not** supposed to be monitored -- as the intelligence services can simply claim that they did know the location of the parties. Additionally, several amendments to U.S. legislation have been enacted in order to broaden the powers of government surveillance to include more people (i.e. inclusion of financial transactions, particularly those involving foreign individuals and entities; broadening the definition of the terrorism to include “domestic terrorism” [8]).

### 2.3.2 *Europe*

The European Union (E.U.) issued a Directive (95/46/EC [9]) that defines the Lawful Interception of Telecommunications with the simultaneously protection of individuals. In general the directive defines the nature of the intercepted data and gives additional details on how the data may be used. Although many countries inside the E.U. were not pleased with this directive, all of the member countries finally accepted it as amended in Directive 2002/58/EC [10]. This acceptance occurred due to political pressure following the increase in terrorist attacks all over the world. Although Europeans are supposed to be more sensitive concerning privacy, their laws concerning data retention can be characterized as even more strict than those in the U.S.A. Finally, Directive (2006/24/EC [11]) specified “the retention of the data that is generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks” [11].

### 2.3.3 *Telecommunication Data Retention*

One of the major parameters in electronic surveillance is data retention (or preservation). Retention refers to “the storage of telephony and Internet traffic” by companies that provide these services. This data specifically includes Call Detail Records (CDRs) (as they provide details of the incoming and the outgoing calls), sent and received emails, visited web sites of the customers of ISPs, and also the location of the customers of telecommunications companies. The preservation of such data gives governments the ability to collect, analyze, and finally monitor the life of hundreds of millions of individuals - allowing in this way mass surveillance of almost the entire population.

The European Union issued a Directive (2006/24/EC [11]) that specifies in detail the period of data preservation and also the kinds of information to be stored. The period varies from 6 months to 2 years, according to the nature of the data. The content of the CDRs can include the following [11]:

- Trace and identify the source of a communication
- Trace and identify the destination of a communication
- Date, time, and duration of a communication
- Identity of the type of communication
- Communication device
- Location of mobile communication equipment

Although all the member-countries of the E.U. must adopt this directive as a national law, countries maintain the right to postpone the application of the directive in the fields of the Internet (visited web sites, sent and received emails, and VoIP) for more than 18 months. Furthermore, each country can extend the storage period of the CDRs according to its own needs, even though the initial proposal of the committee of Civil Liberties, Justice, and Home Affairs (LIBE) was to retain them for **maximum** period of 12 months.

As a consequence of the legislation of the above directive, many organizations and also individuals declared their opposition to the retention of this kind of data - as the investigation of terrorist attacks could be solved by keeping only the legally intercepted data. Moreover, CDRs can assist the police to find people that committed crimes *only after a terrorist attacks* and **not** before. Another point of the opponents to data retention is the excessive power that the state obtains to monitor and control the lives of individuals. As has been all too frequently observed, a common practice is that CDRs will be used by the police against any group or individual who opposes the government's plans or actions. Finally, opponents highlight the time frame of the preservation of the CDRs as excessive. As the retention time can be extended up to 2 years or more, this also causes a serious economical impact upon all the companies that are required to keep all these records, as in many countries they must pay the preservation costs (an exception to this is the country of Finland, where the government pays the storage costs).

#### 2.4 Greek tapping

As described above, the main reason for lawful interception is supposed to be the identification of terrorists and the prevention of their attacks. However, the implementation of the required software and hardware for this legal interception can also be used for illegal monitoring of people (i.e., illegal tapping); even without the knowledge of the network operators or the LEAs that such interceptions are happening. One of the most recent incidents that perfectly illustrate this situation is the illegal interceptions that took place in Greece during the Olympic Games in 2004.

In March 2005 the largest mobile company in Greece – Vodafone Greece – announced that rogue software had activated the lawful interception mechanism implemented in the telephone switches used by the company. The result was that over 100 mobile phone numbers which belonged to a number of important Greek politicians (including the prime minister of Greece) and high ranking military staff had been tapped for at least one year without any warrant from a LEA. Many have asked: How could this happen? It is clear that some “really knowledgeable people managed to infiltrate the network of Vodafone from outside or subverted it from within” [6], as the perpetrators managed to implant the illegal software in four of Vodafone's mobile switching centers. These switches are located in the heart of their mobile phone network. This software took advantage of an upgrade to the switches that took place in 2003. This upgrade to Vodafone's network was to provide the ability to monitor its customers via a remote-control equipment subsystem (RES). The RES can copy a conversation (of a wiretapped phone) to a second stream and send this copy to the LEA. The rogue software activated the RES for specific phone numbers while simultaneously erasing any tracks that might reveal its presence. The key to the illegal interception was “to use the capabilities of the RES without using the

dialog-box of the interception management system (IMS) that would have made auditable logs” [6].

The uncovering of this interception happened accidentally. The intruders upgraded their software in January 2005. However, this caused several problems in the delivery of text messages to other mobile phone companies. Due to this problem, Vodafone with the help of Ericsson’s specialists started an extensive search for the cause of the SMS delivery problem and they finally discovered the existence of the rogue program.

As far as the identity of the intruder, many scenarios have been suggested, but none of them offers sufficient evidence. One of the dominant theories is that a Greek company, Intracom Telecom, which had taken over the programming evolution of a part of Ericsson’s AXE, was involved. Moreover, one of the hacked Vodafone exchanges was located on the campus of the main Intracom facility. Another scenario that appeared in newspapers pointed toward United States agencies, which feared terrorist attacks during and on the Olympic Games in Athens. Moreover, the U.S.A. agencies were believed to have the knowledge necessary to perform the interception. Finally, the location of the monitored phones correlated with apartments and other property under the control of the U.S. Embassy in Athens.

However, the Greek state fined Vodafone €76 million and fined Ericsson €10 million for this interception and their mishandling of the situation (e.g., not preserving evidence). Still there is no final decision on the identity of the perpetrator(s) or the reason for the monitoring.

### 3 LI in telephony (fixed and mobile)

Both criminal investigations and national security concerns oblige governments and their law enforcement agencies (especially in the U.S.A. and in the E.U.) to utilize voice telephone interception systems in order to collect and analyze the information that was produced by telephone calls. Such interceptions have helped police agencies in the investigation of complex criminal affairs and in the prosecution of the culprits – despite frequent accusations of misuse of the interception technology.

However, after some research concerning the credibility of the telephone interception tools – with respect to legitimate security concerns – such interception was proved to be very susceptible to a plethora of countermeasures that a subject<sup>2</sup> can use in order to prevent the accurate tapping of the subject’s calls. These countermeasures mainly exploit “the use of in-band signals between the telephone network and the law enforcement agency” [12]. The use of these countermeasures showed that they can “obscure not only the content of a call, but also the metadata that indicates the presence of a communication and also its endpoints in a way that is sometimes difficult to be detected” [12].

A serious consequence of the above is to undermine lawful interception to an extent that not only negatively impacts the accuracy of the tapped contents, but also to “the acceptability and weight of legal evidence derived from it” [12]. These weaknesses can be alleviated with the application of some recommendations that try to reduce the susceptibility of the lawful interceptions techniques to attacks. These recommendations are detailed in section 3.2.3.

#### 3.1 How does LI work in a fixed telephony setting?

In the U.S.A. there are two federal laws that regulate how telephone wiretapping can be conducted; the Federal Wiretap Act (Title III) and the Foreign Intelligence Surveillance Act (FISA). These two laws specify three categories of wiretaps that law enforcement agencies can use.

The first category is called a pen register or dialed number recorder (DNR) [12] which is an electronic device that can record all the digits dialed from a particular telephone line and also other outgoing signaling information. A pen register tap is only allowed to provide traffic analysis of the targeted line and not the audio contents of calls. In order to be allowed to use this interception technique, a judicial authorization is needed; which is not very hard to acquire – basically the requirement is a suspicion of an illegal activity.

The second category is the “trap and trace device” [17] which is an electronic device that is used in order to show all the incoming calls for a specific targeted telephone number. This device is generally used in concert with a pen register (as the pen register deals with the outgoing calls of a targeted telephone number).

The third category is known as full audio interception. From the name it is clear that this intercept records not only the signaling and the dialed numbers, but also the actual audio (i.e., the call contents). Due to the addition of the audio the authorization for this

---

<sup>2</sup> “Subject” is a term for the person who is target of the LI

kind of interception is more judicially difficult than obtaining a pen register or trap trace authorization. Moreover, the expenses of a full audio interception are high, as it requires “continuous real-time monitoring” [12] by the law enforcement agency.

Apart from the above three categories of interception, law enforcement agencies are casually using telephone records (i.e. CDRs) as a source of information about subjects. The main disadvantages of the telephone records are that they (1) concern a “subject’s past telephone activity” [12] instead of current or potential future activity and (2) they are not practically available to the law enforcement agency until sometime *after* the activity has occurred [12].

### *3.1.1 Wiretapping methods*

As far as the above three categories of telephone intercepts (often referred to as *wiretaps*) are concerned, the same technological equipment can be used in order to implement them (in a pen register intercept the copying of the audio can be disabled). There are two widely known methods that can be used for wiretapping: the loop-extender and the CALEA taps.

A loop-extender (shown in Figure 4) is the oldest form of wiretapping technology. It is used exclusively for wire line (POTS) telephone lines. The main feature of this method is the use of a second line (called a “friendly line”) that connects the subject’s telephone line (called a “target line”) directly with the law enforcement agency’s premises. The friendly line can either be a dedicated leased line or a regular dial-up line that can be placed on the subject’s premises or in the telephone switching center. In order to implement this tapping no special hardware is needed; where the loops are jointed, a small device – called a loop extender – is located in order to “ensure proper isolation and level equalization of the intercepted content” [12]. Note that loop extender is also used for some normal telephony local loops to increase the distance that the subscriber can be from the local exchange. The loop extender captures all the audio (and signaling) on the target’s line and sends it via the friendly line to the premises of the law enforcement agency. At the end of this loop, the pen register equipment which is located in the law enforcement agency building decodes the dialed digits and the call activity signals. Moreover if authorized the equipment can also record the call contents. Note that when using a loop extended the telephony operator has no control over what information the law enforcement agency records for an outgoing call, since the LEA receives all of the signaling and call contents.

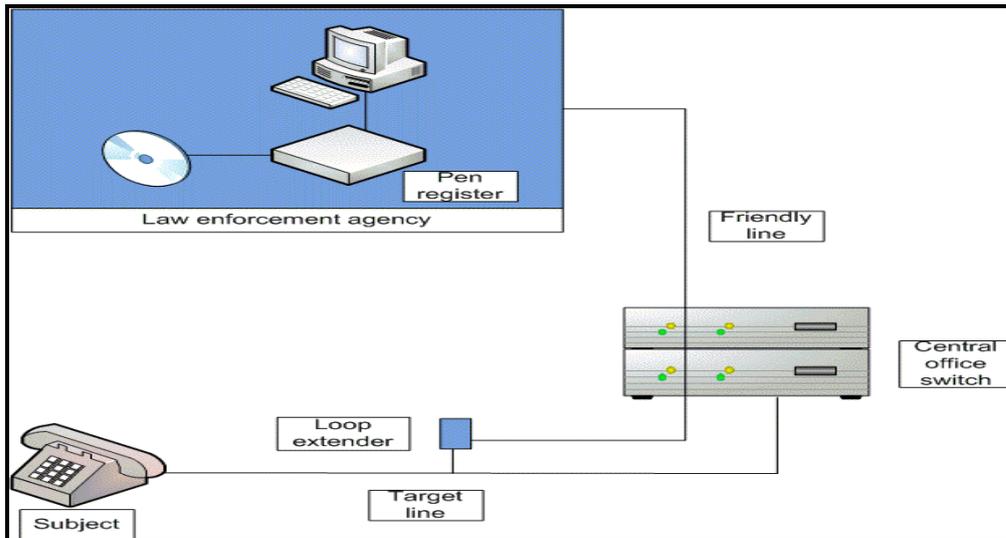


Figure 4: Loop-extender architecture (Adapted from [12])

Apart from the loop extender there is also a newer wiretap method – known as CALEA taps (see Figure 5) – which were designed to comply with CALEA requirements. The CALEA standards specify the existence of a standard interface between the law enforcement agencies and the telephone network providers (both wire line and cellular telephone companies). The main difference between CALEA taps and the loop extender is that the decoding of the signaling information is now performed by the telephone company, rather than the law enforcement agency. Instead, the law enforcement agency is connected to the telephone company through a standard interface (known as J-STD-025A [14] or the newer version J-STD-025B [15] which is used mainly in America). Another difference from the loop extender is that the J-STD-025A standard uses two categories of separate telephone lines: one for the signaling information (dialed digits, on-hook status, call times, line status, etc.) and the other for the call audio. The line that transmits the signaling information is known as a CDC (Call data channel) and is linked with all the telephone lines that the law enforcement agency is monitored. The lines of the second category are known as a CCC (Call content channel); this line transmits the voice content of all the active monitored lines. The CDC can carry information for more than one active interception at a time. In contrast, a particular CCC line can carry the audio information from only one tap at a time, but it can be time multiplexed to carry the call contents from different subjects at different times by dynamically assigning the line for the active targeted lines during a call.

