WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2012; **12**:318–324
Published online 12 May 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.963

RESEARCH ARTICLE

NCTU-VT: a freeware for wireless VoIP performance measurement

Ya-Chin Sung[1], Yi-Bing Lin[1]*, Ren-Huang Liou[1] and Lon-Fon Shieh[2]

[1] Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan
[2] Department of Business Management, National United University, Miaoli, Taiwan

ABSTRACT

Voice over IP (VoIP) is a promising low-cost voice communication over the wireless IP network. To provide satisfactory VoIP services, the Quality of Service (QoS) of the wireless network should be guaranteed. This paper proposes a VoIP performance measurement freeware called NCTU VoIP Testing Tool (NCTU-VT). We compare NCTU-VT with two commercial tools SmartVoIPQoS and IxChariot in terms of packet loss, latency, and Mean Opinion Score (MOS) of the VoIP sessions in Wi-Fi network. Our study indicates that these three tools can accurately measure VoIP performance in Wi-Fi environment.

KEYWORDS
mean opinion score; performance measurement; quality of service; voice over IP

***Correspondence**
Yi-Bing Lin, Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan.
E-mail: liny@cs.nctu.edu.tw

1. INTRODUCTION

Voice over IP (VoIP) is a promising low-cost voice communication over the wireless IP network. In many IP-based telecommunication systems such as IP Multimedia Core Network Subsystem (IMS) [1], Session Initiation Protocol (SIP) [2] is used for call control. A SIP-enabled end device is called a SIP User Agent (UA). In VoIP services, Real-Time Transport Protocol (RTP) [3] is used to transmit voice (or multimedia) packets over UDP. Several codecs are used in RTP under various bandwidth restrictions. For example, G.711 is a 64 Kbps codec where an RTP stream is transmitted at the rate of 50 packets per second [4]. In this paper, G.711 is considered to investigate the performance of a high bit-rate codec.

To provide satisfactory VoIP services, the Quality of Service (QoS) of the VoIP network should be guaranteed. Besides typical QoS measures for Internet packet delivery such as packet loss and delay, a major VoIP performance measure is Mean Opinion Score (MOS) [5]. This measure considers the equipment and impairment factors to subjectively quantify the perceived quality of a voice transmission based on user perception. The MOS values range from 1 to 5, where 1 is unacceptably bad, 2 is poor, 3 is fair, 4 is good, and 5 is excellent.

VoIP performance is particularly important for wireless environment because the radio transmission is not as reliable as wireline transmission. Several commercial products have been developed for VoIP performance measurement; for example, Spirent Communications provides a hardware-based solution called SmartBits [6], which executes SmartVoIPQoS application to conduct the VoIP tests [7]. Another example is software-based IxChariot VoIP measurement tool developed by Ixia [8].

We develop the NCTU VoIP Testing Tool (NCTU-VT), a freeware tool that provides a software-based VoIP performance measurement solution. Then we measure packet loss, delay, and MOS value of a VoIP session, and compare the reported results of SmartBits, IxChariot, and NCTU-VT. Our Study will show that the freeware NCTU-VT can accurately measure VoIP performance just like commercial tools SmartVoIPQoS and IxChariot.

The paper is organized as follows. The experimental environment is described in Section 2. The commercial tools are introduced in Section 3. Details of NCTU-VT are elaborated in Section 4, and the VoIP performance measures are reported in Section 5. Based on the performance measurement results, we compare the characteristics of three tools: SmartVoIPQoS, IxChariot, and NCTU-VT.

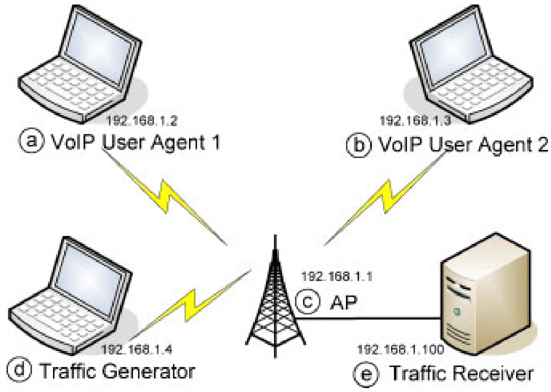


Figure 1. The experimental environment.

