**SUBJECTIVE AUDIO QUALITY OVER A SECURE IEEE 802.11N DRAFT 2.0
WIRELESS LOCAL AREA NETWORK**

THESIS

Benjamin W. Ramsey, Captain, USAF

AFIT/GE/ENG/09-34

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GE/ENG/09-34

SUBJECTIVE AUDIO QUALITY OVER A SECURE IEEE 802.11N DRAFT 2.0 WIRELESS LOCAL AREA NETWORK

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Benjamin W. Ramsey, BSEE

Captain, USAF

March 2009

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT/GE/ENG/09-34

SUBJECTIVE AUDIO QUALITY OVER A SECURE IEEE 802.11N DRAFT 2.0 WIRELESS LOCAL AREA NETWORK

Benjamin W. Ramsey, BSEE

Captain, USAF

Approved:

Barry Mullins

Dr. Barry E. Mullins
Committee Chairman

12 Feb09
Date

Todd Andel, Maj, USAF
Committee Member

12FEB09
Date

Ryan Thomas, Capt, USAF
Committee Chairman

20FEB09
Date

AFIT/GE/ENG/09-34

Abstract

This thesis investigates the quality of audio generated by a G.711 codec and transmission over an IEEE 802.11n draft 2.0 wireless local area network (WLAN). Decline in audio quality due to additional calls or by securing the WLAN with transport mode Internet Protocol Security (IPsec) is quantified. Audio quality over an IEEE 802.11n draft 2.0 WLAN is also compared to that of IEEE 802.11b and IEEE 802.11g WLANs under the same conditions.

Audio quality is evaluated by following International Telecommunication Union Telecommunication Standardization Sector (ITU-T) Recommendation P.800, where human subjects rate audio clips recorded during various WLAN configurations. The Mean Opinion Score (MOS) is calculated as the average audio quality score given for each WLAN configuration. An 85% confidence interval is calculated for each MOS.

Results suggest that audio quality over an IEEE 802.11n draft 2.0 WLAN is not higher than over an IEEE 802.11b WLAN when up to 10 simultaneous G.711 calls occur. A linear regression of the subjective scores also suggest that an IEEE 802.11n draft 2.0 WLAN can sustain an MOS greater than 3.0 (fair quality) for up to 75 simultaneous G.711 calls secured with WPA2, or up to 40 calls secured with both WPA2 and transport mode IPsec. The data strongly suggest that toll quality audio (MOS $\geq$ 4.0) is not currently practical over IEEE 802.11 WLANs secured with WPA2, even with the G.711 codec.

To my boys

AFIT/GE/ENG/09-34

Acknowledgments

Thank you, Dr. Mullins, Barry E.
for the support you have given me.
With equipment galore
and your open door
I've had a blast earning this degree.

Captain Thomas and Major Andel,
your laptops have been awfully swell.
You've answered my questions
and given suggestions
to truly help this thesis go well.

Friends Augustine H. and Luis O,
through thick and thin you never laid low.
Real-time simulations,
stress, and complications
were all fun watching monitors glow.

Table of Contents

Page

List of Figures

Figure Page

List of Tables

SUBJECTIVE AUDIO QUALITY OVER A SECURE IEEE 802.11N DRAFT 2.0 WIRELESS LOCAL AREA NETWORK

I. Introduction

1.1 Motivation

Voice over Internet Protocol (VoIP) and IEEE 802.11 wireless local area networks (WLANs) are two technologies that have experienced rapid growth over the past decade. Many companies now offer VoIP service to homes and businesses for significantly less cost than traditional analog phone service. VoIP can be less expensive because it shares the same infrastructure as IP data networks. VoIP architectures can also scale more easily and offer greater flexibility to customers [BLG07]. Wireless network protocols such as IEEE 802.11 (hereafter referred to as 802.11) provide users with mobile connectivity without the need for expensive and inflexible wiring. The 802.11n extension, for instance, is currently undergoing the final stages of development and will provide bandwidth up to 600 Mbps [Bro06], which could make a complete transition to wireless networking feasible in many situations. The combination of these two technologies, however, creates serious challenges to both audio quality and security.

The most common protocols used for VoIP do not provide encryption or authentication. Many VoIP architectures use Session Initiation Protocol (SIP), Real time Transport Protocol (RTP), and User Datagram Protocol (UDP), all of which are vulnerable to interception and manipulation. Potential solutions to securing VoIP include substituting RTP with Secure RTP (SRTP) or encrypting all IP traffic with Internet Protocol Security (IPsec). WLANs must also be encrypted to prevent unauthorized access to the network and interception of traffic by malicious or curious individuals far beyond the building walls.

The wireless medium introduces considerable challenges to VoIP audio quality. Wireless 802.11 physical and data-link layers are more vulnerable to packet loss, interference, and lower peak transmission rates than wired networks [GaK03]. This is primarily due to the shared wireless medium, where collisions are unavoidable. VoIP audio quality declines when packets are dropped, experience excessive latency (>150 ms), or jitter [CFK06]. All of these factors are potential issues for WLAN traffic. Encryption techniques such as WPA, WPA2, and IPsec increase bit overhead and computation times which can further reduce VoIP audio quality over the WLAN.

When the 802.11n extension is finalized, it is likely to be adopted very rapidly as a viable alternative to completely wired networks due to its high potential throughput. Hardware is already available based on draft 2.0, which is expected to be very similar to the ratified standard [Bro06]. However, before the DoD considers 802.11n WLANs for real-time applications, such as VoIP, the effect that properly securing the WLAN has on audio quality must be well understood.

1.2 Research Goals

The overall goal of this research is to characterize VoIP audio quality over a secure 802.11n draft 2.0 WLAN. This research examines whether or not audio quality is higher over a secure 802.11n draft 2.0 WLAN than over secure 802.11b or 802.11g WLANs when up to 10 calls occur simultaneously. The research also aims to quantify audio quality decline due to additional simultaneous calls and the use of Internet Protocol Security (IPsec). To accomplish these goals, subjective audio scores are collected and used to construct a predictive mathematical model of audio quality over an 802.11n draft 2.0 WLAN.

1.3 Thesis Layout

This chapter introduces the motivation behind the thesis research. Chapter II provides background information on the research topic and also summarizes the results of recent related research. Chapter III details the methodology used to perform the experiments. Chapter IV presents an analysis of the experimental results. Chapter V discusses conclusions from the results and offers suggestions for future research.

II. Background and Literature Review

2.1 Introduction

This chapter provides an introduction to VoIP, its use over secure wireless networks, and previous research into how security protocols affect voice quality over 802.11 WLANs. Section 2.2 provides an overview of VoIP over WLANs. Section 2.3 provides an introduction to VoIP signaling protocols. Section 2.4 describes how the Session Initiation Protocol (SIP) is used to initiate VoIP calls. Section 2.5 describes how voice data is transported via the Real-time Transport Protocol (RTP). Section 2.6 describes different voice codecs and how they are used in a VoIP system. Section 2.7 describes the factors that negatively impact voice quality for VoIP calls. Section 2.8 illustrates methods for obtaining Mean Opinion Score (MOS) and how MOS measures the audio quality of a VoIP system. Section 2.9 describes several 802.11 protocol extensions. Section 2.10 outlines the data link layer security protocols available for 802.11 networks, and Section 2.11 describes additional WLAN measures available to secure VoIP calls. Section 2.12 describes related research efforts into the impact of security on audio quality over 802.11 WLANs.

2.2 VoIP over WLANs Overview

The transition from analog to digital technologies for voice transmission has seen steady progress since the mid 1990s. Many companies now offer VoIP services to homes and businesses for significantly less cost than traditional analog services. VoIP service can be less expensive because it shares the same infrastructure as IP data networks. Instead of separate installations of phone and data lines, only IP network connectivity is required. VoIP architectures can also scale more easily, provide better centralized control, and can integrate with new computer applications [Wal05].

There are difficulties in employing VoIP as a total telephony solution. Voice calls made over the public Internet are not guaranteed a particular quality of service (QoS). However, QoS measures can be taken, particularly on the local network, to help ensure voice calls are of high quality and with low delay. Security is also a major issue with VoIP implementations [GuS07]. Because VoIP shares the IP data network infrastructure and technologies, it also inherits and contributes to the same well-known vulnerabilities. Properly securing a VoIP system is essential, but can be a time-consuming task for large networks. Possible attacks on a VoIP system include call interception, audio injection, data theft, spoofed caller ID, denial-of-service (DoS), and spam over Internet telephony ("spit").

The increasing popularity of 802.11 WLANs is due to their flexibility and low cost of installation. As VoIP rapidly replaces analog phone systems, voice traffic will be delivered more frequently over WLANs. However, WLANs introduce significant challenges to VoIP call quality, particularly under limited-bandwidth or high-utilization conditions. WLANs must also be carefully secured from unauthorized access.

2.3 VoIP Signaling Protocols

Before VoIP end devices exchange real-time traffic they must first use signaling protocols to establish a multimedia session. There are two primary architectures for establishing these sessions: H.323 and Session Initiation Protocol (SIP). When the first draft of H.323 was published by the International Telecommunication Union (ITU) in 1996, it was adopted as an effective means of establishing multimedia services on local area networks. The H.323 architecture was developed with the same philosophy as the Public Switched Telephone Network (PSTN), where calls typically occur between two distant ends.

SIP, on the other hand, was first published by the Internet Engineering Task Force (IETF) in 1999 and was developed from an Internet perspective where communication between users is less structured. The protocol messages in the SIP architecture are sent in plain text, as opposed to the binary formats used with H.323. This makes it much easier to develop multimedia applications using SIP. The most recent version of SIP is described in RFC 3261 [SIP02]. In addition to H.323 and SIP, there are other VoIP signaling protocols including Cisco's proprietary Skinny Call Control Protocol (SCCP), Media Gateway Control Protocol (MGCP), and Megaco/H.248 [NIS05].

While SIP has grown in popularity over the last few years, no one signaling protocol is used in the majority of VoIP systems. Regardless of the signaling protocol used, Real-time Transport Protocol (RTP) or one of its secure variations is used to deliver the real-time media packets.

2.4 SIP Architecture

SIP is a text-based application-layer protocol that is responsible for the establishment, modification, and termination of multimedia sessions [SIP02]. SIP is used in conjunction with other IETF protocols to implement a complete multimedia solution. Furthermore, the SIP architecture is transport-independent and can be used with either TCP or UDP over an IP network. Figure 1 illustrates the different protocols used in the SIP architecture, and these protocols are discussed further in the following sections.

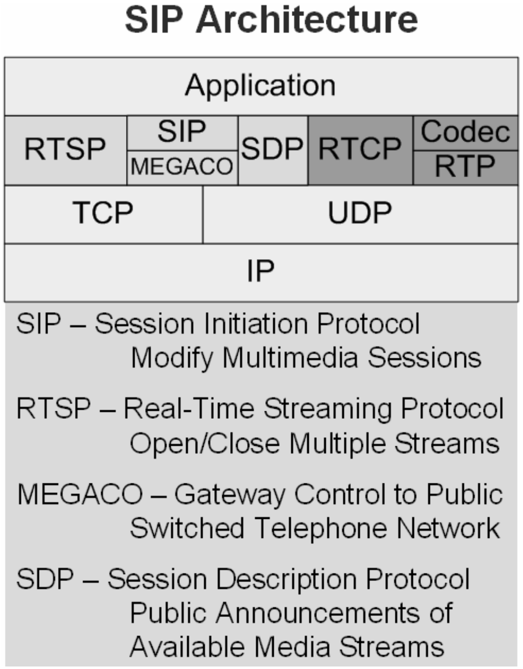

SIP Architecture
Application
RTSP
SIP
MEGACO
SDP
RTCP
Codec
RTP
TCP
UDP
IP
SIP – Session Initiation Protocol
Modify Multimedia Sessions
RTSP – Real-Time Streaming Protocol
Open/Close Multiple Streams
MEGACO – Gateway Control to Public
Switched Telephone Network
SDP – Session Description Protocol
Public Announcements of
Available Media Streams

Figure 1. SIP Architecture Protocol Stack [Sey07]

End-system applications in a SIP implementation are known as User Agents (UA), and the hardware that runs a UA is called a SIP terminal. When requests arrive, the UA receives them and returns responses on behalf of the user.

Call setup under the SIP architecture is relatively simple. Sessions can be arranged through proxies or between UAs directly. Figure 2 illustrates how a typical call takes place between two UAs (User A and User B) that register with a proxy. In this example the UAs both send a REGISTER request to the proxy with the “from” and “to” address fields set to their own SIP address. A SIP address is similar to an email address, and it takes the form *sip:username@domain.com*. User A initiates a session with User B by sending an INVITE message to the proxy. This message includes the username of the entity that User A wishes to invite and indicates the media type to be used. The proxy replies to User A with a TRYING 100 message and forwards the request to User B. Upon

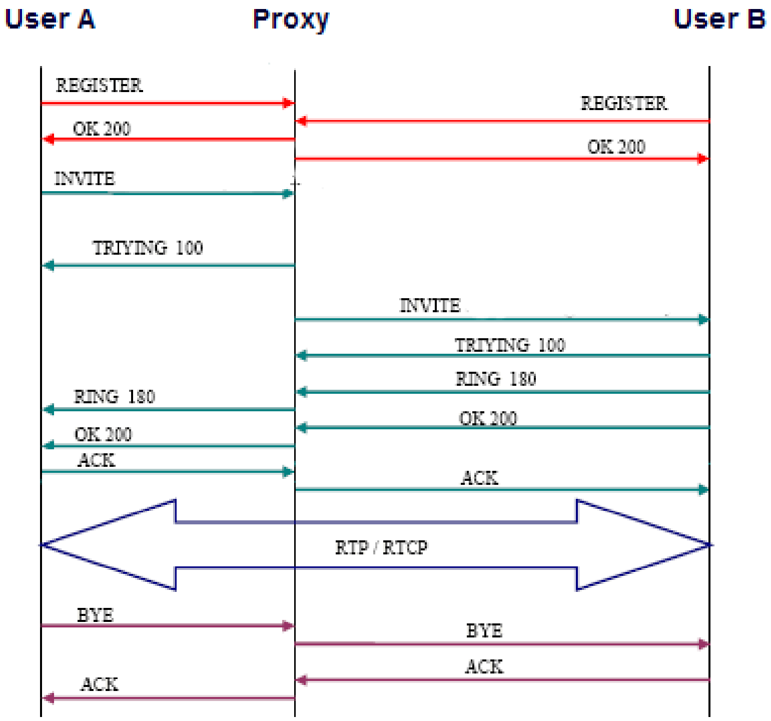


Figure 2. Example of SIP Call Flow [VoI06]

receiving the INVITE message, User B sends a TRYING 100 message to the proxy and then a RING 180 message when the phone begins to ring. When User B picks up the phone or otherwise accepts the session an OK 200 message is sent. When User A receives the OK 200 from User B, it replies with an acknowledgement. The call is then established, and the RTP session begins. Once complete, one of the UAs terminates the call with a SIP BYE message and the other replies with an acknowledgement.

2.5 Real-Time Transport Protocol

Regardless of the signaling protocol, RTP or one of its secure variants is used to stream real-time data. RTP is described in RFC 3550, which has been in its current form since July 2003 [RTP03]. RTP does not operate over a default port, but traditionally uses an even-numbered UDP port. There are no QoS guarantees provided by RTP, and packets can be delivered out of sequence. Proper ordering and parallel processing of delivered packets is possible because each packet includes a sequence number and time stamp. The packet size and transmission frequency of RTP packets is dependent on the audio or video codec used.

Figure 3 illustrates the packet structure of RTP for audio transmissions. The first 20 bytes of the RTP packet is the IPv4 header information, including source and destination IP address. The 8 byte UDP header follows, including source and destination ports and a checksum. Next is the 12 byte RTP header which includes media payload type, sequence number, timestamp, synchronization source identifier, and contributing source identifiers. The audio data consists of the compressed voice signals and can range from 5 to 160 bytes in length.

The Real-time Control Protocol (RTCP) is also outlined by RFC 3550. Its traffic takes place on one port higher than the RTP stream and provides feedback about quality of service by collecting statistics about the data stream such as jitter, round trip time, and dropped packets. Implementation of RTCP is optional, but can be useful for monitoring VoIP performance.

As an example of how RTCP can be used to improve VoIP performance, research by Sfairopoulou, Macian, and Bellalta [SMB06] demonstrates that the impact of data rate changes on a multi-rate WLAN has on audio quality can be corrected using a

IPv4 Header (20 Bytes)	UDP Header (8 Bytes)	RTP Header (12 Bytes)	Audio Data (5-160 Bytes, codec-dependent)

Figure 3. RTP Packet Structure

combination of RTCP and MAC layer feedback. Their simulated VoIP system switches to a lower bit rate codec when the WLAN data rate declines, preventing dropped calls.

2.6 Audio Codecs

Before analog voice signals can be transported over an IP network, they must first be sampled and digitized. The method used to perform this compression and subsequent decompression of the analog data is called a codec. Many codecs have been devised for use in VoIP systems, and they vary greatly in their bandwidth requirements and call quality as shown in Table 1. For example, the G.711 codec operates at 64 kbps while the highly compressed G.723.1 protocol only uses 5.3 kbps.

Table 1. VoIP Audio Codec Data Rates and Typical MOS

Codec	Bit Rate	Sample Period	Frame Size (# Per Packet)	Typical MOS
G.711	64 kbps	20 ms	160 bytes (1)	4.1
G.729	8 kbps	10 ms	10 bytes (2)	3.92
G.723.1	6.3 kbps	30 ms	24 bytes (1)	3.9
G.729a	8 kbps	10 ms	10 bytes (2)	3.7
G.723.1	5.3 kbps	30 ms	24 bytes (1)	3.65

The G.711 codec is also known as Pulse Code Modulation (PCM). It was one of the earliest codecs and is still very popular because it offers the highest audio quality. It samples the audio input at 8 kHz and represents each sample with a 14 bit linear input code. This linear input code is then compressed to 8 bits by replacing all values within 256 ranges with the same number. From this it is straightforward to derive the payload bitrate: (8 kHz x 8 bits = 64 kbps). Packets are transmitted 50 times per second, each representing 20 ms of audio as 160 bytes.

Other codecs decrease the bandwidth requirements by compressing the audio data. One of the most compressed codecs is G.723.1. With a payload bitrate of 5.3 kbps, G.723.1 requires 88% less bandwidth than the G.711 codec. Although lower bandwidth requirements are desirable, the extensive compression used in codecs such as G.723.1 and G.729 comes at the expense of some audio quality.

The call quality provided by VoIP system can be measured as the mean opinion score (MOS). The MOS is a number from 1 (bad quality) to 5 (excellent) and can be derived from actual listeners or by an analytical model, such as the E-model [PMA+05]. An MOS of 4.5 is the highest probable score to result from a large subject pool and is therefore the maximum score predicated by the E-model. As shown in Table 1 the typical MOS of the highly compressed G.723.1 codec is 3.65, compared to a typical MOS of 4.1 for G.711.

2.7 Factors that Impact VoIP Audio Quality

Evaluating the audio quality of a VoIP system is non-trivial because IP networks are nondeterministic. IP networks offer a “best effort” level of service where packets can arrive at their destination after long delays, out of order, or not at all. This is in stark contrast to the circuit switching paradigm of the public telephone network, where end users are granted dedicated connections. All irregularity in VoIP packet delivery between source and destination can negatively impact call quality, while edge devices can also have a significant impact, depending on how they respond to irregular packet delivery [Bro06a].

2.7.1 Network Impact on VoIP Audio Quality

The network can impact VoIP call quality three general ways: packet loss, packet delay, and jitter. RTP relies on the connectionless UDP protocol for end-to-end transport; therefore dropped packets are a distinct possibility. VoIP packets can be lost due to network congestion, excessive latency, or faulty hardware at any point between users. Network congestion can be managed on the local network by keeping the utilization low. Although the audio codec may only require 5-64 kbps, 50-100% more bandwidth is necessary to prevent full network utilization and to ensure timely packet delivery. The National Institute of Standards and Technology (NIST) reports that packet loss in excess of 3% results in intolerable call quality [NIS05].

Packet delay also has a strong impact on VoIP performance. Long delays impede real-time communication because they prevent natural conversation. Significant delay can also introduce echo. In some cases, such as satellite communications, propagation delays in excess of 250 ms occur and are unavoidable. However, the ITU-T recommendation G.114 states that end-to-end delays should be kept to less than 150 ms to prevent performance degradation [ITU03].

Jitter is the variable delay between packet arrivals. Because IP networks are dynamic, a stream of packets can encounter different queuing delays and even follow different routes through the network. Jitter impacts the call quality by introducing variable latency into the conversation. De-jitter buffers are very common in VoIP end devices in order to remove this variability. However, de-jitter buffers also contribute to the end-to-end delay. For example, a 25 ms de-jitter buffer could remove the effect of jitter less than 25 ms, but would also introduce a fixed 25 ms delay into the system. If the de-jitter buffer is too small then excessively delayed packets will be dropped. NIST recommends that jitter be kept below 40 ms [NIS05].

2.7.2 Other Factors that Impact VoIP Call Quality

Voice traffic over 802.11 WLANs faces additional challenges. Compared to wired networks, WLANs experience more events where packet delivery is delayed by several hundred milliseconds [CFK06]. WLANs are also vulnerable to radio frequency interference (RFI). RFI can come from neighboring networks, poorly shielded electronics, or other devices operating in the same frequency range (2.4 or 5 GHz). It has been repeatedly demonstrated [CFK06, GaK03, Sey07] that the true simultaneous VoIP call capacity of an 802.11 WLAN is significantly less than simple bandwidth calculations would indicate.

The 802.11e standard is an amendment to the original 802.11 MAC layer that offers QoS enhancements for delay-sensitive applications. Data priority is specified by an Access Category (AC). The four ACs, in order from highest to lowest priority are Voice, Video, Best Effort, and Background [MLL+07]. Separate queues are maintained for the traffic in each AC so that lower-priority traffic always yields. In the event of a collision, the back off timer for high-priority traffic is also set to be lower than for lower ACs. WLAN hardware that supports 802.11e is widely available and is highly recommended for real-time applications [MLL+07].