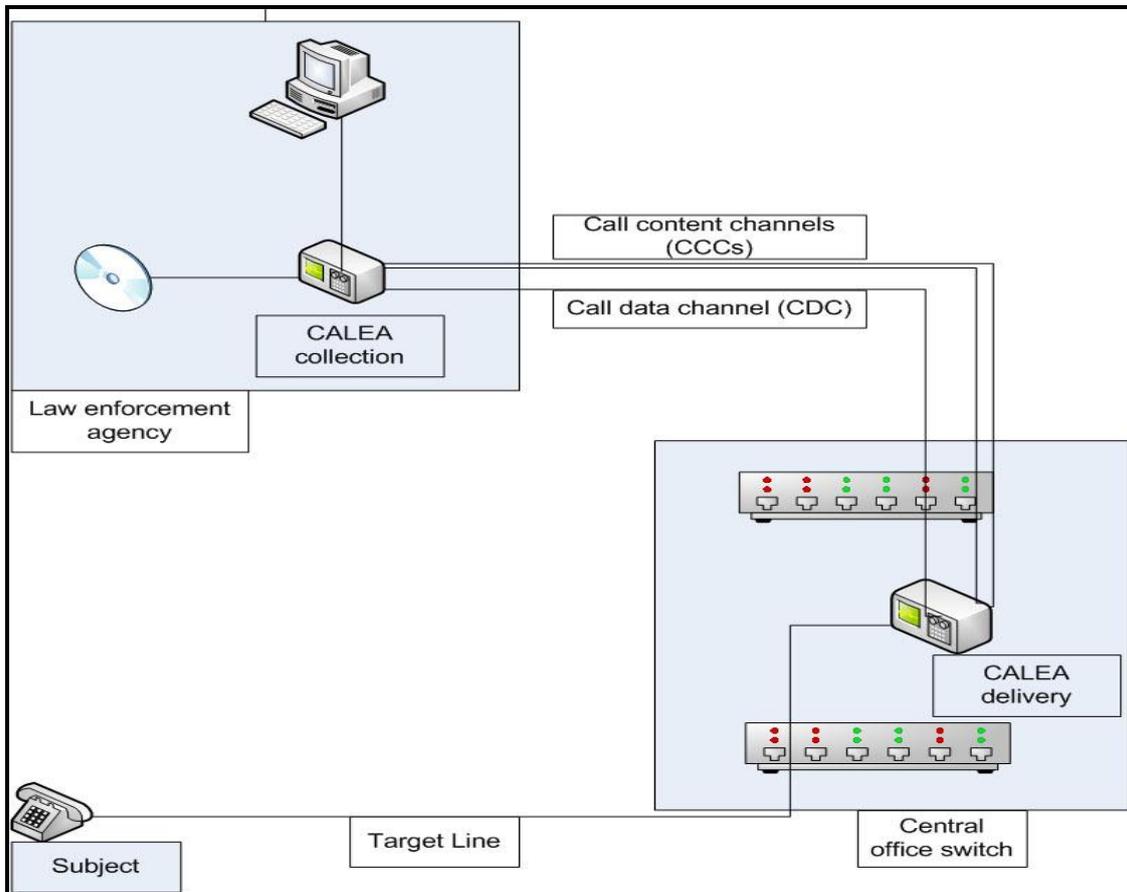


Figure 5: CALEA wiretap architecture (Adapted from [12])

### 3.2 Vulnerabilities in Wiretapping Systems

Even though there is little information (i.e., only limited public available) about the susceptibility of the wiretapping systems, an extensive survey by Sherr, et al. [12], a research team at the University of Pennsylvania, showed that there are a lot of threats that can negatively affect the accurate capturing of telecommunications data.

Apparently, the most prominent threat against a telecommunication tap is the detection of it. If the advantage of the secrecy is lost, then the whole wiretapping procedure is likely to be unproductive or counter-productive. Detection can be performed in many ways. First, if the tapping uses a loop-extender device which is placed near the subject's premises it can easily be detected by physical inspection. Moreover, it is known that taps that change the target's line transmission characteristics and this can sometimes be detected with "electronically means like sensitive loss measurements or time-domain reflectometry" [12]. Another threat is "the penetration of a telephone company's information systems by a computer compromise or physical burglar" [12]. While J-STD-025A specifies that taps should be performed in an undetectable way from the subject's point of view, it does **not** describe how to achieve this.

The most severe threat that a subject can unleash is the encryption of the voice and signaling, as this could lead to content and/or signaling obfuscation. The use of cryptographic techniques is often utilized to provide security between the communicating parties. However, as of today, voice encryption is not widely used by subjects [12] and digital voice systems for encryption of “analog telephones systems are not yet widely available in the market [12]. The major disadvantage of encryption is although it may protect the contents of a call, it generally does not protect the signaling information. However, by forwarding all call through an anonymizing third party, even the call signaling information can be moved to the call contents and hence also be encrypted.

Another public available countermeasure that can affect the CCCs (Call content channels) in CALEA taps is a denial of service attack against CALEA CCCs [12]. This countermeasure exploits the dynamic assignment of CCCs as “the number of different voice channels associated with a monitored line is potentially unbounded if the subject subscribes to a call-forwarding service” [12]. If the subject and their correspondents forward their calls elsewhere, there will be no CCCs available for the monitoring of important calls as every additional call needs a new CCC. It is not clear if the J-STD-025A standard specifies any defense against this countermeasure, as the published literature says little about this type of attack.

Finally, evasion and confusion are two additional countermeasures that a subject can use in order to avoid interception. Evasion means that the subject can prevent legitimate data from reaching the monitoring system. Confusion means that the subject sends additional false data along with the original data to the monitoring system. The result of either of these countermeasures is a serious degradation of the intercepted data’s credibility. In order to avoid evasion or confusion there are some defenses that can be used (however, these lead to the eavesdropper’s dilemma<sup>3</sup>) as by using one defense to avoid evasion the eavesdropping system is more susceptible to confusion and the reverse.

### *3.2.1 Countermeasures against loop-extender taps*

As in-band<sup>4</sup> signaling is used mainly in the loop extender taps, it leads to vulnerabilities that make loop extender tapping very susceptible to attacks. Mainly, there are three kinds of countermeasures that a subject can utilize in order to avoid tapping by the law enforcement agencies: Dialed digit spoofing, incoming calling-number ID spoofing, and line status spoofing and recording suppression.

#### *3.2.1.1 Dialed digit spoofing*

Using this countermeasure the subject can mask the dialed numbers of an outgoing calling by exploiting the weakness of the way that tapping devices decode dialed numbers and audio signals. As the transmission (from the phone device to the call switching center) of a dialed phone number is done in analog form, the use of audio dual-tone multi frequency (DTMF) signals is obligatory.

---

<sup>3</sup> The eavesdropper’s dilemma (See Appendix A) appears whenever the law enforcement agency lacks knowledge of how the network and receiver process traffic or if it destroys information processed at low layers of the protocol stack

<sup>4</sup> In band signaling is the exchange of signaling (call control) information within the same channel that the telephone call itself is using [13]

DTMF digit signals are the outcome of two audio frequency tones: “the ‘low’ tone” [12] which represents the horizontal row of the keypad of a telephone device and “the ‘high’ tone” [12] which represents the vertical column of the keypad. These two tones signals are combined when a telephone user presses a key on a phone device in order to generate a tone that will specify to the call switching center the pressed key. The DTMF standard specifies the existence of the numbers (0-9), characters (\* and #), and a column with the letters (A, B, C, and D) on the keypad.

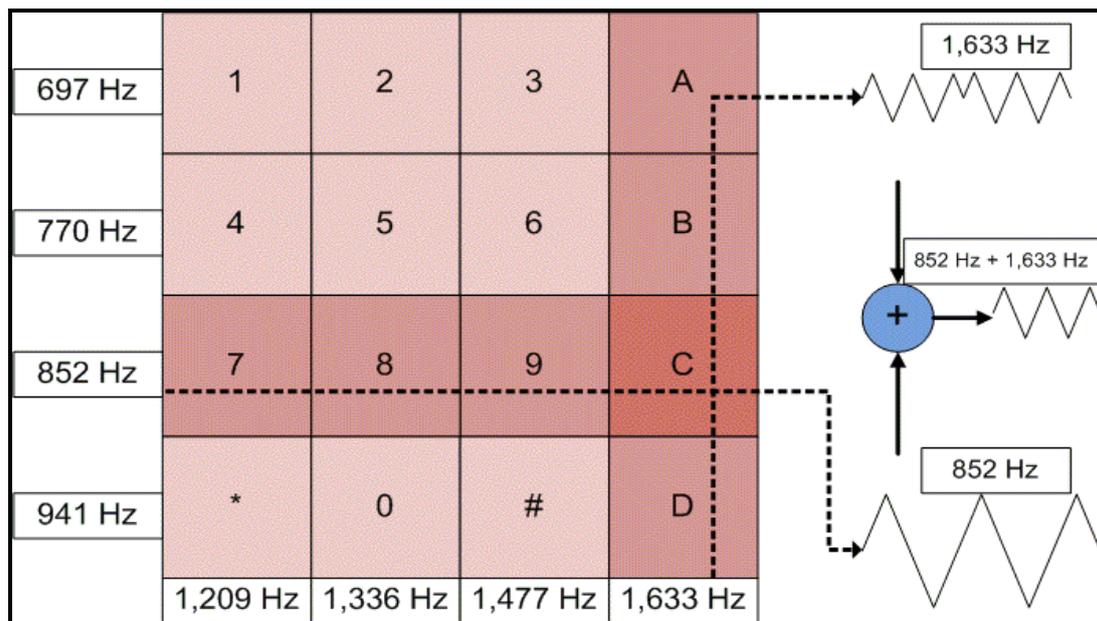


Figure 6: Dual-tone multi-frequency (DTMF) keypad and waveforms of generated tone  
(Adapted from [12])

Even though in the majority of cases the generated tone signals are accepted by the DTMF decoders, there are times that a generated tone signal that is on the edge of the standard can be rejected by some DTMF (but not all) decoders as invalid. The parameters that specify if a generated tone is accepted or not by a DTMF decoder are: the precise frequency of the ‘low’ and the ‘high’ tones, their amplitude, the signal duration, waveform distortion, external noise, etc.

According to the experiments conducted by Sherr, et al. [12] analog telephone subscribers can discover the threshold of their switch’s DTMF decoders, then later uses this knowledge to create signals that will be treated differently by other decoders. In a tap system that uses a loop extender device there are two different DTMF decoders, one at the telephone company’s switch and the other at the law enforcement agency. Every dialed digit from the subject’s line is processed independently by the two separate DTMF decoders. As a result a dialed digit may be accepted as valid by the telephone company’s switch decoder and not the wire tapper or the opposite.

Sherr, et al.[12] showed that after discovering the threshold characteristics of a telephone switch’s DTMF decoder with the help of a laptop, the subject can construct bogus digit encodings in order to be rejected as invalid by the switch’s decoder. Despite

these encodings being rejected it is very possible that they will be accepted by the law enforcement agency's decoder. The bogus signals are used to confuse the monitoring system. On the other hand, "non-standard digit encodings" [12] can be created that will be accepted by the telephone switch's DTMF decoder, but will be ignored by the law enforcement agency's tap. These signals attempt to evade such a tap.

From the above it can be seen that a wiretap can be either liberal or conservative. If a wiretap is liberal, it will accept signals that will be ignored by the switch, thus achieving its goal of confusing the tapping agency. If the wiretap is conservative, the digit signals will only be accepted by the switch and these digits will evade interception.

The effectiveness of this confusion and evasion is highly likely in every wiretapping system that uses its own DTMF decoder and interprets the dialed digits on its own. Even though a subject cannot be sure of how liberal or conservative a law enforcement agency's DTMF decoder is; the subject can unleash both of the threats in concert in a whole digit sequence. This means that "the subject sends n random noise digits" [12] that are dialed among the k 'real' digits in a random sequence. The 'real' numbers are sent using evasion and the noise numbers using confusion. With the above combined method, the real dialed number can be masked from the tapping. An important point is that the recording of **false** data will compromise the ability of a law enforcement agency to present this data as evidence during a trial.

#### *3.2.1.2 Incoming calling-number ID spoofing*

One of the extra options that a subscriber has is to screen the calling-number ID (CNID) when he or she receives a call via his or her phone device. Between the first and the second ring, with the help of in-band signaling, the calling-number and sometimes the related name are transmitted and can be presented on the screen of the phone or a special phone device belonging to the subscriber. When a subject has activated the CNID service it is clear that the law enforcement agency can also monitor this information.

However, if the subject uses some countermeasures lawful interception can lose valuable information. Sherr, et al. [12] showed experimentally that with the use of forged signals the subject can confuse the capturing of the real CNID. However, evasion cannot easily be performed as the central office (CO) is responsible for the transmission of the CNID. During their survey the researchers periodically replayed a counterfeit CNID – with the help of a sound card – whenever the subject's line was on-hook. The result was successful deception of all the tested DNR devices. Another countermeasure is to generate counter-signaling when the CNID is being set – since with suitable signal processing the signal can be received at the subject device, but corrupted for the loop-extender tap(s). Technically this has to do with the need for a given signal to noise ratio to properly receive the data, the likely longer path to the LEA decoder, and the three wire analog telephone circuit.

### 3.2.1.3 *Line status spoofing and recording suppression*

The main disadvantage of the loop extender taps is the use of the in-band signaling by the telephone, the central exchange, and spoofing devices used by the subject. Due to this, a subject can remotely unleash an ‘attack’ by spoofing the on-hook signal which is generated from the loop extender with a result of disabling the audio monitoring equipment.

It is known that in loop extender systems, the signaling data and the audio are transmitted to the law enforcement agency over the same channel – the friendly line – in the “analog voice-band domain” [12]. This is the characteristic that makes the loop extender susceptible to exploitation from the subject; but it is also the feature that makes such a loop-extender inexpensive and low complexity.

A monitoring system gathers many call-processing signals from the subject’s line. These signals can be separated into two categories. In the first category, the signals (DTMF encoded dialed digits, CNIDs, dial tones, and ringing signal) are encoded in “the voice-band audio domain” [12]. So, these signals are simply passed to the friendly line through the loop extender and can be deciphered at the law enforcement agency. However, the second category of signals (including on/off-hook state and incoming-call ringing signals) are not encoded in the voice-band audio domain. Even so, the loop extender device can detect these signals on the subject’s line, but cannot relay them in the same form over the friendly line (to be processed remotely by the law enforcement agency). To solve this problem these signals are encoded as “special audio tones superimposed on the friendly line audio” [12], in order to be recognized and decoded later by the law enforcement agency.

A signal that is very vital and is **not** included in the voice-band audio domain is the on/off-hook status. The on-hook status means that a line is idle, i.e., – no audio or other signals are transmitted at the moment. In order for the law enforcement agency to recognize that the subject’s line status is on-hook, an idle tone is transmitted continuously on the friendly line. If there is a transmission of some signals over the target’s line, then the idle tone is automatically removed from the friendly line. The representation of the idle tone is done using the DTMF C digit that produces a two-frequency audio signal (852Hz+1,633Hz). According to the survey [12] all the loop extenders in the market use the C tone to indicate that the subject’s line is on-hook. As the on/off-hook status signal is sent through the same line, it is easy for the subject to send an “identical-sounding signal” [12] during a call in order to spoof the law enforcement agency’s on-hook detector, thus avoid the call being recorded.

According to Sherr, et al. [12], some loop extenders devices have unexpected behaviors and can perform unexpectedly if a DTMF C idle signal is not followed by new call setup signals. If this happens, then the friendly line will be disconnected and will again try to reconnect. The reestablishing of the connection will take at least 30 seconds. During these 30 seconds no recording of audio can be performed. So, if every 30 seconds the subject sends a DTMF C signal, then no audio can be monitored by the LEA wiretap hardware!