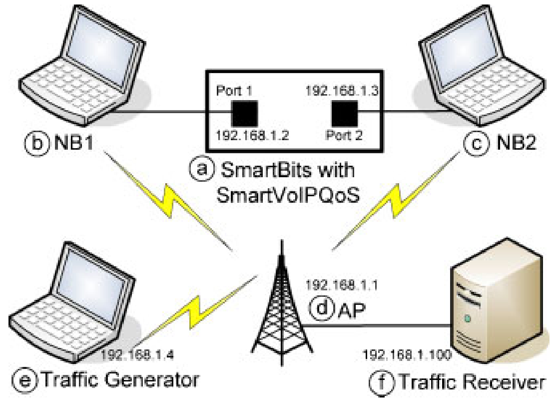


Figure 2. SmartVoIPQoS measurement environment.

2. THE EXPERIMENTAL ENVIRONMENT

As illustrated in Figure 1, we consider an experimental environment that measures VoIP RTP sessions between two VoIP UAs running on notebooks (Figure 1 (a) and (b)). Two RTP streams are transmitted through an Asus WL-500gP WLAN Access Point (AP, see Figure 1 (c)) wirelessly; see paths (a)→(c)→(b) and (b)→(c)→(a). The experiments include a background traffic generator (a notebook; see Figure 1 (d)) that competes the radio resources with the measured VoIP calls. The traffic receiver (Figure 1 (e)) is a PC connected to the AP through Fast Ethernet (100 Mbps). The background traffic is delivered through path (d)→(c)→(e) to provide stringent VoIP measurement environment as described in Reference [9]. This experimental environment is configured as a private network, where the IP address of the AP is 192.168.1.1. The IP addresses of the UAs are 192.168.1.2 and 192.168.1.3; and the IP addresses of the traffic generator and the traffic receiver are 192.168.1.4 and 192.168.1.100, respectively.

The above private network configuration (without involving the public network nodes) is specifically designed for evaluating the Wi-Fi radio link impact on VoIP, where the UAs are connected only through the AP. Therefore, we can clearly identify the packet loss and latency introduced by radio links.

3. COMMERCIAL VoIP PERFORMANCE MEASUREMENT TOOLS

This section overviews two commercial VoIP performance measurement tools SmartVoIPQoS and IxChariot.

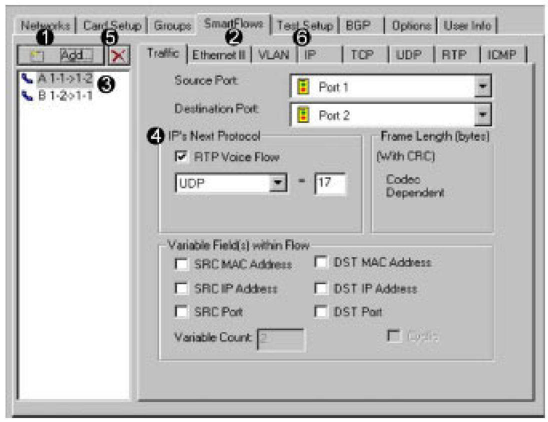
3.1. SmartVoIPQoS

SmartVoIPQoS is an application developed on the SmartBits platform. This tool maps the voice quality to a MOS value based on the Perceptual Speech Quality Measure (PSQM) algorithm [10]. The International Telecommunication Union's Telecommunication Standardization Sector (ITU-T) has withdrawn P.861 because PSQM is not developed to account for the network QoS perturbations such as packet loss and latency [11]. However, PSQM is still utilized in several VoIP performance studies [12,13]. Spirent Communications has enhanced the MOS measurement capability of SmartVoIPQoS by Abacus. Since we do not have access to this commercial product, the MOS feature of Spirent Communications's products will not be investigated in this paper.