2.8 Methods for Obtaining MOS

Speech quality is inherently subjective, as it is established by human perception. Naturally, the most straightforward test of audio quality is to poll a large number of users.

However, subjective tests are very time-consuming and cannot be conducted in real-time. Analytical methods for determining MOS have thus been developed that use VoIP traffic measurements such as packet loss, latency, and codec type to estimate MOS. The advantage of analytical methods is that they can be performed very quickly and in real-time. But unlike subjective tests, analytical models cannot consider conversational factors such as echo and background noise that further degrade audio quality.

2.8.1 Subjective Methods for Obtaining MOS

Subjective testing is considered the most reliable approach to assessing voice quality [DRE+07]. Several widely-used approaches to subjective audio assessment are defined in ITU-T Recommendation P.800 [ITU96]. The P.800 recommendation describes both conversation-opinion and listening-only methods. Conversation-opinion tests are designed to most accurately reproduce in the laboratory the same service conditions as experienced by VoIP users [ITU96]. To achieve this level of fidelity, the experimenter must accurately specify, measure, and setup each experiment to ensure that the test conditions match those of the actual system. The P.800 recommendation also describes listening-option tests. These tests measure audio quality as delivered from a source through the system to a human listener.

The most commonly used listening test described in the P.800 recommendation is called Absolute Category Rating (ACR). For this test, human subjects rate the quality of audio recordings over the VoIP system on a scale from one (bad quality) to five (excellent quality). Table 2 illustrates the relationship between listening quality and subjective score for the ACR test.

Table 2. ACR Quality Rating Scale Recommended by the ITU-T [ITU96]

Listening quality	**Score**
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

Another subjective test described in ITU-T recommendation P.800 is the Comparison Category Rating (CCR) method where test subjects are presented with a pair of speech samples during each trial. The ordering of the original and processed signals is random but ultimately balanced so that the original signal is presented first during half of the trials and second during the other half [ITU96]. Subjects provide judgment on the second audio sample as compared to the first quantify by how much with a number as shown in Table 3.

Table 3. CCR Quality Rating Scale Recommended by the ITU-T [ITU96]

Much Better	3
Better	2
Slightly Better	1
About the Same	0
Slightly Worse	-1
Worse	-2
Much Worse	-3

Test audio clips for both the ACR and CCR tests should consist of between 2 and 5 short, meaningful phrases that are played in random order. For example, “You will have to be very quiet.” and “I want a minute with the inspector.” are possibilities. Because subjective testing results in interaction between the experimenter and subjects, great care must be taken to not influence the outcome. No suggestion should be given to the subjects about technical details, anticipated audio sample quality, or the list of phrases they could hear.

2.8.2 Analytical Methods for Obtaining MOS

The E-model is a commonly cited computational model for determining MOS and is defined by ITU-T recommendation G.107. The E-model estimates MOS based on the principle that network impairments correspond to calculable decreases in perceived audio quality. The E-model’s evaluation of the call quality is given as an R value by

$$R = R_0 - I_s - I_d - I_e \tag{1}$$

where R_0 is the signal-to-noise ratio, I_s represents impairments that occur simultaneously with the voice signal, I_d is the impairment caused by delay, and I_e is the impairment caused by low bit rate codecs. The R-value is then mapped to approximate MOS by

$$MOS = 1 + 0.035R + 7(10^{-6}R)(R\text{-}60)(100\text{-}R). \tag{2}$$

For example, an R value of 40 equates to an approximate MOS of 2.06 [Baj03].

Another objective audio quality model is the Perceptual Evaluation of Speech Quality (PESQ), which estimates perceived one-way audio quality as an MOS value [GLC07]. PESQ is described in ITU-T Recommendation P.862 [ITU01] and estimates speech quality by comparing the original, unprocessed audio signal with the degraded version at the system output. First, a reference speech signal is sent through the system under test. The degraded signal is then recorded at the system output. Then the average power of the reference speech signal and the degraded signal are matched. Finally, software or hardware designed to implement the PESQ algorithm [ITU01] compares the two signals to estimate MOS at the system output.

PESQ was developed to approximate MOS without the need for time-consuming trials with human subjects. It is useful for detecting audio quality decline due to packet loss, codec selection, and real-time testing of prototype networks [Min02].

2.9 IEEE 802.11 Protocol Extensions

The IEEE created the first WLAN standard, 802.11, in 1997. This legacy standard supports a theoretical maximum bandwidth of 2 Mbps. Over time, 802.11 has been expanded to include faster and more capable extensions, including 802.11b (≤ 11 Mbps), 802.11a (≤ 54 Mbps), and 802.11g (≤ 54 Mbps) [Bro03]. The 802.11n (≤ 600 Mbps) [Bro06] extension is also under development, although equipment is already being sold based on the most recent draft (2.0). IEEE 802.11a operates in the 5 GHz band, while 802.11b and 802.11g devices operate in the 2.4 GHz band. The 802.11n can function in either frequency band, depending on the frequencies employed by other clients on the WLAN [Bro06].

2.10 Link Layer Security Measures for 802.11 WLANs

WLANs are inherently more vulnerable to attack than their wired counterparts. Without security measures, anyone within range of the WLAN can access and exploit network resources. Fortunately, several security measures are available.

2.10.1 Wired Equivalent Privacy

The first data link layer algorithm designed to encrypt data on the WLAN is Wired Equivalent Privacy (WEP). WEP was introduced in 1999 to provide confidentiality on par with wired networks using a secret key which can be 40, 104, or 128 bits long. WEP was quickly discovered to be cryptographically weak and current tools can crack its key in minutes or less [BHL06]. Even though WEP offers trivial network protection, it is still in widespread use, particularly in home networks. This can be attributed a combination of legacy hardware support, user ignorance, and indifference with respect to WLAN security.

As shown in Figure 4, WEP uses the RC4 cipher, where plain text data is XORed with a shared secret key. The algorithm begins when the plain-text 802.11 frame is queued for transmission. Then an integrity check value (ICV) is calculated for the header and payload using a 32 bit Cyclic Redundancy Check (CRC-32) and appended to the payload [Gas05]. Meanwhile, a 64 to 256 bit long "WEP seed" is created by appending the 40-232 bit secret key with a 24-bit initialization vector (IV). Then the payload/ICV combination and WEP seed are run through the RC4 algorithm. The RC4 output is sent as the encrypted payload and the IV, key number, and ICV are included in the frame as plain text to allow the recipient to decrypt the frame by following the process in reverse.

The major flaw in WEP is that the IV, integral to the encryption algorithm, has a relatively small number of possible values (2^{24}, or less than 17 million). Moreover, in most implementations of WEP, the secret key is static and the IV is necessarily transmitted in the clear. Therefore two frames that share the same IV are most likely also encrypted with the same WEP seed. Furthermore, the CRC algorithm is not

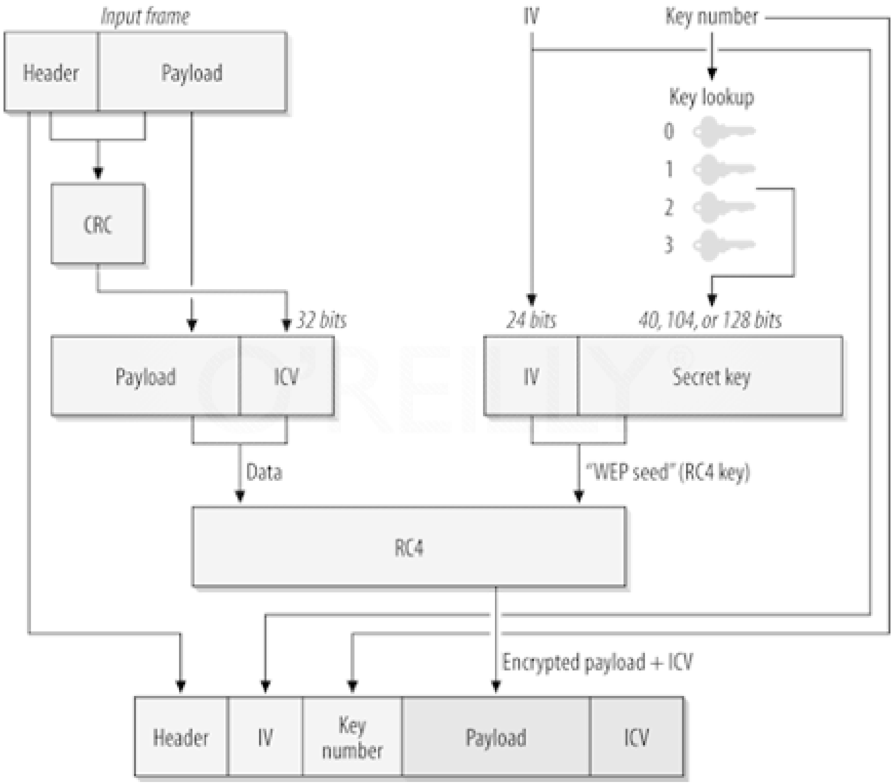


Figure 4. WEP Encryption Algorithm [Gas05]

cryptographically secure [Gas05]. Using these facts, a cracking program such as Aircrack can determine the secret key by examining only 1,000,000 frames [BHL06].

2.10.2 Wi-Fi Protected Access

In order to address the flaws discovered in WEP, the Wi-Fi alliance created the Wi-Fi Protected Access (WPA) protocol as a stop-gap measure while the IEEE 802.11i security standard was developed. WPA incorporates many of the features of the final IEEE 802.11i standard and was designed to be compatible with legacy hardware through firmware upgrades. WPA keys can be distributed using a key authority or set as a fixed Pre-shared Key (PSK). The IV is lengthened to 48 bits, increasing the number of possible values to over 281 trillion. Another major improvement over WEP is the Temporal Key

Integrity Protocol (TKIP), illustrated in Figure 5. Unlike WEP, where the secret key is applied directly, TKIP encrypts the frames with keys that are derived from the master key [Gas05]. The per-frame keys are calculated using a process called key mixing. Furthermore, Message Integrity Check (MIC) is used in place of the more vulnerable CRC and is computed by an algorithm called Michael. Sequence numbers are also added to the frames to protect against spoofing.

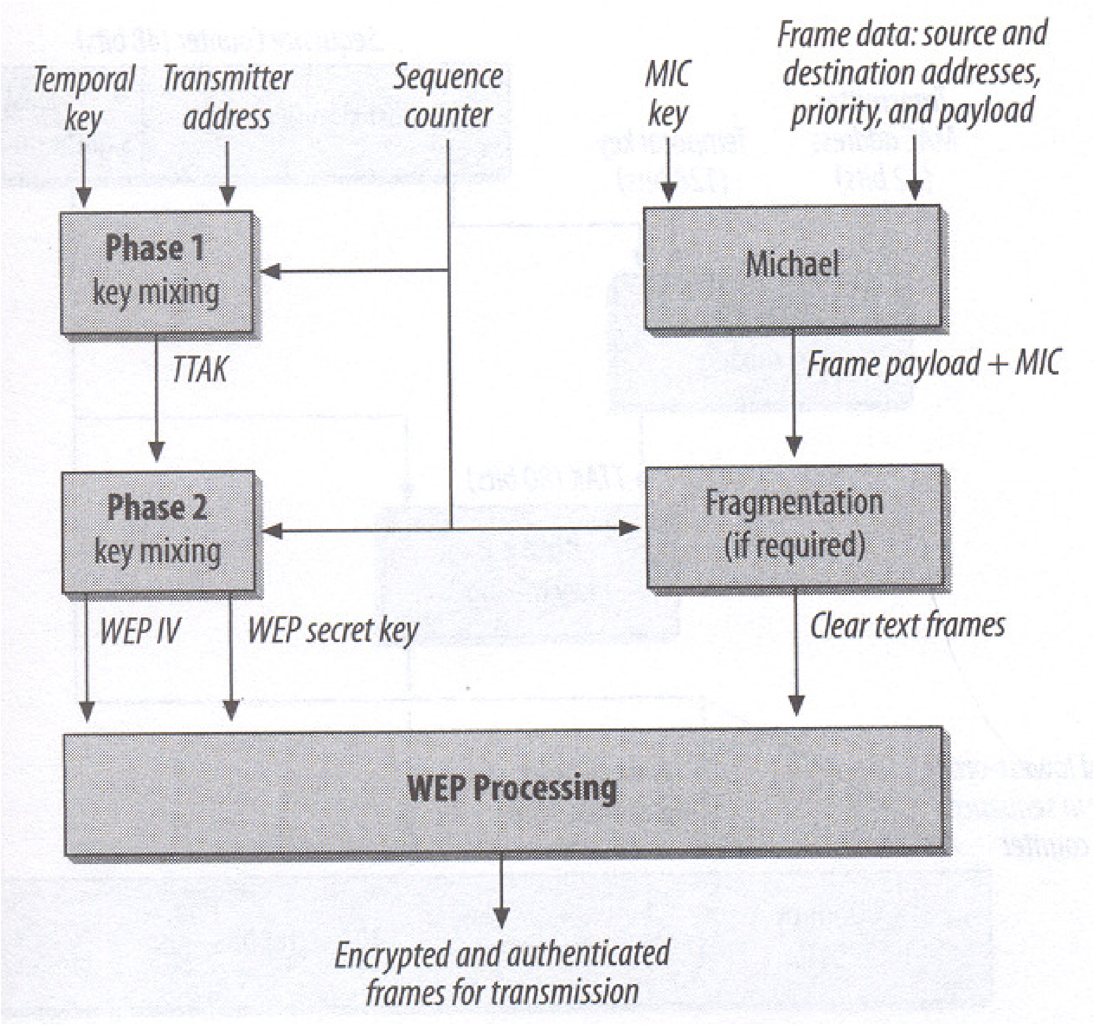


Figure 5. TKIP Encryption Algorithm [Gas05]

2.10.3 Wi-Fi Protected Access 2 (802.11i)

Full implementation of the IEEE 802.11i security standard is also known as WPA2. The WPA2 standard was ratified in 2004 and further increases the confidentiality, integrity, and availability of WLANs beyond that of WEP and WPA. Whereas WPA was meant to strengthen the WEP protocol, WPA2 mandates a completely different encryption standard, Counter Mode with CBC-MAC (CCMP). CCMP incorporates the Advanced Encryption Standard (AES) block cipher with 128 bit keys. WPA2 uses two types of keys - group keys for broadcast and multicast traffic, and pairwise keys for unicast traffic [Gas05]. As a block cipher, AES encrypts and decrypts messages in 128-bit blocks. If the message length is not evenly divisible by 128 bits, it is padded to meet this requirement and the padding is discarded during decryption.

Integrity with WPA2 is provided by XORing every 128-bit cyphertext block with the following 128-bit block until the entire message integrity code (MIC) is computed [SAN04]. Only the first 64 bits of the MIC are used to verify message integrity, but this is sufficient to ensure that the message has not been altered in transit [SAN04].

While WPA and WPA2 provide strong encryption for 802.11WLANs, the PSK implementation of both is more vulnerable than when keys are dynamic and distributed via server. If a weak PSK is chosen, it can be cracked by a dictionary or brute-force attack. A passphrase of at least 20 characters that includes letters, numbers, and special characters is considered secure. However, since PSKs are stored on end devices, they could potentially be retrieved from any authorized device on the WLAN.

2.11 VoIP Security Protocols

2.11.1 Internet Protocol Security

Internet Protocol Security (IPsec) provides authentication and confidentiality for IP networks. IPsec provides security for VoIP by encrypting the entire header and payload of every packet. The encrypted packets are then encapsulated by an unencrypted IP header and Encapsulating Security Payload (ESP) header [XiZ04]. IPsec can use Data Encryption Standard (DES), Triple DES (3DES), or the Advanced Encryption Standard (AES) to perform encryption. Although it provides very strong security, IPsec introduces additional headers and processing requirements and can increase the VoIP bandwidth requirements by 37% [XiZ04]. The additional headers have a significant impact on VoIP traffic because voice packets are only 50 to 200 bytes long. Figure 6 illustrates a packet protected with transport mode IPsec, while Figure 7 illustrates a packet protected with tunnel mode IPsec. In transport mode IPsec, the IP header is left unencrypted, followed by an ESP header that provides authentication, and the transport layer data and higher is encrypted. For tunnel mode IPsec, an unencrypted IP header is added and the original packet and header are encrypted.

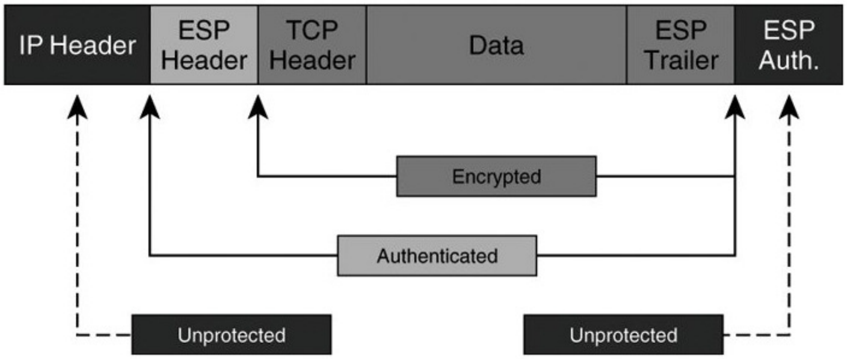


Figure 6. Transport Mode IPsec Packet [Car06]

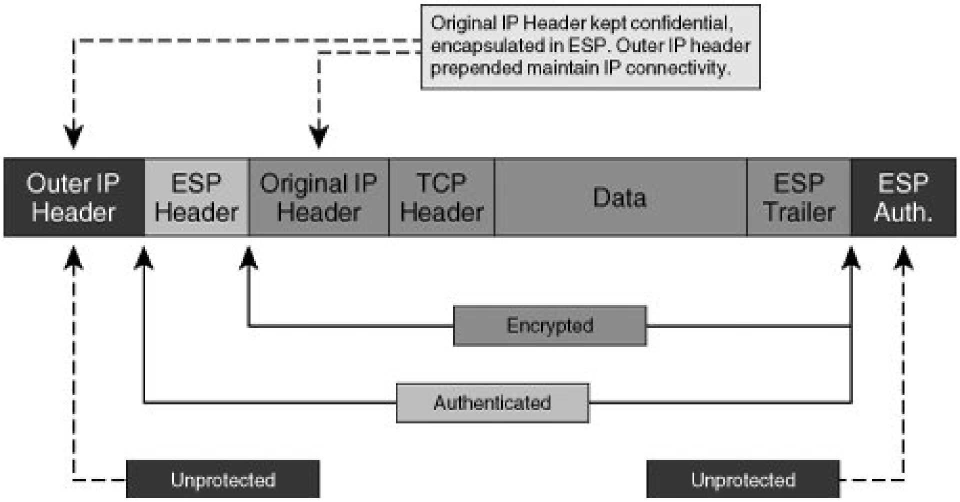


Figure 7. Tunnel Mode IPsec Packet [Car06]

2.11.2 Secure Real-time Transport Protocol

Real-time traffic itself can be encrypted by substituting RTP with a secure real-time protocol such as Secure Real-time Transport Protocol (SRTP). SRTP was introduced in 2004 to provide confidentiality, message authentication, and replay protection for RTP and RTCP [IET04a]. With SRTP the RTP payload is encrypted using AES and message integrity is provided by a SHA1 hash. Two types are keys are used, session keys and master keys. Session keys are derived from the master key, and the master key can be periodically changed through a key management mechanism external to SRTP.

Key distribution mechanisms are not specified in the SRTP standard so several techniques are in use. If SIP is employed, the key can simply be sent inside the SIP message using Session Description Protocol Security (SDPS) [ABW06]. However, SDPS does not encrypt the key so encryption must be achieved through another protocol, such as Secure / Multipurpose Internet Mail Extensions (S/MIME). SDPS is, therefore, not a complete solution for secure key distribution.

Another option for key distribution is the Multimedia Internet KEYing (MIKEY) standard. MIKEY was first published in RFC 3830 in 2004 [IET04]. The three approaches to key distribution as described by MIKEY are PSK, public key encryption, and the Diffie-Hellman algorithm. The use of PSKs is not scalable for large networks, therefore public key encryption and Diffie-Hellman techniques are more prevalent. The public key infrastructure used for DoD networks and maintained by the Defense Information Security Agency is an example of a large public key encryption solution. The Diffie-Hellman key exchange algorithm is scalable and requires less overhead cost than a public key infrastructure.

The Diffie-Hellman algorithm is asymmetric in that the method used to encrypt traffic differs from the decryption. It also relies on the discrete logarithm problem, where the best known algorithms cannot determine the exponential inverses of very large numbers in modulo arithmetic within a reasonable about of time. For example, consider two users, Alice and Bob, communicating privately using the Diffie-Hellman algorithm. First they agree to use the same large prime number *p* and base *g*. Next Alice computes her own private key *a*, and Bob computes his private key *b*. Then Alice sends Bob the value $A = (g^a \bmod p)$. Bob sends Alice the value $B = (g^b \bmod p)$. Alice and Bob arrive at

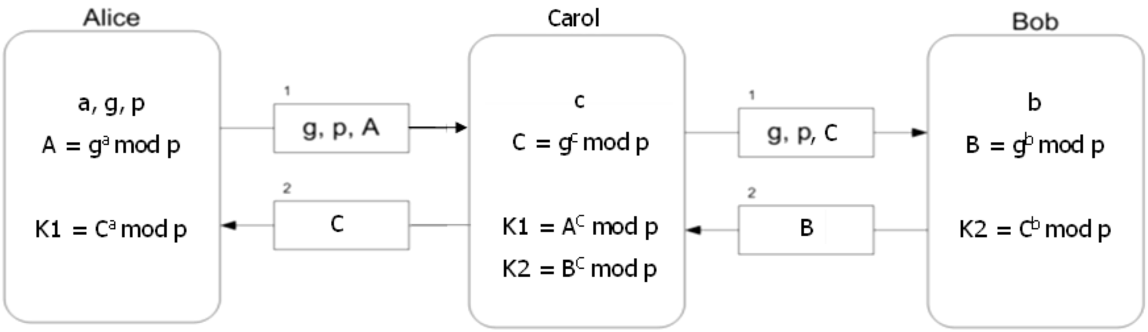


Figure 8. Diffie-Hellman Man-in-the-Middle Attack [Cod07]

the same secret key value K by using their respective formulas (Alice: $K = B^a \bmod p$, Bob: $K = A^b \bmod p$). The two can then exchange traffic encrypted with the key K.