Moreover, the use of the DTMF C idle tone can negatively affect full audio wiretaps. As loop extenders do not record audio when a DTMF C tone-pair is transmitted, it is easy for the subject to continuously send a low volume DTMF C tone-pair during a call in order to spoof the loop extender and avoid the recording of the call. While it might be assumed that only low quality call service can be established if a DTMF C tone-pair is transmitted continuously, however, this can be avoided if the DTMF C tone-pair is not sent at high amplitude. Sherr, et al. [12] showed that the loop extenders could be deceived and would stop recording even when the transmission of a low volume DTMF C tone-pair was present, yet this allowed an acceptable and comfortable conversation to be conducted. Additionally, using signal processing the DTMF C tone-pair could simply be recognized and subtracted from the received signal – as long as it has roughly constant amplitude and the sum of the amplitudes of the actual signal in the pass band for these two tones is within the range of the receiver.

The above countermeasure can be prevented only by filtering the DTMF C tone-pair signal from the point of the loop extender, in order to send the correct information over the friendly line. However, according to Sherr, et al. [12] no loop extenders in the market can perform such filtering.

### *3.2.2 Signaling countermeasures against CALEA taps*

As the loop extender taps face serious problems with the credibility of the wiretapping due to the in-band signaling, the J-STD-025A standard tried to avoid this problem by separating the channels that deliver the subject's signaling (CDC and CCC) with a simultaneous decoding of the DTMF tones at the telephone company's switch instead of by a device at the law enforcement agency. The main idea was that by using a separate channel for the signaling (out-of-band), line status spoofing could be prevented. Moreover, as the DTMF decoding is now done in the telephone company switch's call processing system, it is more probable that the "reported digits" [12] to the law enforcement agency are more accurate than those that would be produced by the loop extender device and that these reported digits correspond to those used by the switching center.

Although CALEA taps are considered more robust than the loop extender taps, CALEA's taps face similar credibility threats. The J-STD-025A standard specified "only a standard interface between the telephone company and the law enforcement" [12]; the problem is that it did not specify any particular implementation of this interface. So, even though the telephone company's switches are supposed to report the real digits that were produced by the call processing system, there is no guarantee that they actually do. As a consequence, the law enforcement agency may gather false calling-numbers. Furthermore, Sherr, et al. [12] showed that 'post-cut-through' digits (for example, extra digits that could be used with a direct inward dialing exchange to reach a specific extension of the destination PBX) which are transmitted through the CDC can still be confused or evaded as they are processed by a "remote endpoint" [12] and not by the switch. Finally, in CALEA taps there is still the countermeasure of recording suppression that results in the evasion of call-conversation recording.

In some CALEA implementations the DTMF C tone-pair signal is used in order to declare that the CCC channel is idle. In many CALEA CCC systems this DTMF C

tone-pair signal will disable recording. The continuing use of the DTMF C tone-pair signal may be motivated in order to achieve backward compatibility with the loop extender systems. Moreover, the U.S. Federal Bureau of Investigation (FBI) and the U. S. Department of Justice requested the use of a continuity tone (not necessarily the DTMF C tone-pair signal) in order to know when a CCC channel is idle or not. However, “the majority of the CALEA vendors” [12] are using the DTMF C tone-pair as an optional feature for the declaration of an idle CCC channel.

The result of the use of the DTMF C tone-pair signal is the same as in the loop extender taps. The subject can transmit a continuous DTMF C tone-pair at a volume that will discontinue monitoring, while allowing a good quality of call-conversation. As the same tone-pair signal is also used in the loop extender taps, the subject does not have to worry if the agency is using CALEA or loop extender methods for the recording, as the countermeasure evades recording using both alternatives.

### *3.2.3 Suggestions for reducing vulnerabilities*

The lack of diverse monitoring systems and the convenience of finding equivalent equipment in the market (by a simple search on the Internet) leads to a quite easy avoidance of lawful interception. In order to make the tapping systems more robust some improvements must be in order to alleviate the architectural and hardware vulnerabilities that make an interception susceptible.

As far as the analog loop extender interception systems are concerned there are not many improvements to be made due to the devices’ inherent design limitations. The main disadvantage of loop extenders is the use of the in-band signaling which can easily be compromised by several different types of attacks.

On the other hand, CALEA systems can more easily be made less susceptible to some interception countermeasures. As one of the most serious problems is the use of the DTMF C tone-pair in order to specify the start or stop of recording, the CCC channel of the law enforcement equipment must be configured properly in order not to be shut off when a DTMF C tone-pair is present. Sherr, et al. [12] suggest the use of the CDC channel, instead of the CCC in order to determine when the recording will stop or start.

Moreover, all lawful wiretaps should be checked by investigators in order to check for signs of signaling countermeasures. This should be performed for both CALEA and loop extenders interceptions. Specifically, the CDRs of the telephone companies should be compared with the records of the dialed numbers and the call times that the law enforcement agency gathered from their monitoring in order to reveal recording discrepancies. Of course obtaining this data from the telephone operator will mean that either more people inside the operator’s network will now be aware of who is the subject of a tap or that an automatic means must be made available for LEA to access the CDRs – without easily being detected.

Finally, all the interception standards (including the J-STD-025A) should be tested against a broad threat model designed to reveal all the weaknesses of these wiretapping models. If a weakness is detected, then the architecture of the monitoring systems should be redesigned by taking into consideration the possible countermeasures that a subject

can unleash in order to avoid them. However, the introduction of a new standard will take a long time, which means the persistence of these vulnerabilities for some time.

## 4 LI in VoIP

With the prevalence of broadband Internet connections (generally fast Internet), another means of communication was added to the plethora of the means that people can use in order to make voice calls to each other. This technology, which has recently become popular, is widely known as Voice over IP (VoIP) and uses the Internet or other packet data network in order to establish a call to another workstation/PC/PDA/... that has the same software installed or via a gateway to mobile or fixed line phones.

Many people including specialists believe that sooner or later VoIP will replace fixed line telephony, as it has significant advantages compared with traditional telephony. Even though VoIP has not replaced telephony yet, millions of people use VoIP for their everyday calls, thus reducing their use of traditional telephony.

However, the spread of the VoIP technology caused serious headaches with regard to making lawful intercepts. The reason is that it is not easy to detect and record a VoIP call to/from a target because the Internet operates in a very different manner than fixed or mobile telephony. In the following sections, a description of the nature of VoIP is given; along with an explanation of why VoIP interception is problematic.

### 4.1 How VoIP works

VoIP can be characterized as a revolutionary technology as it can (and likely will) transform the global phone system. VoIP uses the Internet, which is an existing and widely available network, in order to make calls via a standard Internet connection.

VoIP transforms the analog audio signals into digital data, packetizes this data, and transmits it over the Internet. Today, there are three different ways for individuals to utilize VoIP:

- An analog telephone adapter (ATA) is a device that allows a user to connect a standard phone to his/her computer or router and make a call. The ATA performs the necessary analog-to-digital and digital-to-analog conversion of the signals, performs the signaling necessary to set up a call or to receive one, and provides power to the telephone handset. These devices enable a user to quickly and easily connect their existing telephones to the device and connect it to the Internet.
- IP phones are specialized digital phones that are equipped with a handset, buttons. The main difference from traditional phones is the use of a RJ-45 Ethernet connector instead of a RJ-11 telephone connector. This enables the IP phone to connect directly to an Ethernet hub, switch, or route. Versions exist that utilize IEEE 801.11b wireless local area networks (WLANs), thus the user can connect anywhere there is an IEEE 802.11 access point which will allow the user to connect to the Internet. Extensive networks of open or nearly open WLAN access points exist to facilitate this type of connectivity (see for example FON Wireless Ltd.<sup>5</sup>).

---

<sup>5</sup> <http://www.fon.com/>

- Computer-to-computer VoIP is the most widely known way of using VoIP. All that is needed is to download free or often very low-cost software, a soundcard, a microphone, speakers (or headset) and a fast Internet connection. The main advantage of this for many users is that they can call computer-to-computer without any additional charges - no matter what the distance is. Of course the users have to have Internet connectivity, but this can be by any means (just as in the above cases).

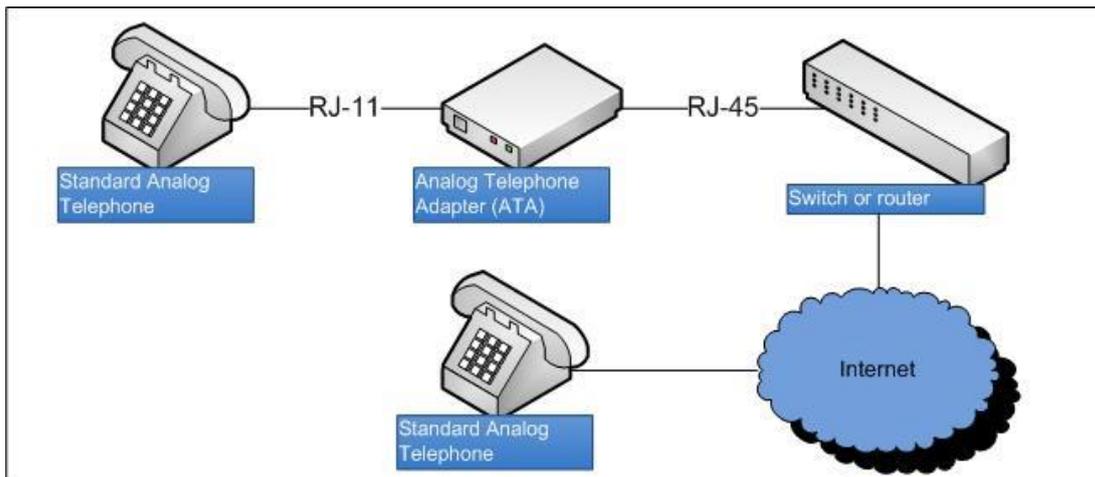


Figure 7: VoIP service with the use of analog telephone adapter (ATA)

By using any of the above VoIP approaches, home & office users obtain two main advantages compared with the traditional phone system: low cost and increased flexibility. By having the ability to make a free (or nearly free) call to another VoIP terminal no matter how far away it is, enables VoIP users to save a lot of money. Moreover, the flexibility that VoIP gives users is almost unique - as they can call (and be called) from anywhere there is broadband connectivity. Furthermore, many extra services such as caller ID, call waiting, call transfer, repeat dial, return dial, return call, and three-way calling are provided by many vendors for free, while traditional telephone companies charge extra for these services. Finally, many advanced call-filtering options are available from many VoIP companies. Based upon the caller's ID the user can decide how he/she will accept/reject/ignore/forward to voicemail/... a call from a specific caller, at a specific time, in a specific place, ... .

As VoIP operates over the Internet, it utilizes packets, in contrast with the circuit switching model which is used by the PSTN. The main difference between the two is that in the circuit-switch network there is a temporary (lasting only for the duration of the call) dedicated connection between the two parties. In a packet-switched network instead of using a dedicated line, each packet is routed over a packet network with thousands of possible paths to the final destination. While in circuit switching the connection is fixed until the parties hang up, in packet switched network a packet<sup>6</sup> is sent from one

<sup>6</sup> Packets are small chunks of data that contain the payload (a piece of clear information) and addresses (source and destination)

workstation to another via a series of switches and routers, thus the path through the network could be different for *each* packet.

The overall process of the sending data through the Internet can be summarized as [20]:

- The source-computer packetizes data, creating packets that contain a payload, source address, protocol, length, ... a **destination address** which the routers along the path utilize to decide where to forward the packet.
- The source-computer sends the packet to a nearby router, this router forwards the packet to another router that is closer to the destination-computer, and so on.
- Finally, the destination-computer may receive these packets (although the packets may have taken different network paths to this destination; some may even have been lost or duplicated) and reassembles the data in the order using sequence numbers contained in the packets (more precisely the sequence number is in the RTP header for the packets which contain the actual digitized audio samples).

The main reason why VoIP calls are cheaper than PSTN calls is that packet switching allows the digitized and encoded contents of multiple telephone calls to be carried the capacity that would be occupied by only a single call in a circuit switched network [20]. Despite the lower cost calls and the flexibility of making calls, why has VoIP not (yet) become the dominant means of making calls? One possible answer is the perception of a lack of reliability, as far as the service is concerned. The quality of a PSTN call generally very predictable; unlike a VoIP call which can vary in quality - as in the Internet, packet losses are quite common which can degrade the quality of a call. Moreover, the broadband connection can have serious delays that can negatively affect the quality of a call to such a point that it may be infeasible to for the called parties to communication via voice. However, today the quality is improving and VoIP is gaining acceptance. – thus it is likely to become **the** preferred means of voice communication for many users.

#### 4.2 Similarities and differences in PSTN and Internet

Both PSTN and Internet are networks that people are using for their daily communication. Nevertheless, each operates in its own way, thus giving each some advantages and disadvantages.

Generally, the PSTN is considered a reliable, secure and moderately expensive network [22] for providing voice communications. One of the main characteristics of the PSTN architecture is that significant investments were made (earlier) in its infrastructure by the telephone companies, the result is a system that is “smart in the center and dumb at the edges” [22]. Moreover, due to the call establishment process and the maintenance of a connection until the parties hang up (due to its circuit switching nature), the call has a high *setup* cost.

On the other hand, because the Internet operates in a different way, the cost of transmission of small packets of data is very low (per packet). Moreover, the investments in the Internet are mainly at the edges with devices that are **computers**. As a consequence

the Internet features a dumb infrastructure with routers and transmission channels (even if some of these devices are computationally powerful) connected to smart edge devices. While the Internet is considered less secure than the PSTN, it is possible to provide end-to-end security by using the appropriate protocols at various layers in the end devices.

Despite the differences between these networks and the different services they offer, there is an interaction between them that cause them to share some characteristics [22].

- Both networks use the same physical transmission medium (optical fibers, coaxial cables, and twisted pairs)
- Both networks are equipped in their central nodes with electronic routing/switching devices that direct bits through the network.
- Both networks serve large numbers of customers and operators try to provide services to these customers by efficiently utilizing their investment in equipment and transmission capacity.
- In both networks there are many companies that operate networks and these companies either provide interconnections between networks or to additional customers, hence the operators must cooperate in order to deliver the traffic that belongs to a user from network A to a user in network B.
- Today both network use digital transmission techniques.

Additionally, the PSTN and Internet differ in the principles upon which they operate [22]:

- Internet uses relatively inexpensive routers for its operation in an attempt to minimize the cost of transferring data. Most of the traffic is forwarded on a *best effort* basis, which does not provide any guarantee of data delivery. In contrast, the PSTN uses expensive switches in order to provide a reliable quality of service.
- Internet is a dumb network with smart edges. The PSTN is a smart network with dumb edges. This means that in the Internet the cost is the same no matter what type of application the user is using. Conversely, in the PSTN when new services are added there is often a high charge for using them – as the high cost of implementing the service via the network must be recovered by the operator.
- In the Internet evolution of services come from unexpected places. Anyone can create a new service and test it. On the other hand, in PSTN innovation was mainly the privilege of telephone vendors.

