Edith Cowan University
Research Online

Australian Information Security Management Conference

Security Research Institute Conferences

2007

Securing VoIP: A Framework to Mitigate or Manage Risks

Peter James
Edith Cowan University

Andrew Woodward
Edith Cowan University

Originally published in the Proceedings of 5th Australian Information Security Management Conference, Edith Cowan University, Perth Western Australia, December 4th 2007
This Conference Proceeding is posted at Research Online.
http://ro.ecu.edu.au/ism/33

Securing VoIP: A Framework to Mitigate or Manage Risks

Peter James and Andrew Woodward
School of Computer and Information Science
Edith Cowan University
Perth, Western Australia
pjames1@student.ecu.edu.au

Abstract

In Australia, the past few years have seen Voice over IP (VoIP) move from a niche communications medium used by organisations with the appropriate infrastructure and capabilities to a technology that is available to any one with a good broadband connection. Driven by low cost and no cost phone calls, easy to use VoIP clients and increasingly reliable connections, VoIP is replacing the Public Switch Telephone Network (PSTN) in a growing number of households. VoIP adoption appears to be following a similar path to early Internet adoption, namely little awareness by users of the security implications. Lack of concern about security by VoIP users is probably due to the relatively risk free service provided by the PSTN. However, VoIP applications use the Internet as their communications medium and therefore the risk profile is significantly different to the PSTN. This paper reviews the risks for two VoIP implementation models now being increasingly used in Australian homes; the PC softphone and the Analogue Telephony Adaptor (ATA). An overview of each of the VoIP implementation models is given together with a description of the respective technologies and protocols utilised. The VoIP security threats, applicable to the two VoIP implementation models considered, are enumerated and vulnerabilities that could be exploited are considered. Available security mechanisms that address the identified vulnerabilities are discussed. A practical and pragmatic VoIP security framework is proposed that will enable a user to mitigate or manage the risks associated with using the VoIP implementation models considered. By applying the VoIP security framework a user will be able to deploy a secure VoIP solution appropriate for residential use.

Keywords

VoIP Security, softphone, analogue telephony adapter (ATA), VoIP threats and vulnerabilities, Risk management

INTRODUCTION

Voice over Internet Protocol (VoIP) is the packaging and routing of voice conversations over an IP-based network, e.g. the Internet. VoIP products have been commercially available for many years but were initially confined to distributed organisations with internal high speed IP networks looking to reduce the cost of intra-company telephone calls. VoIP needs fast IP networks with high throughput, without which VoIP can suffer from poor latency (caused by delays in the time taken for VoIP packets to go from source to destination), jitter (caused by VoIP packets arriving out of sequence) and packet loss. While residential access to the Internet was via dial-up modems VoIP was not feasible. However, the phenomenal growth of residential broadband connectivity has enabled good quality and reliable VoIP services to be available to households.

Most VoIP implementations, like many Internet services, were designed using available standard protocols; with security not being a key design criterion. Adding security to a solution, rather than including it at the design stage, typically leads to an inefficient and sometimes ineffective solution, i.e. adding security functionality can impact a solution's performance and because the security functionality is not integral to the design it may be possible to bypass the functionality. Any impact on performance can be critical to the correct operation of a VoIP solution, i.e. poor latency, jitter and packet loss can occur. Therefore, to be used, VoIP security mechanisms must not impact the performance of VoIP sessions.

A Softphone is application software that runs on a PC and allows telephone calls to be made over an IP network. A Softphone requires no dedicated hardware, it uses the PC soundcard for voice input and output. Whilst a PC's speakers and a microphone can be used, the best results are achieved with a headset, with integrated microphone, or a USB attachable hand phone. There are numerous Softphone applications available; the most popular include Skype, Gizmo and KPhone (distributed with KDE – K Desktop Environment). In Australia, Softphone use started to gain momentum from 2003.

An Analog Telephony Adaptor (ATA), also known as a VoIP adaptor, is a hardware unit that is positioned in-line between an analog telephone and a PSTN line, and connected to an IP network with a broadband modem. Residential use of ATAs in Australia started to gain momentum in late 2004 with rapid growth occurring from mid 2006. An ATA can be packaged in a number of ways including:

- A small box with only ATA functionality: This box, in addition to being connected in-line between a phone and telephone line, must also be connected to an IP network with a broadband modem/router.
- Broadband Modem/Router: ATA functionality is packaged together with a broadband modem/router; both wired and wireless ATA broadband modem/routers are available.
- VoIP Phone: The ATA functionality is packaged within a telephone. The VoIP telephone is connected to both the telephone line and an IP network with a broadband modem/router.

Softphones and ATAs suffer from both common and different security threats; threats can include eavesdropping, denial of service (DoS) and toll fraud. VoIP threats arise because of vulnerabilities in the VoIP technology and/or due to how the technology is configured and deployed. The security framework proposed in this paper identifies the security mechanisms and techniques to prevent the vulnerabilities from being exploited and enable a VoIP solution to be deployed that is appropriate for residential use.

The term ***User Agent (UA)*** is used throughout this paper to refer to both a Softphone and ATA.