A significant vulnerability in the Diffie-Hellman algorithm is the man-in-the-middle (MitM) attack. The encrypted traffic can be read by a third party if the entire exchange between Alice and Bob is intercepted. As shown in Figure 8, Carol can position herself to intercept traffic between Alice and Bob on the network, and then she can establish secure connections with them impersonating the other. It will appear to Alice and Bob that they have a secure connection directly with one another. Meanwhile Carol can decrypt the traffic as it arrives, re-encrypt it, and forward it to the recipient.

2.11.3 Zimmerman Real-time Transport Protocol

In March of 2006 Phillip Zimmerman proposed Zimmerman RTP (ZRTP) which prevents MitM attack through the use of shared keys [Zim06]. The Zfone Project IP phone software uses ZRTP for key negotiation. ZRTP negotiation begins after the users have already used a signaling protocol, such as SIP, and are ready to transmit RTP packets. A Hello message is sent first to determine if the endpoints support the protocol and to see which algorithms are in common. A Short Authentication String (SAS) is calculated by a cryptographic hash of two Diffie-Hellman values and used for key confirmation. The communicating parties confirm this key verbally over the VoIP system by reading the SAS as displayed on a screen [Zim06]. ZRTP provides further protection from MitM attack by using some hashed key material for use in the next call. If the attacker did not intercept the last call, they will be unable to intercept the next.

2.12 Related Research

WLANs are exceptionally vulnerable to unauthorized access unless they are properly secured. VoIP traffic on the WLAN is also vulnerable and should be protected from eavesdropping, manipulation, and DoS. The combination of WLANs and VoIP creates challenges for security, capacity, and call quality that have been the subject of several research efforts.

Xiao and Zarrella [XiZ04] examine the impact of WEP and 3DES IPsec on voice quality over an 802.11b WLAN. They observe that, during a 300 second G.711 call, instantaneous MOS falls below 3.0 three times when no encryption is used, eight times when WEP is used, and five times when 3DES IPsec is used. The method used for determining MOS is not mentioned.

Nascimento, et al. [NPM+06] investigate the impact of AES and 3DES IPsec on call quality over an 802.11b network using the E-Model. Their results indicate that IPsec decreases audio quality, with a MOS approximately 0.25 lower with 3DES than with AES for up to sixteen G.711 calls. The AP-based 802.11b network sustains an MOS above 3.0 for 10 simultaneous G.711 calls secured with a 3DES IPsec VPN, 12 simultaneous calls secured with an AES IPsec VPN, or 14 calls with no encryption.

Fathi, et al [FKC+05] measures the impact of MAC filtering and WEP on latency for an 802.11b AP-based network. Their results indicate that authentication delay roughly increases in proportion to WEP key length.

Rubino, Varela, and Bonnin [RVB05] use a neural network to measure MOS for voice calls made over an 802.11b AP-based network. The results suggest that MOS for a single call can fall below 2 (poor) for 15 seconds at a time during “moderately high” network load. Their data suggest that 802.11b networks can barely support VoIP calls.

Gurkas, Zaim, and Aydin [GZA06] simulate the effects of WEP, WPA, and WPA2 on 802.11b and 802.11g WLAN performance. They conclude that WEP and WPA decrease network throughput by 1%, and that WPA2 decreases throughput by 4%.

Lawrence, Biswas, and Sahib [LBS07] examine the capability of 802.11a and draft 802.11n at handling G.711 VoIP traffic. Their results suggest that draft 802.11n underperformed 802.11a at delivering small packets in real-time.

Seyba [Sey07] evaluates the ability for 802.11g WLANs to transport secure audio and video. Human subjects are used to evaluate subjective MOS provided by different network topologies. The results indicate that true capacity of an AP-based 802.11g WLAN secured with WPA and SRTP is only two simultaneous audio conversations. This is substantially fewer than 802.11b simulation results by Nascimento, et al [NPM+06].

Filho, et al [FFL+07] measure the impact of WEP 64, WEP 128, and WPA on the throughput of 802.11g networks. Their findings indicate that WEP 64, WEP 128, and WPA decrease UDP throughput by an average of 8%, 7%, and 6% respectively.

The number of simulated and modeled studies of voice quality over secure 802.11 networks greatly outnumbers subjective experiments with actual listeners. There appears to be a great deal of disparity between the two as well, with subjective performance tests indicting that 802.11 networks are less capable of supporting VoIP calls than analytical models predict.

2.13 Conclusion

The impact of common security measures on 802.11b WLANs have been repeatedly modeled and simulated. However, the effects of security mechanisms on voice quality on new and higher bandwidth WLAN standards such as 802.11n draft 2.0 have not been thoroughly and subjectively measured. This thesis expands upon this understanding by investigating the impact of security measures on audio quality over an 802.11n draft 2.0 WLAN.

III. Methodology

3.1 Introduction

This chapter explains how the experiments are conducted and details the VoIP WLAN configurations. Section 3.2 discusses the problem definition. Section 3.3 explains the experiment design. Section 3.4 describes the experiment methodology, and Section 3.5 explains in detail how the VoIP WLANs are configured.

3.2 Problem Definition

3.2.1 Goals

This experiment is designed to address three questions regarding VoIP audio quality over encrypted 802.11n draft 2.0 (hereafter referred to as 802.11n) WLANs:

1). Is audio quality higher over an 802.11n WLAN than over 802.11b or 802.11g WLANs when up to 10 simultaneous G.711 calls occur?

2). How significant is the impact of additional calls on audio quality over an encrypted 802.11n WLAN?

3). How significant is the impact of transport mode 3DES IPsec (hereafter referred to as IPsec) on audio quality over a WPA2-encrypted 802.11n WLAN?

3.2.2 Approach

These questions are addressed by building actual VoIP WLANs and having human test subjects subjectively rate audio quality. It is impractical to repeatedly reconfigure the VoIP WLANs for every subject during the audio evaluation, so 42 different recordings are taken under each of 28 network scenarios for a total of 1176 recordings. These recordings are what the human subjects hear and evaluate. As demonstrated in Chapter 4, audio quality of the recordings is statistically indistinguishable from live WLAN audio.

3.2.3 System Boundaries

The system under test (SUT), shown in Figure 9, is a VoIP WLAN consisting of the wireless medium, an 802.11b/g/n capable AP, VoIP private branch exchange (PBX) to direct VoIP traffic, and VoIP softphones on wireless laptops. The VoIP WLAN has a static topology, so physical positioning of the hardware is not part of the SUT. The component under test (CUT) is the 802.11 extension used.

The workload parameter is the number of simultaneous calls occurring on the VoIP WLAN. An increase in the number of simultaneous calls results in more bandwidth utilization, a greater probability of VoIP call distortion, and a greater computational load on the PBX. System parameters include the 802.11 extension used, and whether or not IPsec is implemented. The choice of 802.11 extension impacts the available WLAN bandwidth available for VoIP calls. Only 802.11b, 802.11g, and 802.11n are examined. WPA2 with IPsec or WPA2 without IPsec are the only levels of security for the SUT.

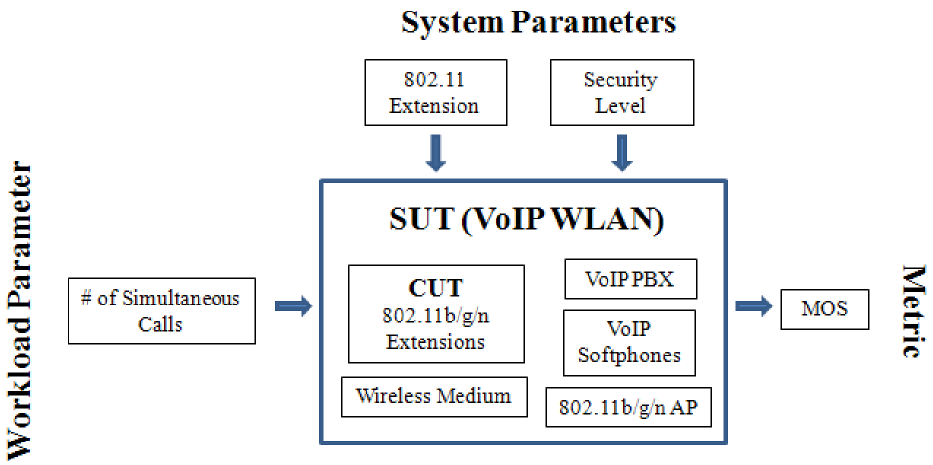


Figure 9. System under Test

The Metric used to evaluate the SUT VoIP call quality is MOS. It is a measurement of average listener perception of the VoIP call quality and is further explained in Section 3.4

3.3 Experiment Design

Twenty-eight network scenarios are divided into three trials to keep the testing periods short, and each trial is conducted with a different group of human subjects. Scenarios examined for Trials I, II, and III are listed in Tables 4, 5, and 6, respectively. Trial I examines audio quality over 802.11b, 802.11g, and 802.11n WLANs in order to evaluate differences in audio quality between the three 802.11 extensions. Trials II and III examine only 802.11n in order to characterize audio quality over a secure 802.11n WLAN during as many scenarios as possible.

Figure 10 illustrates the logical path taken by the audio clips from the source to the human subject. In Scenario 13, the human subjects listen to the source audio as it is directly played into the headphones. In Scenarios 14, 21, and 28, source audio is recorded into X-Lite and played back to the subjects. For all other scenarios, source audio traverses the VoIP WLAN before it is evaluated by the subjects.

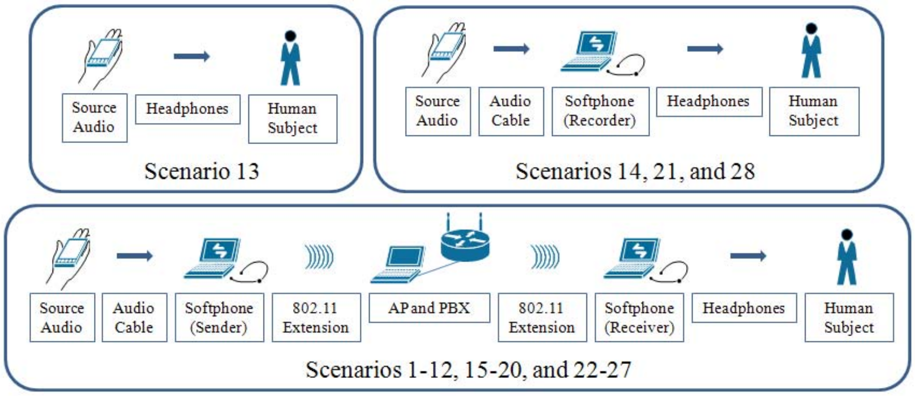


Figure 10. Logical Path of Audio through the SUT

Table 4. VoIP WLAN Scenarios Examined During Trial I

		802.11	Simultaneous	Encryption
Scenario #	Subjects	Extension	G.711 Calls	Technique
1	1-42	802.11b	6	WPA2
2	1-42	802.11b	10	WPA2
3	1-42	802.11g	6	WPA2
4	1-42	802.11g	10	WPA2
5	1-42	802.11n	6	WPA2
6	1-42	802.11n	10	WPA2
7	1-42	802.11b	6	WPA2+IPsec
8	1-42	802.11b	10	WPA2+IPsec
9	1-42	802.11g	6	WPA2+IPsec
10	1-42	802.11g	10	WPA2+IPsec
11	1-42	802.11n	6	WPA2+IPsec
12	1-42	802.11n	10	WPA2+IPsec
13	1-42	*Source Audio File*	N/A	N/A
14	1-42	*X-Lite Recording*	N/A	N/A

Table 5. VoIP WLAN Scenarios Examined During Trial II

		802.11	Simultaneous	Encryption
Scenario #	Subjects	Extension	G.711 Calls	Technique
15	43-84	802.11n	6	WPA2
16	43-84	802.11n	10	WPA2
17	43-84	802.11n	2	WPA2
18	43-84	802.11n	2	WPA2+IPsec
19	43-84	802.11n	4	WPA2
20	43-84	802.11n	4	WPA2+IPsec
21	43-84	*X-Lite Recording*	N/A	N/A

Table 6. VoIP WLAN Scenarios Examined During Trial III

		802.11	Simultaneous	Encryption
Scenario #	Subjects	Extension	G.711 Calls	Technique
22	85-126	802.11n	6	WPA2+IPsec
23	85-126	802.11n	10	WPA2+IPsec
24	85-126	802.11n	14	WPA2
25	85-126	802.11n	14	WPA2+IPsec
26	85-126	802.11n	20	WPA2
27	85-126	802.11n	20	WPA2+IPsec
28	85-126	*X-Lite Recording*	N/A	N/A

The number of subjects (N) is selected to be 42 for each of the trials for the following reasons:

1). (N = 42) is greater than 30, giving a reasonable expectation of distribution normality

2). (N = 42) is greater than 36, the number of subjects in related work by Seyba [Sey07]

3). (N = 42) is small enough to be experimentally practical

4). (N = 42) is evenly divisible by 14 and 7, the number of scenarios in Trials I, II, and III

(This permits balanced listening orders, as explained below)

In order to prevent listening order from influencing the audio scores, every human subject evaluates audio from the network scenarios in a different order. Additionally, listening orders are balanced such that all scenarios are heard first by the same number of subjects. Latin squares are produced (similar to completed Sudoku puzzles) to accomplish this randomization and balanced listening order. Subject listening order for Trials I, II, and III are listed in Appendix C. Dashed lines are included only to illustrate the division between Latin squares. As an example of how to read the tables, Subject 1 in Trial I evaluates audio from Scenario 3 first, from Scenario 14 second, and from Scenario 10 last.

While only two call levels are examined for 802.11b and 802.11g (6 and 10 simultaneous calls), six call levels are examined for 802.11n (2, 4, 6, 10, 14, and 20 simultaneous calls). This is done to more fully characterize audio quality resulting from the use of this new protocol. Additionally, a total of 84 human subjects evaluate audio quality for the 6 (Scenarios 5 and 15) and 10 (Scenarios 6 and 16) simultaneous call scenarios over 802.11n. This is done to narrow the confidence interval (CI) and give the lowest possible upper bound for the decline in audio quality due to IPsec on an 802.11n WLAN.

For validation purposes, one of the scenarios in each trial (Scenarios 14, 21, and 28) consist of source audio directly recorded into X-Lite from the CD player without transmission over the WLAN. This is done to verify that the three human subject pools provide the same MOS to audio of identical quality. Additionally, Scenario 13 consists of the source audio files played directly for the subject (no WLAN transmission) and used to verify that the source audio is of very high quality (MOS > 4.5). The X-Lite recording process must also be eliminated as a possible source of error by verifying that MOS from the direct recordings (Scenarios 14, 21, 28) are not statistically different from the source audio (Scenario 13). Finally, the distributions of subjective scores from all 28 scenarios are examined to verify that they are normally distributed and therefore not skewed.

3.4 Experiment Methodology

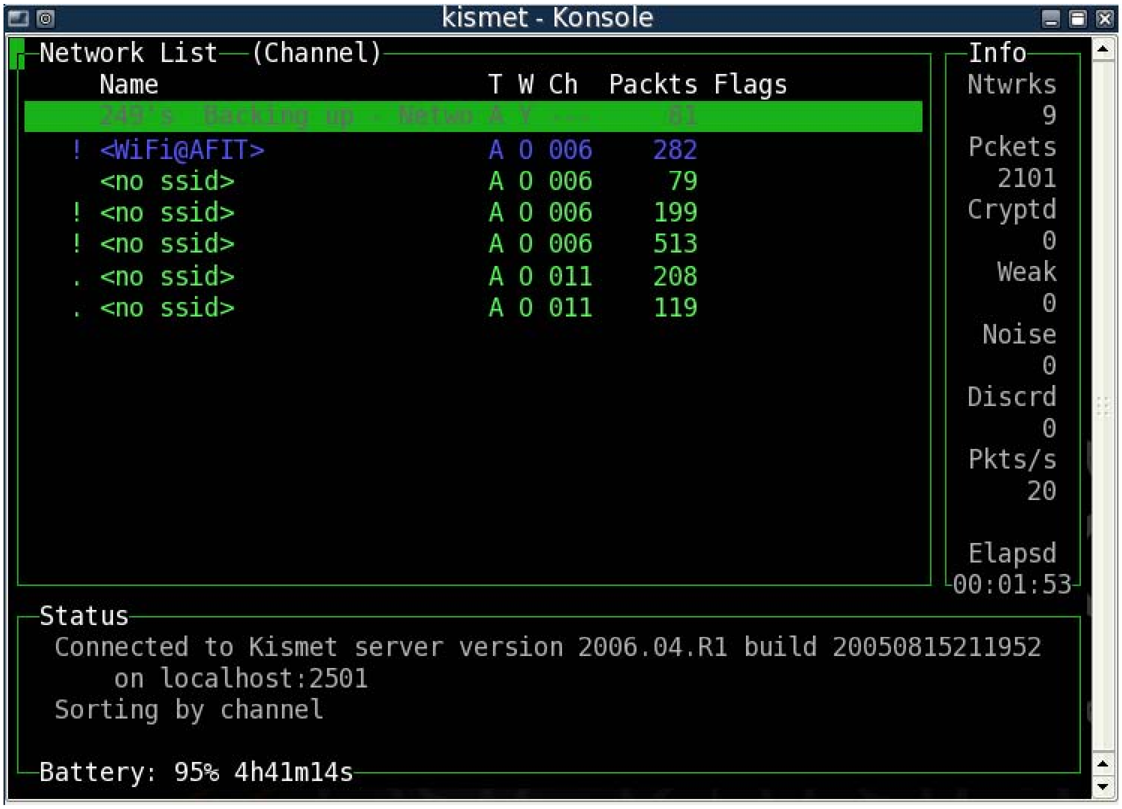
Four studio-recorded source audio files from the ITU-T P.862 source code [ITU01] consisting of two spoken sentences each are used for this experiment. The four files are played in a continuous and random order from a CD player into the microphone input of one wireless laptop running an X-Lite softphone and recorded via the X-Lite softphone program of a recipient laptop. This method allows WPA2 and WPA2+IPsec encrypted audio to be recorded in the same manner.

Audio quality is subjectively rated according to the Absolute Category Rating (ACR) test described in ITU-T Recommendation P.800 [ITU96]. Subjects listen to short audio recordings taken from actual VoIP calls made over the WLAN and rate audio quality with an integer from 1 to 5. The subjective rating scale includes 1 (bad), 2 (poor), 3 (fair), 4 (good), and 5 (excellent) [ITU96]. Subjects individually hear and evaluate the audio recordings for quality in a quiet room. After each recording is played, the subject says aloud the score it deserves so that it can be written down by the experimenter. The Mean Opinion Scores (MOS) are calculated as the average subjective score for each WLAN configuration as given by human subjects.

The potential exists for nearby access points (APs) to create significant interference for the VoIP WLANs. Therefore, prior to data collection the wireless 802.11 environment is monitored using the Kismet wireless network sniffer. Kismet is used to quantify the intensity of traffic from nearby APs and to verify minimal nearby WLAN activity. A representative screenshot of Kismet output is shown in Figure 11. In this figure, data such as neighboring AP SSIDs ("Name"), 802.11 channels in use ("Ch"), packets observed ("Packts"), and the average number of packets per second ("Pkts/s") are visible.

Audio recordings are taken at AFIT within range of several other 802.11g APs as shown in Figure 11. Since 802.11 channel 1 is not active in the vicinity, it is used for the 802.11b and 802.11g configurations. However, 802.11n uses a channel width of 40 MHz (twice as large as for 802.11b/g) so channel 4 is used for those scenarios to fit within the802.11 spectrum. For 802.11n, some operating frequency overlap with the nearby APs on channel 6 is unavoidable. In order to minimize interference from other APs, recordings are only made after 2000 hours on weekends, when ambient WLAN activity at AFIT is at a minimum (< 20 packets/sec across all channels).

As with all research involving human subjects, approval is required from an institutional review board (IRB). Given that the subjects only listen to audio samples and



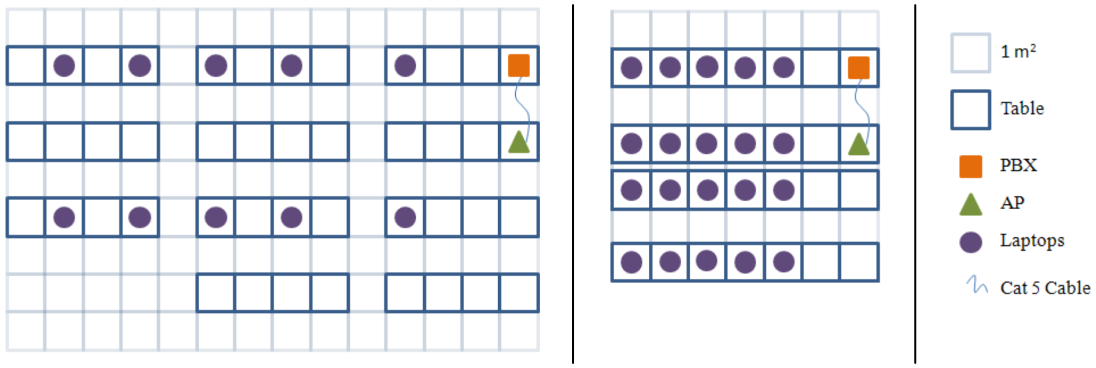
```
Network List—(Channel)                                      Info
    Name                          T W Ch  Packts Flags      Ntwrks
    249's  Backing up - Netwo A Y ---       81                   9
  ! <WiFi@AFIT>                   A O 006    282            Pckets
    <no ssid>                     A O 006     79              2101
  ! <no ssid>                     A O 006    199            Cryptd
  ! <no ssid>                     A O 006    513                 0
  . <no ssid>                     A O 011    208              Weak
  . <no ssid>                     A O 011    119                 0
                                                             Noise
                                                                 0
                                                            Discrd
                                                                 0
                                                            Pkts/s
                                                                20

                                                            Elapsd
                                                          00:01:53
Status
  Connected to Kismet server version 2006.04.R1 build 20050815211952
      on localhost:2501
  Sorting by channel

Battery: 95% 4h41m14s
```

Figure 11. Example Screenshot of Kismet Output

are not exposed to potential harm, an exemption request from human experimentation equirements is obtained and is included in Appendix A. No personally identifiable information is recorded from any of the subjects, as explained on the subject information sheet shown to every subject prior to testing. This sheet is included in Appendix C.