In conclusion, the architecture of these two networks represents the available technology of each epoch. Although digitization and fiber are bringing the two networks closer, there remain some clear differences.

### 4.3 SRTP in VoIP

The Secure Real-time Transport Protocol is considered the secure profile for the Real-time Transport Protocol. RTP is used to carry the digitized real-time audio (and

other media) data over IP networks. SRTP was designed to provide security against the threats that RTP packets confronted. SRTP provides this by data encryption, message authentication and integrity, and replay protection of the RTP data.

As RTP packets carry audio and video data these packets are essential for providing conversational VoIP services. But the security problems that Internet bequeathed to VoIP made VoIP services vulnerable to several types of attacks (forged data, man-in-the-middle attack, etc.). As a consequence SRTP was designed to fill this security gap. RFC 3711 [43] defines in detail the main objectives; the format; and the algorithms that are used for encryption, message authentication and integrity in SRTP. From analysis of this RFC and the proposal [44] & licentiate thesis [45] of Elisabetta Carrara, it is easily understood how SRTP has proved to be important for enabling secure communication over VoIP. Today a plethora of VoIP clients (i.e. minisip [50], Gizmo5 [51], Globe7[52], KPhone[53], etc.) implement SRTP in their software.

#### 4.4 Security Problems in interception of VoIP calls

It is widely known that lawful interception is more successful in the PSTN than in the Internet. Why is this true? In fixed telephony callers use fixed phones which are always connected to the local telephone company's exchange. The switch at the local exchange **must** support wiretapping<sup>7</sup>. This means that the wiretapping hardware and software are quite well secured as only authorized employees have physical access to the interception system. Moreover, this interception is very effective as the switch that serves the target number directly supports wiretapping. So, even if the monitoring has to do with an incoming call that is going to be forwarded somewhere else, the interception can be done since the call first reaches the local switch (or at least the local operator) before being diverted.

What happens in the case of VoIP? Compared with traditional telephony, VoIP has two basic differences that render it problematic for interception. First of all, the target's computer is not owned by the carrier (unlike the switch - in the case of the fixed/mobile telephony). Second, a VoIP call is not associated with fixed location as is generally the case with the phone numbers in the PSTN. Even if some IP addresses that are used by computers are fixed, this is not the norm especially due to the widespread use of wireless networks, Internet cafes, and even home networks that dynamically assign computers new IP addresses every time they connect to the Internet. Dynamic IP address assignment is increasingly common compared with static IP addresses, especially with the increasing use of mobile communication devices. While the details of the dynamic IP address assignment might be recorded in a database or log file, these records are not centralized and easily accessible to law enforcement agencies. Additionally, any logs of calls (if they exist) are also distributed and not easily accessible to law enforcement agencies unlike the CDRs of calls through the PSTN<sup>8</sup>.

---

<sup>7</sup> The requirement to support wiretapping is generally part of the regulations for lawful intercept in each country. In the case of CALEA the law stated the minimum capacity for intercepts as a function of the switch's capacity. Thus while a switch must provide facilities for LI, it may be bounded in the number of simultaneous intercepts which are possible.

<sup>8</sup> Note that with the increase in flat-rate calling plans for telephony, there is no business need for the operator to keep call detail records for these calls – since they do not need this information for billing purposes. Therefore, the availability of CDRs may decrease in the future – unless regulations compel operators to create and keep these records.

In addition, to the problems described above, there is also another problem that can make Internet LI more vulnerable to attacks. The current LI laws do not distinguish between the different types of electronic communications [22]. The U.S.A. is facing this problem with CALEA. There is a high risk to the security of the entire Internet as companies try to adopt a monitoring model that only has to deal with VoIP, while other ways of Internet communication may have security holes. This was the main concern of the IETF Network Working Group when they rejected taking into account wiretapping as part of IETF's standards [23]. Several attacks (man-in-the-middle, capturing of identity, and passwords) [22] can be done, because of the application of a one-dimensional wiretap law. Thus there is a need for a more thorough analysis of what is needed, what is desirable, what is feasible, what is required, ... in order to define laws & standards and to guide implementations.

As far as VoIP calls are concerned, a VoIP provider can help a law enforcement agent make an interception by guiding the target's call to a specified "point" where the tap is installed. However, even if this may work in some cases, the best way to intercept communication between Alice and Bob is to monitor the local router of one of the two persons or both of them. Unfortunately, this is not an easy task. In order to accomplish this, the routers must be under the control of an entity within the jurisdiction of a law enforcement agency. Every time an interception is needed an authenticated message must be sent to the network operators (in this case Internet service providers – ISPs) in order to start the monitoring. This causes a new problem, in the U.S.A. and in many other countries the majority of the ISPs are small companies who do not have "unlimited" resources. Unfortunately, compliance with the resource demanding wiretapping requirements may drive them out of business [22].

However, this is not the only problem that ISPs have to deal with. No matter what the size of a service provider is, VoIP introduces the problem of multiple identities. A VoIP user can very easily create multiple accounts in order to confuse a LEA that is attempting to monitor his or her traffic and/or identify him/her. Multiple identities are something very common in the Internet, but the recognition of the equivalence of these identities is not an easy task. As a consequence, a VoIP user with multiple identities cannot be easily identified by a pen register, since by changing accounts every log entry may have different caller identity information [22].

While the main problems with monitoring a VoIP call are VoIP mobility and multiple identities, there are other security issues that make VoIP interception problematic [22]:

- Physical security of the switching/routing equipment into which wiretap functions are inserted. Small ISPs may not have the expertise for securing their switching/routing equipment.<sup>9</sup>
- The ease of creating and using new VoIP identities on the Internet causes difficulties in identifying the *actual* target.

---

<sup>9</sup> In addition, due to use of co-location facilities the ISP may not have physical access control for this equipment. In co-location facilities many organizations share a physical site and all place equipment at this site. In some sites this equipment simply has a label on it indicating who is responsible for this piece of equipment, while in more secure co-location facilities each operator's equipment may be located in a physically secured cage within the facility. However, in nearly all such shared facilities it will be difficult to carry out operations which can not be observed by others.

- As the Internet is characterized as a dumb network with smart edges, it is easier for a target to discover that he/she is being wiretapped. In contrast the PSTN is not as vulnerable as the Internet because of the smart network and the dumb edges.
- For every interception, the surveillance must be concentrated only on the specified target. However, mobility and the multiple identities that a VoIP user can create both cause extra difficulty in isolating only the targeted VoIP communications. Intercepting untargeted subjects will cause many illegal problems.
- There is no agreement when is best to examine a packet. In a PSTN or in a mobile network telephony interception is generally based on a unique identifier, the phone number<sup>10</sup>. However, Internet packets do not include this kind of information. Instead of a phone number associated with a call (session) in the PSTN, each packet includes source and destination IP addresses. However, these addresses are usually not static and may change every time someone connects to the Internet [24]. As described earlier, to find the mapping between this IP address and a MAC address requires access to the DHCP logs (in the case of dynamic host configuration protocol based IP address assignment). Additionally, many network interfaces enable the user to change the MAC address which this interface uses, so this is not a guaranteed unique identifier for a device – let alone a user.
- Finally the call signaling and the call contents may be encrypted [24]]. There are two categories of VoIP companies. Those who do not encrypt their calls and those who encrypt them. Vonage is one of the most well-known VoIP companies in the first category, as it does not encrypt the packets used for setting up the calls **nor** do they perform encryption of the call contents (in fact they might never see the call contents - as they do not necessarily operate any of the networks over which these contents might be sent!). On the other hand, there are other companies that encrypt their VoIP calls (both call signaling and call contents), but due to their obligation to comply with interception laws, they are obliged to provide the decryption keys if a law enforcement agency asks them to do so. In addition to these two categories, there is a major VoIP company that constitutes an exception leading to headaches for law enforcement agencies. Skype is the world's most popular VoIP company with hundreds of millions users. Both Skype and SIP providers use peer-to-peer communication for the call contents between the callers. Thus Skype is not able to provide interception as the call traffic is not handled by it. Moreover, with strong encryption of packets [26] it is difficult to decrypt them even if the packets were captured at some point in the network. It is considered that only NSA [24] has the necessary computational power to decipher Skype's packets. However, Skype and some SIP operators might be subjected to law enforcement agency requests to (1) weaken the client which a given user uses, (2) disclose the information about the keys which a user is using (if known by

---

<sup>10</sup> Or in some cases an identifier for a specific mobile telephone.

the operator), and (3) disclose the information they do know about a given subscriber (this might include the subscriber's public key, what IP address they have logged in from, information about how the user pays for their service(s), etc.).

It is easily understood that all the above factors render lawful interception of VoIP calls both different and more difficult to perform than in traditional telephony. If a VoIP caller has a fixed location and a fixed IP address from an ISP, then monitoring the call can be performed relatively easily as it does not differ (substantially) from the existing methods of wiretapping in the PSTN. In fact, it may be much easier - since the data seen by the network and by the user's computer are the same (i.e., taking advantage of differences in thresholds for dialed number detection, etc. is much harder). However, if any of the above conditions occurs then the effectiveness of the interception can be substantially reduced.

#### 4.5 Example of problematic VoIP interception

Suppose there are two VoIP users – Alice and Bob - who are connected to different ISPs. Moreover, both Alice and Bob have selected a VoIP service provider associated with another ISP (See Figure 8).

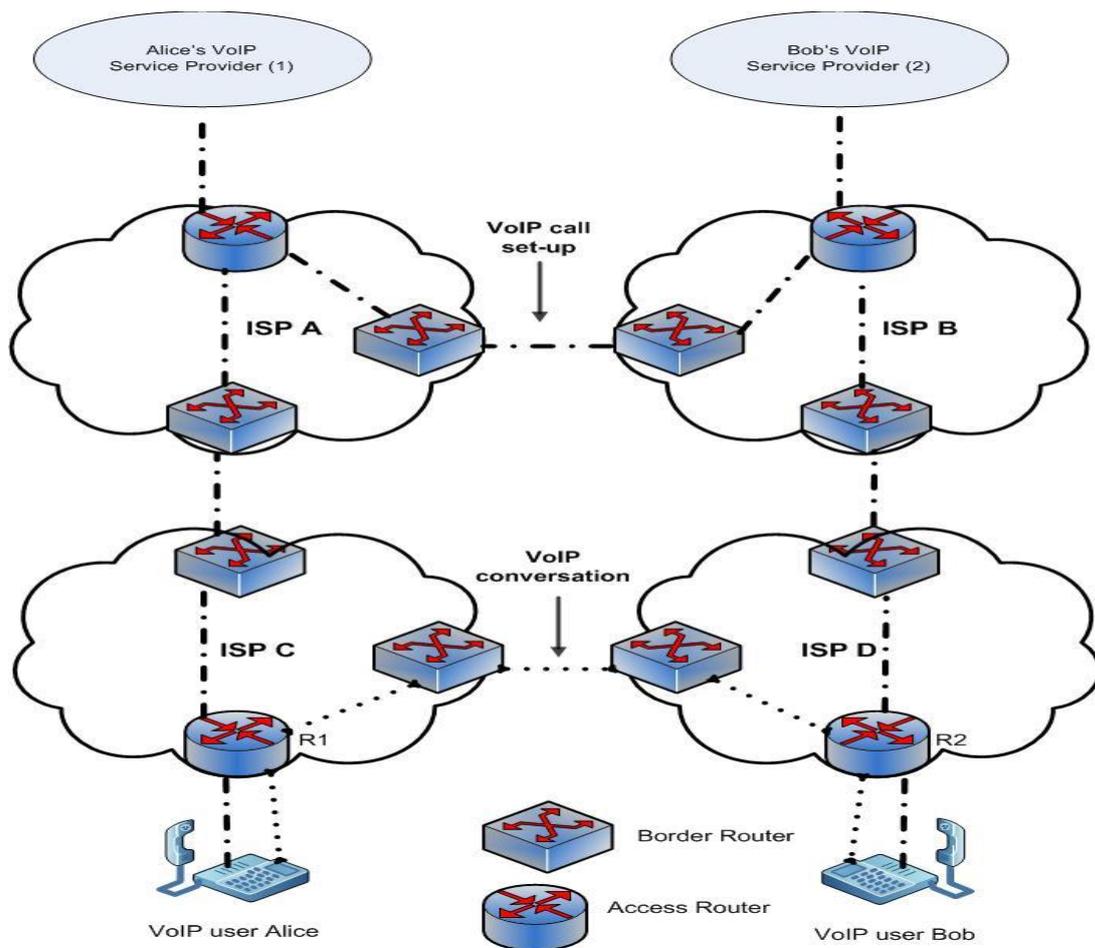


Figure 8: Problematic VoIP interception (Adapted from [22])

Suppose Alice wants to call Bob. Alice's VOIP phone uses the Internet to notify her VoIP provider which informs Bob's provider, which in turn notifies Bob's phone, thus a VoIP call can be set-up.

If an agency wants to start an interception that involves Alice, then the key routers are the access routers R1 and R2. These are the routers that are quite similar to the local exchange in the PSTN. However, routers R1 and R2 do not know the identity of the VoIP callers; hence the interception cannot start here. One of the VoIP providers must give the order to start the monitoring of its customer. Here is where the problems start. The VoIP providers can be located anywhere in the world and they are not obliged to have a "technical relationship" [22] with ISPs other than their own. Thus Alice's VoIP provider might not only have *no control over* router R1, but might not even have a business relationship with ISP C - hence ISP C would have no reason to process a request from Alice's VoIP provider to perform monitoring of Alice's call traffic.

The problem would be simpler if the VoIP provider is owned by the ISP that provides Internet connectivity to Alice (i.e. if ISP A and ISP C are the same). Even if this is feasible, it cannot be expected to be the norm. One of the best known examples is Skype which is not associated with any ISP [22]. As the VoIP provider is not associated with the ISP and the VoIP user can even move from one ISP's network to another (sometimes in the middle of a call), thus effective interception is not an easy task. Even more difficult is the case when the user does not use an external VoIP provider, but instead provide this service themselves - now it become impossible for a LEA to seek the cooperation of the VoIP provider without tipping off the user that their communication is the target of a legal intercept!

#### 4.6 Solutions for VoIP interception

According to the above analysis, law enforcement agencies face serious problems in accomplishing effective monitoring of VoIP calls. They can easily miss a call, they might not have the ability to decrypt a conversation, and they may be unable to intercept targeted calls while avoiding intercepting non-targeted calls. However, some improvements can be made in the way a VoIP interception is performed.