VOIP IMPLEMENTATION MODELS

Concept of Operation of VoIP User Agents

There are a large number of Softphone applications available; one survey identified seventy PC Softphone applications with the majority available as freeware (Wikipedia 2007). There is no standard Softphone architecture although most use a common set of standard protocols and therefore have a common mode of operation. Often freeware applications publish the source code, however a number of popular freeware Softphone applications like Skype (Skype 2007) use unpublished proprietary protocols and algorithms. Configuration and initialisation of a softphone client is a usually trivial exercise. In this paper where a Softphone implementation (product) is required as a point of reference the popular products KPhone (KPhone Website 2007) and Skype are used.

ATA's are manufactured by most of the major network equipment vendors. An ATA can be purchased from a specialist electrical retailer or from a VoIP service provider. If purchased from a retailer an ATA can be configured and tuned by a user. An ATA obtained from a VoIP service provider is usually locked (to a user) and automatically configured/provisioned when connected to the Internet via the VoIP service provider's provisioning server. An important advantage of an ATA is that it can be integrated with the PSTN. An ATA can send and receive calls to and from other VoIP UAs. An ATA can also send and receive calls from phones connected to the PSTN. Typically an ATA handles two telephone numbers for a residence; the designated PSTN number and a VoIP number issued by the VoIP service provider.

Figure 1 presents a simple pictorial model of an ATA. The ATA has three connections; an RJ12 connection to the PSTN, an RJ45 connection to an IP LAN connected via a broadband router to the Internet/VoIP network and an RJ12 connection to a telephone.

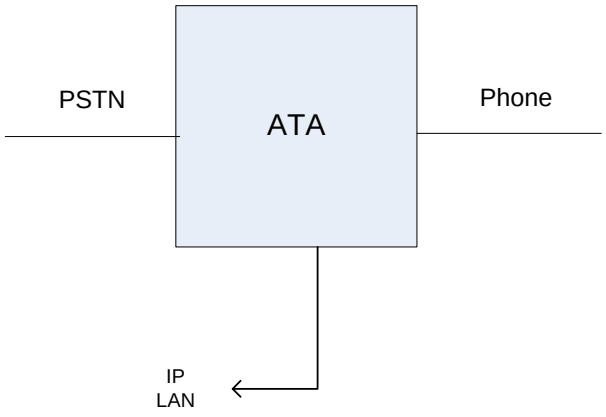

Figure 1: Simple Conceptual Model of ATA

The following set of scenarios show how calls made from and to a residence via either the PSTN or the Internet are handled by an ATA:

9. Call from residence to a phone connected to the PSTN or a UA on the same VoIP network.
    - If the (calling) residence broadband connection is enabled the ATA routes the call onto the LAN and out on to the Internet via the broadband router.
    - If the (calling) residence broadband connection is **not** enabled the ATA routes the call onto the PSTN phone line connected to the ATA, i.e. VoIP is not used to send the call.
10. Call received by residence from a phone connected to the PSTN using the residence PSTN number:
    - The ATA routes the call from the PSTN line straight to the phone, nothing goes through the residence LAN. Calls are received irrespective of whether the residence LAN is working.
11. Call received by residence from a phone connected to the PSTN using the residence VoIP number:
    - The call is received over the Internet (as it has been routed via a PSTN Gateway to a proxy server at VoIP service provider, see Figure 2 below) and therefore the ATA routes the call from the IP connection to the phone. If the residence broadband is **not** enabled the call will not be received.
12. Call received by residence from another UA on the same VoIP network:
    - If the call is made using the PSTN number the call will be routed on to the Internet and then routed on to the PSTN network (via a PSTN Gateway). The call will be received as per scenario 2.
    - If the call is made using the VoIP number the call will be routed over the Internet and through the ATA to the phone.

In this paper where an ATA implementation (product) and VoIP service provider are required as a point of reference the Sipura/Linksys ATA product (Sipura Technology Inc.2004) and the Engin[1] VoIP service (Engin Website 2007) are used.

To enable a concise and consistent analysis to be performed, the two VoIP implementation models considered in this paper utilise the same protocol/technology set; and therefore have the same architecture and concept of operation. The protocol/technology set selected is the set most commonly used in VoIP implementations. The Softphone and ATA products selected as reference points (with the exception of Skype) generally conform to the nominated architecture and protocol/technology set.

Concept of Operation – VoIP Network

VoIP operation is shown using two conceptual models. Figure 2 presents a network model of the different types of 'node' that can exist in a VoIP network while Figure 3 models the sequence of events involved in establishing a VoIP telephone call.

The network model (Figure 2) shows a number of residences, each with different telephony capabilities, and the core elements provided by a VoIP service provider. A description of each of the nodes in the network model is also given.

VoIP Service Provider: The core elements to enable a quality VoIP service to be provided include:

- ***Provisioning Server:*** This server provides configuration details to an ATA when the ATA is first connected to a VoIP network.
- ***Proxy Server:*** The proxy provides two services; registration acceptance and proxying call requests and responses between UAs. When a UA registers it is informing the proxy of its IP address and the port number it can be reached on. Registration is for a finite period and therefore the UA periodically renews registration. When a UA initiates a call it sends a request to the proxy server with details of the intended recipient. The proxy requests the recipient's details from the Location Server.
- ***Billing Server:*** A billing server generates billing data for chargeable calls.
- ***Location Server:*** A location server manages a database of VoIP IP addresses and also contains rules on the most appropriate PSTN Gateway to use to route calls destined for the PSTN.

[1] Engin is currently the largest VoIP broadband service provider in Australia

- ***PSTN Gateway:*** A PSTN Gateway provides an interface to the PSTN to allow a UA to send and receive phone calls from the PSTN. The PSTN Gateway essentially behaves like a UA.

