A Measurement Study of Speex VoIP and H.264/AVC Video over IEEE 802.16d and IEEE 802.11g

Kostas Pentikousis, Jarno Pinola, Esa Piri, and Frerk Fitzek
VTT Technical Research Centre of Finland
Kaitoväylä 1, FI-90571 Oulu, FINLAND
Email: firstname.lastname@vtt.fi

Abstract

We evaluate the capacity of an IEEE 802.16d and IEEE 802.11g testbed to simultaneously carry emulated H.264/AVC video and Speex VoIP and present results from an extensive measurement study. First, we employ two fixed WiMAX subscriber stations and one base station and report results for packet loss and one-way delay under both line-of-sight and non-line-of-sight conditions. In addition, we put these results in perspective by repeating the experiments using an off-the-self IEEE 802.11g router. In particular, we use the IEEE 802.11g access point as an extension to the WiMAX network, a case often considered in simulation studies. Finally, we consider the IEEE 802.11g access network in isolation, and measure its capacity to carry video streams and emulated bidirectional VoIP calls simultaneously and compare it with the WiMAX access network.

1 Introduction

Despite the significant interest in WiMAX technology [1, 2] and deployment (see www.wimaxforum.org), WiMAX equipment is yet to become readily available at affordable prices. In fact, there is lack of publicly reported measurements from testbeds and field trials. As such, most WiMAX studies employ simulation and modeling. This paper takes a different approach and reports testbed measurements, in an attempt to understand what is realistically possible using off-the-shelf fixed WiMAX equipment. In particular, we present results from a measurement study involving both video and voice over a fixed WiMAX testbed. We use two subscriber stations and one base station (BS) and measure performance in terms of the capacity of the WiMAX equipment to handle multiple VoIP flows while delivering a variable load of video streams. We report results for packet loss and delay under line-of-sight (LOS) and non-line-of-sight (NLOS) conditions.

Moreover, in order to put these results in perspective, we repeat the experiments using a COTS IEEE 802.11g router. We experiment with two cases. First, we use the IEEE 802.11g access point as an extension to the WiMAX network, a case often considered in simulation studies. Then, we consider the IEEE 802.11g access network in isolation and measure its capacity to carry video streams and emulated bidirectional VoIP calls simultaneously and compare its performance with the testbed WiMAX access network.

The rest of this paper is organized as follows. In Section 2, we relate our work to previous testbed measurement studies. Our testbed, used tools, and evaluation methodology are introduced in Section 3. We present our results in Section 4 and conclude this paper in Section 5.

2 Related Work

There is a significant amount of work on WiMAX- and VoIP-related topics that we could discuss and relate our work with, but due to space restrictions, we confine ourselves to the most recent results. Pioneering and closely related work to ours has been published by Scalabrino et al. [3, 4] from a fixed WiMAX testbed deployed in Turin, Italy. They focus on VoIP performance over WiMAX in particular when service differentiation is employed in the presence of significant amounts of elastic background traffic. Grondalen et al. [5] also report on fixed WiMAX field trial results, for TCP and UDP transfers from 15 different locations near Oslo, Norway, under both LOS and NLOS conditions. They measure RSSI and maximum throughput, and find that their WiMAX system can deliver 9.6 Mb/s to a single flow in the downlink. Employing the same modulation and FEC in our testbed, we measured an application-level throughput, or goodput, of 5.5 Mb/s for the fixed WiMAX uplink and 9.4 Mb/s for the downlink, using UDP bulk traffic with $<$0.1% packet loss under direct LOS conditions.

Grondalen et al. [5] do not study VoIP performance. However, they note that this throughput level is possible to attain at a distance of up to 5 km from the BS. To some degree, this indicates that the results presented in this paper might be applicable to outdoor environments as well. This, of course, needs to be verified in future work.