##### 4.6.1 *A formal architecture for VoIP interception*

One of the most significant causes for a VoIP call avoiding interception is that no IETF working group supports a standard for lawful interception [23]. However, a widely accepted interception architecture that would specify in detail what the functions are and where in the network they should be done might be successful. This architecture could be used as a base to which new functions can be added or deleted, if something does not work properly. Moreover, if such an architecture was widely known, then many organizations, companies, and individuals might evaluate it, test it, and address open problems. This evaluation could help further development of such architecture.

In recent years Cisco Systems has made a significant effort to support lawful interception of VoIP calls. In 2006, they described an architecture for LI known as Service Independent Intercept [27]. This architecture describes specific Cisco Systems tools that "can be used by the ISPs in order to construct an LI-compliant network" [27]. There is a document [28] that describes in detail how interception can be performed in

Cisco Systems' LI architecture. The main characteristic is that the interfaces and the functions are quite similar to those that exist in the ETSI standard for both traditional and mobile telephony [18]. The Cisco Systems approach shows that there can be an evolution in VoIP interception architectures as standards have been updated since 2003. By incorporating this architecture in products which they have introduced to the market, it has become one of the de-facto standards for VoIP lawful interception.

#### 4.6.2 *Use of a Trojan*

As already stated, Skype (and a number of other clients) uses strong encryption of the media between the parties of a call. The result is that even if a lawful intercept agency managed to intercept these packets it would be difficult to decrypt them by brute force. To overcome this problem a new way of interception has been introduced by DigiTask [29], a German company. The main idea for intercepting Skype calls is the installation of a Skype-Capture-Unit on the target's computer. This unit (which consists of malware software) captures the voice and chat traffic directly and diverts it to an anonymous recording proxy which in turn can forward it to a final recording server. The intercepted data can then be accessed "via mobile evaluation stations" [29]. Also, the intercepted data is encrypted for security reasons and compressed to save bandwidth. The forwarding of the capture-unit's intercept can be done through email or via direct communication with the target's computer. Finally, the analysis of the intercepted data can be done following decryption and decompression using a media player - which can play voice and other forms of data back in real time.

Similar to the above surveillance technique is the use of a "key logger" system [30]. A key logger is surreptitiously installed in the target's computer and starts recording the keystrokes from the keyboard. This means that whatever the user of the computer enters via their keyboard it will be captured by the key logger system (even characters that were typed and then deleted). The captured keystroke data are sent to the law enforcement agency. Key loggers can be placed onto the target computer manually or by sending it remotely as a virus. With this technique encryption does not provide any protection, as the user's entry of a key is captured, hence the key can be entered to a copy of the same software at the LEA and the security circumvented. (note that there can still be inadequate information to decrypt the intercepted encrypted information – for example the user might use an external encryption unit which contains the actual key generation function – thus without access to this information the user entered phrase might not be sufficient)

#### 4.6.3 *Watermark technique*

In 2005, a university team conducted a survey/experiment [31] about how feasible the tracking of VoIP calls through anonymizing networks would be. This team proved that even if a calling party uses a network that anonymities the traffic to make his/her VoIP calls that a call can be tracked. The main idea behind this tracking method is to embed a unique watermark into the encrypted VoIP packets. This watermark is based upon a timing adjustment that can be preserved even if the packet flows through an anonymous network [31]. A key to the success of the experiment was that the time interval for the addition of the watermark (i.e., the added delay) has to be less than 20ms or 30ms, as this is the usual inter-packet delay of a VoIP flow. Due to this constraint the team utilized

3ms adjustments in order to assign unique labels to VoIP packets. A larger watermark delay causes a greater distortion in the original inter-packet timing -- which can be detected by the target. Using this method (See Figure 9) someone who can intercept the traffic can track when two parties are communicating. Note that this interception can now be conducted by access to a very high speed backbone link – rather than being restricted to routers near the caller or callee. The watermark enables the intercept to select only the correctly marked traffic for further processing.

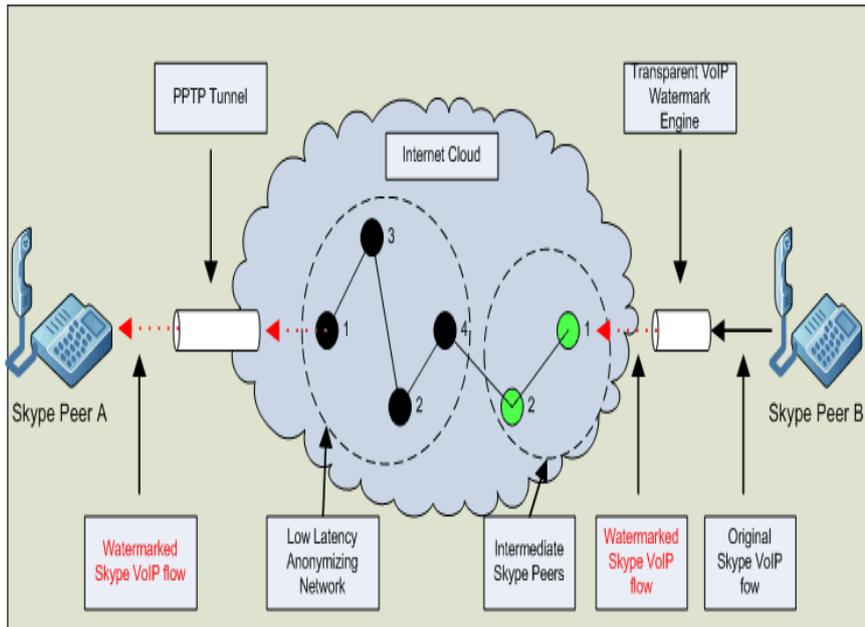


Figure 9: Experimental setup for real-time tracking of anonymous VoIP calls (Adapted from [31])

## 5 Key escrow

During the early of 1990s a new idea for monitoring telephone communications was introduced in the U.S.A. The primary people responsible for this concept were at the U.S. National Security Agency (NSA). This concept is widely known in cryptography as key escrow. The main characteristic of this approach is that the keys needed to decrypt encrypted data were stored by a Trusted Third Party (TTP) which can provide the key to a LEA after a court order or an official warrant. By storing the private keys with a TTP, the government and its LEAs could gain access to all the encrypted conversations because they could decrypt them *after* getting the stored keys, without the knowledge of the target. Of course, storing of the key used to encrypt every encrypted conversation regards everyone as a potential “criminal” whose communication might need to be monitored. One of the important questions is who should be the TTP? The main attributes of the TTP are its widespread trustworthiness and the lack of dependence upon the other TTPs.

In 1993, the U.S.A. government with the assistance of NSA created an encryption device that was supposed to be adopted by telecommunications companies for voice transmissions to implement the key escrow concept. The whole project was widely known as “Clipper chip”. This project was abandoned in 1996.

Despite the attempts that U.S. government and its agencies made to promote key escrow as a solution-system that would help nations avoid terrorism and for organizations to avoid losing encrypted information the reactions of independent cryptographers, companies, and civil-rights organizations were tenacious and negative. This made key escrow systems highly controversial and revealed that establishing a suitable trusted party (or trusted parties) did not result in a feasible solution.

### 5.1 Government’s key escrow goal

As governments know when cryptosystems are well designed it is very difficult to recover encrypted data without the correct keys, thus they understand that they need a widespread system that would provide them easy access to all encrypted data. In order to achieve this, the idea of key recovery systems arose. These systems would be built in such a way that they would provide 24 hour accessibility to the master keys enabling the decryption of encrypted data. However, this availability frightened many people due to a potential (or inevitable) violation of their privacy, hence governments promoted the use of key escrow systems by stating that such systems were also very important for industry which needs a guarantee that it can access its **own** encrypted data to avoid a possible loss (for example, following the death of an employee, a physical disaster or other incident that could destroy the keys or render the keys inaccessible to the lawful owner of the data). However, the real needs of industry are different from those of governments, as far as availability of encrypted data is concerned – thus the government’s argument was not very successful.

As the U.S. government was the global promoter of key recovery systems we will examine what the officials stated as their goal for key escrow systems. In 1996 there was a statement in the U.S. Department of Commerce’s encryption regulations, stating their goal: “envisions a worldwide key management infrastructure with the use of key escrow

and key recovery encryption items” [34]. They cited the main specifications of global key recovery systems (which were based mainly upon law enforcement demands [35]) as:

- **Access of TTP or government without the notice to or consent of the user.** This also affects “self-escrow” systems where companies may deposit their own keys, but they must provide a mechanism that obscures the revealing of the decryption between the key owners and the recovery agents.
- **International adoption of key escrow systems.** In order to succeed and provide real help to law enforcement agencies, key recovery systems must be as widespread as possible. This means that for the majority of the encrypted data and communications a key escrow system must be used, *whether there is an end-user demand or not.*
- **Law enforcement agencies demand high availability to the decryption keys.** In the U.S.A. the time to obtain the keys was specified as two hours (i.e., near real time). Moreover, the recovery process should be available 24 hour per day 365 days of the year.
- **The system should provide access to encrypted communications as well as to encrypted stored data.** This explicit inclusion of encrypted communications was important, as industry only seemed be interested in accessing encrypted stored data as opposed to securing and recovering communication traffic.

The above specifications were widely seen as incompatible with the needs and requirements of commercial encryption users. Unfortunately, governments took into consideration only their own needs, although they tried to appear as if they also considered the requirements of industry and commercial users; but as they failed to convince the industry and commercial users that the government’s needs were the same as or trumped the users’ needs the government’s argument was unsuccessful.

## 5.2 Clipper chip

Perhaps, the most widely known key escrow implementation was the Clipper chip [32]. The clipper chip was a cryptographic device which was designed to encrypt private communications while simultaneously making the conversation’s key available using the keys held by TTPs. The concept behind Clipper was that the session key would be encrypted and transmitted along with the session, thus given access to the device’s keys from the TTP the government agents could easily recover the contents of a telephone conversation – for example to introduce the call as evidence in court. The intention of the U.S. government was that all telecommunications companies would be forced to use this device.

The main idea behind the Clipper chip was that each telephone device would have an embedded Clipper chip which was responsible for the encryption of all data passing through the device. Each chip had a unique key. The government stored “a record of the serial number/encryption key correspondence for every chip manufactured” [32] in order to have the ability to use it later for monitoring/intercepting a conversation. In order to calm the fears that many people had against the government possessing these keys and

the probability of undesired practices, the government decided not to keep the keys only in a single “safe” place. Instead, each key was broken in two quantities and these two parts were to be delivered for deposit and safekeeping with two different agencies. The reason for this was that it would be more difficult to misuse the keys for monitoring purposes as access to the keys would require the cooperation of two separate agencies. In order to enforce this status, the U.S. government claimed that voluntary cooperation *without a court order* to deliver the parts of the key among separate government agencies was improbable [32]. Another strong point of the key escrow system was that the pieces of each deposited key must be XORed in order to produce the actual key. Therefore knowledge of **only** one part of the key is completely useless for the retrieval of the complete key (or even the recovery of any bits of the key!).

For encryption the Clipper chip utilized a classified algorithm known as Skipjack, which was invented by NSA (note that the U.S. government declassified this algorithm in 1998). Furthermore, the Diffie-Hellman key-exchange algorithm was used for the agreement of a session key between the peers [32]. The information that agencies needed to decrypt the packet were included in a field known as the Law Enforcement Access Field (LEAF) which was transmitted during each communication session. An important aspect of the use of the Skipjack encryption algorithm was that the LEAF sent from a caller to a recipient must be valid; otherwise no communication would be permitted. This had the advantage that a strong encryption algorithm could be used, but the government could be ensured that since the LEAF was valid they would be able to decrypt it given the proper key.

Although the Clipper chip escrow system seemed very robust, in 1994 Matt Blaze demonstrated and described in a paper [33] a serious vulnerability in the security of the Clipper chip system. This vulnerability occurred because a 16-bit part of the (128-bit) LEAF contained a hash value that was used as a checksum to ensure the integrity of the LEAF data. If the receiver computes another checksum over the LEAF, then his/her Clipper chip would not decode the message. The problem was that 16 bits was a very small number, thus only  $2^{15}$  random numbers had to be tested (in a brute force attack) before finding the correct checksum. In this case the Clipper chip could be used as an encryption device, but the LEAF field would be invalid and hence the government's key escrow capability could not yield the correct session key.

As a result of the above serious security defect, the Clipper chip was abandon in 1996. Although this key escrow system did work it was not adopted, hence the U.S. government continued to pressure manufacturers to adopt key escrow in order to reduce the difficulty of decrypting encrypted conversations.

### 5.3 Advantages of a key escrow system

The main (and perhaps only) supporters of the key escrow systems were and still are governments and theirs agencies that want an easily controlled method which would offer secure communication while also allowing lawful interception. In order to persuade others (i.e. companies, individual cryptographers, and the average person) they tried to point out the benefits that a key escrow system could offer in society even if the result was not as expected.

The main advantage of a key escrow system is that a government and its LEAs can provide better security for the public, because they can effectively control encrypted communications by knowing all the keys -- but only if every user is forced to deposit their master keys with one of TTPs. This means that a LEA can easily access a target's conversation even if it is encrypted, by using the master keys to decrypt it. Moreover, with the proper participation of the TTPs, they hoped to assured that no one could easily misuse their knowledge to improperly access the master keys and inappropriately gain access to an encrypted conversation. Because each master key would be broken into parts and delivered to different agencies (i.e. as was proposed for the Clipper chip), this was expected to prevent illicit wiretapping as the collaboration of different agencies in an illegal act (i.e., disclosing a key without proper authorization) was supposed to be extremely difficult [32]. The assumption was that separating the key into parts controlled by separate agencies offers a secure way to guarantee that there would not be any violation of human rights and freedom as illegal decryption of the conversation would require a lot of cooperative effort.

Another advantage of a key escrow system – in a company – is that it would allow a company to easily monitor all the communication of its employees. Why would a company need to do this? Basically, there are two reasons for a company to use key escrow system [32]. First, if the employees of a company know that their conversations (even if they are encrypted) can be easily decrypted and read/listened to; then they are more likely to conform to the company's policies. Second, the company can assure that no data will be lost if an employee forgets a password or leaves the company. An additional reason (in the post-Enron era) that a company might need access to a copy of an encrypted conversation is to comply with a legal search and discovery order from a court.

It is understood that in both cases (national and corporate level) a key escrow system's main advantage is the 24hour **availability** of access to encrypted data, thus giving the responsible authorities the ability to access this data. Unfortunately, this does not mean that they can control access to this data (as in the event of law enforcement actions the access to the data might be done without the knowledge or consent of the individuals or company).