3.5 VoIP WLAN Configuration

Figure 12 illustrates the VoIP WLAN physical topologies. In order to maximize WLAN performance, all laptops running software phones are relatively close to the AP (< 15m) and within line-of-sight of the AP. The WLAN topologies for Trials I and II require up to 10 wireless laptops with softphones, while Trial III requires up to 20. Although the physical topology for Trial I and II differs from that of Trial III, the small



Trials I and II

Trial III

Figure 12. VoIP WLAN Topologies

differences in relative hardware location do not have a statistically significant impact on WLAN performance or VoIP audio quality. This conclusion is based on a study of draft 802.11n hardware by Veritest, which concludes that average throughput is essentially constant at distances less than 15 m from the AP [Ver06]. All trials use microphone headsets, X-Lite softphones, an 802.11b/g/n AP, and another laptop running trixbox as the private branch exchange (PBX). The PBX serves as a SIP proxy that manages all VoIP calls on the WLAN. Specifications for the WLAN components are listed in Table 7, and Figure 13 shows a configured wireless laptop with headset and softphone.

The VoIP WLANs are highly configurable in order to create the WLAN conditions examined. Variables include 802.11 extension (selected at the AP), number of simultaneous calls (made between pairs of laptop softphones), and use of WPA2 encryption with or without IPsec (selected at the AP, PBX, and on each laptop). Appendix D explains IPsec configuration for the trixbox PBX, and Appendix E explains IPsec configuration for the wireless laptops.

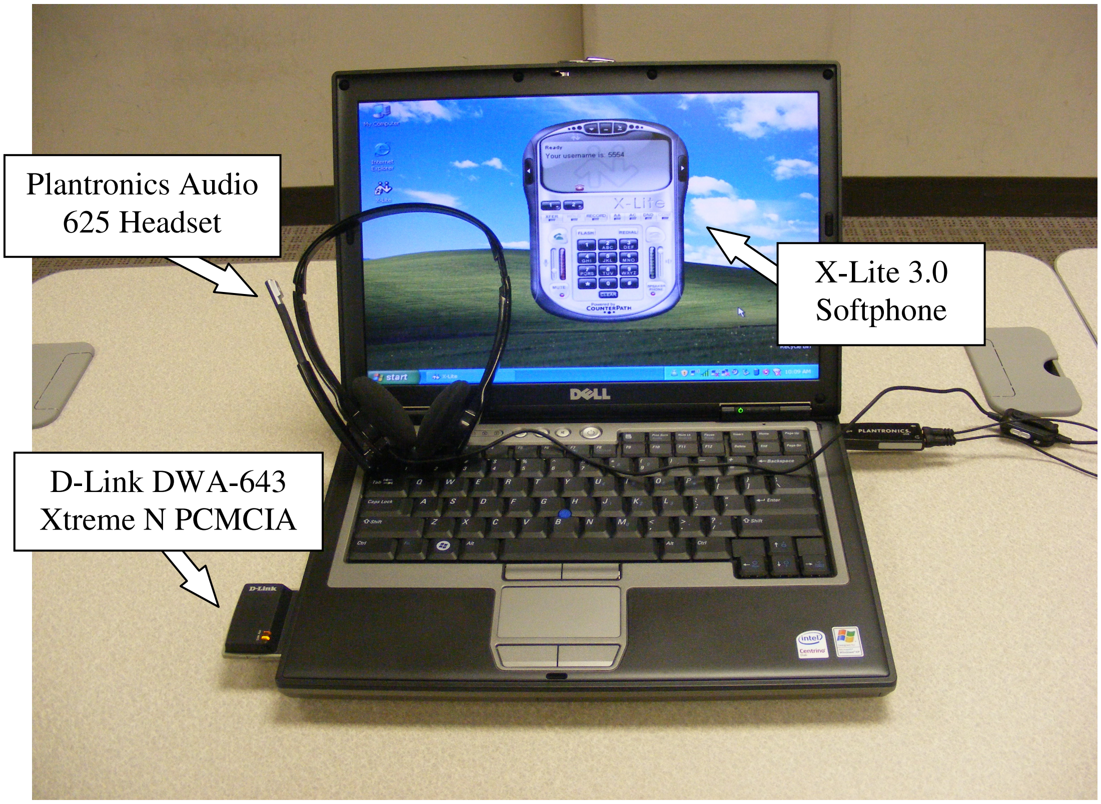


Figure 13. A Configured Laptop with Softphone and Headset

Table 7. VoIP WLAN Component Specifications

Component	Specifications
VoIP PBX	Dell Latitude C840, 2 GB SDRAM, trixbox 2.6.1
802.11 b/g/n AP	D-Link 655 Xtreme N Gigabit Router, Firmware v1.11
Caller/Callee 1-2 Laptops	Dell Latitude C840, 1.8 GHz P4-M, 512 MB SDRAM, Windows XP Professional SP2
Caller/Callee 3-10 Laptops	Dell Latitude D630, T7300 Core 2 Duo, 1 GB SDRAM, Windows XP Professional SP2
Wireless Cards	D-Link DWA-643 Xtreme N PCMCIA
VoIP Softphones	X-Lite 3.0, G.711 20 ms/frame
Microphone Headsets	Plantronics Audio 625 Headset

Advanced AP settings are primarily left to the factory default so only the 802.11 extension ("802.11 Mode") and channel width (in the case of 802.11n) is altered on the AP between scenarios. The full list of AP settings is displayed in Table 8.

Strong WLAN security is enforced with the selection of WPA2 encryption using AES, "Low" transmit power, "Invisible" status, and lengthy passphrase (20 characters). Also, the AP has a built-in QoS engine to help maximize VoIP audio performance.

Table 8. VoIP WLAN AP Settings

Option	Setting
Enable Wireless:	Always
802.11 Mode:	80211b only, 802.11g only, *or* 802.11n only
Transmission Rate:	Best (automatic)
Channel Width: (*for 802.11n*)	Auto 20/40 MHz (*fixed 40 MHz is not an option*)
Visibility Status:	Invisible
Security Mode:	WPA-Personal
WPA Mode:	WPA2 Only
Cipher Type:	AES
Group Key Update Interval:	3600 seconds
DHCP Settings:	Static IP addresses
Transmit Power	Low
Beacon Period:	100
RTS Threshold:	2346
Fragmentation Threshold:	2346
DTIM Interval:	1
WLAN Partition:	No
WMM Enable:	Yes (*helps control latency and jitter*)
Short GI:	Yes (*short 400ns guard interval*)
Extra Wireless Protection:	Yes
Enable Traffic Shaping:	Yes
Auto Uplink Speed:	Yes
Connection Type:	Auto-detect
Enable QoS Engine:	Yes
Auto Classification:	Yes
Dynamic Fragmentation:	Yes

No changes from the default are made to the D-Link wireless network cards used in the laptops. These default driver settings are listed in Table 9.

Table 9. Wireless Network Card Driver Settings

Option	Setting
Driver Provider:	D-Link
Driver Date:	8/28/2006
Driver Version:	6.0.1.75
802.11b Preamble:	Long and Short
Map Registers:	256
Network Address:	Not Present
Power Save Mode:	Normal
Radio On/Off:	On
Scan Valid Interval:	60

In addition to the AP and wireless card settings, the preferences selected for the X-Lite softphones are important to note. The G.711 codec is used in order to maximize MOS provided by the VoIP network, since it has a higher typical MOS (4.1) than any other traditional (< 64 kbps) codec. The default device options used for the trixbox PBX are listed in Table 10. Important X-Lite VoIP settings are listed in Table 11.

Table 10. Softphone Device Options Selected in Trixbox

Option	Setting
Secret:	*blank*
Dtmf_mode:	rfc2833
Can_reinvite:	No
Context:	from-internal
Host:	Dynamic
Type:	Friend
Nat:	Yes
Port:	5060
Qualify:	yes

Table 11. X-Lite Softphone VoIP Settings

Option	Setting
Speaker device:	SigmaTel Audio
Microphone device:	SigmaTel Audio
Use acoustic echo cancellation (AEC):	Yes
Use auto gain control (AGC):	Yes
Use noise reduction:	Yes
Enabled codecs:	G.711 uLaw
Preserve bandwidth during silence periods:	No
Reregister every:	3600 seconds
Send SIP keep-alives:	Yes
Register with domain and receive incoming calls:	Yes
Send outbound calls via:	domain
Presence:	Peer-to-peer

3.5 Summary

This chapter describes the methodology used to quantify the impact of IEEE 802.11 extension, additional calls, and IPsec encryption on audio quality over VoIP WLANs. Human subjects listen to audio samples taken during 28 WLAN scenarios and MOS is calculated with an 85% CI.

IV. Results and Analysis

4.1 Introduction

This chapter presents analysis of the subjective and objective experimental results. Section 4.2 discusses the validation of the MOS results. Section 4.3 examines the MOS data from the VoIP WLANs. Section 4.4 discusses the results of objective measurements of the VoIP WLAN. Section 4.5 provides a summary of the overall analysis.

4.2 Validation of MOS Results

This section explains how the MOS results are validated. Validation is accomplished by verifying that the subjective scores are normally distributed, eliminating the X-Lite recording process as a source of error, and by comparing MOS results from the 802.11b WLAN to an E-Model prediction.

In order to verify that the MOS results are not significantly skewed, subjective scores from each of the 28 scenarios are tested for normality using the Ryan-Jointer test. The null hypothesis for the Ryan-Joiner test is that the data is normally distributed. Data from all experiment scenarios pass the Ryan-Joiner test for normality (p-value > 0.1).

Since the human subjects hear and score recorded audio clips, it is necessary for the MOS for source (original) audio and its direct recording by the X-Lite softphone to be statistically equivalent. If the two are statistically equivalent, then the X-Lite recording process does not introduce error in the form of MOS decline. Figure 14 shows the MOS results for the source audio and the direct recordings into X-Lite for each of the three trials. Source audio receives an MOS of 4.45, and the 85% CI includes an MOS of 4.5, which is the highest possible MOS as explained in Section 2.7. This result demonstrates

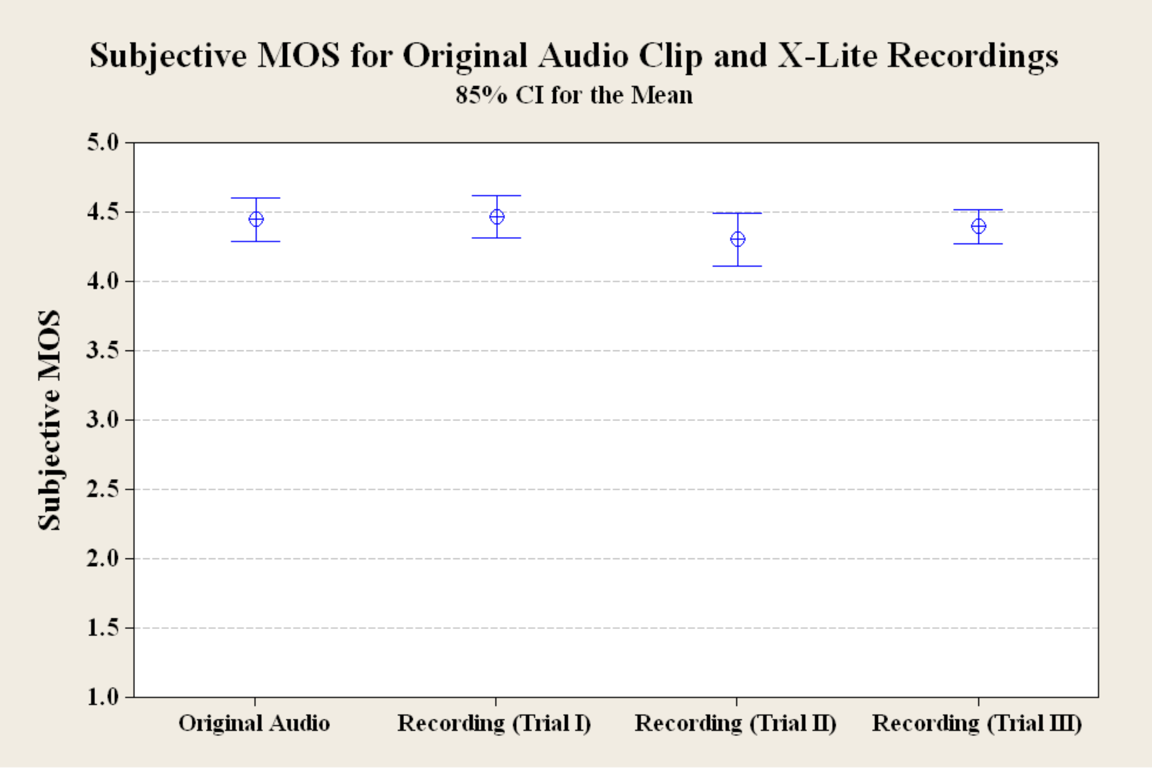


Figure 14. MOS for Original Audio and Direct X-Lite Recordings

that the source audio used for this experiment is of very high quality. Furthermore, the MOS for the direct recordings into X-Lite during Trials I, II, and III are all near 4.5 (4.48, 4.41, and 4.40, respectively), and statistically equivalent to the original audio clip (85% CI). The recording process is therefore eliminated as a possible source of error for the MOS results.

The third method used to validate the subjective audio quality results is to compare them to an E-model prediction for a similar VoIP WLAN. E-model results from Nascimento, et al [NPM+06] predict MOS for G.711 calls made over an 802.11b WLAN with no encryption and an 802.11b WLAN secured with an IPsec Virtual Private Network (VPN). The MOS results from this research should fall between the E-model predictions from [NPM+06]. Figure 15 shows the E-model predictions as grey boxes

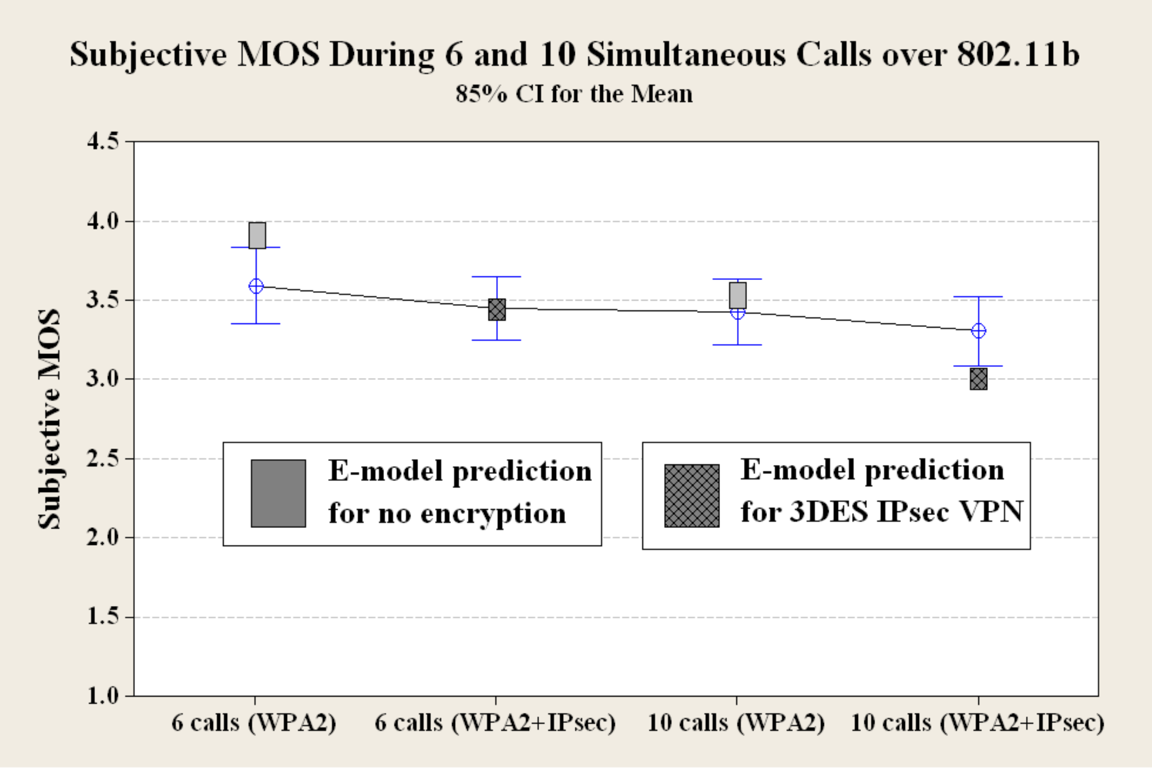


Figure 15. MOS and E-model Predictions for an 802.11b WLAN

and the subjective MOS from this thesis with 85% CIs. The subjective MOS results are statistically equivalent to or lower than the E-model prediction for an unencrypted 802.11b WLAN and statistically equivalent or higher than the E-model prediction for an 802.11b WLAN secured with an IPsec VPN. The computational and bit overhead introduced by encryption levels used for this thesis (WPA2 or WPA2 and transport mode IPsec) fall between those in [NPM+06] (no encryption or IPsec VPN). Therefore, subjective MOS from this experiment over a secured 802.11b WLAN is consistent with the E-model results from [NPM+06]. This fact both validates the subjective MOS results from this experiment and the E-model prediction for MOS presented in [NPM+06].

4.3 Subjective MOS Results

MOS results for 6 and 10 simultaneous calls over an 802.11b WLAN are displayed in Figure 16. The results are plotted in ascending order of utilized bandwidth, from 6 simultaneous calls secured with WPA2 to 10 simultaneous calls secured with both WPA2 and IPsec. Although MOS declines when IPsec is added during both 6 and 10 calls, the declines are not statistically significant. However, the upper bound for possible decline in MOS due to IPsec is calculated by subtracting the lowest value of the WPA2+IPsec scenario CI from the high value of the WPA2 scenario CI. For an 802.11b WLAN, the upper bound for the decline in MOS due to IPsec is 0.59 during 6 calls, and 0.55 for 10 calls (85% CI). The data are consistent with steady MOS decline across the four trials, but this decline is not statistically significant.

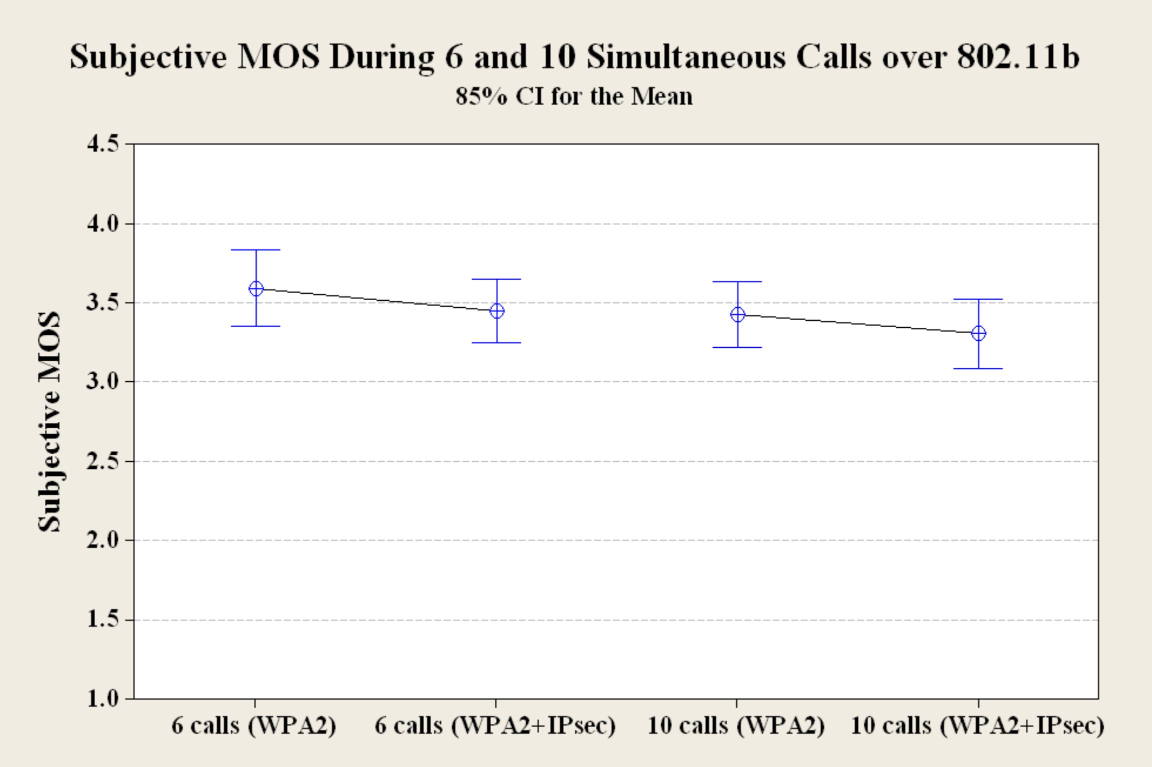


Figure 16. MOS over an 802.11b WLAN (6 and 10 calls)

