Evaluation of VoIP Quality over WiBro

Mongnam Han[1], Youngseok Lee[2], Sue Moon[1], Keon Jang[1], and Dooyoung Lee[1]

[1] Computer Science Department, KAIST
[2] School of Computer Science & Engineering, Chungnam National University

Abstract. In this work, we have conducted experiments to evaluate QoS of VoIP applications over the WiBro network. In order to capture the baseline performance of the WiBro network we measure and analyze the characteristics of delay and throughput under stationary and mobile scenarios. Then we evaluate QoS of VoIP applications using the E–Model of ITU–T G.107. Our measurements show that the achievable maximum throughputs are 5.3 Mbps in downlink and 2 Mbps in uplink. VoIP quality is better than or at least as good as toll quality despite user mobility exceeding the protected limit of WiBro mobility support. Using RAS and sector identification information, we show that the handoff is correlated with throughput and quality degradation.

1 Introduction

Recent emerging wireless networks such as 3G cellular and wireless LAN (WLAN) allow users choices in accessing the Internet based on one's need and cost. WLAN with a high data rate (up to 54Mbps) supports low mobility and limited coverage. Cellular networks support high mobility with low bandwidth. The broadband wireless access (BWA) systems address the market between WLAN and cellular networks. Their goal is to support higher bandwidth than 3G cellular networks, but less mobility for mobile end-user devices. The IEEE 802.16 family of standards specifies the air interface of fixed and mobile BWA systems. WiMax is a subset of the 802.16 standards whose main goal is product compatibility and interoperability of BWA products, just as WiFi is to the 802.11 standards. WiBro has been developed as a mobile BWA solution in Korea, and is generally considered a precursor to WiMax. It is a subset of consolidated version of IEEE Standard 802.16-2004 (fixed wireless specifications), P802.16e (enhancements to support mobility), and P802.16-2004/Cor1 (corrections to IEEE Standard 802.16-2004). The profiles and test specifications of WiBro will be harmonized with WiMAX Forum's mobile WiMAX profiles and test specification, drawing a convergence of the two standards.

Today's Internet users not only write emails and surf the web, but also make Voice over IP (VoIP) calls, play online games, and watch streaming media. These real-time applications have stringent Quality of Service (QoS) requirements on delay and loss. WiMax and WiBro standards have defined multiple service types in order to guarantee different levels of QoS. However, at the initial phase of deployment, often only the best-effort service is made available, while users do

M. Claypool and S. Uhlig (Eds.): PAM 2008, LNCS 4979, pp. 51–60, 2008.

not limit themselves to emails and web surfing over emerging wireless technology networks.

In this work, we conduct experiments to evaluate QoS of VoIP applications over the WiBro network. In order to capture the baseline performance of the WiBro network, we measure and analyze delay, loss, and throughput of constant bit rate streams in both stationary and mobile scenarios. We have measured maximum throuhgputs of 5.3 Mbps on downlink and 2 Mbps on uplink. Packet loss and throughput exhibit more variability in the mobile scenario than stationary. Then we evaluate QoS of VoIP applications using the E–Model of ITU–T G.107 also in both stationary and mobile scenarios. VoIP quality is better than or at least as good as toll quality even in the mobile scenario. By combining the packet traces with physical layer information, we show that the handoff is correlated with throughput and quality degradation on VoIP quality. We note that the deployed WiBro network is lightly loaded.

The rest of this paper is organized as follows. In section 2, we describe the background and related work. In Section 3, we describe our measurement experiment setup in a WiBro network and present the VoIP quality evaluation methodology. We present our analysis results in Section 4 and wrap up the paper with a summary in Section 5.

2 Background and Related Work

Fixed WiMax was first used to assist in the relief effort for the 2004 tsunami in Aceh, Indonesia, and now has more than 350 service providers around the world [14]. WiBro, a mobile BWA service, had its public demonstration in December 2005, and has been in service since June 2006 in Korea. The network architecture of WiBro in the phase I standardization [13] is shown in Figure 1. The WiBro network consists of Access Control Routers (ACR), Radio Access Stations (RAS), Personal Subscriber Stations (PSS), and the network service provider's IP network. An RAS is the interface between PSSs and the core network at the physical layer and it also controls the radio resource at the data link layer in conjunction with an ACR. One of the distinguishing features of WiBro from cellular networks is that Internet Protocol (IP) is used beween an ACR and RASs and also between ACRs. WiBro uses Time or Frequency Division Duplexing (TDD or FDD) for duplexing and Orthogonal Frequency Division Multiple Access (OFDAS) for robustness against fast fading and narrow-band co-channel interference. So far, five service types have been proposed and incorporated into 802.16e: unsolicited grant service (UGS), real-time polling service (rtPS), extended real-time polling service (ertPS), non-realtime polling service (nrtPS), and best effort service (BE). However, only BE is used in current deployment in Korea.