Ng et al. [6] study the performance of GSM 6.10-encoded VoIP traffic and consider the case of a WiMAX cell extended with a wireless LAN ("Wi-Fi") access point. Although they consider only a small number of emulated VoIP flows, they find that in such a topology, it is the Wi-Fi segment that is actually the bottleneck, not the WiMAX segment, at least for VoIP traffic. They propose the introduction of a voice gateway with a multiplexer to increase the capacity in terms of sustained VoIP flows. They call this scheme "multiplex-multicast" VoIP aggregation. However, although this proposed scheme is evaluated over a wireless LAN, it is not actually tested using real WiMAX equipment.

We recently presented results from the tesbed used in this study as well in [7] and [8]. In the former we empirically assessed the merits of VoIP aggregation both at the application layer and at the network layer, with the adoption of a network-side performance enhancing proxy. In the latter paper, we measured WiMAX performance under LOS and NLOS conditions. However, [8] does not compare WiMAX performance with that of Wi-Fi, a topic that was highlighted by reviewers. As such, this paper extends and complements [8], and provides further insights on wireless access networks. In addition, for this paper the setting of our testbed is changed from the previous measurements to resemble more that of a normal everyday scenario, where the data communication of the users inside a WiMAX cell involves remote hosts residing in a different network.

3 Methodology

Fig. 1 illustrates our experimental facility, comprising an Airspan MicroMAX-SoC fixed WiMAX BS, two subscriber stations (SS1 and SS2), and several PCs. SS1 is an Airspan EasyST and SS2 is an Airspan ProST. Symmetrically on the BS and SS sides, we connect GNU/Linux (kernel ver. 2.6.20-16, Ubuntu 7.04) PCs with Broadcom NetXtreme BCM5754 1 Gb/s Ethernet PCI cards. We use an IEEE 1588 Precision Time Protocol (PTP) server to synchronize all PC clocks, as explained below. All tests are performed in our laboratory where conditions are static, even though there can always be a degree of variance on a wireless link. For this reason, we monitor the WiMAX equipment to ensure that key parameters (see Fig. 1) remain unchanged during the entire duration of the tests. We experimented with both direct LOS conditions for both SSs and NLOS for SS1.

We employ JTG [9] to generate synthetic VoIP and trace-driven video traffic in our testbed topology. JTG is a simple, flexible, and configurable open source traffic generator which can be used in a command-line fashion in

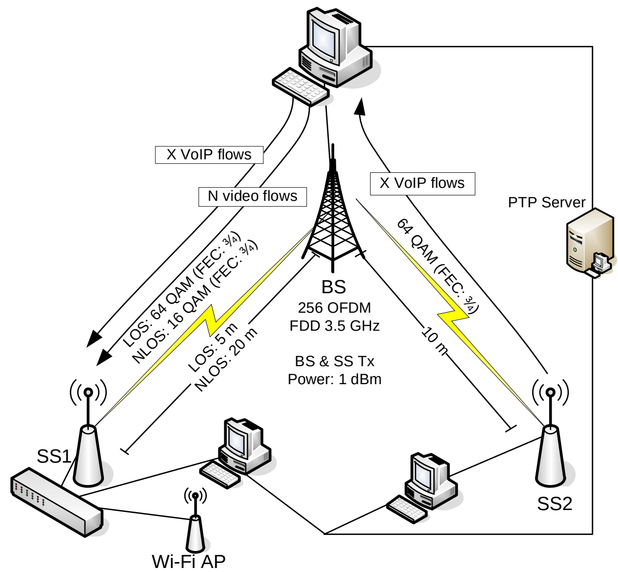

Figure 1. Schematic of the WiMAX testbed

GNU/Linux. Constant bit rate (CBR) traffic is generated by providing the bit rate and the packet size as parameters. JTG can also generate trace-driven traffic. For the tests reported in this paper, we use synthetic Speex [10] VoIP traffic and video traffic based on traces captured from a live IPTV H.264/AVC [11] video server.

3.1 Video and VoIP Traffic Generation

We captured 20 minutes of live IPTV unicast transmission and created a packet trace to be fed into JTG. The captured video stream is in H.264/AVC format [11]; the accompanying audio stream is encoded in MPEG-1 Audio Layer II. The IPTV content was streamed using the Darwin Streaming Server (DSS), an open source RTP/RTSP [12–14] server. The DSS host and the receiving host were set two hops away, with a Gigabit Ethernet switch in between. We collected the video stream packet trace at the receiver side using Wireshark (www.wireshark.org) and observed very low delay and delay variance and no packet loss. We used this packet trace to generate traffic with JTG. Since the IPTV transmission was from a popular music video TV channel, we configured DSS to stream the video at 512 kb/s (360×288, 25 f/s) and the audio at 192 kb/s, emphasizing audio over video quality.