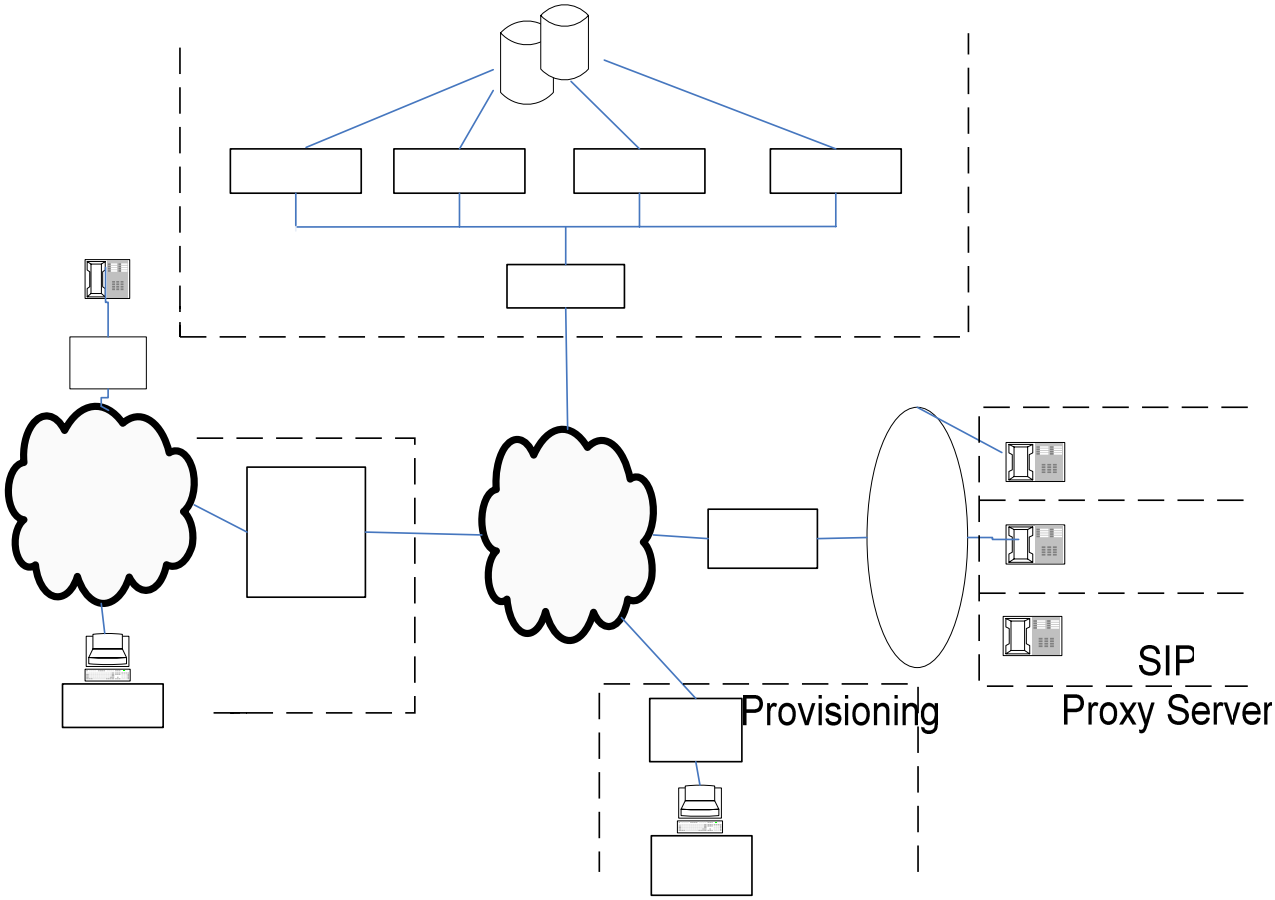

Figure 2: Network Model of Typical VoIP Implementation

Residence A: In this model the residence has both an ATA and a Softphone on an internal IP network. With an ATA the residence can both send and receive calls to/from the PSTN.

Residence B: In this model a Softphone running on a PC is able to make calls to other VoIP users and to the PSTN, however calls from the PSTN may not be able to be received on a Softphone (N.B. some Softphones like Skype can receive calls from the PSTN).

Residence C,D,E: These models represent PSTN phones in homes. These PSTN phones are able to send and receive calls to/from any UA that can access the PSTN via a PSTN Gateway.

Figure 3 models a simplification of the sequence of events in establishing a VoIP call between two UAs.

X can call Y using a PSTN or VoIP telephone number if Y is an ATA; otherwise a VoIP call identifier is required. During registration the VoIP service provider will store UA addressing details in the location database. When X calls Y the following set of events will occur:

13. X's request to call Y is routed to the VoIP service provider's proxy server.
14. The proxy server requests addressing details for Y from the Location Server.
15. Using Y's address the proxy server contacts Y.
16. Y acknowledges and accepts the call from X.

17. The conversation is performed in a 'peer-to-peer' like format, i.e. transmission appears to be direct from UA to UA – in practice transmission is via the VoIP service provider (and/or Internet service provider, if the VoIP and Internet service providers are different).

