Performance Analysis of VoIP over WiMAX

Humberto Véjar Polanco1, Ernesto E. Quiroz M.2, Juan J. Tapia A.2
1 Universidad Tecnológica de Tijuana (UTT) Carr. libre Tijuana-Tecate km.10, Frace. El Refugio Quintas-C., Tijuana, B.C., 22253
Tel. +52 (664) 969-4700, Fax. +52 (664) 969-4700
humberto.vejar@uttijuana.edu.mx
2 Centro de Investigación y Desarrollo de Tecnología Digital (CITEDI-IPN) Avenida del Parque 1310, Mesa de Otay, Tijuana, B. C., 22510
Tel. +52 (664) 623-1344, Fax +52 (664) 623-1388
{eequiroz, jjtapia}@citedi.mx

ABSTRACT

The evolution of WiMAX by introducing Quality of Service provisioning for voice, and uplink/downlink rates in the range of 35/75 MHz respectively, have made it a strong contender to the fourth generation mobile technology LTE (Long Term Evolution). In spite of the constant increase in multimedia traffic carried by mobile networks, voice remains one of the most important sources of revenue to provide sustainability for service providers. This confers great importance to the knowledge of VoIP performance over WiMAX. However efficient transport is hampered by the small size of the frames of digitized voice, and the added volume of WiMAX and IP system's information necessary for VoIP transport. In this context, in this paper, a computational module that provides performance figures of merit of the capability of WiMAX for VoIP transportation is presented.

KEYWORDS: WiMAX, VoIP, 4G, Cell phone Communications.

1. INTRODUCTION

The demand for wireless broadband services is growing rapidly worldwide, in some places even beyond the capability of operators to provide it. According to the Federal Communications Commission of the United States, the use of smart phones has grown nearly 700% in the United States in the last four years [1]. Furthermore, the volume of traffic in AT&T's mobile network has increased by 5,000 % in the last 3 years [2]. One of the drivers of these digital consumption leaps is WiMAX (Worldwide Interoperability of Microwave Access), which is a WMAN (Wireless Metropolitan Area Network) type network, based on the family of broadband wireless access technologies IEEE 802.16. WiMAX is a strong contender to existing last mile access technologies: Cable, DSL (Digital Subscriber Line). [3], as well as LTE (Long Term Evolution: 4th generation UMTS cellular systems).

The mobile WiMAX air interfaces use OFDMA (Orthogonal Frequency Division Multiple Access) to improve multi-path interference in NLOS (Non-Line of Sight) and LOS (Line of Sight) environments, with a range of up to 50 km. Operates with asymmetric duplex transmission of 75 Mbps on the downlink and 35 Mbps uplink, provides high spectral efficiency (up to 2 bps/Hz), multi-channeling, and advanced MIMO (Multiple Input-Multiple Output) antenna technology.

To ensure worldwide application, WiMAX can use unlicensed and licensed spectrum, with variable bandwidth channels (12.5, 1.5, and 1.75 MHz multiples) up to a maximum of 20 MHz. 802.16 system access remains effective even in the presence of multiple connections per terminal, multiple levels of QoS (Quality of Service) per terminal, and a large number of users sharing the medium by statistical multiplexing.

The provision of QoS for real time services has also been integrated (2006, Release 1.0), with the aim to make it competitive with LTE systems. As a technology for broadband provision, WiMAX is applicable to both: subscribers in dense urban areas, and for scattered rural communities, and can be used as a backhaul for Wi-Fi cellular clusters. These characteristics have induced that beginning with the first commercial network in Korea in 2006, until September 2010, 592 WiMAX networks are operating in 149 countries, serving 13 million subscribers [4], number estimated to grow to 18 million in 2011 [5].

Based on the above, this work presents a computational module for WiMAX performance analysis in the transport of VoIP (Voice over IP: Voice over IP).

2. MOBILE WiMAX

WiMAX air interface technology is based on the standard IEEE 802.16 [6]. In particular, the current Mobile WiMAX technology derives from the IEEE 802.16e amendment approved by the IEEE in December 2005, which specifies the OFDMA air interface and provides support for mobility [7].

