Global Information Assurance Certification Paper

GIAC Security Essentials Certification (GSEC)

Practical Assignment Version 1.4

Option One, Research on:-

Eavesdropping an IP Telephony Call

(Submitted by: Tom Long)

1 Introduction

There is a clear and present trend within the enterprise market to replace traditional digital Private Branch Exchange's (PBX) with emerging IP Telephony solutions. Some would say that this trend spell's Danger.

There are numerous excellent technical papers that highlight and discuss the security implications of deploying an IP Telephony solution [1]. This particular paper examines the area of call eavesdropping in a little more detail and outlines how the process of eavesdropping has changed as the PBX technology has evolved. This paper suggests that the art of call eavesdropping may be easier that ever before. Therefore, careful consideration is required before implementing an IP Telephony solution.

2 PBX Evolution

2.1 Legacy Analogue PBX

Before the mid 1980's the PBX mainly consisted of an analogue switch that housed a number of line and trunk modules. The line modules were used to connect the internal staff with the PBX while the trunk modules were used to provide interconnectivity to the Public Switched Telephone Network (PSTN). This solution comprised of simple 2-wire analogue circuits over which both the signalling and audio information would travel (known as in-band signalling). Before a conversation could take place a circuit had to be established between the two telephones, this phase is known as the call setup phase. A call could be eavesdropped with the combination of a standard speaker and access to the PBX switch room or the Main Distribution Frame (DMF).

The only countermeasure to ensure user confidentiality was to restrict access to the PBX room and wiring closets to authorised personnel only.

2.2 Traditional Digital PBX

The PBX, consistent with most other products of the day evolved from analogue technology to digital technology. Even though the principles remained the same (that is, connection oriented systems with call setup and in-band signalling) the analogue modules were replaced with digital modules. For the user, the analogue handset was replaced with a more complicated digital handset. This digital handset would convert the analogue audio signal into a digital format before transmitting the signal over the same simple 2-wire circuit. In addition, the digital handset provided the user with a lot more functionality compared to the analogue set. This new functionality included:

- Programmable keys to support features such as, Speed Dial, Call Forward, Call Pickup and Call Conference.
- A Liquid Crystal Display (LCD) to indicate date, time, incoming caller number and call duration.

- A Light Emitting Diode (LED) to indicate that there is a new message waiting in the user's voice mailbox.

- A speaker option to enable 'hands free' use of the phone.

This new digital era brought with it increased security and hence the process of eavesdropping became much more difficult. As there were no standards available for digital signalling between handsets and PBX's, vendors developed proprietary signalling protocols. The signalling protocol was used to facilitate the following telephony features:

- Call setup
- Call forward
- Call conference
- Calling name display
- Voice Mail LED indication

The digitised voice stream was also encapsulated in a proprietary frame structure to facilitate the transmission of voice between the handset and PBX.

Due to these proprietary features even with access to the MDF one needed sophisticated equipment in order to eavesdrop on a conversation. The digital PBX was definitely a step in the right direction in terms of confidentiality and authentication.

2.3 Emerging IP Telephony Solutions

There is no denying that an IP Telephony solution has a number of advantages over a traditional digital PBX. These advantages includes such things as:

- Single Integrated voice and data network, both voice and data are transported over the same IP network.
- Simplified add, moves and changes.
- Resilient deployment, the different elements of an IP Telephony solution can be located in different communication rooms as long as there is IP connectivity between the locations.
- Excellent application support, for example Unified Messaging. This is an application that enables users to fully integrate their email and voice mail mailbox.
- Reduced fixed wiring cost, typically a user desk is serviced with two Cat 5 cables, one for the phone and the other for the PC. However, with IP Telephony only a single Cat 5 cable is required to service each user's desk as the PC and phone can share the same cable back to the wiring closet.

This paper examines the process of eavesdropping on an IP call. However before understanding the process of eavesdropping it is important to understand the fundamentals of IP Telephony.