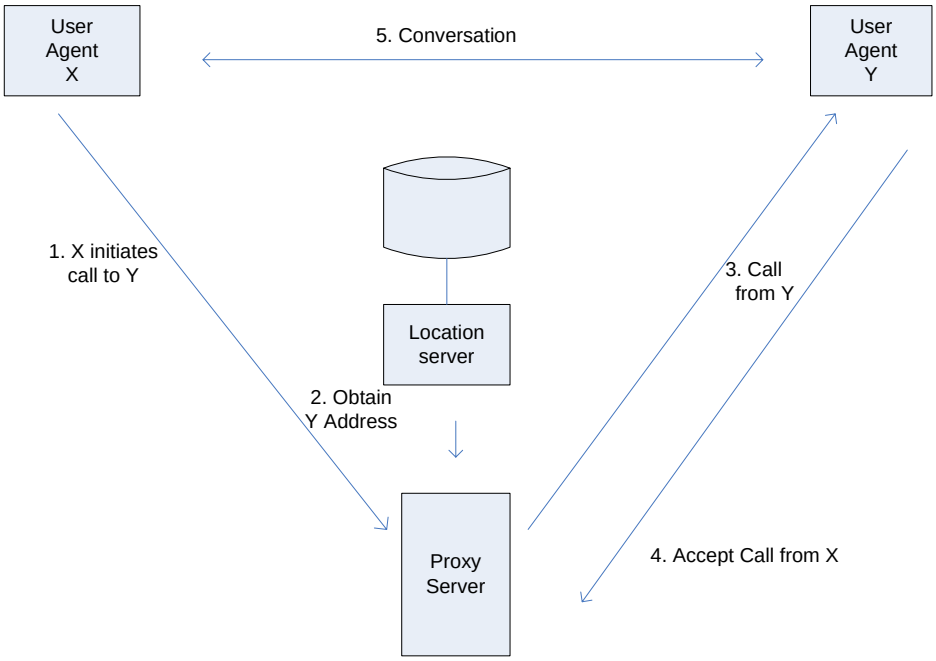

Figure 3: Sequence of Elements Involved in Establishing a VoIP Telephone Call

Typical VoIP User Agent Architecture

The type of UA considered in this paper has both client and server capabilities, i.e. it is 'peer-to-peer' like. The UA architecture is presented as a layered set of components in Table 1; each component is mapped to the relevant layer in the Internet Reference Model and to the particular selected protocol/technology used in this paper for the implementation of the respective component.

Table 1: UA Layered Component Architecture

UA Functional Component	Internet Reference Model	Protocol/Technology
Application GUI (Softphone only)	Application Layer	Typically a Softphone GUI is in the form of a phone.
Call Initialisation (establishing a connection between 2 UAs).		Session Initialisation Protocol (SIP).
Application Quality of Service		Real Time Control Protocol (RTCP)
Analog to Digital Conversion & Digital to Analog Conversion.		A number of standard codecs are available
Packaging of Audio Data	Transport Layer	Real-time Transport Protocol (RTP)
End-to-end connection, delivery assurance		User Datagram Protocol (UDP)
Transport Quality of Service		Resource Reservation Protocol (RSVP)
Packet routing & address resolution	Network/Internet Layer	Internet Protocol (IP)
Data Transmission	Link Layer	Ethernet.

An overview of each of the application and transport layer protocols is given below:

Session Initialisation Protocol: There are a number of protocols available for VoIP call initialisation but SIP is the most popular. SIP can be used with a range of transport protocols but provides a good VoIP solution when coupled with RTP and UDP. SIP is used essentially to introduce the calling parties (UAs) and to inform which IP addresses they are using. SIP requires proxy servers and location servers to perform address resolution, once a connection is established the UAs perform the actual transmission; as shown in Figure 3, i.e. the voice traffic can flow without passing through the proxy servers. A good overview of the application SIP to VoIP can be found in the NIST guidelines (Kuhn et al 2005). SIP uses a default port.

Real Time Control Protocol: RTCP is an application layer protocol used by a UA to assist Quality of Service (QoS) delivery. RTCP does not transport data; its purpose is to supply statistical information to a UA to allow a UA to change parameters and settings to improve QoS.

Real-time Transport Protocol: RTP provides a packet format for delivering VoIP traffic. RTP packets contain the information required to reassemble the packets, upon delivery, into a voice signal. To enable RTP packets to be transmitted across ordinary network nodes on the Internet they are carried as UDP datagrams.

User Datagram Protocol: UDP is one of the key protocols in the IP suite. UDP tends to be used for time sensitive applications and therefore is suitable for VoIP applications; UDP can deliver data faster than other network layer protocols as it does not perform comprehensive transmission error checking.

Resource Reservation Protocol: RSVP is a transport layer control protocol. RSVP does not transport data; it provides QoS information to hosts and routers to enable high quality VoIP sessions to be delivered.

It should be noted that whilst Skype is occasionally referenced as an example product in this paper it does not conform to the protocols defined in the application and transport layers in the above architectural model. Skype uses a proprietary and unpublished protocol set.