2.1. Mobile WiMAX Release 1.0

The Mobile WiMAX System Profile Release 1.0 [8] was developed in early 2006. It belongs to the family of WiMAX Forum standards, and was adopted by the ITU as the 6th air interface of the IMT-2000 family [9]. Support for the allocation of flexible bandwidth and integration of multiple types of QoS in WiMAX network, enables the provision of high-speed Internet access, VoIP, video sessions, multimedia chat and mobile entertainment. WiMAX issued a certification program to ensure interoperability of products from different manufacturers, achieving the first stamps of approval WiMAX Forum Certified for the 2.3 GHz spectrum in April 2008 and later for the 2.5 GHz spectrum.
The Mobile WiMAX Release 1.0 Profile is based on the IEEE air interface (Std. 802.16-2004, 802.16-2004, Cor. 1-2005, 802.16e-2005, 802.16-2004, Cor. 2) and WiMAX Forum's network specifications. Figure 1 shows the five sub-profiles and their components.

![System profile](image)

**Figure 1.- Structure of Mobile WiMAX system profile[7].**

### 2.2. Mobile WiMAX Release 1.5

WiMAX Forum works in the short-term migration to the profile called Release 1.5, which includes the following improvements:

FDD/HFDD efficient operation. FDD/HFDD (Half-duplex Frequency Division Duplex) operations optimization is based on dividing the 802.16 frame into partitions to be used by two different groups of mobiles having separate control channels, such as Uplinks MAPs and Downlink MAPs (downlink/uplink mapping), the fast feedback channel and HARQ ACK channel. This solution enables reuse of chipsets designed for version 1.0 (TDD) without compromising system performance to address FDD markets worldwide.

New bands. New classes of bands to provide a solution to the FDD transmission mode.

Improved MIMO. Closed loop MIMO operation and Beamforming (BF) further enhance the performance and coverage beyond version 1.0, which contains only open-loop MIMO capacity and some BF.

Improved MAC performance (specially improved VoIP capacity). Version 1.0 is highly optimized for data communications such as TCP/IP. The nature of data traffic implies transmission in "bursts". To adequately address this demand, Release 1.0 technology uses the mechanism of Downlink and Uplink MAPs, control messages transmitted in each frame, i.e. every 5 ms. While this is perfect for bursty traffic, support for the flow of data (VoIP, video) needs further optimization.

The idea of optimization is to use persistent allocation so that a simple MAP message provides information on the allocation of periodic resources matching the needs of a specific flow.

### 3. TDD, VoIP and Codecs framing

In order to analyze VoIP handling by WiMAX in Releases 1.0 and 1.5, the way digital information is generated in the voice codecs is explained, and the format in which this information is organized is described.

#### 3.1. Voice coding

Mobile WiMAX does not specify a preferred or base voice encoder. In this paper the AMR and ITU-T G.719 voice codec specifications are applied to carry on the performance analysis.

**AMR Speech Encoder.** The AMR speech codec is one of the standards adopted by the 3GPP for digitization of voice [10]. It is a variable rate encoder, which through an optimized link-adaptation mechanism, selects the best rate according to channel conditions and capacity. Every 20 ms produces one of 14 possible modes (Table 1, 3rd column), where each mode corresponds to a particular bit rate. The lower bit rate is used to transmit background noise during speech absence periods, and is known as Silence Indicator (SID).

<table>
<thead>
<tr>
<th>Mode</th>
<th>Total speech bits</th>
<th>Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMR 4.75</td>
<td>95</td>
<td>FD/HD</td>
</tr>
<tr>
<td>AMR 5.15</td>
<td>103</td>
<td>FD/HD</td>
</tr>
<tr>
<td>AMR 5.90</td>
<td>118</td>
<td>FD/HD</td>
</tr>
<tr>
<td>AMR 6.70</td>
<td>134</td>
<td>FD/HD</td>
</tr>
<tr>
<td>AMR 7.40</td>
<td>148</td>
<td>FD/HD</td>
</tr>
<tr>
<td>AMR 7.95</td>
<td>159</td>
<td>FD/HD</td>
</tr>
<tr>
<td>AMR 10.20</td>
<td>204</td>
<td>FD</td>
</tr>
<tr>
<td>AMR 12.20</td>
<td>244</td>
<td>FD</td>
</tr>
</tbody>
</table>

FD.-Full Duplex  HD.-Half Duplex

In this work we use a simplified On-Off model of the AMR speech codec. During periods of active conversation (Figure 2), the highest bit rate (244 b/20 ms) is used, and during periods of inactivity, the SID rate is used (56 b/160 ms).

**Table 1. AMR Data Rate**

**Figure 2.-Encoded AMR voice packet fields (33 bytes)**

G.719 speech codec. ITU-T specification G.719 describes a low complexity audio codec based on transformation, which operates at a 48 KHz sampling frequency and offers a complete audio bandwidth from 20 Hz to 20 kHz [11].
The coder processes 16-bit PCM linear input signals in 20 ms frames with a 40 ms average delay. G.719 allows for any rate between 32 Kbit/s and 88 Kbit/s in increments of 4 Kbit/s, and 88 Kbit/s to 128 Kbit/s in 8 Kbit/s steps. One byte of the Table of Contents (ToC) is added at the beginning of each frame of compressed audio, along with the frame's length information. Figure 3 shows the format of the G.719 data packet @ 32 Kbps.