Ghosh *et al.* use a link-level simulation to analyze the 802.16 fixed WiMax system [4]. Cicconetti *et al.* analyze the effectiveness of the 802.16 MAC protocol for supporting QoS by simulation and evaluate various scheduling algorithms [1]. In [12], the authors evaluate UGS, rtPS, and ertPS scheduling algorithm in IEEE

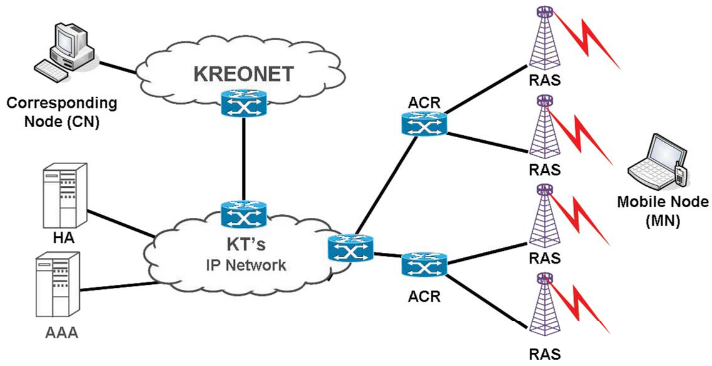


Fig. 1. Experimental environment over WiBro

802.16e system in OPNET simulation. Most of prior work focuses on investigating the performance at the physical layer and MAC layer largely through either simulation or experiment with limited mobility. In this work, we focus on the end-to-end performance at the application layer considering mobility in real life. Our work is unique in that we focus on empirical measurements from a real deployed network, world's first commercial deployment of the mobile WiMax technology.

3 Experiment Setup and Evaluation Methodology

We begin this section with a description of our measurement experiment setup. Then in Section 3.2, we describe the ITU-T E-model for VoIP quality evaluation.

3.1 WiBro Performance Measurement

In Korea, KT launched WiBro coverage for nine subway lines in Seoul on April, 2007. The Seoul subway system moves millions of people a day through an extensive network that reaches almost all corners within the city and major satellite cities outside. The maximum speed of Seoul subway trains is 90 km/h, and it takes about 1∼2 minutes between two stations. We have considered measurement experiments in vehicles moving at or under 60 km/h, the upper limit of WiBro, but chosen the subway, as it presents a more popular scenario with users. Commuters in subway are more likely to use mobile devices than those in moving vehicles, as the measurement experiment on a subway train is easier for us. We have conducted our measurement experiments on subway line number 6. It has 38 stations over a total distance of 35.1 km and six RASs.

We have placed a mobile node (a laptop with a WiBro modem) in the WiBro network and installed a stationary node (a desktop PC) connected to the Internet over a fixed line so that we could focus on the WiBro network performance. We refer to the laptop as the Mobile Node (MN) and the PC as the Corresponding Node (CN), and mark them as such in Figure 1. In order to place the CN as close to the WiBro network as possible, we use a PC directly connected to a router on Korea Research Environment Open Network (KREONET). It is a research network that interconnects super computing centers in Korea and also is used as a testbed for new networking technologies. It peers with KT's IP backbone network at one of KT's exchange points.

For our measurement experiments, we generate two types of traffic: constant bit rate (CBR) and VoIP. The difference between CBR and VoIP traffic lies in the packet sending rate and follow-up analysis. For both types of traffic, we take measurements when the MN is stationary and moving in a subway. We use *iperf* for CBR traffic generation [6], and *D-ITG* [3] for VoIP traffic generation. We configure D-ITG to measure round-trip time (RTT) instead of one-way delay, as we could not instrument the MN in subway and CN at an exchange point to have access to GPS-quality clock synchronization.