The audio stream is effectively CBR traffic with the total packet size fixed at 620 bytes (including codec payload and RTP, UDP, and IP headers). The video stream, often with dramatic changes in scenery, visual effects and so on, has a variable bit rate. The video total packet sizes varied greatly, with the major mode at 1478 bytes. The video and audio parts of the IPTV stream are separated at the streaming server and are transmitted to different port numbers of the receiver. The collected packet trace had no RTSP message exchanges. The separation of audio and video traffic provides an excellent opportunity to study the effects of a congested fixed WiMAX link on IPTV audio and video separately and compare it with VoIP traffic.

We inject synthetic VoIP flows in all our tests. We chose to experiment with Speex [10], an open source variable bit rate audio codec specifically designed for speech compression in VoIP applications over packet switched networks. Speex can be used with three different sampling rates (narrowband at 8 kHz, wideband at 16 kHz, and ultra-wideband at 32 kHz), and has a large range of operational bitrates (2.15-44 kb/s). Speex uses Code Excited Linear Prediction (CELP) for encoding voice samples and is robust to packet loss. Due to its open source and good quality, a number of diverse voice applications use it, including Microsoft's Xbox Live and the U.S. Army Land Warrior system.

We chose to emulate multiple Speex VoIP flows with a wideband codec bitrate of 12.8 kb/s using JTG. For each VoIP flow, JTG generated 50 packets/s with 32 bytes of codec payload and an RTP header of 12 bytes which leads to an application level bitrate of 17.6 kb/s. After adding a total of 28 bytes of headers (UDP+IP), each JTG instance injects 28.2 kb/s of Speex CBR emulated traffic into the testbed. In our configuration, we were able to have 100 such flows in the uplink with negligible ($<0.1\%$) average packet loss (see also [8]). Due to the large header overhead, the cumulative application goodput measured across all flows in the SS to BS direction is only 1.76 Mb/s. This is merely 32% of the goodput achieved by a single UDP flow transmitting Ethernet MTU-sized packets.

3.2 Clock Synchronization

For high-precision delay measurements accurate clock synchronization is necessary, taking care of both absolute time and clock drift at different hosts in the network. When performing only round trip delay measurements, the critical aspect is clock drift. Lack of accuracy in the absolute time is not harmful. However for the kind of one-way delay (OWD) measurements we consider in this paper, both absolute time and clock drift are important. Often, the Network Time Protocol (NTP) [15] is the first choice for synchronizing the clocks at different hosts. NTP uses a hierarchical system of strata, which defines the distance to the reference clock and hence the accuracy. Stratum 0 contains external devices such as GPS-, atomic- or radio-clocks, which are connected to Stratum 1 hosts only, which define the master clock. Typically, one server acts as the master. Stratum 3 devices are end hosts that wish to synchronize their internal clocks with the master clock using request/reply message exchanges with the NTP server. Clearly, synchronization in Strata 0 and 1 is not a big issue. If every PC is equipped with a GPS clock, one can can easily achieve synchronization accuracy in the order of tens of μs. However, synchronization in Stratum 3 depends on the precision of NTP which is in the order of tens of ms.

In [8], we used two PCs with GPS-clocks and found that OWD on the BS-SS2 WiMAX link (Fig. 1) is 8.7 ms, on median, for the downlink and 23.5 ms, on median, for the uplink. Therefore NTP-based synchronization is not sufficient as the one-way delay is in the same order of magnitude as the measured values. Moreover, since we only have two GPS clocks, we cannot perform measurements with multiple VoIP and video clients/servers. As a result, we opted to employ the Precision Time Protocol (PTP) [16], which can synchronize COTS PC clocks with up to 10 μs difference and even better in ideal cases [17]. We used PTPd [18], an open-source implementation of the emerging standard as a software-only system without a Stratum 0 device and obtained clock synchronization precision of hundreds of μs or better (further details are omitted here due to space considerations, but are provided in [8]).