<table>
<thead>
<tr>
<th>Table of Contents</th>
<th>G.719 Voice Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 byte</td>
<td>80 bytes</td>
</tr>
</tbody>
</table>

3.2. VoIP over IP

For real-time applications transport streams like Voice over IP (VoIP: Voice over IP) and video, packets are typically transported using the protocol stack RTP/UDP/IP (Real-Time Transport Protocol/User Datagram Protocol/Internet Protocol) [3]. Each protocol has an associated header, adding up 320 bits (Figure 4) or 40 bytes.

<table>
<thead>
<tr>
<th>Payload</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RTP</td>
<td>12 bytes</td>
</tr>
<tr>
<td>UDP</td>
<td>8 bytes</td>
</tr>
<tr>
<td>IP</td>
<td>20 bytes</td>
</tr>
</tbody>
</table>

This is a huge expense compared to the VoIP packet payload. To reduce the burden of the protocol header, wireless systems use a technique known as Robust Header Compression (ROHC) [3], whereby the header is reduced to 32 bits (Figure 5). Thus VoIP packets for AMR and ITU-T G719 are reduced to two fields, as shown in Figures 5 and 6.

<table>
<thead>
<tr>
<th>MAC Header</th>
<th>Payload (Optional)</th>
<th>CRC (optional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 bytes</td>
<td></td>
<td>4 bytes</td>
</tr>
</tbody>
</table>

3.3. MAC Frames

At the Medium Access Control (MAC) layer, WiMAX organizes information in Packet Data Units (PDU), as shown in Figure 7. Each MAC frame begins with a fixed length MAC header. This header may be followed by the load of the MAC PDU (MPDU). An MPDU may contain a field for Cyclic Redundancy Check (CRC).

Frame structure and allocation of Mobile WiMAX traffic. The 802.16e standard provides different physical layer modes and configurations of radio channel [13]. In the TDD mode, the data mapping is done in two dimensions: time and subcarriers, where an OFDMA frame has a duration of 5 ms subframe comprising a downstream (downlink: DL) subframe and an uplink (UpLink: UL) as shown in Figure 8 [3].
The number of PDUs that can be accommodated in these subframes depends on the modulation scheme, coding rate and channel quality factor (G).

4. COMPUTATIONAL MODULE

The application was developed on the Microsoft Visual Studio 2008 with the C# programming language [14]. The main assumptions adopted are as follows:

- Use the OFDM WiMAX TDD frame with duration of 5 ms, bandwidth of 5 and 10 MHz and the PUSC permutation mode (Partial Usage of the subchannels).
- Apply the operating characteristics of the voice codecs G.719 and AMR to generate voice traffic.
- To calculate the VoIP traffic load in both encoders applies a simplified On-Off model (on-off).
- For AMR, during periods of active conversation, it uses the highest bit rate, this is 244 bits/20 ms, and during periods of inactivity, a rate of 56 bits / 160 ms.
- For G.719, during periods of active conversation, using the highest bit rate, this is 160 bytes/20 ms, and during periods of inactivity rate 80 bytes/20 ms.

Figure 9 shows the capture screen and output of the module for the calculation of operating parameters and efficiency in the transport of VoIP.

The blank fields offer choices, these are detailed below: a) bandwidth (5 or 10 MHz), b) G parameter (1/4, 1/8, 1/16, 1/32), c) modulation schemes (BPSK, QPSK, 16QAM and 64QAM) d) Coding rate (1/2 CTC, 3/4 CTC 2/3 CTC, 5/6 CTC), e) Number of retransmissions (1, 2, 4), f) voice decoder (AMR, G.719), g) Detection of voice (detection or no detection), h) data rate voice packets (12.2 kbps, 10.2 kbps, 7.95 kbps, 7.40 kbps, 6.70 kbps, 5.90 kbps, 5.15 kbps, 4.75 kbps and 1.80 kbps for the AMR codec, 32 kbps and 64 kbps for the G.719 codec).

Pressing the Calculate button, the module calculates: a) The operating parameters of WiMAX, b) length (bytes) of a package of WiMAX Voice, c) the number of slots and OFDM symbols used by the voice packet, d) gross rate (bits/sec) transmission of voice, e) the rate for IP transport, f) rate of speech (without header), g) percent Efficiency, h) the percentage of occupation of a WiMAX frame.