Multiple types of handoff are possible in the WiBro network. An inter-ACR handoff takes longer than inter-RAS or inter-sector handoff. An inter-sector handoff is between two sectors within an RAS. An RAS typically has three sectors. Using a custom tool developed to monitor inter-sector and inter-RAS handoffs, we collect RAS identifiers and corresponding sector identifiers. By aligning the changes in RAS and sector identifiers with the measurement data, we can pinpoint the moments of handoffs in our data.

3.2 VoIP Quality Evaluation

The classic way to evaluate speech quality is Mean Opinion Score (MOS) [9]. However, it is time consuming, costly, and not repeatable, as human experts are involved in the evaluation. Perceptual Speech Quality Measure (PSQM) [10] and Perceptual Evaluation of Speech Quality (PESQ) [11] are the most common objective measurement methods for voice quality. Both still require a reference signal to compare a degraded speech signal against and predict a MOS value. They are called psychoacoustic models. The ITU–T E-model does not depend on a reference signal, but uses a computational model to predict voice quality directly from network measurements. The output of the E-model is a single value, called an "R–factor", derived from delays and equipment impairment factors. The ITU-T G.107 [7] defines the relationship between the R factor and MOS as below:

$$MOS = \begin{cases} 1, & \text{For } R \leq 0 \\ 1 + 0.035R + R(R-60)\cdot(100-R)\cdot 7\cdot 10^{-6} & \text{For } 0 < R < 100 \\ 4.5, & \text{For } R \geq 100 \end{cases} \tag{1}$$

The E-model is based on a mathematical algorithm. Its individual transmission parameters are transformed into different individual "impairment factors" that are assumed to be additive on a psychological scale. The algorithm of the E-model also takes into account the combination effects for those impairments in the connection which occur simultaneously, as well as some masking effects.

The R-factor calculated by the E-model ranges from 0 (poor) to 100 (excellent) and can be obtained by the following expression,

$$R = R_o - I_s - I_d - I_{e-eff} + A, \tag{2}$$

where

R_o :	Basic signal-to-noise ratio
I_s :	All impairments that occur more or less simultaneously with the voice signal
I_d :	Delay impairment factor
I_{e-eff} :	Effective equipment impairment factor caused by low bit-rate codec and by packet loss on the network path
A :	Advantage factor

Cole et al. has reduced (2) to (3) after taking default values for those parameters other than delay and loss [2].

$$R = 94.2 - I_d - I_{e-eff} \tag{3}$$

In this paper, we use (3) in our WiBro VoIP quality and apply Equations (5) and (10) of [2] to translate one-way delay d and loss rate e to I_d and I_{e-eff}.

$$I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3) \tag{4}$$

$$\text{where } H(x) = 0, \text{if } x < 0, \text{ and } H(x) = 1, \text{if } x \geq 0 \tag{5}$$

$$I_{e-eff} = 0 + 30ln(1 + 15e) \tag{6}$$

4 Analysis

On October 5th and 6th, 2007, we took CBR and VoIP measurements in Seoul. For stationary experiments, we placed the MN on KAIST Seoul campus. For mobile experiments, we rode the Seoul subway line 6. For traffic logging, we used *windump* at both the MN and CN. For the VoIP experiments, the MN and CN also dumped log files including sequence numbers, packet departure times, acknowledgement arrival times, and calculated round trip time. The complete set of CBR and VoIP experiments are listed in Table 1

Table 1. The summary of CBR and VoIP experiments (upload/download)

Type	Environment	No of Exps.	Duration (sec)	Rate (Kbps)
CBR	Stationary	55 / 55	120	1500∼2500 / 5000∼6000
	Stationary	10 / 10	300	2000 / 5300
	Mobile	10 / 10	300	2000 / 5300
VoIP	Stationary	10 / 10	300	64 / 64
	Mobile	10 / 10	300	64 / 64

4.1 CBR Traffic Analysis

In order to capture the baseline performance of the WiBro network, we first measure the maximum achievable throughput. We generated 5 Mbps up to 6 Mbps and 1.5 Mbps up to 2.5 Mbps traffic in quantums of 100 Kbps for download and upload, respectively, and found the bandwidth capped at about 5.3 Mbps downlink and 2 Mbps uplink.