VOIP SECURITY THREATS & EXPLOITABLE VULNERABILITIES

VoIP Security Threats

The security threat environment for VoIP is very different to the traditional PSTN security environment (Walsh & Kuhn 2005, Internet Security Systems 2004). PSTN security threats can be summarised as phone tapping (eavesdropping) and cutting the line (denial of service) and are overt actions. To phone tap a PSTN connection without user detection requires both specific equipment and physical access to the line and/or telephone exchange. Similarly, achieving denial of service, cutting a PSTN line, is an overt physical activity. Conversely, a VoIP phone call can be subject to a number of covert security threats (Singhai & Sahoo 2006, Hung & Martin 2006) which can be instigated locally or over the Internet. The set of VoIP security threats can be summarised as:

- ***Eavesdropping:*** Through the use of packet sniffing (more formally known as network traffic analysis), eavesdropping upon a VoIP call is a relatively easy and well documented activity. Tools like Ethereal/Wireshark enable the capture of an IP packet stream (a VoIP conversation). Using software like the publicly available tool 'Voice Over Misconfigured Internet Telephones' (Vomit) allow the captured stream of IP packets to be translated into ".wav" format and replayed through a Media Player.
- ***Call Monitoring:*** Like eavesdropping, packet sniffing technology can be used to capture the calling habits and contacts of a VoIP user.
- ***Denial of Service (DoS):*** Attacking the servers of an Internet service provider (or any node on the Internet) is a common practice to deny service. For example, DoS can be achieved by flooding Internet nodes with packets to either significantly reduce node performance and/or cause a node to fail (crash). As VoIP is a time critical application flooding a node with malformed VoIP packets to reduce the nodes performance and make the conversation incomprehensible is one method of implementing a VoIP DoS attack.
- ***Malicious Software:*** Malicious software may be introduced into PCs running Softphones causing the PC to fail. To enable quality VoIP connections to be achieved a Softphone can require changes to the broadband router and/or firewall configurations (e.g. allowing tunnels, through opening firewall ports to a PC) that expose the PC to viruses, spyware, operating system vulnerabilities and other malicious software.

- ***Cyber Attacks:*** Like malicious software, cyber attacks can exploit exposed PCs (running Softphones) and ATAs due to router/firewall tunnels/open ports on broadband routers and/or firewalls required to achieve QoS.
- ***Toll Fraud:*** Breaking into ('hacking into') a VoIP network can allow an attacker to initiate outbound calls. Whilst this threat is more likely to be exploited in corporate networks than residential, such attacks can occur.

The identified VoIP threats arise due to opportunities that exist to exploit vulnerabilities (Tucker 2004) in the deployed VoIP technology and/or the poor configuration of the VoIP technology and its underlying network infrastructure. Some of the key vulnerabilities in the VoIP implementation model considered in this paper include:

Lack of Signal Confidentiality: As shown in Figure 3, when a call is initiated SIP communicates with proxy and location servers (hosted by the VoIP service provider) to establish the call. SIP sends and receives the call parameters in plain text. It is therefore possible to perform network traffic analysis to monitor calls and enable eavesdropping to be performed.

Lack of Dialogue Confidentiality: Once a call is established, RTP carries the conversation as a peer to peer like transmission (see Figure 3, event 5). RTP does not encrypt its payload therefore it is possible to eavesdrop on the conversation.

End Point to End Point Security: While both SIP and RTP in their standard mode of operation do not provide encryption capabilities both protocols can be configured, using certain Internet encryption standards, to preserve the confidentiality of both VoIP signal and dialogue. However, for an encrypted VoIP stream to occur it requires both UA's to have the UAs encryption capabilities enabled and for the VoIP service provider to support encryption. In practice, the use of encryption is more often only possible between identical Softphones. Softphone to PSTN phone encryption is obviously not possible, and as can be seen below, Softphone/ATA to ATA encryption via a commercial VoIP service provider is unlikely to be available (Enquiry about Engin VoIP Security 2007).

Lack of an Encryption Service from VoIP Service Providers: Information provided by Engin (Enquiry about Engin VoIP Security 2007), currently Australia's largest VOIP service provider, demonstrates that commercial VoIP service providers are not encrypting data from UAs. Engin provided advice on how encryption could be performed for both Softphone and Sipura/Linksys ATA (Sipura Technology Inc 2004), but neither encryption capability was supported by Engin.

End Point Security – Lack of Boundary Security: Many residences connect to the Internet using a basic broadband modem. Lack of any boundary security, (i.e. no firewall or router with firewall capabilities) results in the residence UA being vulnerable to network attacks.

End Point Security – Network Address Translation: As shown in Figures 2 & 3, SIP requires a Proxy Server to enable a call to be established. The SIP proxy server also requires the actual IP address of the UA. If the UA is positioned behind a router or firewall that supports Network Address Translation (NAT) then problems initiating a call occur. NAT is the technique of allowing one IP address to masquerade as a number of addresses behind a router/firewall; each device behind the firewall has its own address that is not visible to the Internet, i.e. the UA's IP address cannot be seen by the proxy server. For SIP to enable the proxy server to communicate with the UA, port forwarding rules (also known as tunnelling) need to be established for the router/firewall if the router/firewall does not support SIP traversal functionality. Port forwarding creates a tunnel that allows network traffic to pass straight through a router/firewall to a specific IP address, i.e. a hole is created in the firewall making it vulnerable to attack. Attackers may be able to exploit the hole/tunnel in the router/firewall to access the residence network, particularly as some VoIP service providers will not specify the source IP address (Engin Website 2007) of proxy servers resulting in the firewall port forwarding rule having to be set to allow forwarding from "ALL" external IP addresses.

Lack of PC Protection: A Softphone should be used on a PC with both an up-to-date operating system, (i.e. the latest vendor patches and updates have been installed) and up-to-date anti-virus and anti-spyware software otherwise it will be vulnerable to malicious software.