The expressions for calculating the above data can be founded in [15].

Efficiency calculations

The most important formulas from the perspective of performance analysis are:

- Efficiency of WiMAX VoIP package is:
  \[ \text{Efficiency} = \frac{\text{Packet Size}}{\text{Voice Packet size adapted for WiMAX}} \]
- Occupation of a VoIP PDU TDD frame is:
  \[ \% \text{Occupancy} = \frac{\text{Packet Size adapted for WiMAX}}{\text{TDD Down Link Frame Size}} \]
- Maximum number of simultaneous calls a single TDD frame can support is:
  \[ \text{No. calls per frame} = \frac{1}{\text{Occupancy Rate}} \]

Substituting the appropriate values in the above formulas, we can calculate the efficiency, the occupancy rate and the number of calls per frame for both AMR to G.719.

5. PERFORMANCE ANALYSIS OF VOIP OVER WIMAX

Because WiMAX was originally conceived as a means of transporting data, WiMAX versions 1.0 and 1.5, show improvements in the treatment of VoIP, which is particularly important for several reasons: (a) The voice is a time service actual maximum tolerance delay of 200 msec. (b) Vocoder digital frames are very small (about 264 bits), compared to data services, leading to proportionally larger control headers compared to the payload.

The module "VoIP over WiMAX" calculates and displays the operating parameters according to Figure 8. To compute the performance of WiMAX systems for voice transportation, we focus on the following figures of merit.

![Figure 9.-Screen to enter the operating parameters of WiMAX](image-url)
- Frame occupation percentage (FOP): The ratio of the number of bits in a PDU (VoIP packet with WiMAX headers, IP headers and payload) to the total number of bits of a TDD frame.
- Packetization efficiency percentage (PEP): The ratio of the payload length (bits) to the PDU's total number of bits.
- Net VoIP Rate (NVR): Is the total number of voice and control bits sent in a one second period by the WiMAX system.
- Number of calls per frame (NCPF): Maximum capacity of VoIP calls a a TDD-DL frame can accommodate, if it were to transport only VoIP traffic. Obtained by dividing the total capacity of a TDD-DL frame by the FOP.

In the tables below FOP, PEP, NVR and NCPF calculations are presented according to the operational characteristics of 1.0 and 1.5 Releases, and AMR/ G.719 voice codecs. Given a bandwidth of 5 MHz for TDD-DL transmission, channel coding rate and G parameter are fixed. Table 2 lists the total capacities in Mbps and the resulting FOP when the modulation scheme is changed.

In the AMR case, it is clear that when upping the levels of modulation, there is a better occupancy of spectrum capacity, which means that the same VoIP rate adopts lower FOP figures. Results vary from 1.287% with BPSK to 0.214% with 64QAM. Similar behavior is observed for AMR in version 1.5, and G.719 in Releases 1.0 and 1.5.

On the other hand, the change in version 1.0 to 1.5 provides a slight gain by reducing the FOP in both AMR and G.719. Since G.719 is a full band audio encoder (up to 20 KHz), while AMR is a narrowband codec (up to 4 KHz), the coded frames of the former are longer, and therefore have a greater FOP.

PEP and NVR parameters are independent of the modulation scheme, and are shown in Table 3 for the same conditions as in Table 2.

The change from 1.0 to 1.5 in terms of transport efficiency is insignificant in both versions, as we can see comparing 32.35 vs. 35.04 and 41.41 vs. 42.68.

### Table 3. Percentage efficiency of bundling and VoIP Average Gross Rate

<table>
<thead>
<tr>
<th></th>
<th>AMR</th>
<th>G.719</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>PEP (%)</td>
<td>32.35</td>
<td>35.04</td>
</tr>
<tr>
<td>TBPV (bps)</td>
<td>11,490</td>
<td>10,050</td>
</tr>
</tbody>
</table>

PEP and NVR parameters are independent of the modulation scheme, and are shown in Table 3 for the same conditions as in Table 2.

As explained above, the improvement lies in that in 1.5 the initial packet carries WiMAX headers that identify the flow, and is omitted in all successive packets.

NVR values of Table 4 consider an activity/silence relationship of 60/40. It is found that Release 1.5 decreases 1,440 bps the data volume transmitted by both the AMR and the G.719 encoders compared to Release 1.0.

Maintaining the same encoding, bandwidth and G conditions, Tables 4 and 5 provide estimates of the number of calls per TDD frame for AMR and G.719 ITU-T, respectively.