3 Fundamentals of IP Telephony

IP Telephony is a lot if different things to different people. For the purpose of this paper IP Telephony is any voice solution that relies on IP to carry voice between two parties. A typical IP Telephony solution consists of the following core components:

Data Network, in order to deploy a successful IP Telephony solution it is a prerequisite that there exists a high performance, resilient and feature rich data network. The data network must be able to support modern Quality of Service (QoS) standards in order to prioritise the voice traffic over data traffic. Standard QoS mechanisms include differentiated services (commonly referred to as 'DiffServ') and 802.1p. DiffServ is a layer-3 QoS mechanism that redefines 6-bits of the Type of Service byte in the IP header. These 6-bits, which are called the DiffServ Code Point (DSCP) are used to classify the priority/importance of an IP packet. 802.1p is a layer-2 QoS mechanism that utilises 3-bits of the 802.1Q frame tag to classify the priority of an Ethernet frame.

IP Handset, the user's handset must be IP enabled so that the audio stream can be digitised and inserted into an IP packet ready for transmission across the IP network. As mentioned above, one benefit of an IP Telephony solution is the reduced fixed wiring cost. This reduction in cost is achieved, by running a single Cat 5 cable to each user to provide both voice and data services. This is possible because the IP handset has an in-built 3-port switch. One port connects back to the ethernet switch, the second port connects to the user's PC while the third port connects directly to the phone (this is usually an internal connection).

Call Server, an application running on a dedicated workstation that provides all call signalling and call control functionality. In essence, the core operating code of a PBX has been ported onto a standard workstation. Cisco have developed their call server application to run on a Windows™ 2000 platform while Avaya have ported their PBX code onto a Linux platform.

Gateway, a purpose built networking device to enable voice connectivity between the IP network and public carrier network. Typically, a gateway consists on an ethernet interface that connects directly to the IP network and a G.703 interface that connects to a primary rate ISDN interface.

As previously discussed, voice is based on a connection oriented technology which implies there must be a call setup phase before any voice traffic is carried across the IP network. There are two standards based signalling protocols:

- H.323
- Session Initiated Protocol (SIP)

3.1 H.323 Protocol

H.32X is an umbrella of recommendations from the ITU-T to support voice, video and data in one application over various media. More specifically H.323 [2] is a standard to support multimedia over a non-guaranteed bandwidth network such as a Local Area Network (LAN). The H.323 standard introduces the following terminology [3].

Terminals, are devices used by the end-users to send and receive voice, video and/or data across a H.323 enabled network. Terminals connect into the LAN and enable real time bi-directional communications. Examples of H.323 terminals include an IP handset or a workstation running Mircosoft's™ NetMeeting application.

Multipoint Control Unit (MCU), for a standard two-way H.323 session terminals establish a direct connection. However, to support multi-terminal conferences (for example a four-way conference call) all terminals must establish a direct connection to an MCU. The MCU in turn controls the conference call between the different terminals.

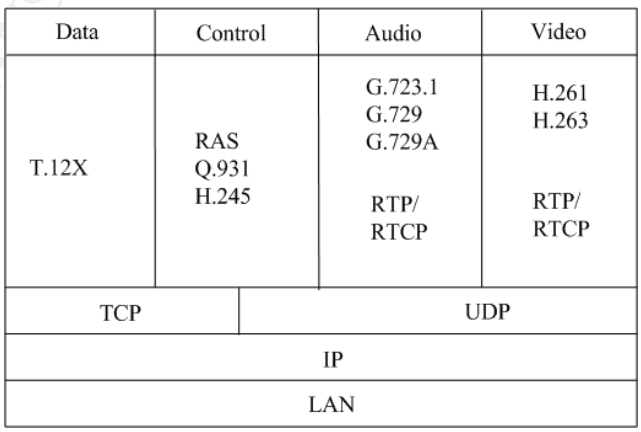
Gatekeeper, a gatekeeper provides access control to the H.323 network for terminals, gateways and MCU's. A gatekeeper is responsible for access control, bandwidth control and call management. In simple terms the gatekeeper is the device that assists the terminals in establishing connections.