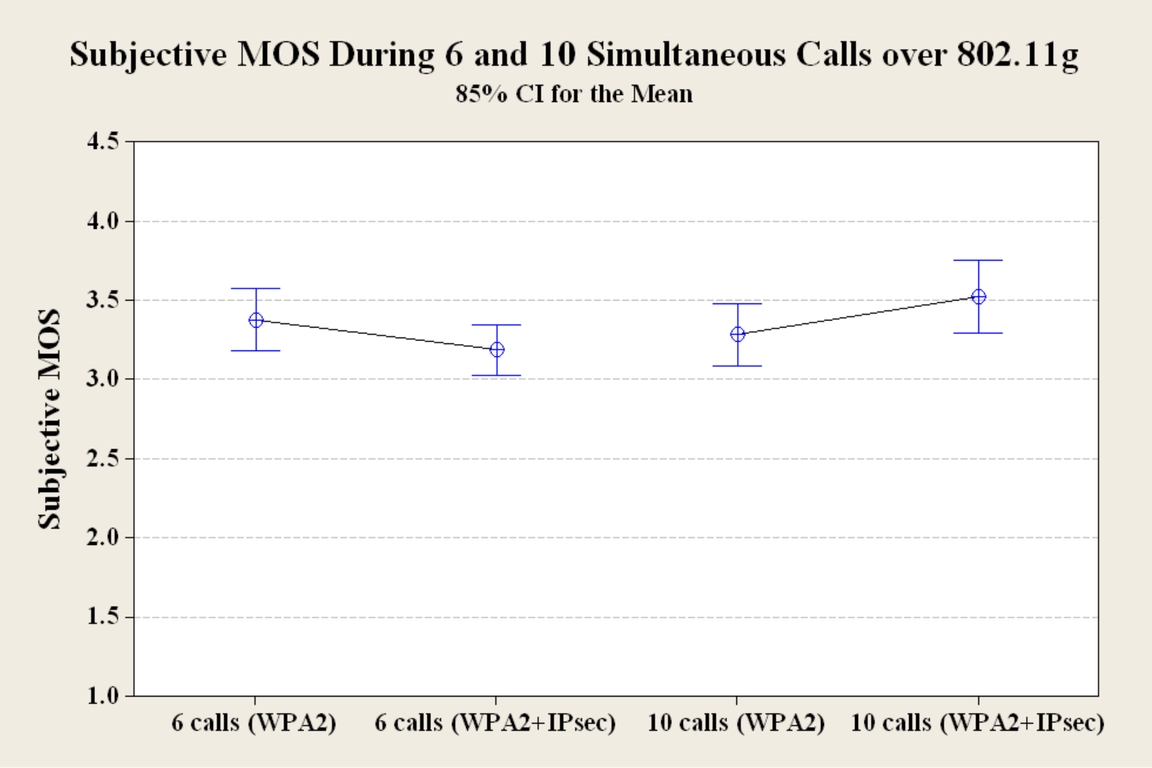


Figure 17. MOS over an 802.11g WLAN (6 and 10 calls)

Subjective MOS results for an 802.11g WLAN are shown in Figure 17. The results are plotted in ascending order of utilized bandwidth, from 6 simultaneous calls encrypted with WPA2 to 10 simultaneous calls encrypted with both WPA2 and IPsec. Unlike the steady decline in audio quality indicated by the 802.11b results, the 802.11g MOS results do not consistently decline across all four scenarios. During 6 calls secured with WPA2+IPsec, the MOS CI nearly falls below 3.0, while at 10 calls secured with WPA2+IPsec the MOS CI unexpectedly rises higher than for the other three 802.11g MOS CIs. The erratic nature of these MOS results suggests that the 802.11g WLAN experienced interference that randomly lowered the subjective scores. This problem is further investigated through objective measurements presented in Section 4.4

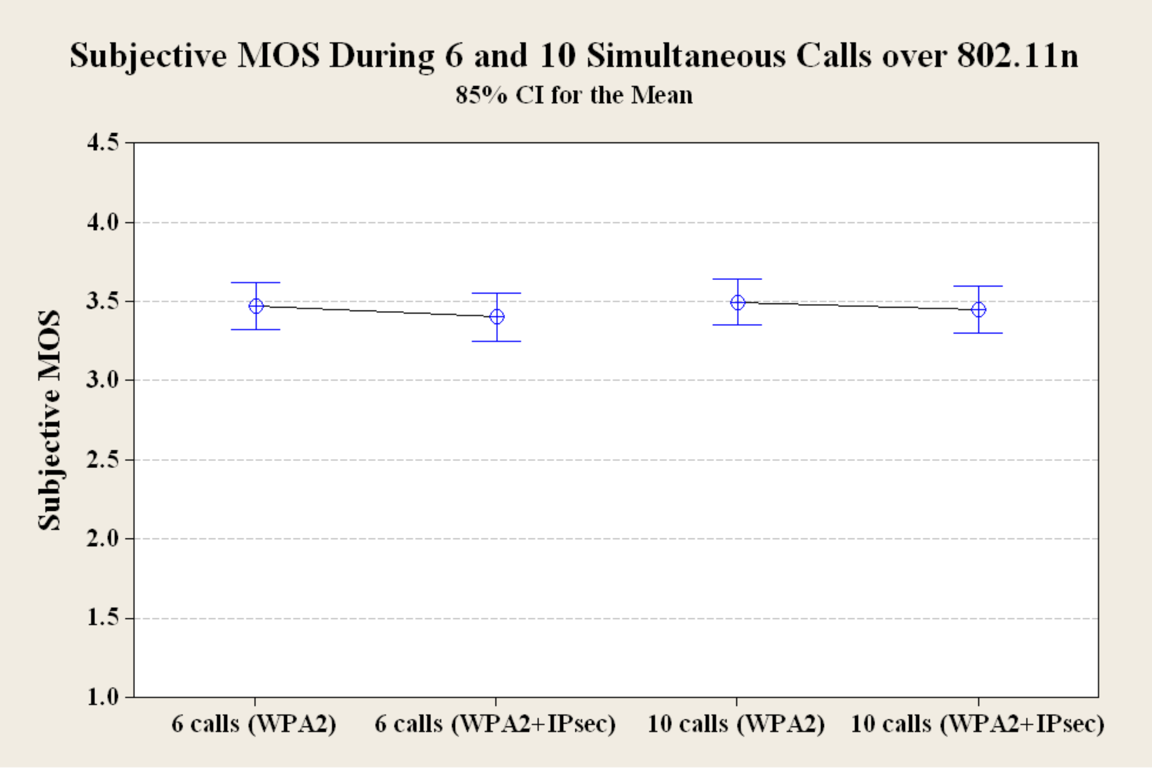


Figure 18. MOS over an 802.11n WLAN (6 and 10 calls)

Figure 18 shows the subjective MOS results for 6 and 10 simultaneous calls over 802.11n. The number of subjects used to evaluate these scenarios is increased to 84 to provide the most precise MOS results for the 802.11n WLAN. Doubling the number of human subjects from 42 to 84 decreases the 85% CI widths from approximately 0.4 to 0.3. This small improvement in CI width illustrates the primary limitation of subjective testing: that a large number of subjects are required for precise results.

The data suggest that an increase from 6 to 10 calls or the addition of IPsec encryption does not result in a statistically significant MOS decline. This result is reasonable due to the large 802.11n bandwidth available to handle additional packet overhead and increased VoIP traffic. Upper bound for MOS decline due to IPsec is 0.38 for 6 simultaneous calls and 0.35 for 10 calls over 802.11n.

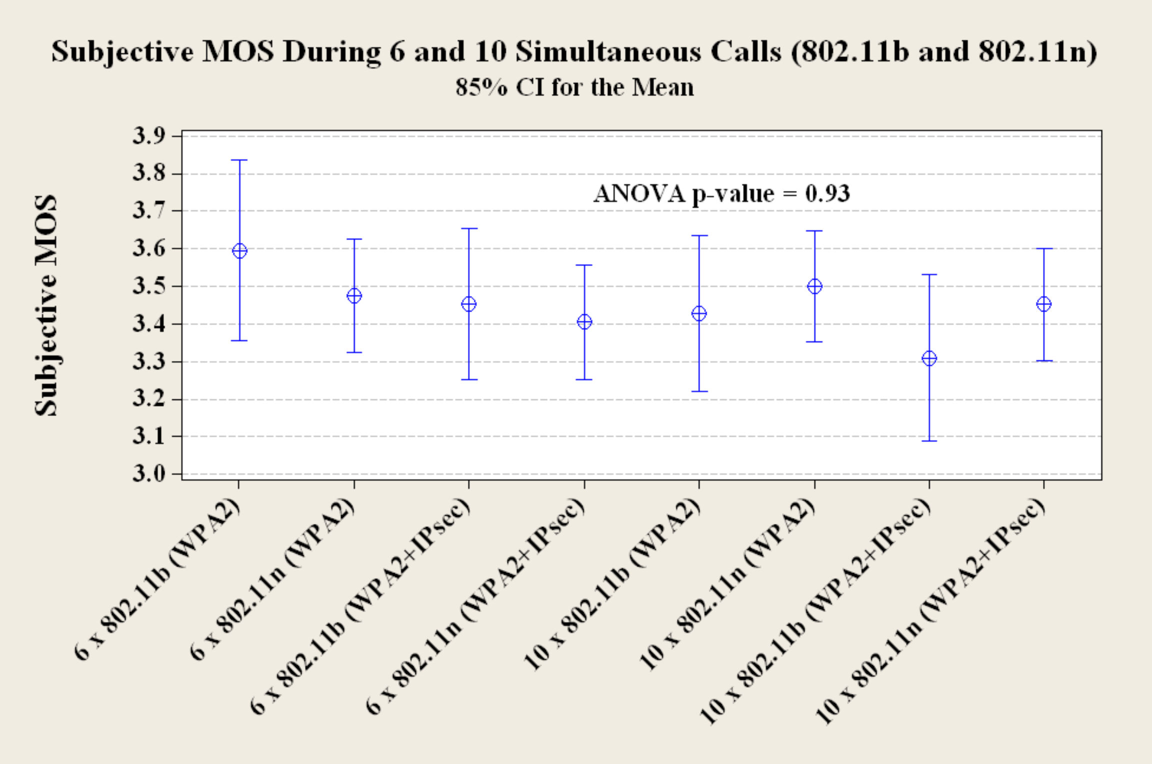


Figure 19. MOS over 802.11b and 802.11n (6 and 10 calls)

When the subjective MOS results for 6 and 10 calls over an 802.11n WLAN are compared to results from an 802.11b WLAN, they are statistically equivalent. Figure 19 shows the 85% CI for all 6 and 10 simultaneous call scenarios examined for 802.11b and 802.11n. Note that the 802.11n MOS CIs are identifiable due to their relative narrowness. The p-value from an ANOVA test on the eight scenarios in Figure 19 is 0.93, indicating no statistically significant difference in MOS. The data suggest that although 802.11n offers significantly greater bandwidth than 802.11b, audio quality over an 802.11n WLAN is not significantly higher (or lower) than over an 802.11b WLAN for up to 10 simultaneous G.711 calls encrypted with WPA2+IPsec, utilizing 1.872 Mbps (1872 bits x 100 packets/sec x 10 calls). MOS results from 802.11g are not included in this comparison due to possible interference effects observed over the 802.11g WLAN.

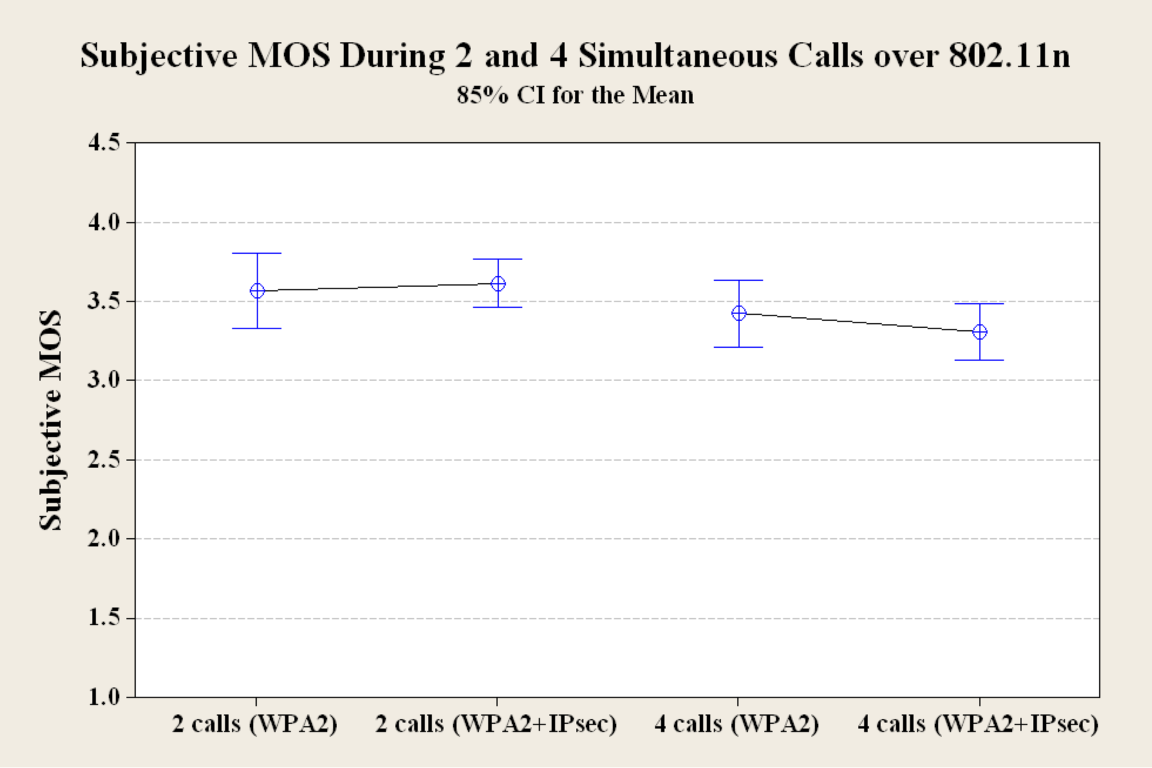


Figure 20. MOS over an 802.11n WLAN (2 and 4 calls)

In addition to the 6 and 10 call scenarios, 2, 4, 14, and 20 simultaneous calls are subjectively rated for the 802.11n WLAN. Figure 20 shows the MOS results for 2 and 4 simultaneous calls over an 802.11n WLAN. No statistically significant MOS decline is observed when the number of calls increases from 2 to 4 or when IPsec is added. This result is as expected due to the low utilization (< 750 kbps) of these scenarios (1872 bits x 100 packets/sec x 4 calls). During 2 calls, the upper bound for the decline in MOS as a result of IPsec encryption is 0.34 for 2 simultaneous calls and 0.51 for 4 simultaneous calls over 802.11n.

Although no significantly significant MOS decline is observed for 2 or 4 calls over 802.11n, the results are consistent with a small but steady decline across the 4 scenarios shown in Figure 20.

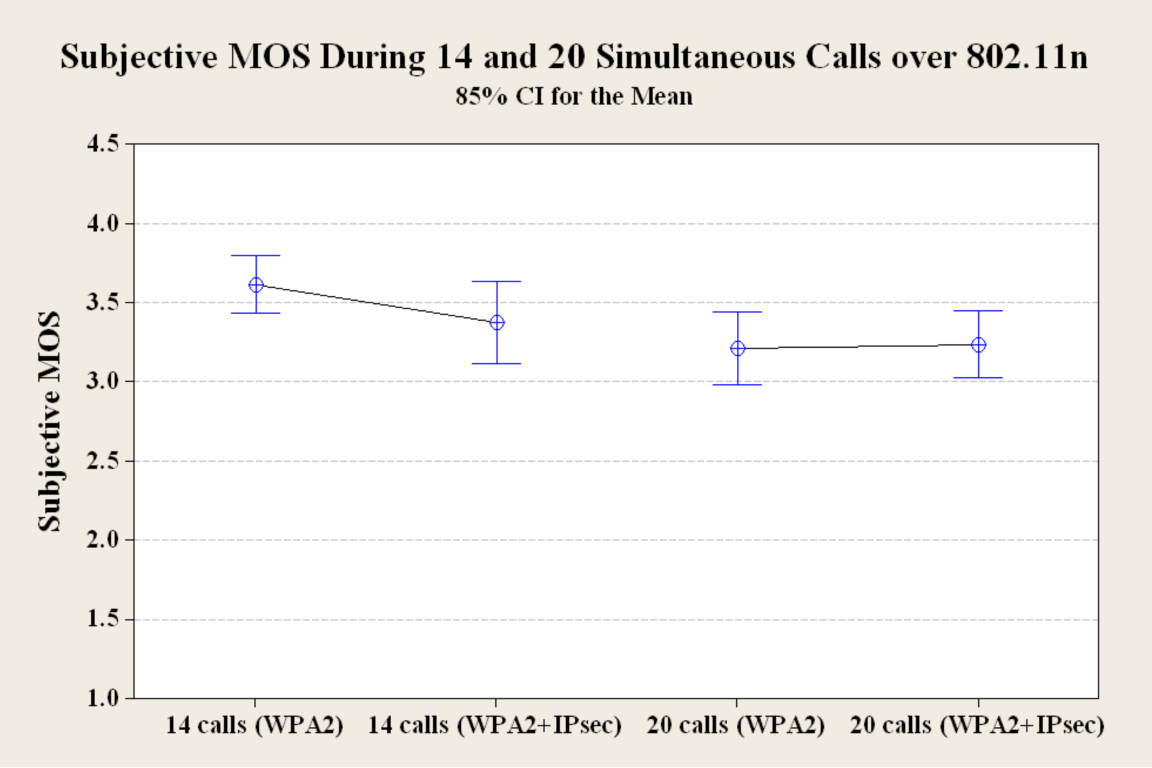


Figure 21. MOS over an 802.11n WLAN (14 and 20 calls)

Figure 21 shows the subjective MOS results for 14 and 20 simultaneous calls over an 802.11n WLAN. The data do not suggest a statistically significant decline in MOS when IPsec is added at either call level. Upper bound for MOS decline due to IPsec encryption is 0.68 for 14 simultaneous calls and 0.42 for 20 calls over 802.11n. Results suggest that MOS over an 802.11n WLAN with 20 simultaneous calls and secured with both WPA2 or WPA2+IPsec encryption are significantly lower (85% CI) than for an 802.11n WLAN with 14 simultaneous calls secured with WPA2. This is the first subjective result to indicate a definitive MOS decline due to additional calls. Furthermore, the MOS data confirms that the relatively small bandwidth utilized by 20 G.711 calls (< 4 Mbps) is sufficient to cause MOS decline over an 802.11n WLAN.

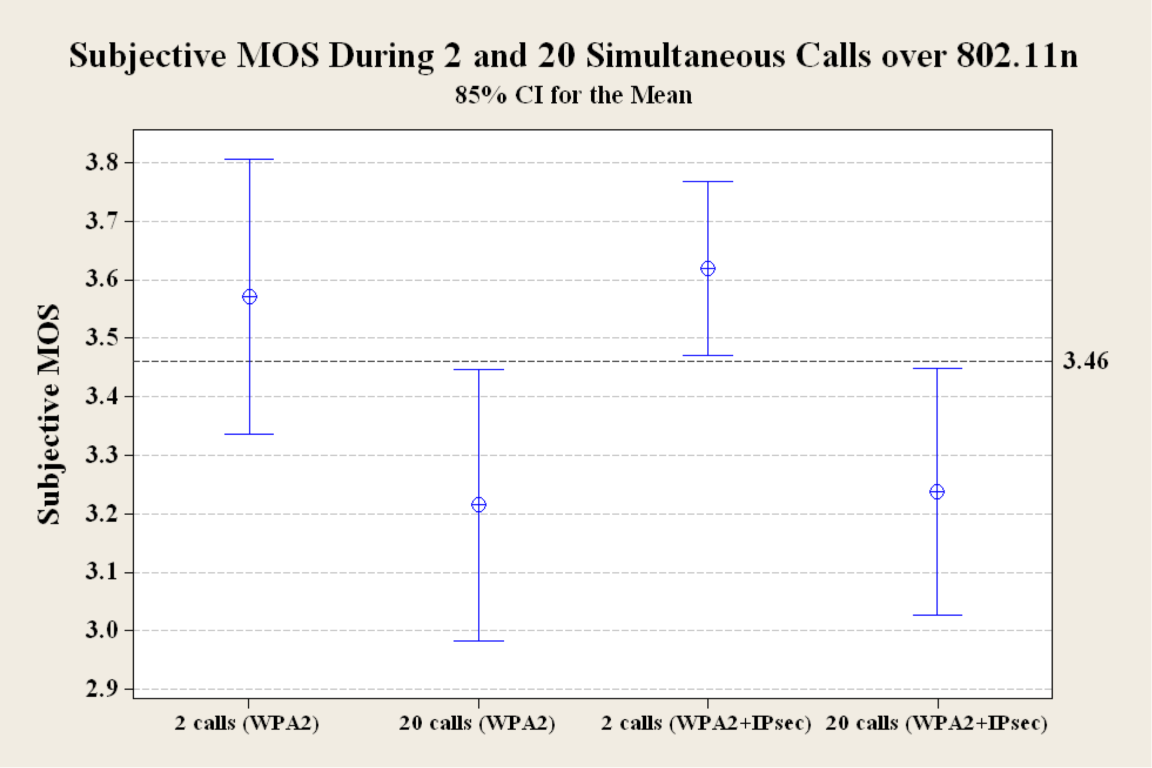


Figure 22. MOS over an 802.11n WLAN (2 and 20 calls)

Figure 22 illustrates the subjective MOS results for 2 and 20 calls over an 802.11n WLAN. From Figure 22 it is clear that MOS for 2 calls secured with WPA2+IPsec over 802.11n is significantly higher than for both 20 call scenarios over 802.11n. This result suggests that the large bandwidth available on an 802.11n WLAN does not prevent a statistically significant decline in MOS when the number of calls increases from 2 to 20. The results also suggest that MOS decline due to IPsec MOS is much lower than that caused by additional calls. This result is as expected, since the bit overhead required by IPsec is less than that generated by additional calls. For instance, securing 20 calls with IPsec adds 544 kbps (272 bits x 100 packets/sec x 20 calls), while adding 20 unencrypted calls generates 3.2 Mbps (1,600 bits x 100 packets/sec x 20 calls).