Based on the experimental environment described in Figure 1, Figure 2 illustrates the SmartVoIPQoS measurement environment. SmartVoIPQoS does not implement SIP UAs. Instead, the RTP streams are created between Port 1 and Port 2 residing in the SmartBits (Figure 2 (a)). These two ports connect to NB1 (Figure 2 (b)) and NB2 (Figure 2 (c)) with RJ-45 interfaces (Figure 2 (a) Port 1 and Port 2). Each NB is configured as a bridge between the SmartBits and AP. Our measurements indicate that the NBs do not cause any packet loss in the RTP streams, and therefore this bridge configuration does not introduce extra inaccuracy in our experiments. The VoIP packets are delivered to the AP (Figure 2 (d)) through the wireless interfaces equipped in NB1 and NB2. SmartVoIPQoS (Figure 2 (a)) generates packets that simulate the call patterns of voice traffic and produces the output statistics based on the packets collected by the UAs. The traffic generator and the receiver (Figure 2 (e) and (f)) provides wireless background traffic for stress tests.

The SmartVoIPQoS configuration Graphical User Interface (GUI) is shown in Figure 3 (a) and the parameter setup are summarized in Figure 3 (b). The 'Networks' Tab (1) in Figure 3 (a) configures the IP addresses of the SmartBits RJ-45 Ports. In our experiment, NB1 connects to the RJ-45 Port 1 with the IP address of 192.168.1.2. NB2 connects to the Port 2 with the IP address 192.169.1.3. The 'SmartFlow Tab' (2) in Figure 3 (a) sets the traffic flows (3) in Figure 3 (a) and the protocols used by

(a)



(b)

Port 1 IP Address		192.168.1.2
Port 2 IP Address		192.168.1.3
Flow 1	Source	Port 1
	Destination	Port 2
Flow 2	Source	Port 2
	Destination	Port 1
Protocol		RTP/UDP
Codec		G.711
Test Duration		600 Sec.

Figure 3. Configurations used in SmartVoIPQoS: (a) configuration GUI and (b) summary of parameter setup.

the traffic flow (4) in Figure 3 (a) of the experiments. Figure 3 defines two RTP flows for UA 1 and UA 2 (paths (a)→(b)→(d)→(c)→(a) and (a)→(c)→(d)→(b)→(a) in Figure 2). The 'Card Setup' Tab (5) in Figure 3 (a) selects the codec, (i.e., G.711 in our experiment). The 'Test Setup' Tab (6) in Figure 3 (1) sets the test duration (600 s in our experiment).

3.2. IxChariot

The network performance test tool IxChariot consists of a console and several endpoint programs. The console can be installed in a computer with an endpoint. Two endpoints define a traffic flow, where Endpoint 1 is the source and Endpoint 2 is the destination. The traffic flow of an endpoint pair is configured by the console. This configuration is used by Endpoint 1 for generating the test packets and collecting the statistic results. Then the measured statistics are sent to the console for analysis display.

Figure 4 illustrates the IxChariot measurement environment, where the VoIP UAs reside in NB1 (Figure 4 (a)) and NB2 (Figure 4 (b)). The IxChariot console (Figure 4 (c)), NB1, and NB2 are connected to a Buffalo LSW-TX-8NP switch (Figure 4 (d)). The RTP streams are transmitted between NB1 and NB2 via the AP (Figure 4 (e)), and the background traffic is delivered through path (f)→(e)→(g) in Figure 4.

Figure 5 shows the configuration setup in our experiment. NB 1 and NB 2 are configured as Endpoint 1 and Endpoint 2 of RTP Flow 1, respectively (Figure 5 (1)). RTP Flows 2 is similarly configured (see Figure 5 (4)). The parameter setups (i.e., (2) and (3) in Figure 5) are the same as those for Smartbit in Figure 3.