Possible Flaws in Security Enhanced Protocols: Ongoing work to improve the capabilities of protocols like SIP and RTP have resulted in security enhanced versions becoming available. When security is added as an afterthought, (i.e. security was not an initial key design criterion and therefore is unlikely to be pervasive throughout the implementation) exploitable vulnerabilities can be introduced.

How Significant are the Vulnerabilities?

The identified set of vulnerabilities will present a different risk level depending upon the environment in which VoIP is used. If the opportunity to exploit a vulnerability is high then the threat may be realised. The impact from threat realisation in the residential environment may be low compared with the commercial environment, e.g. the impact of eavesdropping upon a social conversation is likely to be considerably lower than the eavesdropping upon a commercial in confidence conversation. In table 2 the opportunities to exploit a vulnerability are considered together with impact of the threat being realised in a typical home environment.

Table 2 – Assessment of Threat Impact & Opportunity to Exploit Vulnerabilities

Threat	Likely Impact	Vulnerability	Opportunity to Exploit Vulnerability
Eavesdropping & Call Monitoring	Whilst most residential users would object to conversations being listened to, the worst case impact is probably a breach of privacy.	Lack of signal and dialogue encryption. Service provider or receiving UA does not support encryption. Allows packet sniffing to be performed.	In the residential environment, assuming access to the home LAN is restricted to trusted users, the only point where packet sniffing can occur is at the VoIP or Internet service provider.
Denial of Service	Loss of telephony.	Lack of boundary security.	DoS attacks are more likely to be made against a VoIP service provider – which will have defensive infrastructure. However, 'Internet vandals' will perform DoS attacks against residential VoIP connections.
Malicious Software	Infection of residence PC(s) with viruses and/or spyware	Lack of PC protection.	All PCs accessing the Internet are potential targets for spyware and viruses.
Cyber Attacks	Gain access to residence PCs resulting in breach of privacy.	Lack of boundary Security Network Address Translation	Monitoring services like DShield show how prolific and real network intrusion has become. The opportunities to exploit vulnerabilities are greater than users realise.
Toll Fraud	Use of VoIP service to make expensive overseas phone calls		

AVAILABLE VOIP SECURITY MECHANISMS FOR THE VOIP IMPLEMENTATION MODEL

Security & Quality of Service

The main inhibitor to the implementation of strong VoIP security mechanisms is the impact upon QoS. The following two examples show how some security mechanisms impact QoS:

- ***Encryption:*** Delays caused by the time to encrypt and decrypt VoIP packets can cause poor latency and jitter resulting in a poor or incomprehensible dialogue.
- ***End Point Security:*** Whilst firewalls provide essential end point boundary security they can also introduce delays in the delivery of packets. For data transfers the delay is either not noticeable by the end user or does not impact upon the activity being performed. However, for VoIP any delay to a VoIP packet stream can result in incomprehensible dialogue due to poor latency and also possibly jitter and packet loss.

It becomes a trade-off between the application of security mechanisms and the quality of VoIP service that can be delivered. A number of security mechanisms are available to improve the security of VoIP. The mechanisms considered below are the security mechanisms most suitable for the two VoIP implementation models presented in this paper. An overview of each mechanism is given to provide an appropriate reference point in the security framework; if the mechanism is used as a risk counter measure.

Use of Available Security Mechanisms to Preserve Signal & Dialogue Confidentiality

A number of encryption mechanisms can be applied at the Application (SIP), Transport (RTP and UDP) and Network (IP) layers to provide protocol security. Some of the mechanisms currently used within each of the three aforementioned layers of the Internet Reference Model are considered below:

Application Layer – SIP:

- *SIP Digest Authentication:* SIP does allow a challenge-response mechanism to be enabled. To ensure the password is never sent in the clear instead a digest/checksum of the username and password is used for the response. This mechanism is considered too weak (Kuhn et al 2005) and obviously cannot be used with ATA's as an interactive challenge-response is not possible.
- *Use of Secure Multi-purpose Internet Mail Extensions (S/MIME):* SIP messages have MIME bodies and S/MIME (Tucker 2004) defines public key encryption and authentication mechanisms for MIME content. Therefore SIP can use the S/MIME encryption, public key distribution and authentication mechanisms to protect its signalling data. One disadvantage is that S/MIME requires Public Key Infrastructure (PKI); therefore a certificate must be generated and signed for use with a UA. Another problem is that S/MIME is best suited to TCP than UDP.
- *Use Transport Layer Security (TLS) to Protect SIP:* The standard Internet encryption protocol TLS can be used with SIP however TLS cannot be used with UDP.

Transport Layer - RTP:

- *Secure RTP (SRTP):* A secure version of RTP, SRTP, allows a VoIP RTP/UDP packet stream to be encrypted (Muncan 2006). SRTP uses Multimedia Internet KEYing (MIKEY) for key management. MIKEY supports either PKI or Diffie-Hellman for key generation (Orrblad 2005).

Network Layer – IP:

- *IPsec:* Network layer encryption can be applied to secure a VoIP packet stream, but due to the key management issues and problems integrating with higher level protocols its appears not to be used in commercially available UAs.

The Sipura/Linksys ATAs provide support for encrypting the VoIP payload using SRTP. However, the VoIP service provider Engin has not enabled the capability in its provisioned ATAs. KPhone also supports SRTP and provided valid key distribution has been performed an encrypted VoIP session can be established.