Then we set the transmission rate of our CBR traffic at 5.3 Mbps for downlink and 2 Mbps for uplink with the packet size of 1460 bytes and saturated the link. We conducted 10 sets of 300-second-long uploads and downloads. Due to limited space, we present only the downlink performance. We first plot the throughput of CBR traffic over time and plot it in Figure 2. From 10 300-second-long sets, we get a time span of 3000 seconds, which is the range on the x-axis. We use a 5-second interval to compute the throughput. In Figure 2(a), we see that the throughput remains almost constant when the MN is stationary. When the MN is mobile, the throughput fluctuates. We plot the inter-quartile dispersion of throughput of both the stationary and mobile experiments in Figure 2(b). In the stationary experiment, the inter-quartile range is so small that most 5-second throughput values converge to 5.3 Mbps. In the mobile experiment, the inter-quartile range spans from 4.1 Mbps to 5.1 Mbps, and has noticeably more points below 3 Mbps. To view the dispersion of throughput in a more visually intuitive way, we plot the variability in Figure 2(c) using the second-order difference plot. The difference between two consecutive throughputs are plotted against that between next two consecutive values. In this figure, the median from the center of the mobile station is about 13 times larger than that of the stationary station. The throughput of MN still remains consistently above 1 Mbps.

Next, we analyze the jitter and loss rates of CBR traffic. For this work, we define jitter as the difference between sending intervals and arrival intervals. Figure 3(a) depicts a cumulative distribution function (CDF) of CBR traffic jitter. In both stationary and mobile experiments, more than 90% of jitters are below 15 milliseconds. Given that our traffic by itself saturated the link, this result is encouraging for real-time applications. Now we look at the loss rate of our CBR traffic. In Figure 3(b) the loss rate in the mobile environment is much higher than in stationary. In a WiBro network, a MAC layer retransmission mechanism called Hybrid Auto Repeat reQuest (HARQ) is adopted to reduce loss rate at the cost of increased delay. As our CBR traffic used UDP as an underlying transport protocol and saturated the link, we expect the loss rates to decrease once we lower the sending rate. We revisit the discussion of the loss rate in the next section.

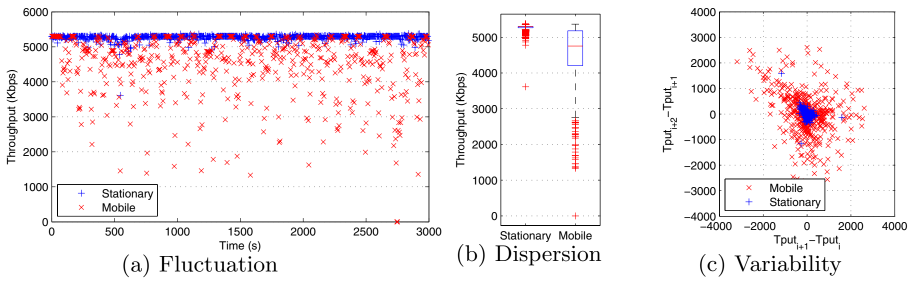
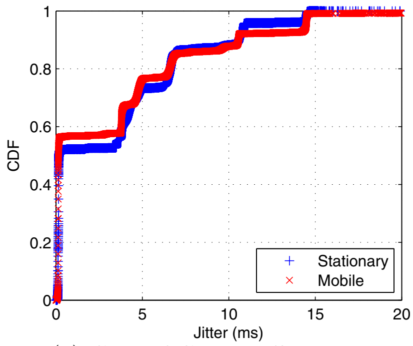
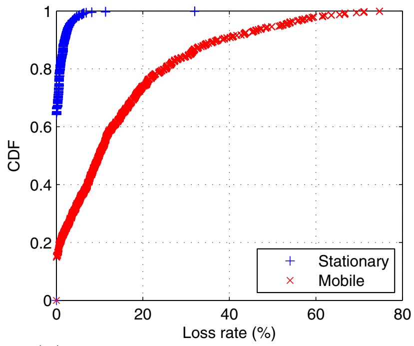