To compute the MOS value, IxChariot employs the E-model defined in ITU-T Recommendation G.107 [9,14]. The E-Model assumes that various kinds of transmission impairments are additive on the scale of a rating

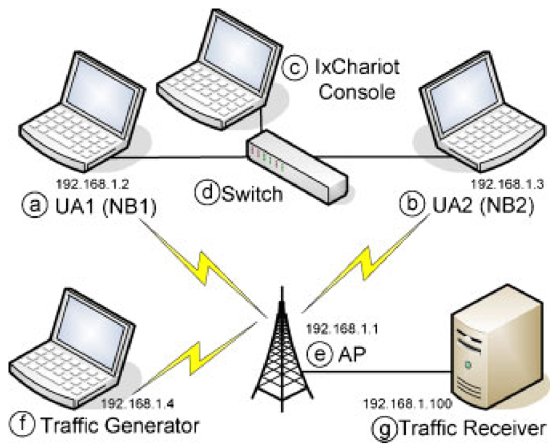


Figure 4. IxChariot measurement environment.

factor R. The considered impairments include the noise sources (such as circuit noise and room noise), the quality decreases caused by overall loudness, non-optimum sidetone and quantizing distortion, the delay impairment, and equipment impairment. After accumulating these transmission impairments, the E-model then converts the R value into a MOS scale to quantify the overall conversational quality. The estimated MOS values range from 1 to 4.5 (not 5).

4. NCTU VoIP TESTING FREEWARE

NCTU VoIP Testing Freeware (NCTU-VT) is a software-based solution that supports both Perceptual Evaluation of Speech Quality (PESQ) [15] and E-Model to evaluate the MOS value of the VoIP sessions. PESQ is defined in ITU-T Recommendation P.862, which takes packet loss into consideration.

Wirel. Commun. Mob. Comput. 2012; **12**:318–324 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/wcm

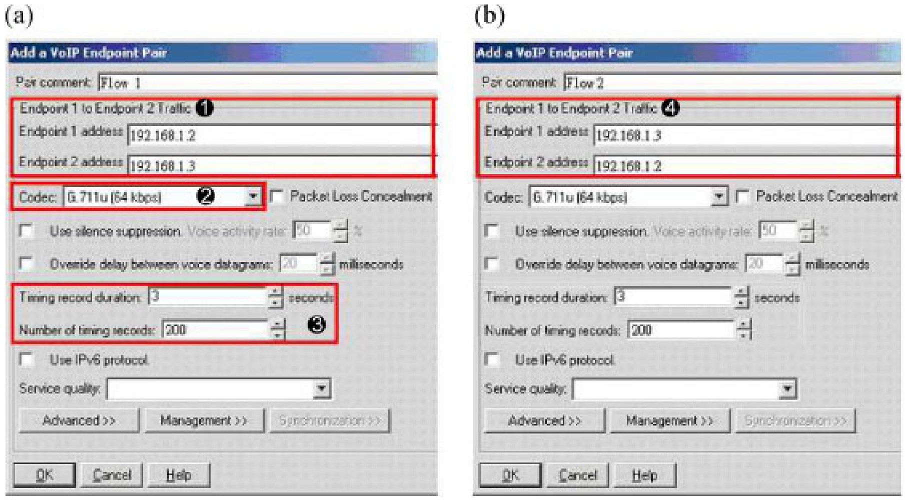


Figure 5. VoIP session configurations set in IxChariot: (a) RTP flow setup from UA1 to UA2 and (b) RTP flow setup from UA2 to UA1.

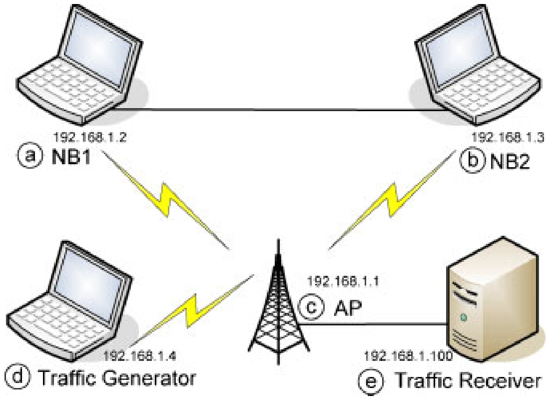


Figure 6. NCTU-VT measurement environment.