4 Results

We saturate the WiMAX uplink with synthetic Speex VoIP traffic, in order to quantify system performance and behavior at the limit of its capacity. Experimentation showed that the testbed fixed WiMAX uplink is capable of sustaining 100 Speex-encoded VoIP flows before packet drops start to occur. All 100 flows originate from the SS2 domain and via the BS terminate at the wired part of the testbed (see Fig. 1). This limit on uplink capacity is irrespective of whether one or two SSs are employed. That is, we can apportion the 100 flows evenly between PCs connected to SS1 and SS2, or split the same number of flows in any random manner between the two SSs while observing negligible packet loss. In [8] we showed that our fixed WiMAX equipment can sustain 50 emulated bidirectional Speex calls within the same cell (calls originate and terminate within the SS1/SS2 domains). In this paper, we have 100 unidirectional emulated Speex flows originating from the WiMAX segment (SS2) and another 100 unidirectional flows terminating at the (SS1) WiMAX access network.

In addition to the VoIP flows, we inject N streams of emulated synchronous H.264/AVC-encoded video and MPGA-encoded audio. We make 10 replications of the experiment for each N, which was gradually increased while keeping the VoIP traffic at the same level (100 flows). Each measurement lasts 60s and each time we start IPTV streaming from a different point in the captured trace. We report the measurements for each N using box plots.