### Table 4. Number of calls per TDD frame with the AMR encoder

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Data Rate (Mbps)</th>
<th>% Occup.</th>
<th>NLLM</th>
<th>% Occup.</th>
<th>NLLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>3.168</td>
<td>1.287</td>
<td>77</td>
<td>1.188</td>
<td>84</td>
</tr>
<tr>
<td>QPSK</td>
<td>6.336</td>
<td>0.643</td>
<td>155</td>
<td>0.594</td>
<td>168</td>
</tr>
<tr>
<td>16QAM</td>
<td>12.672</td>
<td>0.321</td>
<td>311</td>
<td>0.297</td>
<td>336</td>
</tr>
<tr>
<td>64QAM</td>
<td>19.008</td>
<td>0.214</td>
<td>467</td>
<td>0.198</td>
<td>505</td>
</tr>
</tbody>
</table>

Viewed from the perspective of call capacity, the TDD-DL frame could accommodate 77 simultaneous conversations in the AMR lower level of modulation case, up to 467 in the highest level. In this last comparison it is necessary to mention the fact that capacity is being measured in a single frame, which is not equivalent to the total capacity, as voice frames are generated with a 20 ms periodicity and TDD frames every 5 ms.

### Table 5. Number of calls per TDD frame with the G.719 encoder

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Data Rate (Mbps)</th>
<th>% Occup.</th>
<th>NLLM</th>
<th>% Occup.</th>
<th>NLLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>3.168</td>
<td>2.5</td>
<td>40</td>
<td>2.425</td>
<td>41</td>
</tr>
<tr>
<td>QPSK</td>
<td>6.336</td>
<td>1.25</td>
<td>80</td>
<td>1.212</td>
<td>82.5</td>
</tr>
<tr>
<td>16QAM</td>
<td>12.672</td>
<td>0.625</td>
<td>160</td>
<td>0.606</td>
<td>165</td>
</tr>
<tr>
<td>64QAM</td>
<td>19.008</td>
<td>0.417</td>
<td>239</td>
<td>0.404</td>
<td>247</td>
</tr>
</tbody>
</table>

The increase in the number of G.719 calls is less notable than in the AMR case, which is attributable to the denser digital volume of G.719 compared to that of AMR.

### Table 2. Occupancy rate of the TDD-DL frame for a VoIP package

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Encoding</th>
<th>G</th>
<th>Data Rate (Mbps)</th>
<th>POM (%)</th>
<th>POM (%)</th>
<th>POM (%)</th>
<th>POM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>½</td>
<td>1/8</td>
<td>3.168</td>
<td>1.287</td>
<td>1.188</td>
<td>2.5</td>
<td>2.425</td>
</tr>
<tr>
<td>QPSK</td>
<td>½</td>
<td>1/8</td>
<td>6.336</td>
<td>0.643</td>
<td>0.594</td>
<td>1.25</td>
<td>1.212</td>
</tr>
<tr>
<td>16QAM</td>
<td>½</td>
<td>1/8</td>
<td>12.672</td>
<td>0.321</td>
<td>0.297</td>
<td>0.625</td>
<td>0.606</td>
</tr>
<tr>
<td>64QAM</td>
<td>½</td>
<td>1/8</td>
<td>19.008</td>
<td>0.214</td>
<td>0.198</td>
<td>0.417</td>
<td>0.404</td>
</tr>
</tbody>
</table>

1. AMR @ 12.2 Kbps  2. G.719 @ 32 Kbps
6. CONCLUSIONS

- The VoIP_over_WiMAX module calculates the performance in the transport of VoIP for any combination of settings (bandwidth, modulation technique, channel condition, etc.) considered in the WiMAX standards 1.0 and 1.5. Includes AMR and G.719 ITU-T coders parameters. And the possibility of integrating other codec formats if necessary.
- The transport efficiency of VoIP over WiMAX 1.0 is 32.32, which is slightly improved by Release 1.5 to 35.04. A similar improvement is obtained in the case of the G.719 coder. The figures allow us to affirm that WiMAX is not efficient in terms of the relation payload to control headers. This situation can be improved with the addition of more voice frames in a single PDU, with the risk of increasing the BER in case of packet loss.
- The capacity calculation of calls that can be transported within the same TDD-DL frame shows the enormous flexibility and capability of WiMAX, since its base number is 77 and can grow up to 467 (with conditions indicated in Section V).
- Issues to consider in the improvement of this tool include the calculation of loading and efficiency of VoIP sessions at cell level. Consider the effect of VoIP packet loss and eventual re-transmission.

Acknowledgments. This work was supported by the IPN through grant SIP-IPN 2010-0060 and COFAA Exclusivity Scholarship.

7. REFERENCES