Gateway, a gateway is the interface between the IP Telephony domain and the traditional circuit switched domain. The gateway is responsible for providing transparent connectivity between a packet switched telephony network and circuit switched telephony network.

The H.323 protocol suite has four major components:

- Data
- Control
- Audio
- Video

The positioning and relevant protocols can be viewed in the following graphic [4],



The first component, the Data component enables the terminals to transfer data across the IP network. As this transfer must be reliable this component uses Transport Control Protocol (TCP) as the layer-4 protocol. The driver behind this component is to allow users to exchange information during a call. For example, this component could be used to transfer a document to all parties of a conference call.

The second component, the Control component is responsible for

- Registration, Admission and Status (RAS)
- Call Signalling (Q.931)
- Control Channel (H.245)

Every time a terminal joins a H.323 network it must register with the gatekeeper and provide constant status update messages to ensure that the gatekeeper is always aware of the status of the terminal. RAS relies on User Datagram Protocol (UDP) instead of TCP to satisfy this requirement, as TCP would place a large overhead on the gatekeeper with either a large number of established TCP connections or a large frequency of new TCP connections.

Call signalling is used to setup a connection between two terminals. Call signalling is used to indicate to the called terminal that there is an incoming call (i.e. make the far-end terminal ring), to update the calling terminal as to the status of the call (i.e. generate ringing tone or engaged tone) and issue the calling terminal with the IP address of the called terminal.

The Control Channel is used to exchange terminal messages that indicate the capability of each terminal.

Both Call Signalling and Control Channel are tolerant of delay and intolerant of packet loss and therefore use TCP for reliable transport across the IP network.

The third component, the Audio component comprises of voice encoders such as G.711 (64 Kbps encoded voice) and G.729 (8 Kbps encoded voice) to digitise and compress the analogue voice signal and utilises the Real-time Transport Protocol (RTP) to transport for the audio stream between terminals. As voice is tolerable of a small amount of packet loss but intolerant of delay RTP is transported over UDP.

The fourth component, the Video component comprises of video encoders that compress and encapsulate the video signal prior to transmission across the IP network.

3.2 SIP Protocol

The Session Initiation Protocol has been developed by the Internet Engineering Task Force (IETF) as an alternative signalling protocol to H.323.

"SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions (conferences) such as Internet telephony calls. SIP can also invite participants to already existing sessions, such as multicast conferences. Media can be added to (and removed from) an existing session. SIP transparently supports name mapping and redirection services, which supports personal mobility - users can maintain a single externally visible identifier regardless of their network location." [5].

Due to the simple and more flexible implementation it is commonly agreed that in the future SIP will be the preferred signalling protocol of choice. SIP is a peer-to-peer protocol and within SIP terminology each peer is called a User Agent (UA). Each user agent can operate in two different modes:

- User Agent Client (UAC), an agent that initiates a SIP call and generates a SIP request.
- User Agent Server (UAS), an agent that receives a SIP request and generates a SIP response.

Every user agent can operate in both modes but must operate as either a client or server on a per session basis. The actual mode that the user agent operates as depends on whether the user agent initiates the session.