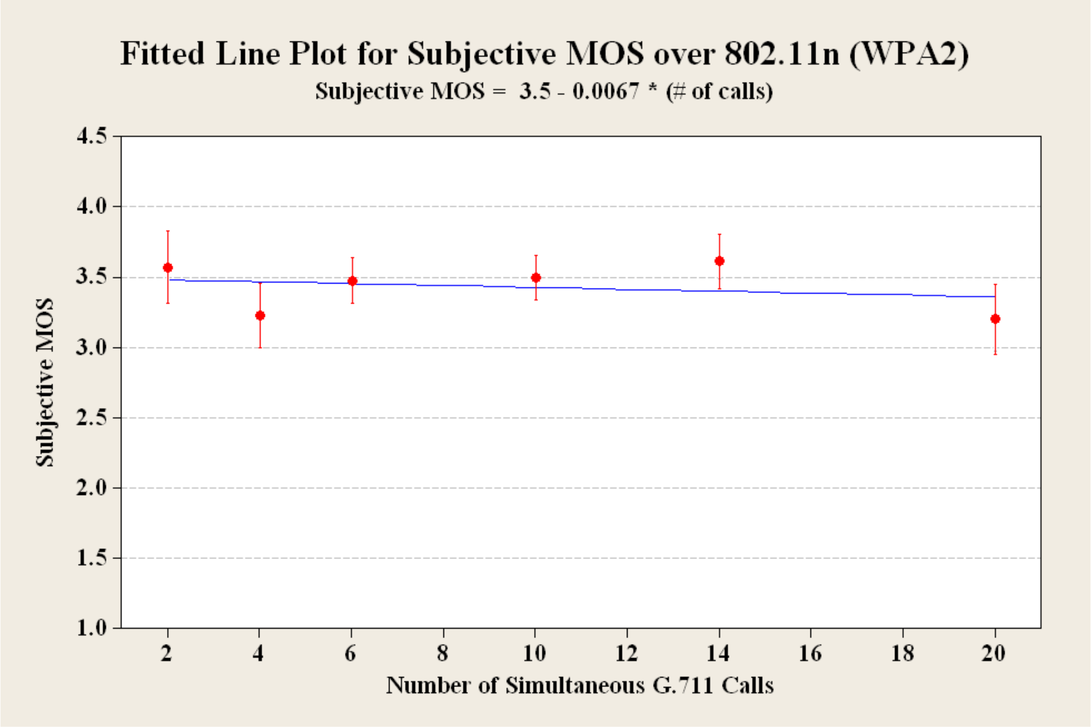


Figure 23. Fitted Line Plot for Subjective MOS over 802.11n (WPA2)

With MOS results for six different call levels over 802.11n, it is possible to calculate a regression line of best fit to estimate MOS for call levels not examined during this experiment. Figure 23 shows the fitted line plot for subjective MOS over a WPA2 secured 802.11n WLAN. The formula for the regression line is

$$\textit{Subjective MOS} = 3.5 - 0.0067(\#\ \textit{of calls}) \qquad (3)$$

and predicts that MOS falls below 3.0 (fair quality) during 75 simultaneous G.711 calls encrypted with WPA2. Predictions by (3) are statistically equivalent to all 6 subjective MOS results for a WPA2 secured 802.11n WLAN, confirming its accuracy. The highest possible MOS predicted by (3) is 3.5 for a single G.711 call on the 802.11n WLAN secured with WPA2. This strongly suggests, as do the subjective MOS results, that toll quality audio (MOS ≥ 4.0) is not practical over a WPA2-secured 802.11n WLAN, even with the G.711 codec.

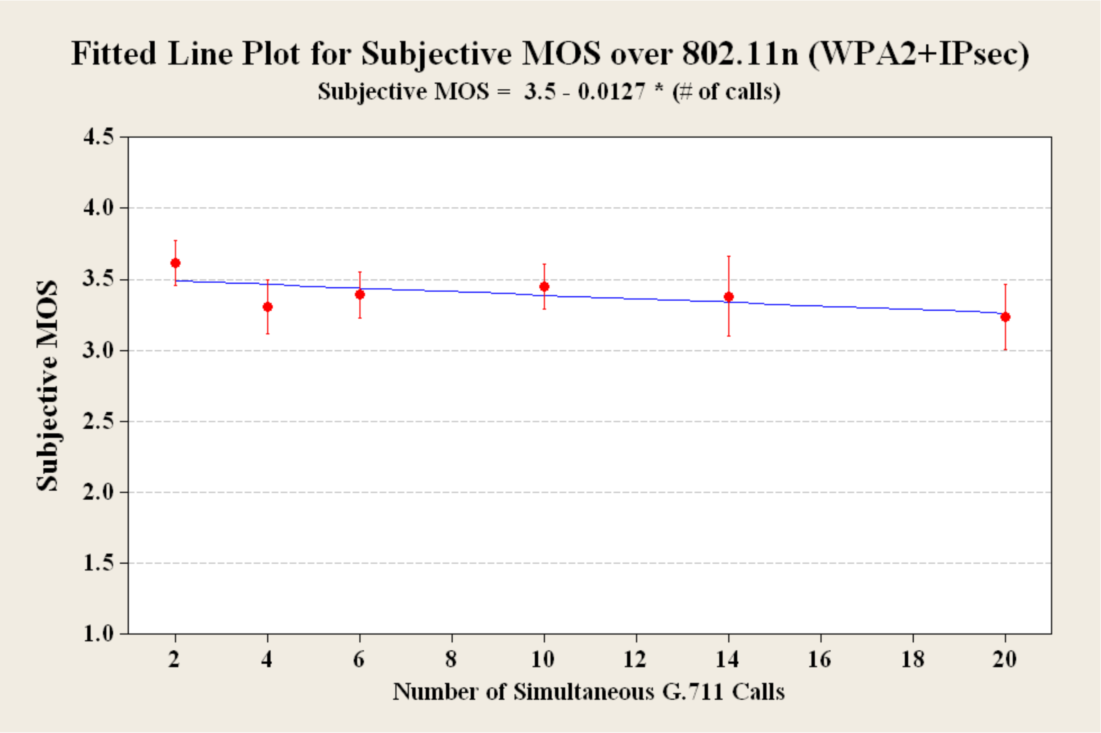


Figure 24. Fitted Line Plot for Subjective MOS over 802.11n (WPA2+IPsec)

Figure 24 shows the fitted line plot for subjective MOS over an 802.11n WLAN encrypted with both WPA2 and IPsec. The formula for the regression line is

$$\textit{Subjective MOS} = 3.5 - 0.0127(\#\ \textit{of calls}) \qquad (4)$$

and predicts that subjective MOS falls below 3.0 (fair quality) during 40 simultaneous G.711 calls encrypted with WPA2 and IPsec. Predictions by (4) are statistically equivalent to all 6 subjective MOS results for an 802.11n WLAN secured with WPA2 and IPsec, confirming its accuracy. These predicted declines in MOS are also well within the upper bounds presented in Figures 18, 20, and 21. The highest possible MOS predicted by (4) is 3.5 for a single G.711 call on the 802.11n WLAN secured with WPA2+IPsec. This strongly suggests, as do the subjective MOS results, that toll quality audio (MOS ≥ 4.0) is not practical over an 802.11n WLAN secured with WPA2 and IPsec, even with the high quality G.711 codec.

MOS decline due to IPsec is predicted by subtracting (3) from (4). The result is

$$\textit{MOS Decline due to IPsec} = 0.006(\textit{\# of calls}) \tag{5}$$

which predicts a very small MOS decline over an 802.11n WLAN due to securing G.711 calls with IPsec. MOS decline predicted by (5) is well within the upper bounds calculated by comparing MOS CIs between WPA2 and WPA2+IPsec scenarios at every call level.

4.4 Objective VoIP WLAN Measurements

In order to investigate the cause of the anomalous MOS results for the 802.11g WLAN, objective measurements of VoIP traffic performance are collected. The open-source network packet analyzer Wireshark is used to report statistics on packet loss, maximum latency, mean jitter, and maximum jitter. Wireshark can only provide these statistics if it can interpret the RTP stream of a VoIP call, so when IPsec is implemented Wireshark can no longer readily analyze VoIP traffic. Therefore 10 simultaneous calls encrypted with WPA2 is selected as the scenario to evaluate the statistical traffic differences between 802.11b, 802.11g, and 802.11n WLANs. Thirty-four additional 60-second calls are analyzed for each 802.11 extension to collect the following data. The number of objective trials is selected to be greater than 30 in order to give a reasonable expectation of distribution normality. Calls are 60 seconds long in order to scrutinize WLAN performance for anomalies over a significant length of time.

For all three 802.11 extensions, the average packet loss is less than 1%. This value is well below the maximum packet loss of 3% advised by NIST [NIS05].

Figure 25 shows mean jitter during a 60-second call for the three WLANs. Mean jitter is low for all three 802.11 extensions, with means less than 6 ms (85% CI). Although the mean jitters are all low, mean jitter over the 802.11g WLAN is significantly

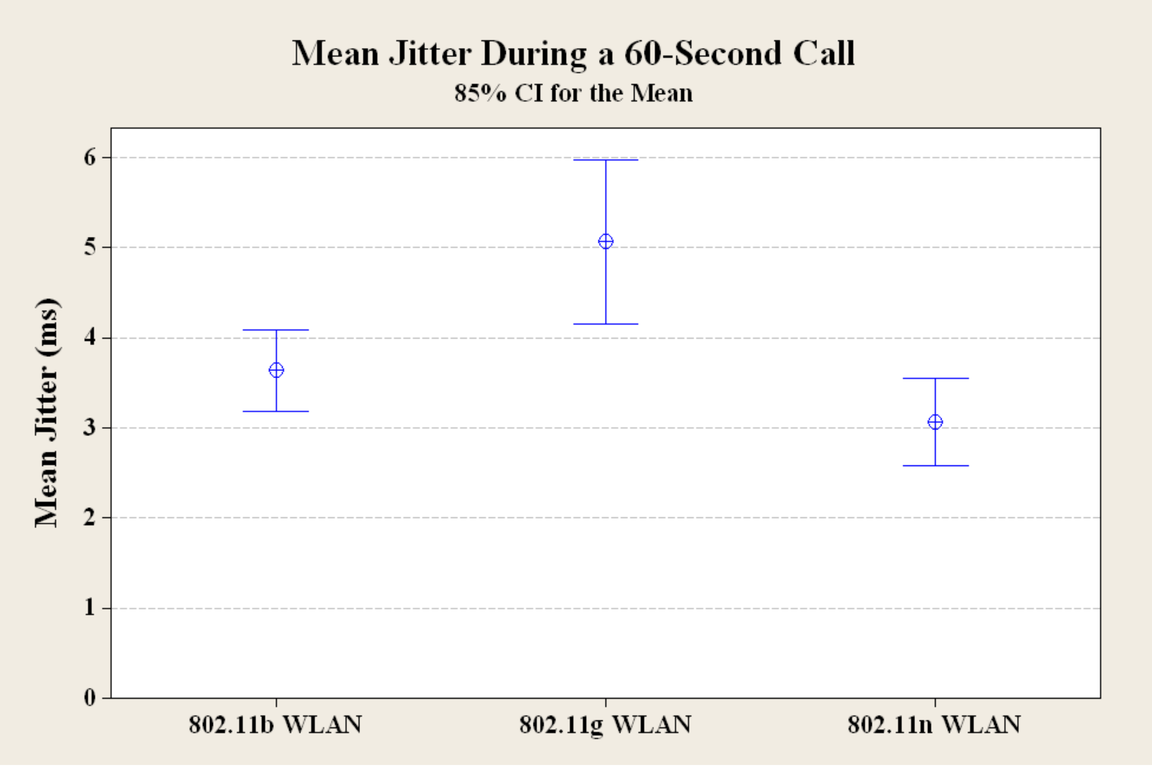


Figure 25. Mean Jitter During a 60-Second Call over the 802.11 WLANs

greater (85% CI) than the 802.11b and 802.11n WLANs. The results over 802.11b and 802.11n are statistically equivalent. The 802.11g data has a wider CI due in part to two calls where mean jitter is over 14 ms. The mean jitter over 802.11b and 802.11n does not exceed 10 ms. This data is evidence that the 802.11g WLAN is more affected by RFI than the other WLANs, since in every other respect the WLANs are identical.

Figure 26 shows maximum jitter during a 60-second call over the three 802.11 extensions. The data indicate that maximum jitter over 802.11g is significantly higher than for both 802.11b and 802.11n. High variability in maximum jitter over 802.11g, including 4 calls where jitter exceeds 50 ms, results in a noticeably wider 85% CI. Average maximum jitter performance over the 802.11n WLAN is measurably superior to the 802.11b WLAN, albeit by a short time of between 1.4 and 4.4 ms (85% CI).

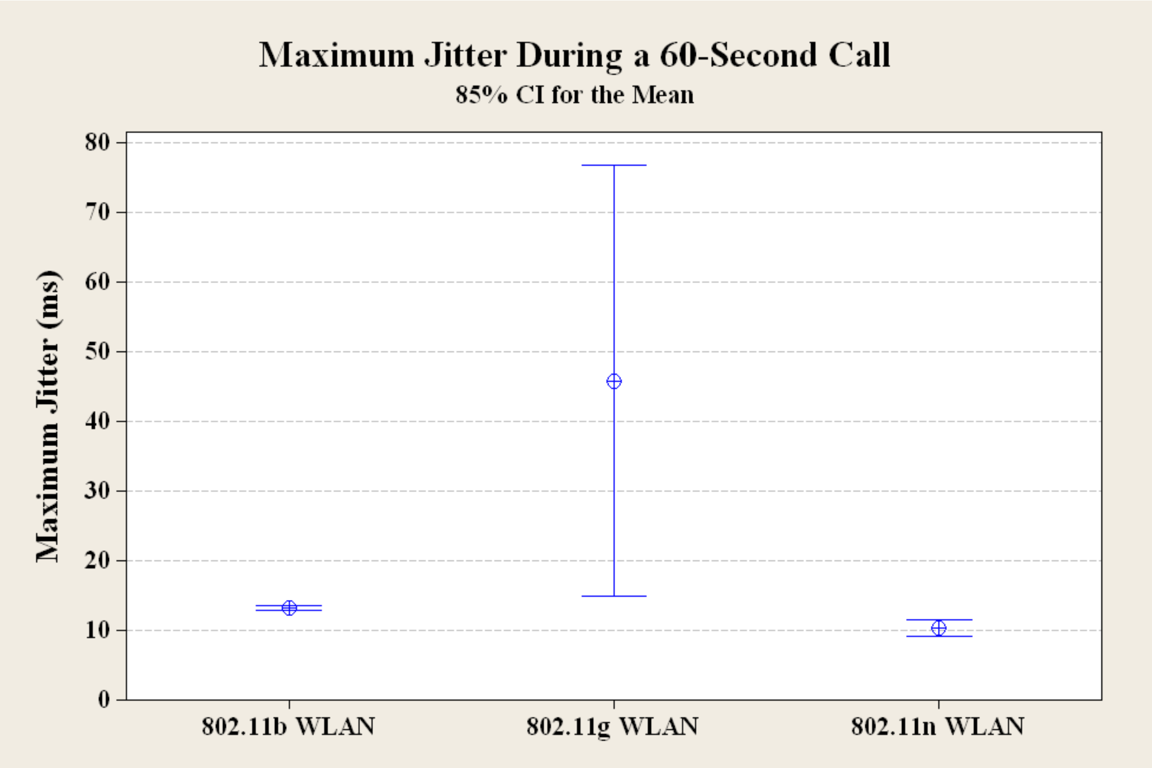


Figure 26. Maximum Jitter During a 60-Second Call over the 802.11 WLANs

Figure 27 shows the maximum latency during a 60-second call over the three 802.11 WLANs. As with the jitter measurements, average maximum latency over the 802.11g WLAN is greater than over both 802.11b and 802.11n by at least 100 ms (85% CI). Maximum latency over the 802.11b and 802.11n WLANs are not statistically different from one another and are both below 111 ms (85% CI). Objective measurements of jitter and latency therefore support the subjective MOS data that suggest equivalent audio quality of 802.11b and 802.11n when 10 simultaneous G.711 calls are present.

From the objective data it is evident that the 802.11g WLAN is significantly affected by RFI, while the 802.11b and 802.11n WLANs are not. In turn, this RFI randomly lowers subjective audio quality scores for 802.11g WLAN scenarios. The most likely cause of this interference is the presence of other 802.11g APs around the

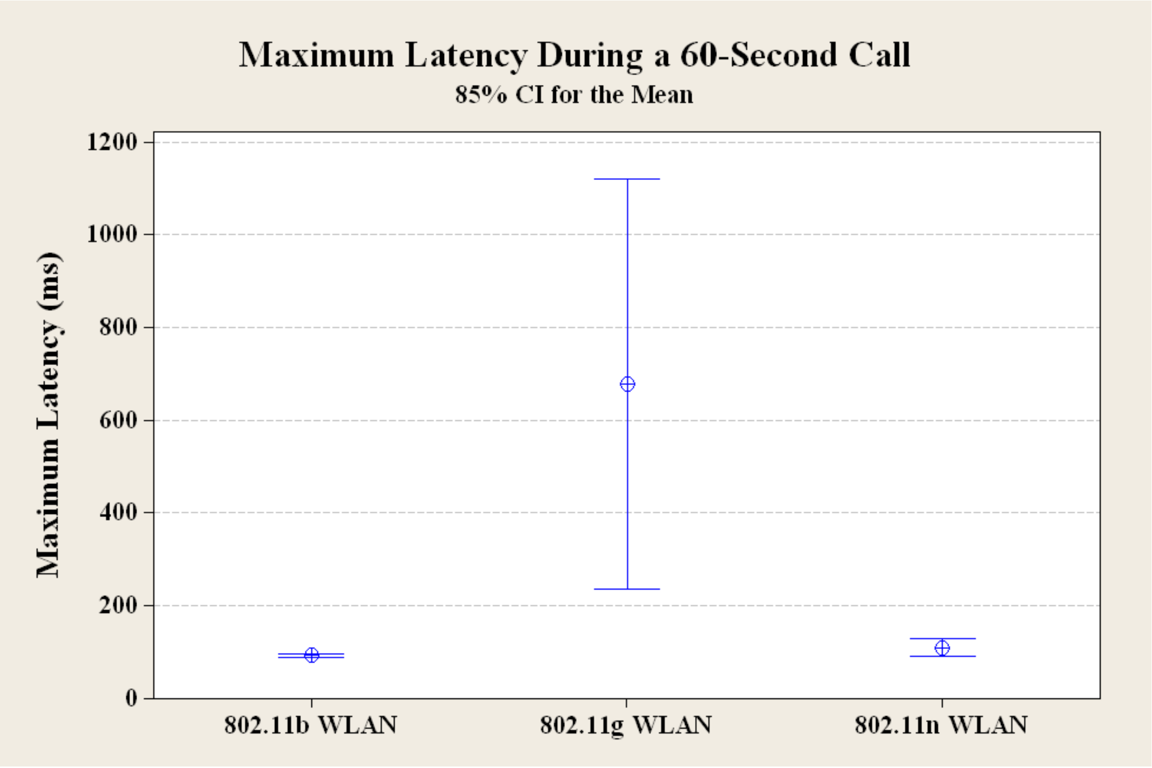


Figure 27. Maximum Latency During a 60-Second Call over the 802.11 WLANs

experimental facility. The different wireless encoding scheme of 802.11b (DSSS) and the wider channel width of 802.11n (40 MHz) make those two WLANs less vulnerable to interference from nearby 802.11g (OFDM) APs.

4.5 Summary

This chapter presents an analysis of the subjective and objective data collected for this experiment. The subjective audio scores are validated and compared to fulfill the experimental goals. Finally, objective measurements of WLAN VoIP traffic are presented to explain the anomalous MOS results from the 802.11g WLAN.

V. Conclusions and Recommendations

5.1 Introduction

This chapter summarizes the overall conclusions of the research. Section 5.2 presents conclusions derived from the experimental data. Section 5.3 discusses the significance of this research. Finally, Section 5.4 outlines recommendations for future research.

5.2 Conclusions of Research

5.2.1 Goal 1: Determine if MOS is higher over 802.11n than 802.11b or 802.11g

Subjective MOS over an 802.11n WLAN is shown to be no higher than over an 802.11b WLAN for up to 10 simultaneous G.711 VoIP calls. The large available throughput of 802.11n therefore does not result in higher subjective MOS than the much slower 802.11b extension under low WLAN utilization (< 2 Mbps). Subjective results from the 802.11g WLAN cannot be reliably compared to those from the 802.11n WLAN because interference affects from neighboring 802.11g APs is detected. Toll quality voice (MOS ≥ 4.0) is not attained using any of the three 802.11 extensions (85% CI).

5.2.2 Goal 2: Quantify MOS Decline over 802.11n Due to Additional Calls

Subjective MOS results from the 802.11n WLAN decline as the traffic increases from 2 to 20 simultaneous G.711 calls. Linear regression models of the subjective MOS results suggest that MOS decline is directly proportional to the number of simultaneous calls. These models are statistically consistent with the experiment results and predict that MOS falls below 3.0 (fair quality) during 75 simultaneous calls when WPA2 is used or 40 calls when WPA2 and IPsec are both used.

5.2.2 Goal 3: Quantify MOS Decline over 802.11n Due to IPsec

MOS decline due to IPsec is too small to be determined directly through subjective measurements at each call level. However, MOS decline due to IPsec is quantified by (5), which is the difference between MOS models (3) and (4). MOS decline due to IPsec is predicted to be only 0.006 for one call but increase to 0.24 during 40 simultaneous calls. This prediction is well within the upper bounds calculated by comparing subjective MOS results at each call level.

5.3 Significance of Research

WLANs and VoIP are both technologies with serious inherent security risks. Without strong encryption, such as WPA2, intruders can gain remote access into a corporate WLAN without needing to even step foot inside the building. However, WPA2 does not protect the privacy of VoIP calls from other authenticated clients on the WLAN. VoIP must therefore be encrypted separately from the WLAN, with protocols such as SRTP or IPsec.

This thesis presents models, based on and verified by subjective trials, which quantify call quality over the new draft 802.11n standard. MOS decline due to IPsec is also quantified. The research suggests that VoIP over 802.11n WLANs is not appropriate for DoD applications where toll quality voice (MOS $\geq$ 4.0) is essential. The research also suggests that IPsec encryption does not cause a dramatic decline in audio quality and is a viable method for securing WLAN VoIP calls. The results from this thesis validate E-model predictions by Nascimento, et al [NPM+06] for MOS over an 802.11b WLAN and should also be used to validate future models and simulations of audio quality over secure 802.11n WLANs as well.