#### 5.4 Disadvantages of a key escrow system

Despite the government's attempt to persuade companies, independent cryptographers, and other people of the usefulness of the key escrow systems, the negative reaction was significant from all sides. From the very beginning key escrow systems were criticized due to the severe security flaws they suffered and the great potential for possible violation of human rights and individual privacy. There are a number of risks that make key escrow highly insecure [36]:

- Introduction of new vulnerabilities
- *Potential illegal access to data:* As the TTPs have the ability to use the master keys in order to intercept communication (even without a court order) it is possible for them to perform illegal acts. Moreover, because the companies or individuals will not know that the TTPs has used their master keys for

monitoring them, the company or individual cannot protest this (potentially) illegal access to their communications. Access to such data may give a government or its agencies a strong advantage against targets, so there is an incentive to circumvent court orders and simply access the target's conversations. Examples of this type of potential abuse include the U.S. government's ability to persuade telecommunications companies to give them unrestricted access to international trunks [40], [41], and [42].

- *Insider misuse:* One of the most dangerous threats that a key escrow system introduces is insider misuse, for example to produce records that appear to be legitimate, but which in fact have been forged. As the employees of TTPs may not be as trustworthy as he/she should be, they might be “motivated by greed or ideology” [35] to compromise the secrets of individuals or companies (sometimes this has been part of a scheme to blackmail the individual or company). This has happened many times in the recent past; in these cases someone exploited his/her position in order to violate others privacy for his/her own advantage. However, revealing secret information of companies or individuals is only one type of insider misuse. The other type of misuse is more severe, in this case a person who has the session keys of a conversation may “fabricate” contents [37] (i.e. counterfeit contents) that a user never sent in order to incriminate him/her. This practice has also been used a lot in the past. However, creating forged contents is a double sided sword as it may impair the ability to use wiretapped communications as evidence in the court.[37].
- *New targets for attack:* One of the main characteristics of key escrow systems is the depositing of the master keys in central databases. This feature makes these databases a high value target as “the theft of only one private key (or a small set of keys) which are held by a recovery agent may disclose all the data or part of them of an individual or a company” [35] or even worse from a “broader array of communication” [35]. This problem will be highlighted by the key escrow systems themselves, as for every encrypted communication a “pointer” must exist in order to direct the LEA how to retrieve the key information. Thus this “pointer” becomes a target upon which attackers can focus their efforts for intrusion. Even though, the risk may be decreased by splitting the keys and delivering them to different agencies, doing so will create several other problems (i.e. increased cost, longer response times in order to assemble the keys and provide the plaintext).
- *Destruction of Forward Secrecy:* As key escrow systems enable access to the session keys of every conversation encrypted by the device, forward secrecy is not available. Forward secrecy offers two characteristics. First of all, its design is simpler. Secondly, the system has enhanced security and lower cost. This happens because if a system uses forward secrecy, even if one key of a conversation is compromised there is no threat that other communications will be revealed. For example, in an encrypted telephone call the session key(s) exists only during the conversation. When the call ends the key(s) are destroyed (sometimes they are destroyed several times during the same call)

so later decryption of this conversation is not possible - since the keys have not been retained. But with a key escrow system the ability to restore the original communication exists, as the keys which the TTPs have stored enable one to access the encrypted session keys which were transmitted as part of the original communication, hence later access to this encrypted communication is both feasible and simple for a possible subsequent use voiding any thought of forward secrecy – since all traffic which has been captured from this device in the past or in the future are accessible.

- *Storage of the keys in the Key Recovery Center (KRC):* Another problem that a key escrow system faces is what kinds of keys must be deposited for subsequent recovery. As there are many different kinds of encryption keys, it is difficult to decide which keys must be stored in the TTPs. For the purpose of key escrow systems, encryption keys can be divided into three categories. First there are the keys that are used for encrypting stored data. It is easy to understand that these keys must be available for the lifetime of the data, as the owner has an interest in access to his/her information. The second category of the keys includes those that “are used for the encryption of a real-time conversation” [38], such as calls. Here the interest in storing these keys is solely on the part of LEAs or someone else who wants to have access to this communication. Generally the communicating parties do not want recoverability of their conversation (although as noted above they may be compelled to provide this for regulatory or other legal or business purposes). During the call the keys may be destroyed and new ones can established without any loss of information between the communicating parties.; thus these keys are only used for a very limited period of time (the communication session duration or less). Lastly, there are the keys used for authentication and signatures which “insure that messages originated from a particular party” [38]. For these keys there is no need for anyone (either owner or LEAs) to recover them as they do not prevent lawful access to data, but rather protect the **integrity** of the data. If an owner loses a signature key, he/she can establish a new one very easily (although perhaps at some cost of time and money). Therefore there is no legitimate need for any of these key categories to be stored. However, as these keys are usually “indistinguishable from one another outside of the application in which they are used” [38] all of them might be stored and consequently they risk exposure in key escrow systems.

➤ Complexity

- *Design complexity:* According to many professional cryptographers there is a general admission that building a practical secure cryptographic system is extremely difficult. The introduction of new parameters in a cryptographic system may create several new security flaws. Even in non-key recovery systems which have fewer requirements, many exploitable flaws are still discovered [35]. Key escrow systems are considered more complex due to the necessity of key storage. This feature adds a lot of complexity in the overall design of the system. One of the best examples of the weaknesses (and subsequent failure) which can result is well illustrated by the Clipper chip.

Even though the Clipper chip was designed by NSA (an agency with a lot of experience in cryptography) – security deficiencies still existed. It is important to note that the weakness was not in the algorithms, but rather the implementation choice of a short checksum.

- *Scale factors:* Governments and their law enforcement agencies have envisaged a key recovery system that would extend worldwide. This means that a very large number of companies, LEAs, and TTPs must cooperate in order to succeed in their “mission”. However, considering only the numbers of the involved people and organizations it is easy to understand that a worldwide system is almost impossible (simply due to scaling issues). Nowadays, millions of users every day use encrypted communications (i.e. access secure pages via the Internet). Moreover, there are thousands of products in the worldwide market that provide encryption. There are several tens of thousands of TTPs and LEAs worldwide. It is hard to image the secure and successful depositing of hundreds of billions of recoverable session keys (for every encrypted telephone call, stored encrypted file, e-mail, web session) [35]. Finally, the infrastructure to support and manage the system will be vast. As a result of the overall magnitude of a worldwide key escrow system, it is clear that a system is not viable.
- *Operational complexity:* According to the U.S. government, a key escrow system in order to provide good services must operate on a worldwide scale. However, this requires a lot of complexity, not only due to the enormous number of the entities involved, but also due to the number of the key recovery requests - each of which requires a speedy response. The result is that the system could be vulnerable to fraudulent key requests and keys might be exposed to the wrong people. Consequently, there will be partial compromises of the system. While a well designed system, well trained staff, and technical controls can reduce the risks of a key escrow system (to some extent), “the operational vulnerabilities in the process of key recovery cannot be eliminated entirely” [35]. Thus there are going to be some leaks - which may have either little or major impact on individuals, companies, and governments. One only need to look at the leaks of confidential and secret personal information which already takes place today, with frequently several reports per week of disclosures of personal data (ranging from medical records to tax records).
- *Authorization for key recovery:* One of the requirements of a key escrow system is that only authorized people can request a session key. However, it is widely known that identity documents, such as passports and birth certificates can easily be forged. Using such forged identification, someone might be able to illegally obtain a key in order to eavesdrop or even forge conversations. It is possible that a key escrow system that would operate locally (i.e. inside a company) may not face this risk of forged personal identification. This is because the key administrator might personally know everyone who has “the rights to which keys” [35]. However, this “knowledge” is not readily available when the scale of the key escrow system exceeds a very small scale.

➤ Costs

- *Operational costs*: Even though “cryptography is an intrinsically inexpensive technology”, the scale of a key escrow system that would extend beyond the national borders of U.S.A. would cost a lot for its operation. The reason for high operational costs is that government requires availability around the clock (365 days per year), which requires to a lot of employees. Moreover, staff should be well trained in order to perform their job properly. High-assurance hardware and software should be purchased in order that the government could be sure that the system will not have any bugs.
- *Product design costs*: For the operation of such a large scale key escrow system the products (software and hardware) should be designed to support user-level encryption. Doing so will increase the product’s design costs as correctly adding encryption is not a simple job. The result could be poor implementations on the vendors’ part, i.e. poorly design products. While the product might be designed well, if it is mis-configured by the user, then the security might be weak. In addition, if encryption is needed for older products, this might entail significantly greater costs.
- *Government oversight costs*: Finally there will be the expenses necessary for the government and its LEAs to “test and approve key recovery products” [35]. The plethora of key recovery products that will exist in the market should all be tested to see that they comply with the government’s requirements and if they suffer from any security flaws. This can be a time-consuming and costly process. Moreover, government also needs to certify and audit the recovery agents and their ability to provide the services they are to perform. This also means extra cost as there must be an appropriate inspection mechanism to do so.

Based on the above analysis, it is clear that key escrow systems suffer from many vulnerabilities (i.e. security problems, cost, difficulties in their design and operation) that render key recovery unsuccessful, at least as far as it concerns large scale systems. However, a key escrow system implemented on a much smaller scale – such as inside a single organization or company – might eliminate many of these risks, this is not the intension of the U.S. government, as they want to have the ability to intercept the majority of communications and decrypt the majority of stored data.

## 6 SRTP/MIKEY and Key Escrow

According to chapter 5 analysis it is understood that key escrow systems suffer from a very serious flaw – forged data – that can undermine the whole credibility of the system. An authorized person that has access to the session keys of a conversation may forge traffic in order to create “evidence” that will incriminate one (or both) call parties. This is a very problematic situation, as a court cannot easily decide if the recording of a conversation is legitimate or not.

The idea of using digital signatures was suggested as a solution to this problem. The idea is that the signature can be used to verify the real source of the data. As many VoIP companies are using SRTP in order to transmit data in a secure way, one means to prove who the source of the data was is to compute signatures over groups of packets. The signature will be produced using the private key of the user – as this key is **not** stored in any key escrow system. This means that even if a malevolent employee of a TTP organization knows the session key that was used to encrypt the SRTP data, he/she may generate forged data, but will not be able to correctly fabricate the digital signatures. Hence with the help of digital signatures it can easily be proven (by anyone with access to the source’s public key) that the contents were not produced by the alleged source.

### 6.1 Secure Real-time Transport Protocol

The secure real-time transport protocol (SRTP), defined in RFC 3711 [43], is a secure profile for RTP. SRTP adds confidentiality, message authentication, and integrity – through replay protection – to the RTP traffic and to the control traffic for RTP (Real-time Transport Control Protocol - RTCP). RFC 3711 defines the format of an SRTP packet, the algorithms that can be used in order to encrypt/decrypt RTP packets, and also the mechanism for key derivation. Generally, there are two types of keys in SRTP: session keys and master keys. Session keys are used for the encryption of the data or for message authentication. Master keys are random bit strings that are provided by a key management protocol, (i.e. Multimedia Internet KEYing, MIKEY) which are used for the production of the session keys in a cryptographically secure way.

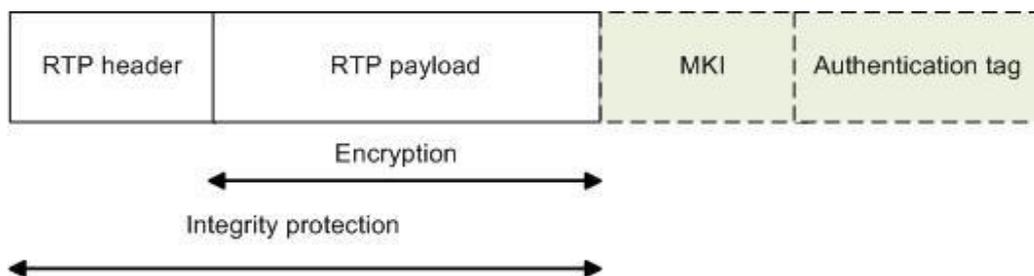


Figure 10: Format of SRTP packet (Adapted from [44])

Figure 10 shows the structure of an SRTP packet. Encryption is only applied to the RTP payload (to ensure confidentiality of the message). However, integrity protection includes both the RTP header and payload by adding an authentication tag to the end of each packet. There is also a Master Key Identifier (MKI) field in the SRTP packet which indicates to the receiver which master key was utilized for the derivation of the session

key(s) that was/were used in this particular SRTP packet. Both the MKI and Authentication tag fields are optional.

SRTP needs six session keys to protect the media. The first triplet concerns the security of RTP packets (a session encryption key, a session authentication key, and a session salt key) and the second triplet the security of RTCP packets with the equivalent session keys. To generate all of these session keys, SRTP uses a key derivation function which needs only a single master key. The master key was exchanged via a key management protocol (i.e. MIKEY). The production of the six session keys is known as key splitting and it works with the help of a pseudo-random function (PRF), the master key, the derivation rate, the master salt (which is exchanged also via the management protocol) and a label. Different values of the label generate different the session keys.

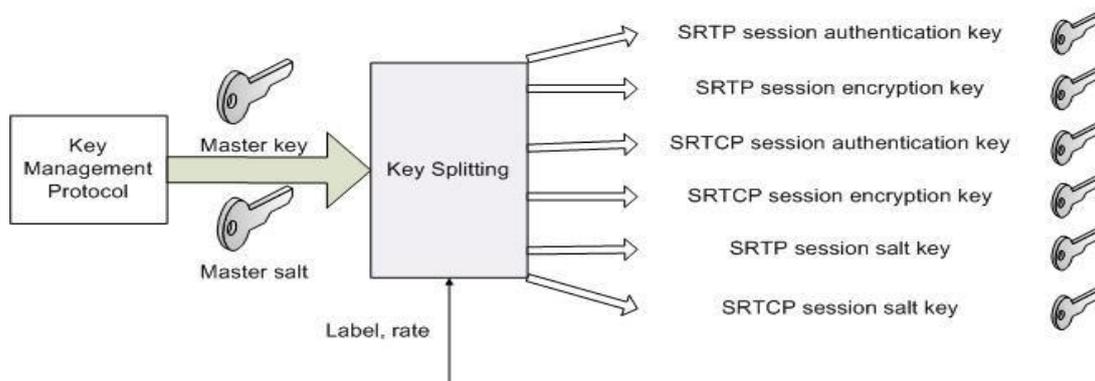


Figure 11: SRTP Key Splitting (Adapted from [44])

The previous description shows that SRTP is an independent protocol, as far as the generation of the sessions keys are concerned. With the embedded key splitting function SRTP only needs one key (the master key and its related parameters) for the derivation of all the other keys. However, SRTP must provide some mechanism(s) to change keys in order to provide better security:

- *Key refresh*: The key derivation function can also be used to regenerate session keys. This is very helpful as with the same software new keys can be produced. The use of new session keys can provide increased security as small pieces of data will be encrypted with the same session key. This means that if an attacker managed to compromise a session key, only the data that are encrypted under this specific key can be revealed. The other session keys do not have any relation with the compromised key as the derivation function prevents this from happening. By using multiple session keys, SRTP achieves perfect forward secrecy for all of the separate streams. The key derivation rate defines how often the key refresh will take place. However, key refresh is an optional feature in SRTP as although it works well in unicast sessions such a mechanism is not suitable for a multicast sessions -- as every participant would have to maintain a potentially large number of keys (6 times the number of participants) which could have a serious impact on performance -- especially if these keys are refreshed frequently. However, if the key was only

refreshed at a rate driven by the user's transmission; this mechanism might even be applicable to multicast sessions with a modest number of users.