User agents can be grouped into two different categories:

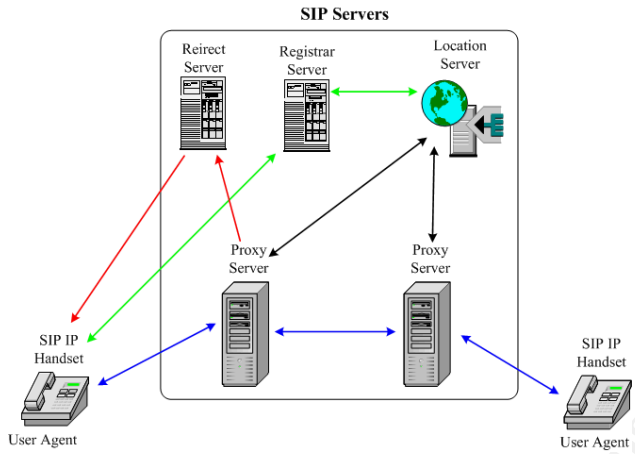
- SIP Clients, any IP phone or gateway with SIP support.
- SIP Servers, a range of servers that are required in order to establish an end-to-end SIP session.

SIP Servers
A Proxy Server is an intermediate device that receives a SIP request, checks the destination SIP client and forwards the request to the appropriate agent.

A Redirect Server instead of forwarding a SIP request to the next agent a redirect server instructs the SIP client to directly forward the SIP request to the appropriate agent.

A Registrar Server processes requests from SIP clients that wishes to join a SIP network. The registrar server will in turn update the location server with the client information.

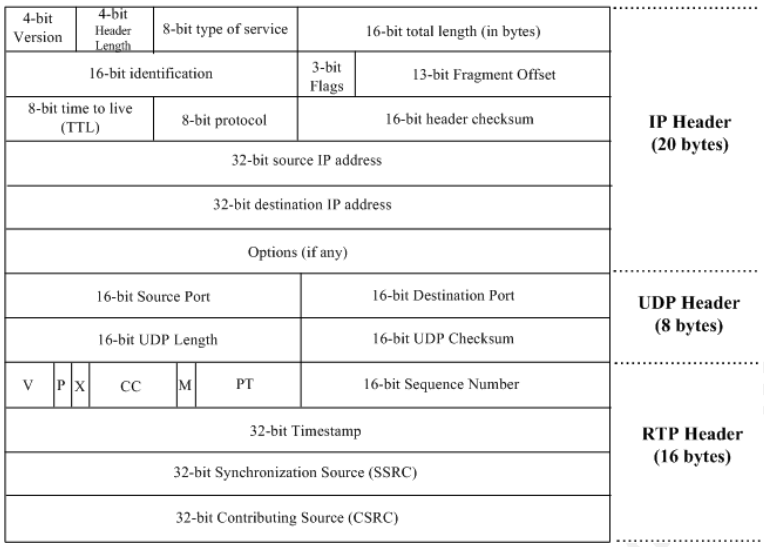
A Location Server is basically a database that contains information regarding the location of SIP clients. The proxy server will reference this server to select the appropriate agent to forward the request to.



The size and scale of a SIP network will dictate the number of servers required for an implementation. However, it should be noted that for small implementations a number of services can exist on the same physical server.

3.3 Introduction to RTP

RTP was developed by the IETF to transport audio and video across an IP network. A complete definition of the RTP protocol can be found in RFC 1889 [6]. With an IP Telephony solution, one important point to note is that independent of the signalling protocol deployed (H.323 or SIP) RTP is used to encapsulate the audio stream. For each IP Telephony call, there is a call setup phase and it is during this phase that both terminals are instructed to use a particular port number. Having completed the call setup phase each terminal samples the input audio stream, digitises it and encapsulates the payload into a RTP header. The RTP header contains sufficient information so that the far end terminal knows what codex was used to digitise the voice and if the voice packet is in the correct sequence.



The graphic [7] above shows the IP, UDP and RTP headers that are used to transport voice traffic across an IP network. Some of the important fields are identified below:-

IP Header
8-bit Type of Service Field, this field has also been redefined as the 'Diffserv' code point field and is used to identify the priority of the IP packet.

8-bit Protocol Field, this field is used to identify the layer-4 protocol. If this field is set to a value of 6 decimal then the layer-4 protocol is TCP while if this field is set to 17 decimal then the layer-4 protocol is UDP.