Skype: Skype is a Softphone and VoIP service. Skype will not work other types of Softphone, although calls can be made to, and received from the PSTN. All Skype sessions are encrypted but Skype uses a proprietary protocol set and will not disclose details on how it performs encryption (Garfinkel 2005). For an encrypted dialogue to be established both UAs obviously need to be Skype.

ZRTP: To simplify the key management required to use SRTP the ZRTP protocol (McLaughlin 2006) has recently been released. ZRTP does not require a PKI. SRTP requires both UAs to agree on what key to use for the SRTP encryption; however ZRTP negotiates the SRTP session key using opportunistic encryption.

Implement Boundary & PC Security

To protect a UA a firewall (the more sophisticated the better) can be established. Firewalls with features like stateful packet filtering to prevent VoIP DoS attacks, traffic shaping to support VoIP QoS and protocol support to traverse NAT are able to provide good protection and enable quality VoIP dialogues to be successfully established. Where Softphones are used the host PC should be 'hardened' with all the latest operating system patches and updates, and have up to date anti-virus and anti-spyware applications.

Resolving NAT Issues

As discussed above, SIP does not work well with NAT and therefore requires port forwarding rules to be established. However, port forwarding creates a hole in the firewall presenting a potentially exploitable opportunity for attackers. More recently functionality to resolve the SIP/NAT issue has become available that

avoids the use of port forwarding. One of two different mechanisms that can be used to avoid port forwarding are:

- ***Siproxd*** is a router/firewall capability that allows a UA to work behind the router/firewall that is implementing NAT. Siproxd allows UAs behind a firewall to register with the VoIP service provider proxy. When initiating a call Siproxd rewrites SIP message bodies to function within a NAT environment; or
- ***Simple Traversal of UDP over NATs (STUN)*** is a client server protocol that allows a UA to find out its public address and the Internet-side port associated by NAT, with a particular local port. To enable SIP messages to traverse firewalls running NAT, both the UA and the VoIP service provider's proxy server to need to support STUN.

Separate Voice and Data

Although not a security mechanism in itself, the separation of voice from data traffic can assist with QoS and NAT. Separating VoIP traffic onto its own network segment can improve performance particularly when coupled with traffic shaping and a firewall running Siproxd. Achieving good performance and high QoS can allow other security mechanisms like encryption to be successfully used.

VOIP SECURITY FRAMEWORK

Overview

Using the available VoIP security mechanisms outlined above a VoIP risk management framework is presented. The framework is a practical and pragmatic risk reduction strategy that can be applied within a home. It is important that any framework can be applied within the constraints imposed by a service provider, therefore only available security mechanisms are utilised as countermeasures to threats.

Risk is generally defined as a factor of the threat and the likelihood of the threat being realised. Where the threat is technology related, the threat is most likely to be realised by the exploitation of a vulnerability. Therefore, if the likelihood of a vulnerability being exploited is minimised or mitigated the resultant risk is reduced.

The VoIP Security Framework has two risk reduction categories; ***Secure Infrastructure*** and ***Preserving Confidentiality***. The goal of each category is to:

- ***Secure Infrastructure:*** Reduce the risk of DoS attacks, network intrusion (cyber attacks), toll fraud and problems caused by malicious software.
- ***Preserving Confidentiality:*** To reduce the risk of call monitoring and eavesdropping.

A user can decide whether to implement one or both categories. Implementing the Secure Infrastructure category should be performed (irrespective of how concerned a user is about VoIP security) because the set of activities within the category secure the user's infrastructure and limit the opportunities for many types of Internet attack. The Preserving Confidentiality category involves utilising security mechanisms to encrypt the stream of call signalling and dialogue packets. For a user wishing to implement the Preserving Confidentiality category the following factors need to be considered:

- Is encryption supported by the VoIP service provider?
- Does the receiving UA support encryption?
- Can the transmitting UA support encryption?

Within each category of the framework a set of activities is defined; application of the activities will provide a level of security to mitigate and/or mange the security risks associated with residential VoIP usage.

Securing Infrastructure Category

While most of the methods outlined in this category are reasonably generic to any good security model, specific means for protecting networks running VoIP services are provided (Table 3).

Preserving Confidentiality Category

The fastest growing segment in the Australian VoIP market (Engin Website2007) is the use of ATA in homes. However, as has been outlined above Engin, the largest Australian VoIP service provider does not enable ATAs

to perform encryption. Unless a user can identify a VoIP service provider who supports encryption for ATAs, the only option to keep a VoIP conversation private is through a Softphone to Softphone connection. Selecting one of the following will enable an encrypted and private conversation to be performed (Table 4)

Table 3: Suggestions for securing infrastructure for VoIP activities and their associated threats