Fig. 1 summarizes the test configuration. In the NLOS

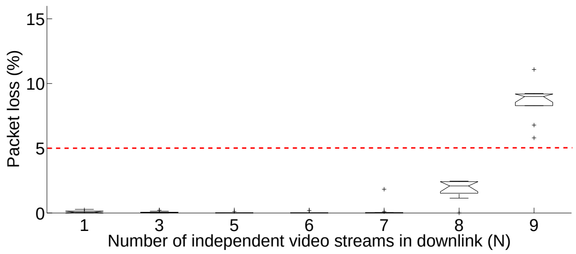

Figure 2. Packet loss for IPTV H.264/AVC video in WiMAX LOS

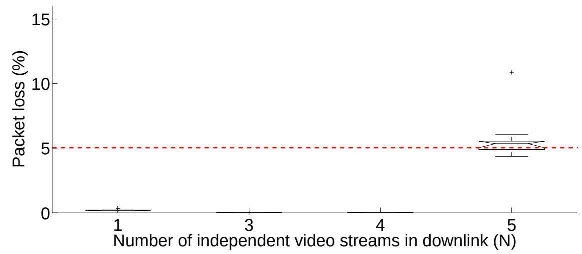

Figure 4. Packet loss for IPTV H.264/AVC video in WiMAX NLOS

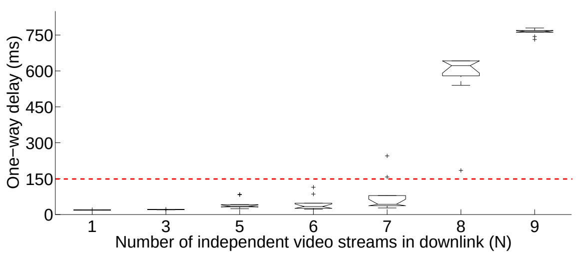

Figure 3. One-way delay for IPTV H.264/AVC video in WiMAX LOS

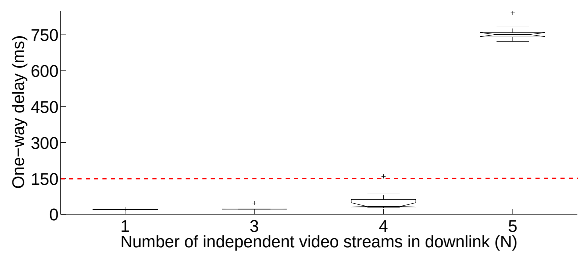

Figure 5. One-way delay for IPTV H.264/AVC video in WiMAX NLOS

measurements, the direct path from the BS to SS1 was blocked by two solid steel reinforced concrete walls and one concrete wall including a doorway with a wooden door. SS2 and BS were always in LOS. Before starting the measurements, we were interested in seeing whether any of the different traffic types (IPTV video/audio and VoIP) would be affected unevenly given the lack of traffic differentiation support in our WiMAX testbed. During the measurements, the uplink VoIP traffic was monitored and no excess packet losses or delays were observed, thus, only downlink measurements are presented below. After the measurements, it became evident that the variation in drop rates and delay between the different traffic types was marginal, and are thus omitted in this paper due to space restrictions. We instead focus our attention to the IPTV video stream results.

4.1 WiMAX LOS and NLOS

Fig. 2 illustrates the packet loss rates for N video streams. With $N = 8$ the median video packet loss rate is 2%, degrading the received video quality to some extent. For the accompanying audio stream, the median packet loss rate is slightly over 1%, which is not severe and can be concealed. Overall, the VoIP flows have to deal with a median loss rate of 2% (not shown), which Speex can handle without service degradation. Overall, and with respect to packet loss alone, when $N = 8$ our testbed can provide all application data to their receivers with an acceptable quality.

On the other hand, with respect to packet delay for video traffic (Fig. 3), the vast majority of inter-packet delays are less than 90 ms, for $N \leq 7$. For $N = 8$, the median jumps to >600 ms for all traffic types. Such delays can be acceptable only when receiver-side buffering can be used, and for near-real time applications. This is not the case for VoIP calls, which suffer severely. Effectively, for VoIP, although the packet loss rates are not prohibitive for $N = 8$, the one-way packet delays are too high for a satisfactory end-user experience.

Figures 4 and 5 illustrate the video stream packet loss and delay in our NLOS measurements. Again, 100 emulated Speex flows are injected in both directions of the WiMAX access network, as with the LOS measurements above. With respect to total bitrate, packet loss and delay, the results are consistent with what was measured under LOS conditions. As expected, fewer video streams can be supported and the deteriorating effects when capacity is exceeded are more rapid. Nevertheless, although the BS uses 16 QAM modulation, due to NLOS and the associated decreased signal quality, up to $N = 4$ IPTV streams can be handled with excellent quality.

4.2 Extending WiMAX with Wi-Fi

Using fixed WiMAX to wirelessly backhaul traffic from a number of Wi-Fi access points (APs) is an often cited scenario. In order to empirically study such a topology, in particular for VoIP and IPTV traffic, we connected a COTS IEEE 802.11g AP directly to SS1 (see Fig. 1). The mea-

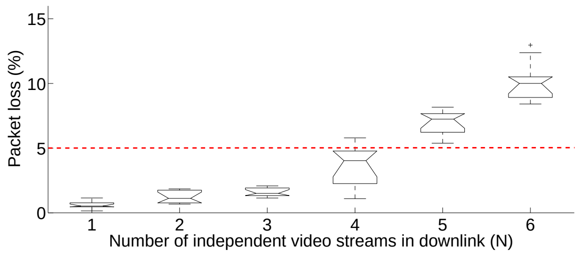

Figure 6. Packet loss for IPTV H.264/AVC video in Wi-Fi extended WiMAX

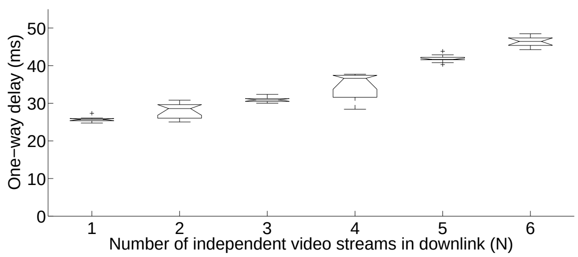

Figure 7. One-way delay for IPTV H.264/AVC video in Wi-Fi extended WiMAX

surement PCs were then connected to the AP over IEEE 802.11g. The distance between the end host and the AP was one meter and all VoIP and IPTV traffic went through the Wi-Fi link. Such a short distance between the AP and the end hosts defines a "best case scenario" with relatively static signal conditions. The received signal strength as reported by the PC IEEE 802.11g-compatible network card was -29 dBm. In other words, for these experiments, we replace the Ethernet switch between SS1 and the measurement PCs with an IEEE 802.11g AP and repeat the WiMAX LOS measurements described in the previous subsection.