32-bit Address Fields, these fields represent the IP addresses of the end devices, in this case the IP address of the terminals.

UDP Header
16-bit Port Fields, these fields contain the port numbers that are used for any given call. As these ports are selected from a range of ports a network analyser will not instinctively know that a particular packet is carrying an RTP payload.

RTP Header
16-bit Sequence Number, this field is used by the far-end terminal to detect packet loss and ensure that packets are received in the correct order.

32-bit Timestamp, this field is used by the far-end terminal to monitor the latency (packet delay) and jitter (delay variation) of an IP Telephony call. Latency and Jitter are very important parameters that are used to indicate the quality of a voice call.

32-bit SSRC Field, this field is used to uniquely identify the source of an RTP stream.

As with any voice solution, the quality of the voice call is key when determining the success of the technology. For every established call, IP telephony relies on the RTP Control Protocol (RTCP) to periodically update terminals and gateways with voice

quality information.

4 Eavesdropping an IP Telephony Call.

The main objective of this paper is to highlight the ease at which an IP Telephony call can be eavesdropped. In order to successfully eavesdrop a call there are a number of simple steps:

- Understand the fundamentals of IP Telephony.
- Obtain the necessary tools.
- Connect the tools into the IP network.

Step One:- Understanding the fundamentals of IP Telephony
This paper provides a brief introduction to IP Telephony but in terms of eavesdropping the most important fact is, independent of vendor and signalling protocol all IP Telephony implementations use RTP to transfer the digitised voice packets across an IP network. It should be noted that RTP does not use a fixed UDP port but instead relies on the call server to dynamically select a port from a pre-defined range [8]. Each IP phone learns of this UDP port via the signalling protocol.

Step Two:- Obtain the necessary Tools
The Fireberd DNA-323 [9] commercially available software application is an easy to use H.323 analysis tool that enables the user to capture and play back IP Telephony conversations. Fireberd DNA-323 is a Windows™ based application however as an alternative there is a UNIX based application available called VOMIT [10] (Voice Over Misconfigured Internet Telephones)

Step Three:- Connect the tool into the IP Network

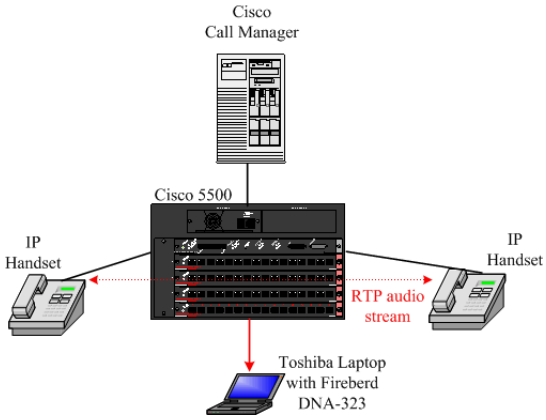
For the purpose of this paper the tool was connected into the same ethernet switch as the IP phones and the switch was configured to copy all traffic to/from the IP phone to the tool.

4.1 Test Setup

This test setup comprised of a Cisco™ AVVID IP Telephony solution connected directly to a Cisco Catalyst™ 5500 Ethernet switch. The IP phones were directly connected into two fast ethernet switch ports. The IP phones were configured to be a member of the same domain and all calls within the domain were configured to use the G.711 encoding algorithm.

The Fireberd DNA-323 application was installed on a Toshiba™ Tecra 8100 laptop running the Windows™ NT4 (SP6a) operating system.

The Catalyst™ 5500 ethernet switch was configured to copy all frames to/from the IP phone to the port with the Fireberd analysis tool connected.