- *Re-keying*: This mechanism concerns the change of the master key. Re-keying can be achieved with the execution of a key management protocol, which makes the whole procedure complicated and computationally expensive. Why is there a need for a new master key? Because compromising of the master key would reveal all the session keys, re-keying is obligatory for the following reasons [45]:
  1. When the lifetime of the master key has expired.
  2. When there is a compromise of the master key.
  3. Depending upon the application policy (i.e. the application may trigger rekeying every time there is departure from a session or a new participant added to a session, or due to a conservative security policy that limits the amount of ciphertext encrypted with the *same* master key).

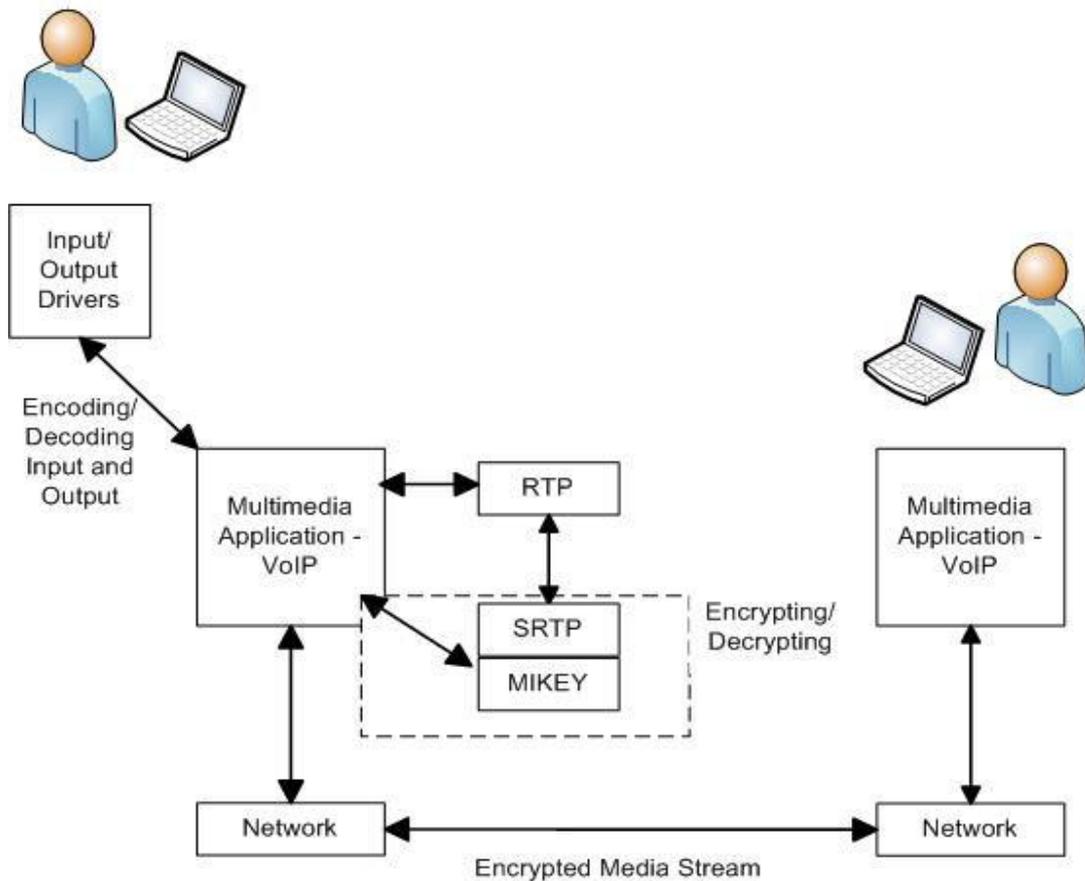


Figure 12: SRTP encoding/decoding process (Adapted from [46])

Figure 12 depicts the SRTP encoding/decoding process during a VoIP conversation. First the application “captures the input from a device (i.e. microphone) and encodes the signal” [46]. Then it creates the RTP packet payload which is encrypted by an encryption

algorithm (i.e. AES). Before RTP packets are sent to the network they must be secured. SRTP converts the RTP packets to SRTP packets. Now, the SRTP packets are ready to be transmitted through the network. When the data arrives at the destination, then the multimedia application of the other party (i.e. calling party) decrypts the SRTP packets, producing RTP packets that are ready to present (heard or watched) by the other party.

## 6.2 Multimedia Internet KEYing (MIKEY)

Multimedia Internet KEYing is a key management protocol that has its main goal to efficiently generate and distribute keys and their related parameters. The keys can be used as master keys in other protocols (i.e. SRTP).

The MIKEY protocol supports three different variants of key agreement [47]:

- *Pre-shared key (PSK)*: This method presumes that the peers share a pre-shared key which has been exchanged by some other means. In this variant, symmetric cryptography based upon the pre-shared key is used to derive keys for the encryption and the integrity of the MIKEY message. MIKEY is considered the best solution for key transport as it requires only half or one roundtrip, but it suffers from a lack of scalability as the key has to be shared with every other peer that a user wants to communicate with.
- *Public key*: This method is similar to the PSK method, but here the initiator chooses a pseudo-random key which will be used for encryption and integrity. This key is encrypted using the responder's public key and sent to the responder. Although this approach consumes greater resources than the PSK method, it can be scalable using a Public Key Infrastructure (PKI).
- *Diffie-Hellman*: This method provides perfect forward secrecy. The main difference from the other methods is that in this method the key is not sent to the recipient, but instead both parties contribute to the generation of the key. This method is considered the most resource consuming method and also requires the existence of a PKI.

## 6.3 Scenario: Minisip and Key Escrow

One of the most serious weaknesses that key escrow has is insider misuse. An authorized employee of the TTP who is motivated by greed or ideology might use his/her authorization in order to access the session keys of a conversation for malevolent reasons. This is a double threat in that he/she can reveal both the session contents and (even worse) he/she can forge the contents of a conversation.

This scenario is very frightening as it will destroy the reliability of evidence coming from systems which have made use of the key escrow systems. The threat of insider misuse is a real scenario, as this has happened several times. Thus this risk is a real threat to individuals or organizations making use of a key escrow system.

Many VoIP clients (i.e. Minisip) utilize MIKEY/SRTP in order to provide protection to the transmitted audio/video packets in a session. These protocols are used to give some

protection to the contents of a conversation (as the exchange of the keys is secured and there is also encryption and integrity control). However, if key escrow is used, then all the master keys from which the session keys are derived will be automatically available to the authorized TTP's employees. This has as a consequence that an insider misuse (revealing or forging the contents of the conversation) is possible.

Therefore, the main problem when using key escrow is how a session participant can ensure that even if a key escrow system has stored the conversation's master keys, it is impossible for the session's contents to be forged. One possible solution utilizes digital signatures. Public key cryptography using public/private keys can ensure that when someone sends data he/she is the only possible owner send. Thus an insider can forge contents, but not forge the correct digital signatures – hence protecting the real source from fraudulent or misleading “evidence”. One possible implementation of this solution is depicted in Figure 13.

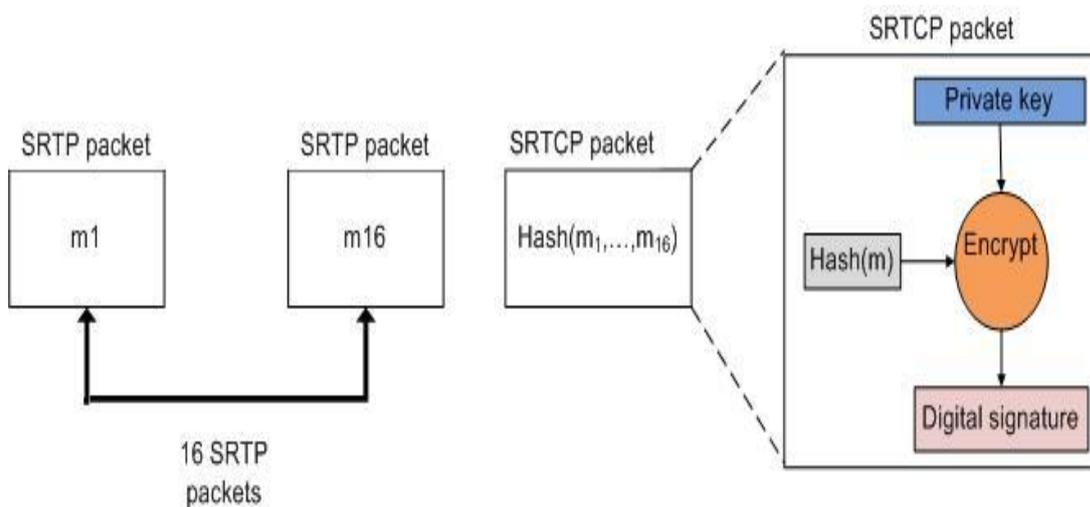


Figure 13: Digital signatures for SRTP packets

In this solution, digital signatures can be used for groups of SRTP packets (i.e. 16 SRTP packets). Every SRTP packet will be encrypted and integrity protected as defined by the SRTP protocol. Following the transmission of 16 SRTP packets an SRTCP packet will be sent which contain a hash of the payloads of the 16 SRTP packets ( $Hash(m_i)$ ). As the sender of the SRTP packets is the only entity to know his/her private key, (i.e., the private key will **not** be stored by the key escrow system), thus the  $Hash(m_i)$  will be signed with this private key producing a digital signature over these packets. This means that anyone with knowledge of the sender's public key can easily verify if the contents came from the real owner of the data. Tampering with data is easily detected as only the owner of the private key can produce a valid signature. Each individual SRTP packet is also protected by encrypting it and with integrity protection. However, as the digital signatures are independent of the session keys (which were derived in a known way from the master keys - which a key escrow system will have stored), the private key (of the public-private key pair) is only known by the real owner. Hence only the owner can

correctly sign the hash. Note that the number of SRTP packets which are grouped before the signed hash is an implementation matter – as this represents a tradeoff between increased computation and traffic and the security which this solution offers.

Digital signatures can provide **security and credibility** as to who is the real owner of the contents is, but it does not protect the contents from being revealed by a malevolent TTP's employee. The advantage is that anyone can detect this forgery.

#### 6.4 Problems to consider

The solution with the digital signatures can provide the wanted security and credibility but in order to be implemented some possible problems must be taken into consideration. The main problem is the **rate** of computing the signature. If the signing of the SRTP packets is very often it will add extra traffic as more packets must be produced for the inclusion of the digital signatures. This may cause delay problems in the performance of a VoIP call. However, the delay problem can be mitigated as digital signatures will be transmitted through SRTCP packets which are sent **periodically** to control information of the SRTP flow.

Furthermore, there is the problem of a possible **leakage** of a private key. If for the production of the digital signatures the same private key is used several times, the possibility of someone to find the private key increases a lot. The continuous encryption of a text with the same key may reveal information of the identity of the key. One solution that can be used is the use of more than one private key. By using several private keys the can be confusion to the person who will try to find out the identity of the private key.

Both the above problems must be taken into consideration as they can affect negatively the implementation of the digital signatures.

## 7 Conclusions

Safe & secure communication is very critical nowadays. However, citizens' communications – at a worldwide level –not only face a threat of privacy violation from malevolent individuals but also from governments. In (nearly) every country, the government has decided to allow lawful intercept of the communications of their citizens (and often most especially non-citizens) for security reasons. Moreover, the tendency is to enact laws that broaden the range of interception by keeping the intercepted data for a longer period of time, by intercepting all the means of communication, and by trying to circumvent court orders in order to perform an interception without proper judicial review.

This thesis examined lawful interception in fixed and mobile telephony and Internet networks. It also presented some of the difficulties in lawful interception due to the specifics of the operation of the various communication networks. The thesis also discussed several countermeasures that a target can utilize in order to avoid (or minimize the value of) interception. Furthermore, an analysis of key escrow systems was given. In addition, there was an examination of a possible elimination of one of the key escrow's main disadvantages - tampering of the contents.

This analysis showed that even if many VoIP clients use SRTP and MIKEY protocols to provide security, a key escrow system that stores the master keys that are used for the production of the session keys might allow a TTP's employee to forge contents could be detected. The use of digital signatures to detect tampering with the data can be a very powerful technique for preventing the introduction of fabricated evidence. This means that evidence from a key escrow system could be reliable as it can be proved if a particular person is the real owner of the data or not. However, the possibility of forged data is only one of the problems that a key escrow system faces. Even if digital signatures can circumvent the problem of forgery, there remain many problems in implementing a successful & secure key escrow system.

## 8 Future work

This thesis tried to present how lawful interception is performed in fixed, mobile, and Internet telephony by government law enforcement agents. It also presented the difficulties that exist for the different means of interception as targets can utilize countermeasures to avoid being monitored or to create misleading evidence. Finally there was an analysis of a means of overcoming the problem of the insider misuse that exists in key escrow systems.

Even if key escrow systems are difficult to implement and some attempts in the past to create large scale key escrow systems failed, the idea of using a key escrow system to store master keys of all the kinds of communication still exists. This thesis proved that by using digital signatures it is very easy to detect if a conversation has been tampered with. However, as there remain many other problems for key escrow systems, such system should be further analyzed in order to be more effectively & securely implemented. A good start should be the analysis of each key escrow problem separately and the description/implementation of a possible solution for each one.

Moreover, as technology – in the field of telecommunication – is changing rapidly new threats may appear that reduce the effectiveness of lawful interception. Targets may find new countermeasures in order to avoid tracking and monitoring. Also, new documents may be published that present the existing or planned means of intercept or the difficulties that the existing mechanism confronts. So, there will be a need for a more extensive analysis of the new ways of interception and new countermeasures.