Figure 6 illustrates the NCTU-VT measurement environment where the NCTU-VT software (including SIP UAs) are installed in NB1 (Figure 6 (a)) and NB2 (Figure 6 (b)). Two RTP streams, (a)→(c)→(b) and (b)→(c)→(a), are transmitted through the AP (Figure 6 (c)) over the wireless links. NB1 also connects to NB2 through Fast Ethernet (wireline link (a)–(b) in Figure 6) to synchronize their clocks. The background traffic is delivered through path (d)→(c)→(e).

Figure 7 illustrates the NCTU-VT software architecture [16]. The Input/Output Module (including user interface (Figure 7 (1)) allows the users to configure the parameters used in the experiments and shows the statistic results. The NCTU-VT interface configures the parameters, including the IP address of the other call party (Figure 7 (10)), VoIP codec (e.g., G.711 in Figure 7 (12)) and the test period (e.g., 600 s in Figure 7 (11)). The test results are shown in Figure 7 (14). After parameter setup, the user presses the 'Test' button (Figure 7 (13)) to measure the VoIP performance by sending the parameters to configure VoIP UA (Figure 7 (2)). The UA consists of the SIP Module for call control and the RTP Module for voice packet transmission.

Each RTP packet contains a timestamp to indicate the packet sending time obtained from the Time Synchronization Module (Figure 7 (3)). This module utilizes the win32 API function QueryPerformanceFrequency [17] for timer resolution in number of ticks per second. In our environment, the timer resolution is 3 579 545 ticks per second (i.e., 2.79×10^{-7} s for a tick).

Suppose that the RTP Module of UA1 sends a RTP packet to UA2 through path (a)→(c)→(b) in Figure 6. After the RTP Module of UA2 has received this packet, it passes the packet back to UA1's Time Synchronization Module (through the Ethernet link (b)→(a)) to compute the latency. Measured packet loss and latency are sent to the Performance Evaluation Module (Figure 7 (4)), and then forwarded to the E-Model Processor Module (Figure 7 (5)) for computing the E-Model-based MOS value. Also, packet loss is forwarded to the Pattern Processor Module (Figure 7 (9)) to compute the PESQ-based MOS value with 20 Reference Files (Figure 7 (7)) as the input voice samples, which works as follows:

Consider the voice pattern (represented as wave data) in a Reference File (Figure 8 (a)). The Performance Evaluation Module provides the packet loss information in Figure 8 (b), where packets 3, 8, and 9 are lost (marked by '×'). Since every G.711 RTP packet contains 20 ms audio data, the voice in 40–60 ms and 140 –180 ms are deleted as shown in Figure 8 (c) and are stored in the Degraded File (Figure 7 (8)).

Then the PESQ Processor Module (Figure 7 (6)) generates the MOS value by comparing the Reference File (Figure 8 (a)) and the Degraded Files (Figure 8 (c)), and returns the MOS value to the Performance Evaluation Module. Finally, the Performance Evaluation