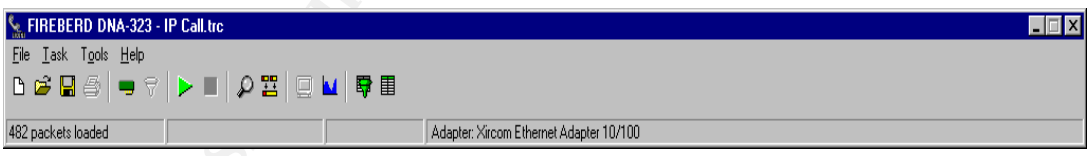
If configuration access to the switch was unavailable other techniques such as ARP spoofing could have been used to ensure that the tool captured the RTP packets. ARP spoofing is a technique used by hackers to alter the normal flow of IP packets. The technique involves generating a spoofed ARP response packet in response to a valid ARP request packet. This process enables the hacker to direct the flow of IP packets through a compromised workstation from where the packets can be captured.

4.2 Fireberd DNA-323 Installation

The Fireberd DNA-323 application is a 3.3 Mbyte application that is downloaded as a single compressed file. When this file is uncompressed simply double click on the 'Setup' icon to install the application.
The evaluation license is valid for 45 days from the date of installation. The latest version that was available for this analysis was version 2.2. The application does require a reboot of the operating system before it can be used.

4.3 Fireberd DNA-323 Operation

Having installed the application the next step is to capture frames, this is achieved by selecting 'Start Capture' under the Task Menu option.

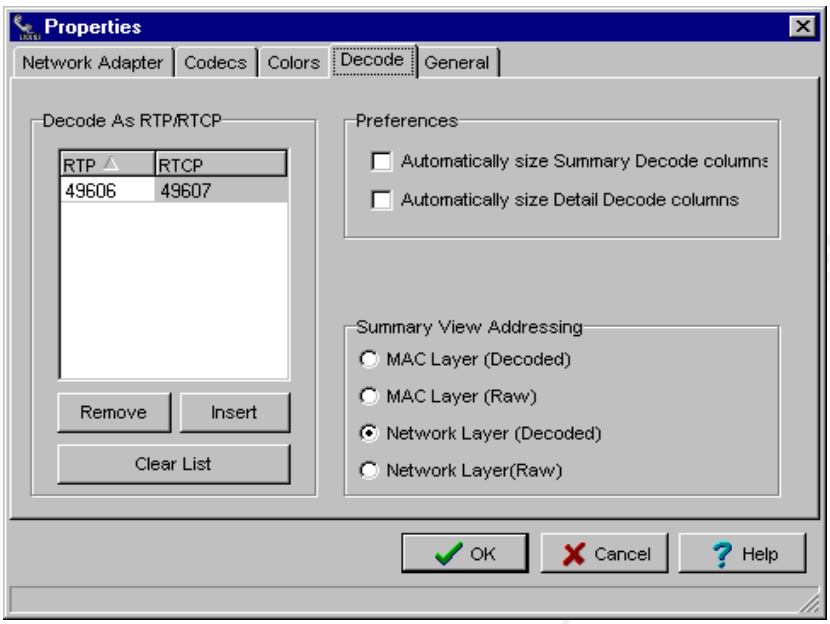


Assuming that there is an IP Telephony call in place and that the switch is correctly configured to forward the IP telephony frames, the application should start collecting frames.

Stop the capture process and open the Decode Display, by selecting 'Display Capture' under the Task menu option. In this window it will be possible to identify the RTP frames by looking for a high rate of UDP packets with a fixed frame size. At this point make a note of the destination UDP port number.

As discussed earlier each IP Telephony call is assign a unique UDP port number and as the Fireberd application does not automatically obtain this information from the signaling protocol we must manually configure the application with the UDP port in use. This is achieved by returning to the main application window and selecting 'Properties' under the Tools menu option.

Under the Decode tab enter the UDP port number that was noted in the previous step.



Now when the decode window is re-opened the IP Telephony call packets are displayed as RTP packets. Ensure the Layer View is set to filter and click on one of the RTP packets. At this point select the 'Filter Conversation' button at the bottom of the window. This action will open a Conversation Display window.