First, we identify the "break-point" for the new setup. We found that we can inject only 75 Speex unidirectional flows in the uplink with negligible loss. Then, we continue with the mixed VoIP/IPTV measurements by injecting $X = 75$ VoIP flows and gradually increasing the number of IPTV streams (N), until the packet loss rate exceeds 10%.

By comparing Figures 2 and 6, we conclude that in this set of measurements it is the Wi-Fi link that is the bottleneck, not the WiMAX backhaul. As we saw above, without the Wi-Fi extension, our fixed WiMAX can sustain $N = 7$ IPTV streams and 100 VoIP flows. With the Wi-Fi extension, only $N = 3$ IPTV streams and 75 VoIP flows can be sustained with negligible loss. The cutoff loss rate of 10% is exceeded with $N = 6$ IPTV streams which is a third less than in the Ethernet-extended WiMAX LOS experiments. We could argue that for such a setup our fixed WiMAX testbed can backhaul traffic from at least two IEEE 802.11g APs for average traffic loads.

The IEEE 802.11 MAC layer uses Carrier Sensing Multiple Access/Collision Avoidance (CSMA/CA), which is well-known to underperform when a large amount of small VoIP packets are injected into the access network. Fixed WiMAX, even with the best effort profile employed in our testbed, appears to be more efficient for VoIP, despite the significant difference in nominal bitrates between the two technologies. In 802.11, packets wait in the AP's buffer as carrier sensing is performed before every transmission. This can cause buffer overflows when a large number of closely-spaced packets arrive. Of course, if larger packet payloads are the norm, 802.11 can demonstrate a much larger effective capacity, closer to its nominal value.

Interestingly though, as illustrated in Fig. 7, the Wi-Fi extension does not significantly affect the end-to-end one-way delay, even when the AP gets severely congested. This implies a small AP transmission buffer. This is in contrast with the sharp increase in one-way delay (see Fig. 3) when crossing the capacity threshold for the Ethernet-extended WiMAX experiments. In Fig. 7, the end-to-end one-way-delay remains below 50 ms even as the drop rate exceeds 10%. An important difference between IEEE 802.11 and 802.16 is the link resource allocation process. In the latter, the uplink and downlink capacities are differentiated and transmission is tightly coordinated by the BS. In 802.11g, the total capacity of an AP is shared between all data streams, without differentiating between uplink and downlink, and there are no bandwidth guarantees.

Finally, to gather further comparison points, we empirically measured the capacity of the COTS Wi-Fi AP to carry Speex VoIP with negligible loss in another two scenarios. First, we used two PCs, instead of one, both associated with the same AP and found that only a total of $X = 42$ unidirectional flows can be sustained. This is a significant decrease when compared to the $X = 75$ flows when a single host transmits towards the AP. Then, we considered the case where the Speex flows originate and terminate in the same Wi-Fi AP cell. In this case, only $X = 22$ bidirectional VoIP flows can be sustained with negligible loss. This confirms that the total capacity of the IEEE 802.11g AP is shared evenly between competing sources. It also provides evidence that, although our measurements were conducted with synthetic traffic and only a handful of PCs, our results should be applicable, at least to some extent, when multiple hosts with real VoIP traffic are considered.

5 Conclusion

We evaluated IPTV streaming and VoIP over a fixed WiMAX testbed with simultaneous use of two subscriber stations in a single WiMAX cell, using both Ethernet and Wi-Fi extensions. We experimented with both LOS and NLOS conditions and reported accurate one-way delay and packet loss measurements from independent replications. We found that our fixed WiMAX testbed can cope with VoIP and IPTV traffic more efficiently than a COTS Wi-Fi access point, and that it can backhaul such traffic from two Wi-Fi APs proficiently, despite the disparity between the nominal capacity rates of the two technologies. We expect that our measurement results will be of interest to both network practitioners and simulationists.