Activity	Description	Threats Addressed
Activity 1: Install Strong Firewall	A good quality firewall should be installed that supports the following features: • Stateful packet filtering • Network address translation • Traffic shaping to prioritise VoIP traffic • Siproxd.	DoS, Cyber Attacks, Toll Fraud.
Activity 2: Position UA on Separate Network Segment	If the UA is an ATA it should be positioned on a separate network segment, i.e. configure a separate network port on the firewall with its own subnet addressing range and attach the ATA. Where possible a Softphone should be placed on a separate network segment, however it is recognised that in practice a Softphone will reside on a multi-purpose PC and therefore separating voice and data traffic may not be possible.	Cyber attack, Toll Fraud
Activity 3: Harden Softphone PC	Where a Softphone is used as a UA the PC on which Softphone resides should be hardened by: • Application of all patches and updates (as soon as they become available) for the PC operating system, Internet applications and Softphone. • Use of up to date anti-virus and anti-spyware software.	Malicious Software
Activity 4: Accommodate SIP with NAT	NAT provides an important security feature, i.e. the ability to masquerade PCs and devices behind a firewall. However, because SIP has problems traversing a firewall running NAT either of the following mechanisms will allow SIP to be accommodated within a NAT environment: • Use a firewall with Siproxd; or • Use UAs that support STUN.	Cyber attacks, Toll Fraud

Table 4: Suggestions preserving confidentiality and integrity for VoIP activities and their associated threats

Activity	Description	Threats Addressed
Activity 1: Use Encrypting Softphone	Select a Softphone (probably a Softphone recommended by the VoIP service provider) that uses SRTP to secure a VoIP conversation, e.g. KPhone utilises SRTP. The VoIP service provider will need to support SRTP and therefore have the infrastructure to issue public/private keys and certificates. Both parties will need to have been issued with keys/certificates before the an encrypted dialogue can occur.	Lack of Signal Confidentiality, Lack of Dialogue Confidentiality, End Point to End Point Security.
Activity 2: Use ZRTP with Softphone	ZRTP allows a Softphone to perform encryption without the need for a PKI (Muncan 2006). Both parties need to use a SRTP Softphone with ZRTP integrated. ZRTP will work with different types of Softphone, i.e. a caller can be using a different type of Softphone to the receiver.	Lack of Encryption Service from VoIP Service Provider
Activity 3: Use Skype	Skype is the largest VoIP service provider in the world. A Skype to Skype conversation is encrypted.	Lack of Encryption Service from VoIP Service Provider

CONCLUSION

The increasing use of VoIP technology in Australian homes, as well as elsewhere in the world, creates a potential new category of attack vectors. Decreased call costs, the ability to make calls from different places mean that VoIP is not likely going away any time soon. Additionally, most ISPs in Australia are now bundling VoIP services with home broadband plans. The security implications and risk of using such technology must be considered by all users, not just corporations, as the risks will apply equally to all. This paper has proposed a framework that can be implemented by an IT literate user to mitigate and manage the VoIP security risks. It has been shown that some relatively simple measures can be applied that will reduce the security risks when using VoIP in a residential environment.

Future work would look at expanding this framework and using it to create guidelines that could be followed by users who lack the technological background to implement adequate security. Also of use would be to examine how well current users understand the risks of using VoIP, and also at ways in which general SOHO users can be made more aware of the risks.

REFERENCES

Engin Website, URL http://www.engin.com.au, Accessed 13 September 2007.

Enquiry about Engin VoIP Security (2007), Question & Answer email to Engin Support Desk, October 2007.

Kuhn, D.R. Walsh, T.J. & Fries,S. (January 2005). Security Considerations for Voice Over IP Systems: Recommendations of the National Institute of Standards and Technology. Department of Commerce, National Institute of Standards and Technology.

Garfinkel, S. L. (2005) VoIP and Skype Security, Computer Science & Artificial Intelligence Laboratory, MIT.

Hung, P. C. K. and M. V. Martin (2006) Through the Looking Glass: Security Issues in VOIP Applications, University of Ontario Institute of Technology.

Internet Security Systems Inc. (2004) VoIP: The Evolving Solution and the Evolving Threat.

Kphone Website, URL http://sourceforge.net/projects/kphone, Accessed 12 September 2007.

McLaughlin, L. (2006). Philip Zimmermann on What's Next after PGP. IEEE Security & Privacy**:** 10 - 13.

Muncan, M. (2006) Secure telephony: SIP/SRTP (PKI) vs. Zfone vs. Skype, Universitat Konstanz.

Orrblad, J. (2005). Alternatives to MIKEY/SRTP to secure VoIP. Telecommunication Systems Laboratory (TSLab), Department of Microelectronics and Information Technology (IMIT). Stockholm/Kista, Royal Institute of Technology (KTH)

Simon, M. and J. Slay (2006) Voice over IP: Forensic Computing Implications, Enterprise Security Management Lab, University of South Australia.

Singhai, R. and A. Sahoo (2006) VoIP Security, School of Information Technology, Indian Institute of Technology, Bombay.

Sipura Technology Inc. (2004) Implementing Residential Voice over Broadband Services with the Sipura Phone Adaptor (SPA).

Skype (2007) About Skype URL http://www.skype.com, Accessed on 23 September 2007.

Tucker, G. S. (2004) Voice Over Internet Protocol (VoIP) and Security, SANS Institute.

Walsh, T. J. and D. R. Kuhn (2005). "Challenges in Securing Voice over IP." IEEE Security & Privacy: 44-49.

Wikipedia (2007) Comparison of VoIP software. URL, http://en.wikipedia.org/wiki/Comparison_of_VoIP_software Accessed 1 October 2007

COPYRIGHT

Peter James and Andrew Woodward ©2007. The author/s assign SCISSEC & Edith Cowan University a non-exclusive license to use this document for personal use provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive license to SCISSEC & ECU to publish this document in full in the Conference Proceedings. Such documents may be published on the World Wide Web,

CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.