The last task is to select ' Voice Drop' under the RTP menu option, the application will open a new window titled Voice Drop. From this window the IP conversation can be simply played back by clicking on the Play button.

5 Conclusion

In conclusion, there are two observations that this paper presents.
Firstly, there is an inherent vulnerability with current implementations of IP Telephony solutions as the voice stream is transported unencrypted across an IP network. There is also a threat that this vulnerability could be exploited due to the number of tools readily available (such as; DNA-323, VOMIT, ethereal and arpoison) to eavesdrop on an IP Telephony call. The combination of this vulnerability and threat suggest that there is a risk associated with making an IP Telephony call.

Secondly, the risk of confidentiality being compromised can be minimised by

implementing a number of simple procedures:

- Ensure that access to wiring closets is restricted to authorised personnel only.
- Consider implementing port based MAC address security on any vulnerable network points, for example a reception courtesy phone.
- Initiate a procedure to regularly scan the network for devices running in promiscuous mode [11].
- Both the H.323 and SIP protocol can support a configuration where all voice communication is encrypted however this will introduce some adverse effects, such as; increased packet delay, increased complexity and increased cost.

In summary there is a risk associated with deploying an IP Telephony solution however, this risk can be minimised by understanding the risk and implementing a number of countermeasure actions.

6 References

[1] Ackermann,Ralf Schumacher, Markus Roedig, Utz Steinmetz, Ralf "Vulnerabilities and Security Limitations of Current IP Telephony Systems" URL: http://www.kom.e-technik.tu-darmstadt.de/publications/abstracts/ASRS01-1.html
Section 3 of this paper highlights a number of real attacks that can be instigated against an IP Telephony solution.

[2] Open H323 Project. OpenH323. URL: http://www.openh323.org
This is a very useful site for obtaining information pertaining to the H.323 signalling protocol.

[3] Leppanen, Marko "Voice over IP" URL: http://www.hut.fi/~mjleppa2/voip/voip.pdf
Section 3.1 of this paper contains a good introduction to the major components of the H.323 protocol.

[4] Voye, Dave " H.323 – Multimedia over LAN" Motorola University

[5] Rosenberg, J. Schulzrinne, H. Camarillo, G. Johnston, A. Peterson, J. Sparks, R. Handley, M. Schooler, E. "SIP: Session Initiation Protocol" June 2002 URL: http://www.ietf.org/rfc/rfc3261.txt?number=3261

[6] Schulzrinne, H. Casner, S. Frederick, R. Jacobson, V. "RTP: A Transport Protocol for Real-Time Applications" January 1996 URL: http://www.ietf.org/rfc/rfc1889.txt?number=1889

[7] Arkin, Ofir "VoIP The Next Generation of Phreaking" URL: http://www.blackhat.com/html/win-usa-02/win-usa-02-spkrs.html
Slide 42 from this powerpoint presentation displays the RTP header format.

[8] Chown, T. De Roure, D. Juby B. Thompson M. "The Potential threat of interception, monitoring and changes to the content of H.323 conferences" March 2001 URL: http://www.ja.net/development/video/vip/reports/south2.pdf
This paper produced by the University of Southampton provides excellent information regarding the

process of eavesdropping on an IP Telephony call using the Fireberd DNA-323 application.

[9] Fireberd DNA-323 URL:
http://ttcweb1.ttc.com/ttc/prod.nsf/vwPDF/FIREBERD+DNA-323+H.323+Analyzer/$File/fbDNA-323.pdf?OpenElement
This pdf provides a product overview of the Fireberd DNA-323 application.

[10] "VOMIT – voice over misconfigured internet telephones" URL:
http://vomit.xtdnet.nl

[11] Packet Storm "AntiSniff" URL: http://packetstormsecurity.nl/sniffers/antisniff/
This site contains information regarding a product that can be used to identify any device running in promiscuous mode.