Fig. 2. Analysis of CBR traffic throughput over WiBro



(a) CDF of CBR traffic jitter



(b) CDF of CBR traffic loss rate

Fig. 3. Analysis of CBR jitter and loss over WiBro

4.2 VoIP Traffic Analysis

For VoIP experiments, we have generated voice traffic that has the same characteristics of the G.711 voice codec without Packet Loss Concealment (PLC) [8]. The payload size is set to 160 bytes and the sending interval to 20 ms in G.711 codec without PLC. The resulting throughput of VoIP traffic is 64 Kbps. We collected 10 300-second-long data sets after transmitting voice packets between the MN and the CN. Because the clocks on the MN and CN were not synchronized, we could not measure the one-way delay accurately. Instead, we took round-trip measurements of VoIP traffic, and halved the delay. Due to the difference in uplink and downlink bandwidths, half the round-trip delay is likely to be larger than the one-way uplink delay. However, the WiBro link was very lightly loaded

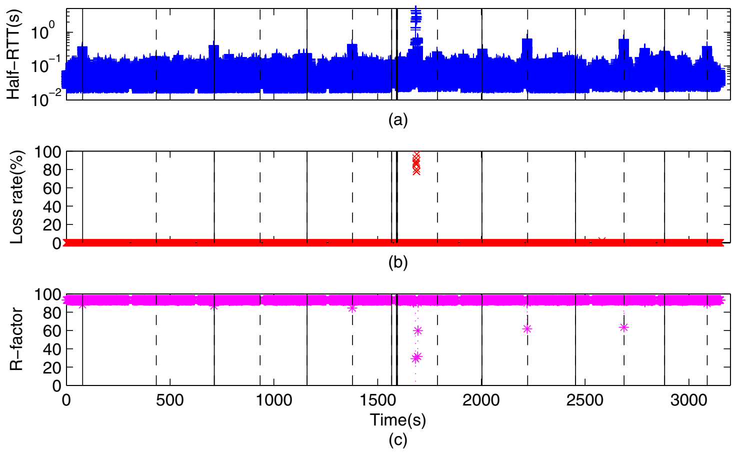


Fig. 4. The time-series plots of (a) half-RTT, (b) loss rate, and (c) R-factor. Each vertical line indicates the time when either inter-RAS(solid line) or inter-sector(dashed line) handoff occurs.

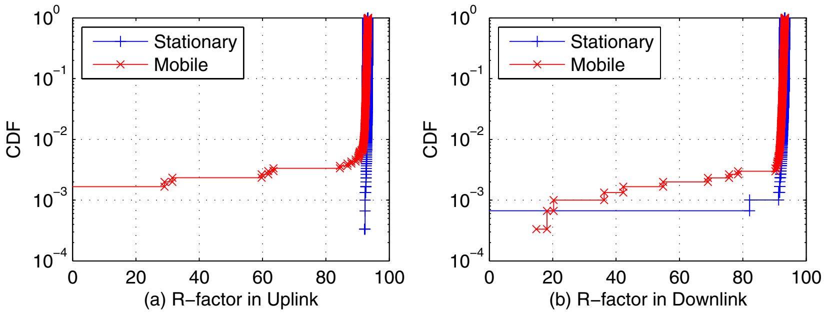


Fig. 5. CDF of R-factor in uplink and downlink

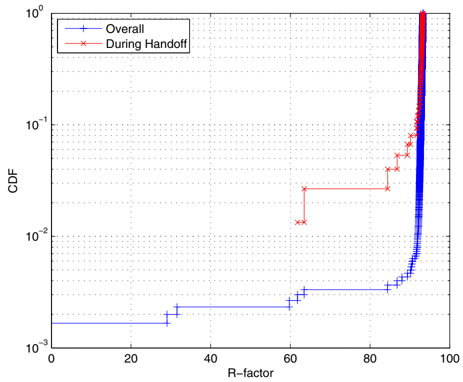


Fig. 6. CDF of R-factor in 5-sec-intervals vs. during handoffs

as we have seen in the previous section and thus we assume the difference in transmission delay to be minimal. In the rest of this section, all the delays we use are calculated as described.

Figure 4 plots the delay, loss, and R-factor of mobile VoIP traffic measurements. We plot loss rates and R-factors calculated over 5 seconds as before. During the mobile experiment, handoffs occurred 17 times and most of delay spikes and burst losses occur near handoffs. We mark the points of the inter-RAS handoffs in solid lines and inter-sector handoffs in dashed lines in Figure 4. Almost all packets experiences delay below 200 ms, but during handoffs some packets experiences delay over 400 ms. We note a delay spike of 5 sec about half way through the measurement experiment between 1600 and 1700 seconds on the x-axis. Delay spikes and burst losses have been reported in both wired and wireless networks, and there are many possible causes, such as cell or sector re-selection, link-level error recovery, wireless bandwidth fluctuation and blocking by higher priority traffic [15,5]. However, this delay spike does not coincide with a handoff, and needs further investigation.