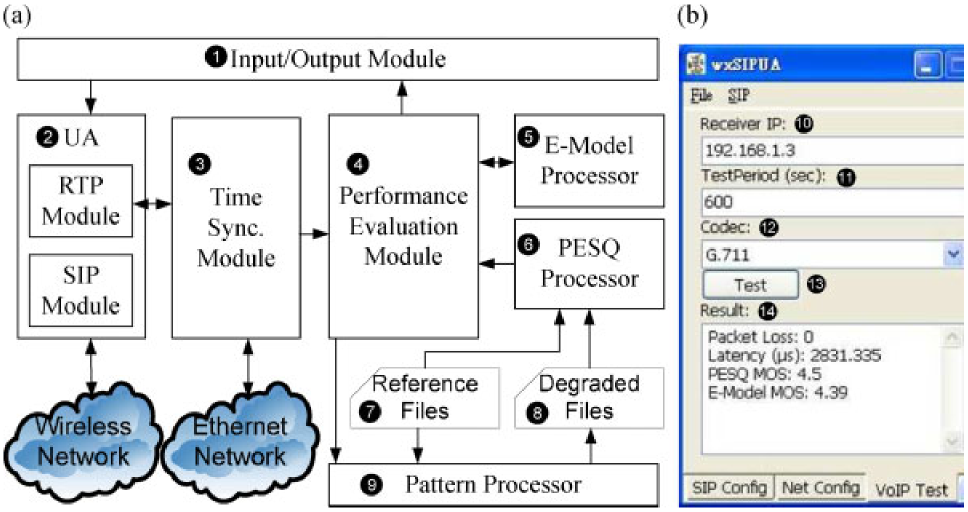


Figure 7. NCTU-VT software architecture: (a) software architecture and (b) user interface.

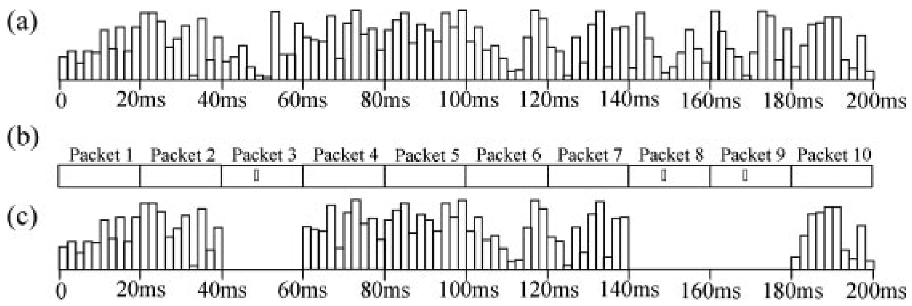


Figure 8. Generation of degraded file: (a) reference file, (b) packet loss information, and (c) degraded file.

Module provides the VoIP performance measures (i.e., packet loss, latency, and PESQ/E-Model MOS) to the Input/Output Module for statistic display.

5. COMPARISON OF MEASUREMENT RESULTS

This section compares the SIP-based VoIP measurement functions of SmartVoIPQoS, IxChariot, and NCTU-VT, and uses these tools to investigate the wireless VoIP performance.

Both SmartVoIPQoS and NCTU-VT measure the delay at the precision of 10^{-6} s. IxChariot supports precision of 10^{-3} s. Low resolution of delay measure is due to the fact that IxChariot allows the Endpoints to reside in different locations. Therefore, it is necessary to perform time synchronization between the Endpoints, which is not an easy task. For the experiment environment in Figure 4, the clock error between the IxChariot Endpoints is about 3 ms [18]. In NCTU-VT, on the other hand, the UAs are linked through wired Ethernet for time synchronization. In SmartVoIPQoS, the UAs are run in the same SmartBits hardware.

Table I. Number of packet loss.

Background traffic (KB/s)	0	100	200	300	400	500	600	0-600
(1) SmartVoIPQoS	5.6	2.7	3.1	3.7	2.6	4.1	10.1	31.9
(2) IxChariot	5.1	4.3	4.5	1.8	4.3	2.1	10.6	32.7
(3) NCTU-VT	3.7	3.9	2.8	3.3	3.4	3.0	11.0	31.1

Table I lists packet loss measured by SmartVoIPQoS, IxChariot, and NCTU-VT. We control the background traffic from 0 to 600 KB/s by increasing 100 KB/s for every 10 m trial, and repeat the trials for 10 times (i.e., 700 trials in total). Sixty thousand RTP packets are delivered in each trial. It is noted that when the background traffic is less than 600 KB/s, the Wi-Fi bandwidth is not significantly consumed, and therefore the number of packet loss is less than 6. The number of packet loss is larger than 10 when the background traffic is 600 KB/s. Similar phenomenon is observed in Reference [19]. The table indicates that SmartVoIPQoS, IxChariot, and NCTU-VT report the same average packet loss statistics (i.e., 31 or 32 lost packets).