5.4 Recommendations for Future Research

This thesis examines audio quality over AP-based WLANs, but more complex topologies such as multi-hop and ad hoc are also capable of supporting VoIP calls. The impact of encryption on MOS over other WLAN topologies should be thoroughly investigated. Future work should also focus on quantifying the impact of encryption on MOS over other wireless protocols used for VoIP calls, such as Bluetooth (IEEE 802.15) and WiMax (IEEE 802.16).

G.711 is selected as the codec for this thesis research, but more compressed codecs are also popular for VoIP, including G.729. The impact of strong WLAN encryption on compressed codecs is likely to be more severe than on G.711 because of their greater sensitivity to packet loss and latency. Future research should quantify these effects.

Finally, this research examined WLANs that support only VoIP traffic. In workplace WLANs, it is possible that VoIP calls could share wireless bandwidth with other activity such as email and file transfers. The complex impact that network congestion and QoS have on WLAN audio quality warrants further study.

5.5 Summary

This chapter presents conclusions from the research. The research significance and recommendations for future work are also discussed.

Appendix A. Scenario Listening Orders by Subject

Trial I:

	Scenario Listening Order													
Subject #	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th
1	3	14	4	8	6	5	1	9	12	7	2	13	11	10
2	11	8	12	2	14	13	9	3	6	1	10	7	5	4
3	5	2	6	10	8	7	3	11	14	9	4	1	13	12
4	2	13	3	7	5	4	14	8	11	6	1	12	10	9
5	13	10	14	4	2	1	11	5	8	3	12	9	7	6
6	12	9	13	3	1	14	10	4	7	2	11	8	6	5
7	6	3	7	11	9	8	4	12	1	10	5	2	14	13
8	9	6	10	14	12	11	7	1	4	13	8	5	3	2
9	1	12	2	6	4	3	13	7	10	5	14	11	9	8
10	7	4	8	12	10	9	5	13	2	11	6	3	1	14
11	4	1	5	9	7	6	2	10	13	8	3	14	12	11
12	14	11	1	5	3	2	12	6	9	4	13	10	8	7
13	8	5	9	13	11	10	6	14	3	12	7	4	2	1
14	10	7	11	1	13	12	8	2	5	14	9	6	4	3
15	1	10	11	13	9	4	8	6	14	3	12	2	7	5
16	12	7	8	10	6	1	5	3	11	14	9	13	4	2
17	14	9	10	12	8	3	7	5	13	2	11	1	6	4
18	9	4	5	7	3	12	2	14	8	11	6	10	1	13
19	13	8	9	11	7	2	6	4	12	1	10	14	5	3
20	6	1	2	4	14	9	13	11	5	8	3	7	12	10
21	3	12	13	1	11	6	10	8	2	5	14	4	9	7
22	5	14	1	3	13	8	12	10	4	7	2	6	11	9
23	8	3	4	6	2	11	1	13	7	10	5	9	14	12
24	10	5	6	8	4	13	3	1	9	12	7	11	2	14
25	2	11	12	14	10	5	9	7	1	4	13	3	8	6
26	11	6	7	9	5	14	4	2	10	13	8	12	3	1
27	7	2	3	5	1	10	14	12	6	9	4	8	13	11
28	4	13	14	2	12	7	11	9	3	6	1	5	10	8
29	1	2	5	10	4	1	7	6	3	12	8	14	11	9
30	14	3	6	11	5	2	8	7	4	13	9	1	12	10
31	12	1	4	9	3	14	6	5	2	11	7	13	10	8
32	2	5	8	13	7	4	10	9	6	1	11	3	14	12
33	6	9	12	3	11	8	14	13	10	5	1	7	4	2
34	1	4	7	12	6	3	9	8	5	14	10	2	13	11
35	5	8	11	2	10	7	13	12	9	4	14	6	3	1
36	9	12	1	6	14	11	3	2	13	8	4	10	7	5
37	11	14	3	8	2	13	5	4	1	10	6	12	9	7
38	10	13	2	7	1	12	4	3	14	9	5	11	8	6
39	8	11	14	5	13	10	2	1	12	7	3	9	6	4
40	4	7	10	1	9	6	12	11	8	3	13	5	2	14
41	3	6	9	14	8	5	11	10	7	2	12	4	1	13
42	7	10	13	4	12	9	1	14	11	6	2	8	5	3

Trial II:

	Scenario Listening Order						
Subject #	1st	2nd	3rd	4th	5th	6th	7th
43	16	18	19	15	20	17	21
44	21	16	17	20	18	15	19
45	18	20	21	17	15	19	16
46	17	19	20	16	21	18	15
47	20	15	16	19	17	21	18
48	19	21	15	18	16	20	17
49	15	17	18	21	19	16	20
50	16	17	19	18	21	20	15
51	17	18	20	19	15	21	16
52	18	19	21	20	16	15	17
53	21	15	17	16	19	18	20
54	15	16	18	17	20	19	21
55	20	21	16	15	18	17	19
56	19	20	15	21	17	16	18
57	16	20	17	19	21	15	18
58	15	19	16	18	20	21	17
59	21	18	15	17	19	20	16
60	18	15	19	21	16	17	20
61	20	17	21	16	18	19	15
62	17	21	18	20	15	16	19
63	19	16	20	15	17	18	21
64	19	21	15	17	18	16	20
65	21	16	17	19	20	18	15
66	18	20	21	16	17	15	19
67	16	18	19	21	15	20	17
68	20	15	16	18	19	17	21
69	17	19	20	15	16	21	18
70	15	17	18	20	21	19	16
71	19	20	18	21	15	17	16
72	16	17	15	18	19	21	20
73	15	16	21	17	18	20	19
74	18	19	17	20	21	16	15
75	21	15	20	16	17	19	18
76	20	21	19	15	16	18	17
77	17	18	16	19	20	15	21
78	16	21	18	17	19	15	20
79	19	17	21	20	15	18	16
80	21	19	16	15	17	20	18
81	18	16	20	19	21	17	15
82	20	18	15	21	16	19	17
83	15	20	17	16	18	21	19
84	17	15	19	18	20	16	21

Trial III:

	Scenario Listening Order						
Subject #	1st	2nd	3rd	4th	5th	6th	7th
85	23	25	26	22	27	24	28
86	28	23	24	27	25	22	26
87	25	27	28	24	22	26	23
88	24	26	27	23	28	25	22
89	27	22	23	26	24	28	25
90	26	28	22	25	23	27	24
91	22	24	25	28	26	23	27
92	23	24	26	25	28	27	22
93	24	25	27	26	22	28	23
94	25	26	28	27	23	22	24
95	28	22	24	23	26	25	27
96	22	23	25	24	27	26	28
97	27	28	23	22	25	24	26
98	26	27	22	28	24	23	25
99	23	27	24	26	28	22	25
100	22	26	23	25	27	28	24
101	28	25	22	24	26	27	23
102	25	22	26	28	23	24	27
103	27	24	28	23	25	26	22
104	24	28	25	27	22	23	26
105	26	23	27	22	24	25	28
106	26	28	22	24	25	23	27
107	28	23	24	26	27	25	22
108	25	27	28	23	24	22	26
109	23	25	26	28	22	27	24
110	27	22	23	25	26	24	28
111	24	26	27	22	23	28	25
112	22	24	25	27	28	26	23
113	26	27	25	28	22	24	23
114	23	24	22	25	26	28	27
115	22	23	28	24	25	27	26
116	25	26	24	27	28	23	22
117	28	22	27	23	24	26	25
118	27	28	26	22	23	25	24
119	24	25	23	26	27	22	28
120	23	28	25	24	26	22	27
121	26	24	28	27	22	25	23
122	28	26	23	22	24	27	25
123	25	23	27	26	28	24	22
124	27	25	22	28	23	26	24
125	22	27	24	23	25	28	26
126	24	22	26	25	27	23	28

Appendix B. Request for Exemption from Human Experimentation Requirements

20 August 2008

MEMORANDUM FOR AFIT IRB REVIEWER

FROM: AFIT/ENG (Dr. Mullins)
2950 Hobson Way
Bldg 640
Wright-Patterson AFB, OH 45433-7765

SUBJECT: Request for exemption from human experimentation requirements (32 CFR 219, DoDD 3216.2 and AFI 40-402) for Research on Secure Wireless Voice Networks

1. The purpose of this study is to determine how Internet Protocol Security (IPSec) and the number of simultaneous Voice over Internet Protocol (VoIP) calls on a wireless network affect the perceived voice call quality. A group of 42 volunteers listen to voice recordings taken from the wireless network and rate the call quality of each on a scale from 1 (bad) to 5 (excellent).

2. This request is based on the Code of Federal Regulations, title 32, part 219, section 101, paragraph (b) (2) Research activities that involve the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior unless: (i) Information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) Any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

3. The following information is provided to show cause for such an exemption:

a) Equipment and facilities: The wireless VoIP network consists of 10 laptop computers with wireless cards, a wireless router, and microphone headsets. Data collection takes place in building 640, room 245. Subjects listen to the audio samples in the AFIT library.

b) Subjects: Subjects are 42 volunteers from among the approximately 750 personnel assigned to AFIT. Recruitment is performed via an email invitation to AFIT personnel.

c) Timeframe: The study takes place from 20 August 2008 to 31 December 2008.

d) Data collected: This study will not collect personal identifiers or specific demographic information. Subjects are presented with 14 voice recordings taken from actual traffic on the wireless network. Upon listening to each recording, the subject rates the call quality on a scale from 1 (bad) to 5 (excellent). The instructions given to subjects are included as Attachment 1.

e) Risks to Subjects: Subjects could be exposed to uncomfortably loud recordings if the volume is not set correctly. This threat is mitigated by calibrating the volume before each test.

f) Informed consent: All subjects are self-selected to volunteer to participate in the interview. No adverse action is taken against those who choose not to participate. Subjects are made aware of the nature and purpose of the research, sponsors of the research, and disposition of the survey results.

g) If a subject's future response reasonably places them at risk of criminal or civil liability or is damaging to their financial standing, employability, or reputation, I understand I am required to immediately file an adverse event report with the IRB office.

4. If you have any questions about this request, please contact 1st Lt Benjamin Ramsey – phone (919) 604-0956; e-mail – benjamin.ramsey@afit.edu

BARRY E. MULLINS, Ph.D., P.E.
Assistant Professor of Computer Engineering
Principal Investigator

Appendix C. Human Subject Information Sheet

Subject Information Sheet

For Research on Encrypted Wireless Voice Networks

You are invited to participate in a research study on secure wireless voice networks. This research is to be conducted by 1st Lt Benjamin Ramsey, USAF. This research is in partial fulfillment of a Master's Degree program at the Air Force Institute of Technology (AFIT). The objective of this study is to determine ways to improve secure voice communication links over wireless networks.

You will hear a series of voice recordings from a wireless communication network, and will evaluate your impression of the voice quality by answering the following question about each recording:

LISTENING QUALITY SCALE

Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

Your participation is COMPLETELY VOLUNTARY. However, your input is important to improving analytical models for wireless secure voice capacities. You may withdraw from this study at any time without penalty, and your personal information will not be used in the research. Your decision to participate or withdraw will not jeopardize your relationship with your organization, the Air Force Institute of Technology, the Air Force, or the Department of Defense.

If you have any questions about this request, please contact 1st Lt Benjamin Ramsey at benjamin.ramsey@afit.edu

Appendix D. IPsec Configuration Files for Linux

The following configuration files are used to enable transport-mode IPsec on the trixbox PBX. The following four files (*psk.txt*, *setkey.config*, *raccoon.log, and ipsecstart*) are placed in the */etc/raccoon* directory after ipsectools has been installed. IPsec is enabled by the command *"./ipsecstart"* from the raccoon directory.

psk.txt

```
# Private Shared Keys for IPsec with 20 softphone laptops

# IP address             KEY

192.168.0.10   123456789012345678901234
192.168.0.11   123456789012345678901234
192.168.0.12   123456789012345678901234
192.168.0.13   123456789012345678901234
192.168.0.14   123456789012345678901234
192.168.0.15   123456789012345678901234
192.168.0.16   123456789012345678901234
192.168.0.17   123456789012345678901234
192.168.0.18   123456789012345678901234
192.168.0.19   123456789012345678901234
192.168.0.20   123456789012345678901234
192.168.0.21   123456789012345678901234
192.168.0.22   123456789012345678901234
192.168.0.23   123456789012345678901234
192.168.0.24   123456789012345678901234
192.168.0.25   123456789012345678901234
192.168.0.26   123456789012345678901234
192.168.0.27   123456789012345678901234
192.168.0.28   123456789012345678901234
192.168.0.29   123456789012345678901234
```

setkey.config

```
# IPsec rules for incoming and outgoing communication
# IP address of PBX is 192.168.0.2
# IP address range of softphone laptops is 192.168.0.10-29

flush;
spdflush;

#192.168.0.10
spdadd 192.168.0.2 192.168.0.10 any -P out ipsec
     esp/transport//require;
spdadd 192.168.0.10 192.168.0.2 any -P in ipsec
     esp/transport//require;

#192.168.0.11
spdadd 192.168.0.2 192.168.0.11 any -P out ipsec
     esp/transport//require;
spdadd 192.168.0.11 192.168.0.2 any -P in ipsec
     esp/transport//require;

#192.168.0.12
spdadd 192.168.0.2 192.168.0.12 any -P out ipsec
     esp/transport//require;
spdadd 192.168.0.12 192.168.0.2 any -P in ipsec
     esp/transport//require;


#192.168.0.13
spdadd 192.168.0.2 192.168.0.13 any -P out ipsec
     esp/transport//require;
spdadd 192.168.0.13 192.168.0.2 any -P in ipsec
     esp/transport//require;

                              .
                              .
                              .


#192.168.0.29
spdadd 192.168.0.2 192.168.0.29 any -P out ipsec
     esp/transport//require;
spdadd 192.168.0.29 192.168.0.2 any -P in ipsec
     esp/transport//require;
```

racoon.conf

```
path pre_shared_key “/etc/racoon/psk.txt” ;

remote anonymous
{
     exchange_mode main ;
     proposal {
          encryption_algorithm 3des ;
          hash_algorithm sha1 ;
          authentication_method pre_shared_key ;
          dh_group 2 ;
          }
}

sainfo anonymous
{
     encryption_algorithm 3des ;
     authentication_algorithm hmac_sha1 ;
     compression_algorithm deflate ;
}
```

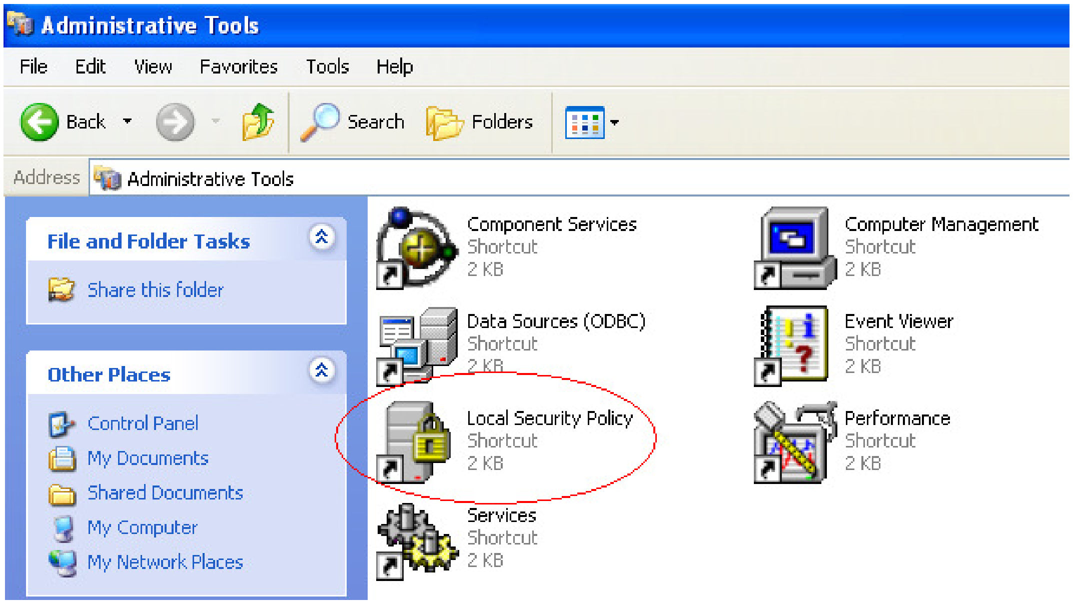
ipsecstart

```
#!/bin/sh

clear
ifconfig eth0 mtu 1400
killall racoon
sleep 1s
chmod 0600 /etc/racoon/psk.txt
setkey -F
setkey -FP
setkey -f /etc/racoon/setkey.config
setkey -D
setkey -DP
racoon -f /etc/racoon/racoon.conf
```

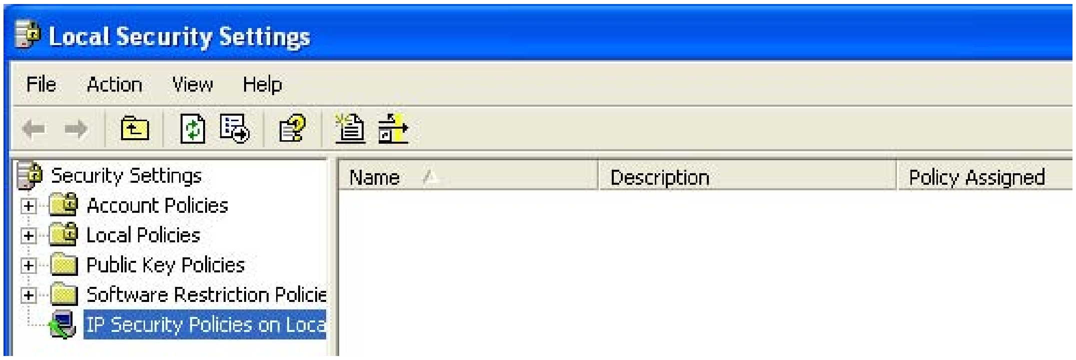
Appendix E. IPsec Configuration Guide for Windows XP

This guide demonstrates how transport mode 3DES IPsec is configured for the softphone laptops running Windows XP.

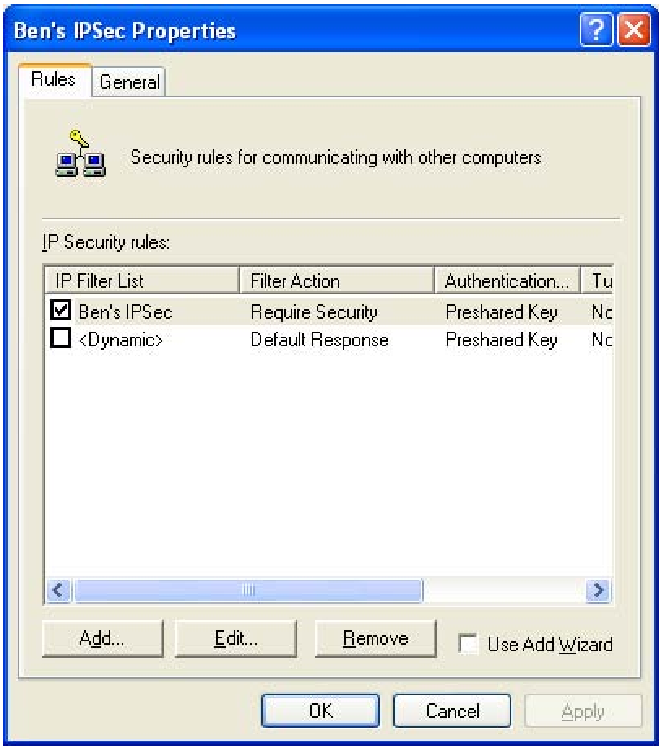
1. Open *Administrative Tools* and click on *Local Security Policy*


Administrative Tools
File Edit View Favorites Tools Help
Back Search Folders
Address Administrative Tools
File and Folder Tasks
Share this folder
Other Places
Control Panel
My Documents
Shared Documents
My Computer
My Network Places
Component Services Shortcut 2 KB
Computer Management Shortcut 2 KB
Data Sources (ODBC) Shortcut 2 KB
Event Viewer Shortcut 2 KB
Local Security Policy Shortcut 2 KB
Performance Shortcut 2 KB
Services Shortcut 2 KB

2. Create a new policy



3. Select *Require Security* and *Preshared Key* under the security rules of the new policy

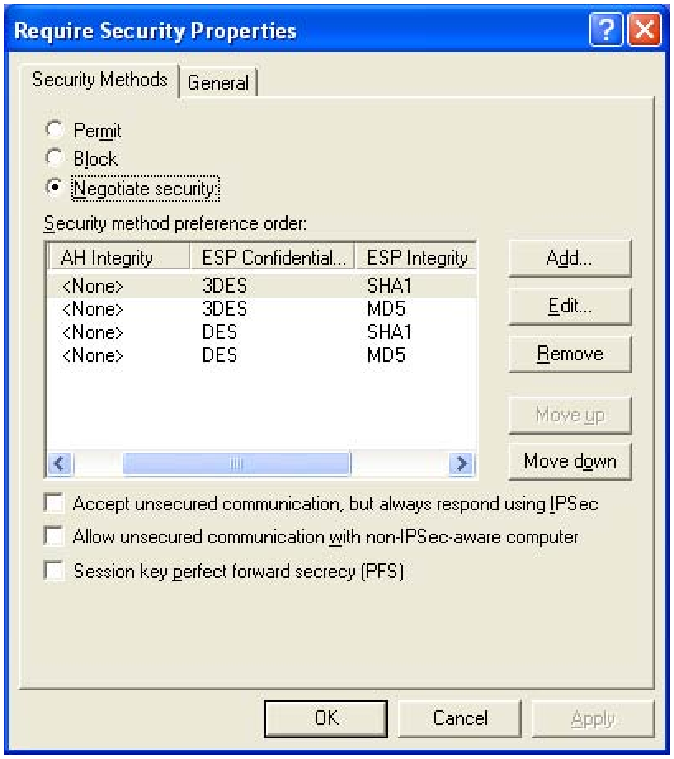


Security rules for communicating with other computers

IP Security rules:

IP Filter List	Filter Action	Authentication...	Tu
☑ Ben's IPSec	Require Security	Preshared Key	Nc
□ <Dynamic>	Default Response	Preshared Key	Nc

4. Select *3DES* for ESP Confidentiality and *SHA1* for ESP Integrity



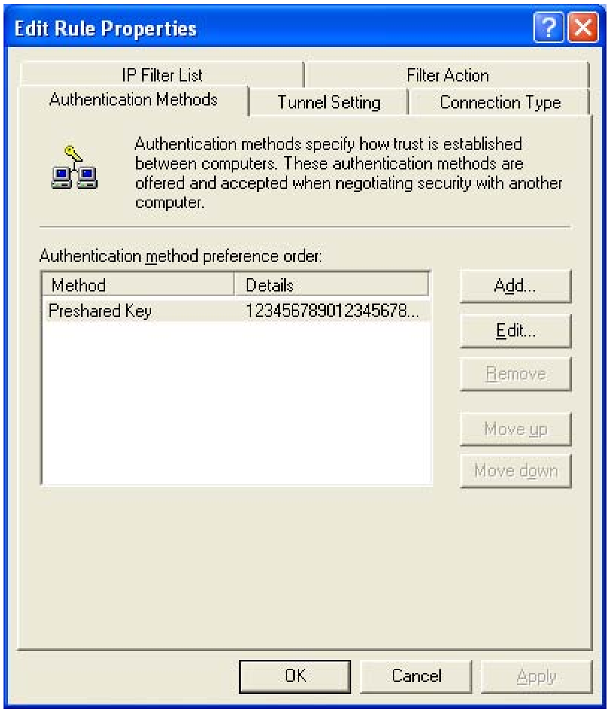
○ Permit
○ Block
● Negotiate security:

Security method preference order:

AH Integrity	ESP Confidential...	ESP Integrity
<None>	3DES	SHA1
<None>	3DES	MD5
<None>	DES	SHA1
<None>	DES	MD5

□ Accept unsecured communication, but always respond using IPSec
□ Allow unsecured communication with non-IPSec-aware computer
□ Session key perfect forward secrecy (PFS)

5. Enter the 24-character *Preshared Key*

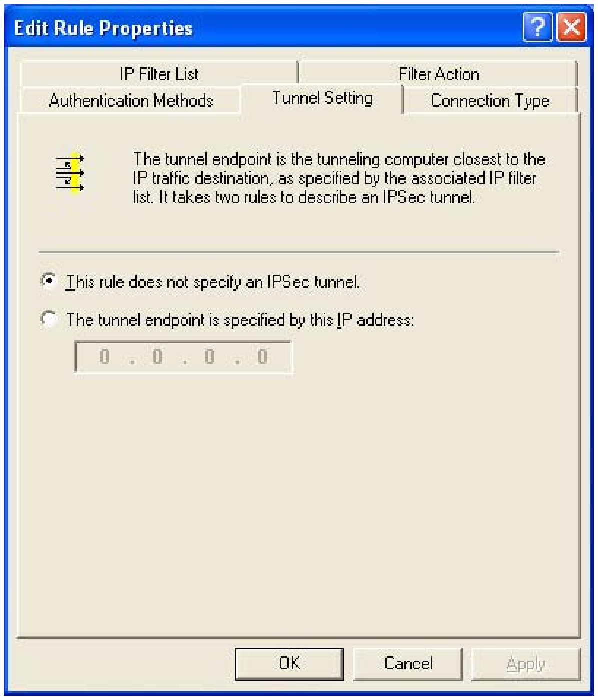


Authentication methods specify how trust is established between computers. These authentication methods are offered and accepted when negotiating security with another computer.

Authentication method preference order:

Method	Details
Preshared Key	123456789012345678...

6. Transport mode IPsec does not require a specified IPsec tunnel.



The tunnel endpoint is the tunneling computer closest to the IP traffic destination, as specified by the associated IP filter list. It takes two rules to describe an IPSec tunnel.

● This rule does not specify an IPSec tunnel.

○ The tunnel endpoint is specified by this IP address:

0 . 0 . 0 . 0

7. Start IPsec by assigning the newly created policy under *Local Security Settings*.

Appendix F. Raw Subjective Audio Quality Scores

Scenario #																											
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**
4	3	4	3	4	4	2	3	3	2	4	3	5	4	5	3	4	5	4	4	5	4	4	4	3	2	2	4
3	3	2	2	2	3	2	5	3	3	4	4	4	5	3	3	2	3	2	3	3	4	2	3	3	2	3	4
4	4	4	3	4	5	4	5	4	4	3	4	5	5	4	5	3	4	3	4	5	4	4	5	4	3	3	5
4	2	3	3	3	2	4	3	1	4	2	1	4	5	4	4	4	5	3	3	4	4	3	4	1	3	3	4
3	3	4	4	3	3	3	4	3	5	4	4	5	5	4	3	5	3	5	4	5	4	5	3	4	4	4	5
3	4	3	3	3	4	3	3	3	4	3	4	5	5	2	2	4	2	1	2	5	3	4	4	4	3	4	4
4	5	2	5	4	3	3	4	4	3	3	3	5	5	4	4	5	4	4	4	4	2	3	4	3	5	4	5
4	2	3	2	2	3	3	2	3	4	3	3	3	4	3	2	4	3	3	3	4	5	2	3	3	3	4	4
1	3	3	3	3	4	3	3	3	3	3	4	4	4	5	5	4	3	3	4	5	3	2	4	4	3	3	5
4	1	5	2	3	3	2	2	3	1	2	2	4	4	3	4	4	3	4	4	5	4	3	3	3	4	4	5
3	3	3	4	4	2	3	4	2	3	2	4	4	5	3	2	4	3	3	2	3	3	4	4	3	5	4	5
4	3	3	2	3	3	2	1	2	2	3	2	3	3	4	3	4	4	5	2	5	4	5	4	5	4	3	4
3	3	4	3	4	4	5	5	3	4	5	4	4	5	4	3	4	4	5	4	5	4	3	3	5	5	2	4
3	3	2	3	5	4	3	4	3	4	4	4	5	4	3	2	3	3	2	2	3	3	4	3	5	4	4	4
3	4	2	3	3	4	3	2	2	2	2	3	4	5	4	4	4	4	5	3	5	4	4	5	5	3	3	5
4	4	3	3	4	5	3	3	4	5	3	4	5	5	4	3	3	4	2	2	3	5	5	5	5	5	5	5
5	2	3	4	4	4	3	4	2	2	2	2	5	3	4	5	5	4	4	3	4	4	3	3	3	2	4	4
4	3	5	3	5	3	4	2	4	2	2	3	5	4	4	5	3	4	4	4	5	3	5	1	4	1	3	5
4	2	3	4	4	4	3	3	3	4	3	4	4	5	3	4	4	3	4	3	3	4	5	4	5	4	3	4
4	3	2	3	3	3	5	4	3	4	3	5	4	3	2	3	4	3	4	3	4	4	3	4	1	3	3	4
5	3	4	4	2	4	3	3	4	5	5	4	5	5	3	4	4	4	4	5	5	5	2	3	3	3	4	5
5	4	4	4	4	4	5	5	4	4	5	4	5	5	3	2	4	3	2	2	4	4	3	4	3	2	1	5
4	4	3	4	4	3	3	4	3	4	5	4	5	5	3	3	5	4	3	5	4	4	3	3	3	4	4	4
4	3	3	3	3	3	4	3	3	4	3	4	4	4	1	2	1	3	2	4	5	3	3	3	4	5	3	4
5	4	2	2	3	4	2	3	4	3	4	3	5	5	3	4	2	4	3	3	4	4	5	4	5	4	3	5
4	3	3	3	3	3	3	3	3	3	3	3	4	4	4	5	5	3	4	3	5	4	3	4	1	3	3	4
4	3	4	3	3	3	3	3	2	4	3	3	4	3	3	4	4	3	4	3	5	5	2	3	3	3	4	5
4	4	3	5	3	3	3	4	3	4	3	4	4	4	3	4	3	4	3	3	5	4	4	4	3	2	2	5
4	4	4	4	4	5	4	3	4	4	3	4	5	5	3	4	3	4	4	4	5	3	3	2	4	3	1	3
1	3	3	4	1	1	4	3	1	3	4	3	4	4	3	4	4	3	3	3	3	1	2	3	4	3	1	4
5	4	3	4	3	4	4	4	4	4	4	4	5	5	4	4	4	5	3	4	5	3	4	3	3	3	3	4
4	3	3	3	3	3	5	4	4	3	4	3	5	5	3	5	4	4	4	3	5	2	3	4	2	3	3	4
4	5	4	3	4	4	3	4	3	4	2	3	5	5	2	3	2	4	3	3	3	4	2	4	1	3	3	4
4	4	5	2	3	3	3	3	4	5	3	3	2	4	4	3	4	4	4	3	4	1	5	4	3	2	4	5
4	5	5	4	5	5	4	4	4	5	4	4	5	5	4	4	4	4	3	3	4	1	4	4	3	3	3	4
2	5	5	4	4	3	3	5	3	3	4	2	5	4	4	4	5	3	3	3	4	2	5	3	4	4	3	4
3	5	3	2	3	3	4	3	2	2	4	3	5	5	2	3	3	3	4	3	4	4	4	3	3	1	4	4
1	3	3	2	4	3	4	2	1	4	2	5	4	5	3	3	1	4	3	3	2	3	3	4	2	2	5	5
2	3	4	3	3	3	4	2	2	2	3	3	4	4	3	4	2	4	4	4	5	3	4	5	5	4	4	5
5	4	4	4	4	4	5	3	4	5	4	5	5	5	4	5	3	4	3	4	5	4	2	4	3	3	3	4
4	4	4	5	4	4	4	3	3	4	3	3	5	4	4	3	2	3	5	4	5	4	4	4	4	4	4	5
2	4	3	4	5	3	5	2	3	4	4	3	5	5	3	4	4	4	3	4	5	4	3	4	3	3	3	4

Appendix G. Raw Objective Measurements of Wireless VoIP Traffic Performance

802.11b WLAN			802.11g WLAN			802.11n WLAN		
Max Latency	**Max Jitter**	**Mean Jitter**	**Max Latency**	**Max Jitter**	**Mean Jitter**	**Max Latency**	**Max Jitter**	**Mean Jitter**
78.59	11.57	2.78	69.41	6.23	1.47	274.04	21.41	5.42
80.46	13.87	3.07	72.13	8.14	1.89	71.68	6.64	1.68
86.63	12.02	2.51	1925.34	517.35	15.56	71.81	6.73	1.79
84.45	13.56	4.25	1925.34	517.35	15.56	79.14	7	1.73
84.8	16.97	9.7	7148	68.53	7.48	81.03	12.36	6.02
79.92	14.51	4.7	76.39	7.78	1.68	78.29	13.54	8.04
81.76	12.28	2.75	68.49	6.27	1.78	70.64	7.13	1.83
93.89	12.95	2.57	74.1	7.51	2.03	74.47	6.62	1.73
99.3	14.92	4.11	81.83	12.4	6.18	361.28	22.33	2.42
143.64	16.25	6.3	87.99	11.49	4.51	79.01	11.67	7.22
98.83	12.53	2.53	81.57	11.85	7.21	93.51	11.23	3.79
104.44	13.92	2.59	70.43	7.69	2.3	77.01	7.23	2.04
80.89	12.01	2.9	77.46	11.11	4.63	74.2	7.39	1.8
92.02	11.84	2.71	77.9	8.69	2.94	74.04	6.95	1.88
112.57	12	2.51	77.8	8.75	2.42	74.78	7.01	1.88
84.23	13.1	2.69	75.02	6.82	1.79	71.96	6.83	1.8
116.25	12.9	2.55	77.58	6.94	1.95	73.69	8.15	2.06
103.84	13.12	3.34	71.18	7.59	1.95	74.26	11.44	7.03
82.2	12.73	5.34	83.55	11.17	3.77	76.81	10	3.82
102.82	12.9	3.12	76.57	11.57	6.89	68.59	6.26	1.82
76.28	12.92	2.72	76.65	7.12	1.89	72.9	7.18	1.87
92.5	11.64	2.49	40.19	8.97	4.56	71.53	6.5	1.91
81.5	12.11	2.52	2223.33	140.14	7.6	75.2	6.75	1.9
104.01	12.15	2.4	89.84	14.06	4.41	81.56	13.36	6.77
81.26	12.22	2.37	95.16	11.77	5	198.97	17.37	2.86
110.3	11.92	2.5	82.37	12.59	5.36	76	9.14	4.64
101.05	13.09	3.39	91.15	11.82	5.26	83.78	9.7	3.93
93.26	15.26	7.26	81.03	11.99	5.4	79.47	7.33	1.9
95.06	16.47	7.22	91.08	13.23	5.41	75.6	7.16	2.01
81.87	12.47	3.3	90.39	11.58	5.54	243.1	16.24	2.2
98.35	13.28	2.86	82.22	11.7	5.54	75.12	7.08	1.96
77.73	12.53	2.55	87	11.14	1.74	167.66	14.6	2.07
76.34	12.65	2.74	7278.97	17.34	9.75	189.16	14.03	2.1
104.53	15.57	6.37	378.39	20.14	10.68	256.4	17.41	2.18

Bibliography

[ABW06] F. Andeasen, M. Bauer, and D. Wing, "Session Description Protocol (SDP) Security Descriptions for Media Streams", RFC 4568 (2006). Available: http://www.faqs.org/rfcs/rfc4568.html.

[Baj03] C. Bajorek (2003), "R-Value vs. MOS," *Call Center Magazine,* Available: http://www.callcentermagazine.com/shared/printableArticle.jhtml?articleID=8701338.

[BHL06] A. Bittau, M. Handley, and J. Lackey, "The final nail in WEP's coffin," *2006 IEEE Symposium on Security and Privacy,* pp. 21-24, 2006.

[Bro03] Broadcom (2003), "IEEE 802.11g: The New Mainstream Wireless LAN Standard," Available: http://www.54g.org/pdf/802.11g-WP104-RDS1.pdf.

[Bro06] Broadcom (2006), "802.11n: Next-Generation Wireless LAN Technology," Available: http://www.broadcom.com/docs/WLAN/802_11n-WP100-R.pdf.

[Bro06a] S. Broom, "VoIP Quality Assessment: Taking Account of the Edge-Device," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.14, no. 6, pp. 1977-1983, 2006.

[BLG07] D. Butcher, X. Li, and J. Guo, "Security Challenge and Defense in VoIP Infrastructures," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews,* vol.37, no.6, pp. 1152-1162, 2007.

[Car06] J. Carmouche, *IPsec Virtual Private Network Fundamentals*. Indianapolis: Cisco Press, 2006.

[CFK06] A. da Conceicao, J. Li, D. Florencio, and F. Kon, "Is IEEE 802.11 ready for VoIP?," *Proceedings of the 2006 IEEE 8th Workshop on Multimedia Signal Processing,* pp. 108-113, 2006.

[Cod07] The Code Project (2007). Available: http://www.codeproject.com/KB/vista-security/ECDH/Diffie-Hellman.png.

[DRE+07] L. Ding, A. Radwan, M. El-Hennawey, and R. Goubran, "Performance Study of Objective Voice Quality Measures in VoIP," *12th IEEE Symposium on Computers and Communications,* pp. 197-202, 2007.

[FFL+07] E. Filho, P. Fonseca, M. Leitao, and P. Barros, "Security versus Bandwidth: The Support of Mechanisms WEP e WPA in 802.11g Network," *IFIP International Conference on Wireless and Optical Communications Networks,* pp. 1-5, 2007.

[FKC+05] H. Fathi, K. Kobara, S. Chakraborty, H. Imai, and R. Prasad, "On the impact of security on latency in WLAN 802.11b," *IEEE Global Telecommunications Conference,* vol.3, pp. 1752-1756, 2005.

[GaK03] S. Garg, and M. Kappes, "Can I add a VoIP call?," *Proceedings of the IEEE International Conference on Communications*, pp.779-783, 2003.

[Gas05] M. Gast, *802.11 Wireless Networks: The Definitive Guide*. Sebastopol: O'Reilly Media, 2005.

[GLC07] GL Communications (2007), "ITU Algorithms." Available: http://www.gl.com/ITUalgorithms.html.

[GuS07] P. Gupta, and V. Shmatikov, "Security Analysis of Voice-over-IP Protocols," *20th IEEE Computer Security Foundations Symposium,* pp. 49-63, 2007.

[GZA06] G. Gurkas, A. Zaim, and M. Aydin, "Security Mechanisms and Their Performance Impacts On Wireless Local Area Networks," *2006 International Symposium on Computer Networks,* pp. 1-5, 2006.

[IET04] RFC 3830 – MIKEY: Multimedia Internet KEYing (2004). Available: http://www.ietf.org/rfc/rfc3830.txt.

[IET04a] RFC 3711 – The Secure Real-time Transport Protocol (2004). Available: http://www.ietf.org/rfc/rfc3711.txt.

[ITU96] ITU-T Recommendation P.800 (1996). Available: http://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-P.800-199608-I!!PDF-E&type=items.

[ITU01] ITU-T Recommendation P.862 (2001). Available: http://www.itu.int/rec/T-REC-P.862-200102-I/en.

[ITU03] ITU-T Recommendation G.114 (2003). Available: http://www1.cs.columbia.edu/~andreaf/new/documents/other/T-REC-G.114-200305.pdf.

[LBS07] S. Lawrence, A. Biswas, and A. Sahib, "A comparative analysis of VoIP support for HT transmission mechanisms in WLAN," *27th International Conference on Distributed Computing Systems Workshops,* pp. 22-29, 2007.

[Min02] Mindspeed Technologies (2002), "Measuring Voice Quality," Available: http://www.mindspeed.com/web/download/download.jsp?docId=25026.

[MLL+07] A. Mowlaei, B. Lee, T. Lim, and C. Yeo, "Service Differentiation in Wireless Networks: Adaptive Ad Hoc Qs Versus IEEE 802.11e," *Proceedings of the 15th IEEE International Conference on Networks,* pp. 306-311, 2007.

[NPM+06] A. Nascimento, A. Passito, E. Mota, E. Nascimento, and L. Carvalho, "Can I add a secure VoIP call?," *International Symposium on World of Wireless, Mobile and Multimedia Networks,* pp. 26-29, 2006.

[NIS05] National Institute of Standards and Technology, *Security Considerations for VoIP Systems*, 2005. Available: http://csrc.nist.gov/publications/nistpubs/800-58/SP800-58-final.pdf.

[PMA+05] A. Passito, E. Mota, R. Aguiar, L. Carvalho, E. Moura, A. Briglia, and I. Bids, "Using an E-model implementation to evaluate speech quality in voice over 802.11b networks with VPN/IPSec," *Wireless Communications, Networking and Mobile Computing,* vol.2, pp. 1123-1127, 2005.

[RVB05] G. Rubino, M. Varela, and J. Bonnin, "Wireless VoIP at Home: Are We There Yet?," *Measurement of Speech and Audio Quality in Networks,* 2005.

[SAN04] SANS Institute (2004), "802.11i (How we got here and where we are headed)," Available: http://www.sans.org/rr/whitepapers/wireless/1467.php.

[Sey07] J. Seyba, "Voice and Video Capacity of a Secure Wireless System," Masters Thesis, Dept. of Elect. and Comp. Eng., Air Force Institute of Technology, Wright-Patterson AFB, OH, 2007.

[SIP02] Session Initiation Protocol – RFC3261 (2002) http://www.faqs.org/rfcs/rfc3261.html.

[Ver06] Veritest, *NETGEAR: Wireless N Performance Study*, 2006. Available: http://www.netgear.com/upload/marketingasset/competitivecharts/wireless_n_performance_study_12-12-06.pdf.

[VoI06] http://www.voipforo.com/en/SIP/SIP_example.php – 2006.

[Wal05] T. Wallingford, *Switching to VoIP*. Sebastopol: O'Reilly Media, 2005.

[Zim06] P. Zimmerman, "ZRTP: Extensions to RTP for Diffie-Hellman Key Agreement for SRTP" (2006). Available: http://www.tools.ietf.org/html/draft-zimmermann-avt-zrtp-01.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)* 20-03-2009	**2. REPORT TYPE** Master's Thesis	**3. DATES COVERED** *(From - To)* May 2007 - Mar 2009

4. TITLE AND SUBTITLE Subjective Audio Quality over a Secure IEEE 802.11n Draft 2.0 Wireless Local Area Network	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Benjamin W. Ramsey, Capt, USAF	**5d. PROJECT NUMBER** ENG 09 310
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GE/ENG/09-34
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Communications Agency / Dynamic Network Analysis Division Attn: Mr. Wade Farrar 203 W. Losey St. Scott AFB, IL 62225 (618) 229-6795; wade.farrar@us.af.mil	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFCA
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approval for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This thesis investigates the quality of audio generated by a G.711 codec and transmission over an IEEE 802.11n draft 2.0 wireless local are network (WLAN). Decline in audio quality due to additional calls or by securing the WLAN with Internet Protocol Security (IPsec) is quantified. Audio quality over an IEEE 802.11n draft 2.0 WLAN is also compared to that of IEEE 802.11b and IEEE 802.11g WLANs under the same conditions. Audio quality is evaluated as Mean Opinion Score (MOS), calculated as the average subjective audio quality score given for each WLAN configuration. Results suggest that audio quality over an IEEE 802.11n draft 2.0 WLAN is not higher than over an IEEE 802.11b WLAN when up to 10 simultaneous G.711 calls occur. A linear regression of the subjective scores also suggest that an IEEE 802.11n draft 2.0 WLAN can sustain an MOS greater than 3.0 (fair quality) for up to 75 simultaneous G.711 calls secured with WPA2, or up to 40 calls secured with both WPA2 and transport mode IPsec. The data strongly suggest that toll quality audio (MOS > 4.0) is not practical over WPA2-secured IEEE 802.11 WLANs.

15. SUBJECT TERMS

voice communications, network security, wireless computer networks, performance tests, information security

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Barry E. Mullins
U	U	U	UU	89	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-3636 x7979; barry.mullins@afit.edu

Reset