Acknowledgement

This work was conducted at the VTT Converging Networks Laboratory within the framework of the IST Sixth Framework Programme Integrated Project WEIRD (IST-034622), which is partially funded by the Commission of the European Union. Study sponsors had no role in study design, data collection and analysis, interpretation, or writing the report. The views expressed do not necessarily represent the views of VTT, the WEIRD project, or the Commission of the European Union.

References

[1] IEEE 802.16 Working Group. *IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems*. IEEE Std. 802.16-2004, October 2004.

[2] IEEE 802.16 Working Group. *IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Amendment 2: Physical and Medium Access Control Layer for Combined Fixed and Mobile Operation in Licensed Bands*. IEEE Std. 802.16e-2005, December 2005.

[3] Nicola Scalabrino, Francesco De Pelegrini, Imrich Chlamtac, Andrea Ghittino, and Sandro Pera. Performance evaluation of a WiMAX testbed under VoIP traffic. In *Proc. ACM WiNTECH*, pages 97–98, September 2006.

[4] Nicola Scalabrino, Francesco De Pellegrini, Robert Riggio, Andrea Maestrini, Cristina Costa, and Imrich Chlamtac. Measuring the quality of VoIP traffic on a WiMAX testbed. In *Proc. TRIDENTCOM*, pages 1–10, May 2007.

[5] Ole Grondalen, Pal Gronsund, Tor Breivik, and Paal Engelstad. Fixed WiMAX field trial measurements and analyses. In *Proc. 16th IST Mobile and Wireless Communications Summit*, pages 1–5, July 2007.

[6] Ping Chung Ng, Soung Chang Liew, and Chinlon Lin. Voice over wireless LAN via IEEE 802.16 wireless MAN and IEEE 802.11 wireless distribution system. In *Proc. IEEE WirelessCom*, July 2005.

[7] Kostas Pentikousis, Esa Piri, Jarno Pinola, Frerk Fitzek, Tuomas Nissilä, and Ilkka Harjula. Empirical evaluation of VoIP aggregation over a fixed WiMAX testbed. In *Proc. Fourth International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TRIDENTCOM)*, Innsbruck, Austria, March 2008.

[8] Kostas Pentikousis, Jarno Pinola, Esa Piri, and Frerk Fitzek. An experimental investigaion of VoIP and video streaming over fixed WiMAX. In *Proc. Fourth International Workshop on Wireless Network Measurements (WiNMee)*, Berlin, Germany, March 2008.

[9] J. Manner. *Jugi's Traffic Generator (JTG)*. Available: http://hoslab.cs.helsinki.fi/savane/projects/jtg.

[10] J.-M. Valin. *Speex: A Free Codec for Free Speech*. Available: http://www.speex.org.

[11] International Telecommunication Union. *Advanced video coding for generic audiovisual services*. ITU-T Recommendation H.264, 2005.

[12] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. *RTP: A Transport Protocol for Real-Time Applications*. IETF RFC 3550, July 2003.

[13] S. Wenger, M.M. Hannuksela, T. Stockhammer, M. Westerlund, and D. Singer. *RTP Payload Format for H.264 Video*. IETF RFC 3984, February 2005.

[14] H. Schulzrinne, A. Rao, and R. Lanphier. *Real Time Streaming Protocol (RTSP)*. IETF RFC 2326, April 1998.

[15] D. L. Mills. *Network Time Protocol (Version 3)*. IETF RFC 1305, March 1992.

[16] International Electrotechnical Commission. *Precision clock synchronization protocol for networked measurement and control systems*. Reference numbers: IEC 61588:2004(E), IEEE 1588-2002(E), September 2004.

[17] Kendall Correll, Nick Barendt, and Michael Branicky. Design Considerations for Software Only Implementations of the IEEE 1588 Precision Time Protocol. In *Proc. Conference on IEEE 1588*, Winterthur, Switzerland, October 2005.

[18] Kendall Correll. *PTP daemon (PTPd)*. Available: http://ptpd.sourceforge.net.