Table II lists the latencies measured by SmartVoIPQoS, IxChariot, and NCTU-VT. SmartVoIPQoS and NCTU-VT report similar results and the discrepancy between them

Wirel. Commun. Mob. Comput. 2012; **12**:318–324 © 2010 John Wiley & Sons, Ltd.
DOI: 10.1002/wcm

Table II. Latency performance.

Background traffic (KB/s)	0	100	200	300	400	500	600
(1) SmartVoIPQoS	2168.669	2807.352	3100.451	3455.964	3893.675	4268.377	4901.059
(2) IxChariot	2000	3000	3000	3000	4000	4000	5000
(3) NCTU-VT	2208.785	2831.335	3144.880	3487.230	3993.095	4328.360	4967.065
Discrepancy (1) and (2)	7.77%	6.86%	3.23%	13.19%	2.73%	6.28%	2.01%
Discrepancy (1) and (3)	1.84%	1.00%	1.43%	0.90%	2.55%	1.40%	1.34%
Discrepancy (2) and (3)	9.45%	6.96%	4.61%	13.97%	0.17%	7.59%	0.66%

Table III. MOS performance.

Background traffic (KB/s)	0	100	200	300	400	500	600
(1) NCTU-VT PESQ	4.499485	4.499428	4.499561	4.499488	4.499571	4.499499	4.498329
(2) NCTU-VT E-model	4.383	4.378	4.383	4.382	4.383	4.382	4.354
(3) IxChariot E-model	4.384	4.386	4.387	4.39	4.386	4.388	4.383
Discrepancy (2) and (3)	0.02%	0.18%	0.09%	0.18%	0.23%	0.06%	0.66%

is smaller than 2.55%. The latency statistics of IxChariot is restricted by the supported precision and thus introduce higher discrepancy (i.e., up to 13.97% as compared with SmartVoIPQoS/ NCTU-VT).

In our experiment, the latency increases when the background traffic becomes heavy. When the background traffic increases from 0 to 100 KB/s, the latency increases about 30%. When the background traffic is larger than 100 KB/s, the latency increases about 10% for every background traffic increase of 100 KB/s.

Table III lists the MOS values measured by NCTU-VT and IxChariot. NCTU-VT reports higher PESQ MOS than E-Model MOS because PESQ only considers the packet loss factor while E-Model accommodates both packet loss and latency factors. The MOS value discrepancy between the NCTU-VT E-Model and IxChariot is smaller than 0.23%. The table also indicates that the MOS value is not significantly affected by the background traffic.

6. CONCLUSION

This paper described the development of NCTU-VT, a freeware tool that provides a software-based solution for VoIP performance evaluation. NCTU-VT is compared with two commercial tools SmartVoIPQoS and IxChariot by a case study on VoIP performance in a Wi-Fi environment.

Both NCTU-VT and SmartVoIPQoS provide high accuracy of latency measure as compared with IxChariot. SmartVoIPQoS supports PSQM-based MOS measurement, which is out of date. IxChariot supports E-model for MOS measurement. NCTU-VT supports both PESQ (an enhanced version of PSQM) and E-Model. An enhanced version of SmartVoIPQoS, called Abacus, supports both PESQ and E-Model just like NCTU-VT. We do not have access to Abacus at the time of writing, and therefore do not have performance assessments of this commercial tool. In summary, the freeware NCTU-VT can accurately measure VoIP performance just like commercial tools SmartVoIPQoS and IxChariot.

ACKNOWLEDGEMENTS

Y.-B. Lin's work was supported in part by NSC 97-2221-E-009-143-MY3, NSC 98-2221-E-009-059-MY2, NSC 98-2219-E-009-016, Intel, Chunghwa Telecom, IBM, ITRI and NCTU Joint Research Center, and MoE ATU plan. The IxChariot is provided by ITRI through the ITRI-NCTU JRC project.