A clear need is to implement the method for using digital signatures that is proposed in this thesis in order to understand the practical performance of such a solution in terms of both computation resources needed and the additional traffic which is generated.

## 9 REFERENCES

- [1] Office of the Inspector General, “The Implementation of the Communications Assistance For Law Enforcement Act, U.S. Department of Justice, Audit Division, Audit Report, 06-13 March 2006
- [2] MobileIN.com, “Lawful Intercept”, Mobile in a Minute mini-tutorials, MOBILEIN.COM <http://www.mobilein.com/LL.htm>, 28/02/2008
- [3] “White Paper – Lawful Intercept Overview”, Newport Networks, <http://www.newport-networks.com/whitepapers/lawful-intercept1.html>, 28/02/2008
- [4] U.S. House of Representatives, “50 U.S.C. Chapter 36 – Foreign Intelligence Surveillance” - amendments, Jan. 2, 2006 – Aug. 1, 2008 [http://www.law.cornell.edu/uscode/html/uscode50/usc\\_sup\\_01\\_50\\_10\\_36.html](http://www.law.cornell.edu/uscode/html/uscode50/usc_sup_01_50_10_36.html), 07/04/2008
- [5] Steven M Bellovin, Matt Blaze, Whitfield Diffie, Susan Landau, Peter G. Neumann, and Jennifer Rexford, “Risking Communications Security: Potential Hazards of the “Protect America Act””, Security & Privacy, IEEE Computer Society, Jan./Feb 2008, October 22, 2007
- [6] Vassilis Prevelakis and Diomidis Spinellis, “The Athens Affairs”, Spectrum, <http://www.spectrum.ieee.org/print/5280>, 28/02/2008
- [7] “Public Safety and Homeland Security”, Mobile in a Minute, [http://www.mobilein.com/public\\_safety\\_homeland\\_security.htm](http://www.mobilein.com/public_safety_homeland_security.htm), 28/02/2008
- [8] “Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA Patriot Act) Act Of 2001”, U.S.A. Congress, Public Law 107-56-OCT. 26, 2001. 115 STATUTORY 272, [http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107\\_cong\\_public\\_laws&docid=f:publ056.107.pdf](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ056.107.pdf), 31/03/2008
- [9] “Directive 95/46/EC Of The European Parliament And Of The Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data”, Official Journal of the European Union, 23/11/1995
- [10] “Directive 2002/58/EC Of The European Parliament And Of The Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)”, Official Journal of the European Union, 31/7/2002
- [11] “Directive 2006/24/EC Of The European Parliament And Of The Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC”, Official Journal of the European Union, 13/4/2006

- [12]Micah Sherr, Eric Cronin, Sandy Clark, and Matt Blaze, “Signaling Vulnerabilities in Wiretapping Systems”, Security & Privacy, IEEE Computer Society, Nov./Dec. 2005, pp.13-25.
- [13]Denis Howe, “In-band signalling”, The Free On-line Dictionary of Computing <http://dictionary.reference.com/browse/in-band%20signalling>, 07/04/2008
- [14]“TR-45 J-STD-025 Rev. A Lawfully Authorized Electronic Surveillance”, Telecommunication Industry Association, May 31, 2000
- [15]Scott W. Coleman, “DoJ Files Deficiency Petition with FCC over J-STD-025B”, TMCnet, June 7,2007, <http://blog.tmcnet.com/lawful-intercept/doj-files-deficiency-petition-with-fcc-over-jstd025b.asp>, 02/04/2008
- [16]U.S. Congress, Foreign Intelligence Surveillance Act of 1978 Amendments Act of 2008, 3 January 2008 <http://thomas.loc.gov/cgi-bin/query/z?c110:H.R.6304>
- [17]Jim X. Dempsey, “CDT's Analysis of S. 2092: Amending the Pen Register and Trap and Trace Statute in Response to Recent Internet Denial of Service Attacks and to Establish Meaningful Privacy Protections”, Center for Democracy & Technology, Security & Privacy, Wahsington, DC, April 4, 2000, <http://www.cdt.org/security/000404amending.shtml>, 07/04/2008 – see also The "Carnivore" Controversy: Electronic Surveillance and Privacy in the Digital Age, Hearing before the Committee on the Judiciary United States Senate, Committee on the Judiciary, September 6, 2000, Serial No. J-106-105, U.S. Government Printing Office Washington, 2001, pages 47-61. <http://www.loc.gov/law/find/hearings/pdf/00089583263.pdf>
- [18]Aqsacom, “Lawful Interception For 3G Networks”, Document No. 040450, v .4, November, 2005, Aqsacom Inc. Washington, DC
- [19]Shana K. Rahavy, “The Federal Wiretap Act: the Permissible Scope of Eavesdropping in the Family Home”, The Journal of High Technology Law, vol. II, No 1, 2003, pages 95.-98.
- [20]Eric Cronin, Micah Sherr, and Matt Blaze. “The eavesdropper's dilemma”, University of Pennsylvania Technical Report, number MS-CIS-05-24. August 2005. <http://www.crypto.com/papers/internet-tap.pdf>
- [21]Robert Valdes and Dave Roos, “How VoIP Works”, 09/05/2001, HowStuffWorks.com <http://communication.howstuffworks.com/ip-telephony.htm>, 15/05/2008
- [22]Steven Bellovin, Matt Blaze, Ernest Brickell, Clinton Brooks, Vinton Cerf, Whitfield Diffie, Susan Landau, Jon Peterson, and John Treichler, “Security Implications of Applying the Communications Assistance to Law Enforcement Act to Voice over IP”, Information Technology Association of America, June 13,2006
- [23]Fred Baker, Brian E. Carpenter, NWG, “RFC 2804 – IETF Policy on Wiretapping”, May 2000

- [24]P. Branch, "Lawful Interception of IP Traffic," Australian Telecommunications, Networks and Applications Conference (ATNAC), Melbourne, December 8-10, 2003
- [25]The Economist, "Bugging the cloud", March 6<sup>th</sup> 2008, [http://www.economist.com/printedition/displaystory.cfm?story\\_id=10789393](http://www.economist.com/printedition/displaystory.cfm?story_id=10789393), last access 22/05/2008
- [26]Berson Tom, "Skype Security Evaluation", Anagram Laboratories, 18 October 2005
- [27]Cisco Systems, "Lawful Intercept Architecture", [http://www.cisco.com/en/US/docs/ios/12\\_2sb/feature/guide/ht\\_ssi.html](http://www.cisco.com/en/US/docs/ios/12_2sb/feature/guide/ht_ssi.html), 29/05/2008
- [28]Cisco Systems, "Cisco Service Independent Intercept Architecture Version 1.0", version 2, 15 March 2006, <http://www.cisco.com/technologies/SII/SII.pdf>
- [29]WikiLeaks, "Bavarian trojan for non-germans", [http://wikileaks.org/wiki/Bavarian\\_trojan\\_for\\_non-germans](http://wikileaks.org/wiki/Bavarian_trojan_for_non-germans), last modified 28 April 2008, last accessed 29/05/2008
- [30]Privacy International, "PHR2006 – Privacy Topics – Surveillance of Communications/ Internet Surveillance: Black Boxes and Key Loggers", 18/12/2007, <http://www.privacyinternational.org/article.shtml?cmd%5B347%5D=x-347-559085>, 07/08/2008
- [31]Xinyuan Wang, Shiping Chen, and Sushil Jajodia, "Tracking Anonymous Peer-to-Peer VoIP Calls on the Internet", 7-11 November 2005, Alexandria, Virginia, USA
- [32]Charlie Kaufman, Radia Perlman, Mike Speciner, "Network Security Private Communication in a Public World", Prentice Hall PTR, 2002
- [33]Matt Blaze, "Protocol Failure in the Escrowed Encryption Standard", Proceedings of Second ACM Conference on Computer and Communications Security, Fairfax, VA, November 1994.
- [34]Dept. of Commerce, "Interim Rule on Encryption Items," Federal Register, Vol. 61, p. 68572 December 30, 1996
- [35]Hal Abelson, Ross Anderson, Steven M. Bellovin, Josh Benaloh, Matt Blaze, Whitfield Diffie, John Gilmore, Peter G. Neumann, Ronald L. Rivest, Jeffrey I. Schiller, Bruce Schneier, "The Risks of Key Recovery, Key Escrow, and Trusted Third Party Encryption", Sun Microsystems, Menlo Park, 27 May 1997
- [36]Erland Jonsson, "KEY ESCROW – a System for Law-Enforced Covert Surveillance and its Risks", Department of Computer Engineering, Chalmers University of Technology, 01 December 2004
- [37] Martin Rex, "RE: [TLS] draft-housley-evidence-extns-00 worse than key escrow", IETF-TLS, 8 January 2007

- [38]Matthew Blaze, “Testimony Of Dr. Matthew Blaze Before The Senate Committee On Commerce, Science, And Transportation, Subcommittee On Science, Technology, And Space”, Congressional Hearings Intelligence and Security, June 26 1996, [http://www.globalsecurity.org/intell/library/congress/1996\\_hr/960626\\_blaze\\_test.htm](http://www.globalsecurity.org/intell/library/congress/1996_hr/960626_blaze_test.htm), 04/06/2008
- [39]Simson L. Garfinkel, “VoIP and Skype Security”, Creative Commons, 3/12/2005
- [40]Mark Klein, “Wiretap Whistle-Blower’s Account”, Wired News, April 6, 2006, <http://www.wired.com/science/discoveries/news/2006/04/70621>, 08/08/2008
- [41]Ryan Singel, “NSA Must Examine All Internet Traffic to Prevent Cyber Nine-Eleven, Top Spy Says”, Wired News, January 15, 2008, <http://blog.wired.com/27bstroke6/2008/01/feds-must-exami.html>, 08/08/2008
- [42]Ryan Singel, “NSA’s Lucky Break: How the U.S. Became Switchboard to the World”, Wired News, October 10, 2007, [http://www.wired.com/politics/security/news/2007/10/domestic\\_taps](http://www.wired.com/politics/security/news/2007/10/domestic_taps), 08/08/2008
- [43]M. Baugher, D. McGrew, M. Näslund, E. Carrara, and K. Norrman, "The Secure Real-time Transport Protocol", IETF RFC 3711, March 2004
- [44]Elisabetta Carrara, “Security for IP Multimedia Applications over Heterogeneous Networks”, Licentiate thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, August 31 2004, <http://web.it.kth.se/~carrara/licproposal.pdf>
- [45]Elisabetta Carrara, Security for IP Multimedia Applications over Heterogeneous Networks, Licentiate thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, May 2005, <http://web.it.kth.se/~carrara/lic.pdf>
- [46]Peter Thermos, Ari Takanen, “Securing VoIP Networks Threats, Vulnerabilities, and Countermeasures”, Addison-Wesley Professional, August 01,2007
- [47]J. Arkko, E. Carrara, F. Lindholm, M. Naslund, and K. Norrman, “MIKEY: Multimedia Internet KEYing”, IETF RFC 3830, August 2004
- [48]John Leyden, "Italy tops global wiretap league: State of the surveillance nation", The Register, Wednesday 7th March 2007 18:15 GMT [http://www.theregister.co.uk/2007/03/07/wiretap\\_trends\\_ss8/](http://www.theregister.co.uk/2007/03/07/wiretap_trends_ss8/)
- [49]Lewis Page, "VOIP and the web baffle Brit spook wiretappers: MI5, GCHQ bemoan 'biggest change since telephones', The Register, Wednesday 30th January 2008 15:27 GMT [http://www.theregister.co.uk/2008/01/30/gchq\\_mi5\\_baffled\\_by\\_ip\\_wiretapping/](http://www.theregister.co.uk/2008/01/30/gchq_mi5_baffled_by_ip_wiretapping/)
- [50]Erik Eliasson, et al., Minisip, <http://www.minisip.org/>, last accessed 2008.09.07.
- [51]SIPphone.com, Gizmo5, <http://gizmo5.com>, last accessed 2008.09.07.
- [52]Axill Europe Ltd., Globe7, <http://www.globe7.com/>, last accessed 2008.09.07.
- [53]Billy Biggs, Wirlab, and others, KPhone, <http://sourceforge.net/projects/kphone>, last accessed 2008.09.07.

## 10 Appendix A

### *Eavesdropper's dilemma*

Generally, it is considered that in order for a tap to be effective the recording must have 'fidelity'. Here, fidelity means that the recording is free from all the factors (like excessive noise or distortion) that can decrease the quality of the reproduced call contents. For an interception system in order to achieve fidelity it must be free from some serious threats that can seriously undermine its performance. The above threats are known as: obfuscation<sup>11</sup>, confusion<sup>12</sup>, and evasion<sup>13</sup>.

The main problem that faces an eavesdropper is how sensitive<sup>14</sup> and selective<sup>15</sup> a recording has to be. The above two characteristics are in collision. If the eavesdropper is very sensitive and so captures all the messages in order to avoid the attack of evasion, the interception system is vulnerable in confusion as it may record a lot of noise that will decrease the quality of the intercepted contents. On the other hand, if the eavesdropper is very selective in order to avoid the attack of confusion, many of the transmitted messages can elude from the tapping as they can be ignored as insignificant messages.

The above problem exists also in POTS especially when the analog signal is converted to digital (i.e. the signaling between the subscriber and the switch). In order to perform, POTS uses a touch-tone system which "is an international standard known as DTMF" [20] that all the manufacturers of telephone equipment accepted in order to exist a global compatibility. The DTMF's specification "lists acceptable ranges for tone duration, spacing, frequency, amplitude, and twist" [20].

However, in reality the production of a DTMF decoder that can precisely follows the above specifications proved to be harder and more expensive than it was expected. The majority of the decoders are tolerant in some of the standard's specifications. As a result every type of decoder has a unique range of accepted tones that leads to production of out-of-specification tones and in the rejection of them by some other decoders. The negative consequence of the above is the exploitation of this knowledge by a subject with the performing of the attacks of evasion and confusion.

The above dilemma remains in every effort to intercept. The solution for on how selective or sensitive a tap has to be is not an easy task. However, suggestions and decisions can be taken according to the specific characteristics of each individual interception.

---

<sup>11</sup> When an interception system cannot correctly interpret a message (the message's contents, its headers, or both) [20].

<sup>12</sup> When an interception system records messages that are sent on purpose from one party in order to confuse the law enforcement agency as the second telephone party knows to reject them [20].

<sup>13</sup> When an interception system cannot capture and record all the messages which are sent between two telephone parties [20].

<sup>14</sup> When an eavesdropper considers all the messages as important and records them [20].

<sup>15</sup> When an eavesdropper must recognize the real call contents [20].