Figure 5 shows the R-factor of VoIP quality of uplink (MN to CN) and down-link (CN to MN) in the stationary and mobile cases. The R-factor is calculated every 5 seconds. From the figure we have found that more than 99% of R-factors are above 90 in both the stationary and the mobile case and only 0.2% of R-factors is below 70 in the mobile case. The R-factor of 70 or above is considered toll-quality, and thus mobile devices attached to the WiBro network are likely to experience toll-quality using VoIP applications.

To quantify the impact of handoffs on QoS of VoIP, we have computed R-factors using average delay and loss rate for the interval of 5 seconds before and after the handoff and compared with the overall cumulative distribution function (CDF) of R-factors in Figure 6. Here again we observe that the about 99% of R-factors during handoffs are more than 85, which translates to good quality for voice communication.

5 Summary and Future Work

In this work, we have conducted experiments to evaluate QoS of VoIP applications over the WiBro network. In order to capture the baseline performance of the WiBro network we measure and analyze the characteristics of delay and throughput under stationary and mobile scenarios. Then we evaluate QoS of VoIP applications using the E–Model of ITU–T G.107. Our measurements show that the achievable maximum throughputs are 5.3 Mbps in downlink and 2 Mbps in uplink. VoIP quality is better than or at least as good as toll quality despite user mobility exceeding the projected limit of WiBro mobility support. Using RAS and sector identification information, we show that the handoff is correlated with throughput and quality degradation.

The WiBro network is in its early phase of deployment and about 70,000 subscribers are reported to have signed up. As our CBR traffic analysis shows, the network is very lightly loaded, allowing near maximum throughputs for many 5-second-long periods. In future, we plan to conduct more experiments with cross-traffic injection and TCP traffic.

Acknowledgement

Sue Moon, Keon Jang, and Dooyoung Lee were supported by KOSEF through AITrc (Advanced Information Technology Research Center). Youngseok Lee was supported by the Ministry of Information and Communication, Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement) (IITA-2008-(C1090-0801-0016)).

References

1. Cicconetti, C., Lenzini, L., Mingozzi, E., Eklund, C.: Quality of service support in IEEE 802.16 networks. IEEE Network 20, 50–55 (2006)
2. Cole, R.G., Rosenbluth, J.H.: Voice over IP performance monitoring. SIGCOMM Comput. Commun. Rev. 31(2), 9–24 (2001)

3. D-ITG, http://www.grid.unina.it/software/itg/
4. Ghosh, A., Wolter, D.R., Andrews, J.G., Chen, R.: Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential. IEEE Communications Magazine 43, 129–136 (2005)
5. Hornero, R., Abasolo, D.E., Jimeno, N., Espino, P.: Applying approximate entropy and central tendency measure to analyze time series generated by schizophrenic patients. In: Proc. of the 25th Annual International Conference of the IEEE, September 2003, pp. 2447–2450 (2003)
6. iperf, http://dast.nlanr.net/projects/iperf/
7. ITU–T standard G.107. The e-model, a computational model for use in transmission planning (March 2005)
8. ITU–T standard G.711. Pulse code modulation (PCM) of voice frequencies (November 1988)
9. ITU–T standard P.800. Methods for subjective determination of transmission quality (August 1996)
10. ITU–T standard P.861. Objective quality measurement of telephoneband (300–3400 hz) speech codecs (August 1996)
11. ITU–T standard P.862. Perceptual evaluation of speech quality (pesq): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs (February 2001)
12. Lee, H., Kwon, T., Cho, D.-H., Lim, G., Chang, Y.: Performance analysis of scheduling algorithms for VoIP services in IEEE 802.16e systems. In: Proc. of Vehicular Technology Conference, vol. 3, pp. 1231–1235 (2006)
13. TTAS.KO-06.0065R1. WiBro standard phase–I (December 2004)
14. WiMax, http://en.wikipedia.org/wiki/wimax
15. Xin, F., Jamalipour, A.: TCP performance in wireless networks with delay spike and different initial congestion window sizes. Computer Communications 29, 926–933 (2006)