REFERENCES

1. Lin Y-B, Pang A-C. Wireless and Mobile All-IP Networks. Wiley: New York, 2005.
2. IETF. SIP: session initiation protocol. *RFC* 3261, June 2002.
3. IETF. RTP: a transport protocol for real-time applications. *RFC* 3550, July 2003.
4. ITU-T. Pulse code modulation (PCM) of voice frequencies. *ITU-T G.711 Recommendation*, November 1988.
5. ITU-T. Methods for subjective determination of transmission quality. *ITU-T P.*800 *Recommendation*, June 1998.
6. SmartBits. Available at: www.paralink.com.tw/product/spirent/smartbits-index.htm
7. SmartVoIPQoS User Guide. Available at: www.spirentcom.com/documents/438.pdf

8. IxChariot. Available at: www.ixchariot.com/products/datasheets/ixchariot.html
9. Lin Y-B, Sung Y-C, Liu C-L, Rashvand HF, Liao Y-J. (in press). WiMAX–based VoIP experience in M-Taiwan. Accepted and to appear in *IET Communications*.
10. ITU-T. Objective quality measurement of telephone-band (300-3400Hz) speech codecs. *ITU-T P.861 Recommendation*, February 1998.
11. ITU-T P Series Recommendation. Available at: www.itu.int/rec/T-REC-P/e
12. Lee W, Lee M, McGowan J. Enhancing objective evaluation of speech quality algorithm: current efforts, limitations and future directions. *European Transactions on Telecommunications*, published online, February 2009; **20**(6): 595–603.
13. Počta P, Turský J, Palúch P. Design and evaluation of novel VoIP link adaptation algorithm based on GTS function. *International Journal of Electronics and Communications*, published online, March 2010; **64**(3): 191–201.
14. ITU-T. The e-model, a computational model for use in transmission planning. *ITU-T Recommendation G*.107, April 2009.
15. ITU-T. Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. *ITU-T Recommendation P*.862, February 2001.
16. Chan Y-H. VoIP quality measurement for multiple codes. *Master Thesis*, National Chiao Tung University, Taiwan, 2007.
17. MSDN Library. Available at: msdn.microsoft.com/en-us/library/ms644905(VS.85).aspx
18. IxChariot User Guides. Available at: www.ixiacom.com/library/user_guides/index.php
19. Sung Y-C, Lin Y-B. IPsec-based VoIP performance in WLAN. *IEEE Internet Computing* **12**(6): 77–82. 2008.

Authors' Biographies

Ya-Chin Sung received the BSCSIE and MSCSIE degrees from National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 2002 and 2003, respectively. She is currently working toward the Ph.D. degree at NCTU. Her current research interests include design and analysis of personal communications services networks, network security, and performance modeling.

Yi-Bing Lin (M'96-SM'96-F'03) is Chair Professor and Dean of the College of Computer Science at National Chiao Tung University. Lin received his EEBS degree from National Cheng Kung University in 1983 and the Ph.D. degree from the Department of Computer Science and Engineering in University Washington, Seattle in 1990. His current research interests include wireless communications and mobile computing. He has published more than 200 journal articles, 200 conference papers, and is coauthor of Wireless and Mobile Network Architecture (Wiley, 2001), Wireless and Mobile All-IP Networks (Wiley, 2005), and Charging for Mobile All-IP Telecommunications (Wiley, 2008). Lin is a fellow of the ACM, the American Association for the Advancement of Science (AAAS), the IET, and the IEEE.

Ren-Huang Liou received the B.S. and the M.S. degrees in Computer Science from National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 2007 and 2009, respectively. He is currently working toward the Ph.D. degree at NCTU. His current research interests include design and analysis of personal communications services networks, mobile computing, and performance modeling.

Lon-Fon Shieh is an assistant professor of Department of Business Management, National United University, Taiwan, R.O.C. He is Chief of Industrial Development Bureau, Ministry of Economic Affairs, Taiwan, R.O.C.

DOI: 10.1002/wcm