An Analytic and Experimental Study on the Impact of Jitter Playout Buffer on the E-model in VoIP Quality Measurement

Olusegun Obafemi
School of Information Technology
Illinois State University
Normal IL 61790 USA
oeobafe@ilstu.edu

Tibor Gyires
School of Information Technology
Illinois State University
Normal IL 61790 USA
tbgyires@ilstu.edu

Yongning Tang
School of Information Technology
Illinois State University
Normal IL 61790 USA
ytang@ilstu.edu

***Abstract*—Over the years, the quality in Voice over IP (VoIP) applications has been defined from the perspective of VoIP service providers but not necessarily from the perspective of end-users. Conversational quality as perceived by end-users is affected by a wide variety of factors that are not exclusive to network performance. While service providers focus their efforts on Quality of Service (QoS) in VoIP, end-users expect improvement in Quality of Experience (QoE). The measurement of QoE is significantly challenging due to the diverse nature of factors that collectively determine QoE. It is highly desired to have a more comprehensive and accurate quality measurement model that can provide objective measurement of QoS in real-time while simultaneously offering QoE measurement as well. Thus, unveiling the relevance of various factors in a quality measurement model is crucial. In this paper, we study a commonly used ITU-T recommended quality measurement model: the E-model with a focus on how an important but ignored parameter jitter playout buffer affects the measurement accuracy by conducting an analytical and experimental approach. Our study shows that adaptive jitter playout buffering does affect the measurement of end user perceived VoIP quality. Thus, the E-Model that excludes the impact of the dynamic jitter playout buffer cannot provide accurate measurement on how an user perceives conversational quality in VoIP application.**

Keywords- Quality of Service, Quality of Experience, Jitter Buffer, Adaptive Playout Buffering, Conversational Quality, VoIP.

I. INTRODUCTION

Quality of Service (QoS) in VoIP prioritizes voice traffic at the expense of traffic from other packet types. The factors of QoS mainly include jitter, packet loss and latency. Jitter reflects the delay variation that affects voice IP packets as they travel through a network. Packet loss occurs when voice packets are dropped due to various reasons (e.g., congestions) in the network. Latency or delay is due to several factors, including physical distance between a sender and a receiver, the number of router hops the packets have to pass through, and the packet processing time in the network.

The QoS provisioned by VoIP service providers is not necessarily proportional to the Quality of Experience (QoE) perceived by end users [4]. VoIP QoE describes the satisfaction of end users with respect to conversations using VoIP technology. The tremendous increase of VoIP users in the recent years makes the satisfaction of users an equally important metric, compared to the QoS related network performance metrics.

The measurement of QoE is significantly challenging due to the diverse nature of factors that collectively determine QoE. Several ITU recommended methods are commonly used, namely Mean Opinion Score (MOS) [4], Perceptual Evaluation of Speech Quality (PESQ) [5] and E-Model [9], which we will discuss later. However, the MOS based user satisfaction measurement has several inherent shortcomings [1], including the unrepeatable nature of the tests and the fact that the tests are not scalable. Similarly, the PESQ and E-Model cannot provide accurate objective measurement on user perceived quality as well as user satisfaction [8]. PESQ does not consider several factors such as end-to-end delay in its computations. On the other hand, the E-Model, although it includes conversational voice quality measurements perceived by end-users, it still ignores the dynamic nature of IP networks and subsequently excludes the impairment caused by inappropriate playout buffering.

It is highly desired to have a more comprehensive and accurate quality measurement model that can provide objective measurement of QoS in real-time while simultaneously offering QoE measurement as well. Thus, unveiling the relevance of various factors in a quality measurement model is crucial.

In this paper, we study a commonly used ITU-T recommended quality measurement model: the E-model with a focus on how an important but ignored parameter, the jitter playout buffer, affects the measurement accuracy by conducting an analytical and experimental approach. Our study shows that adaptive playout buffering does affect the measurement of end user perceived VoIP quality. Thus, the E-Model that excludes the impact of the dynamic playout buffer cannot provide accurate measurement on end user perceived conversational quality in VoIP applications.

The rest of the paper is organized as follows. Section II first presents a comprehensive literature review. Then, Section III discusses the three ITU-T recommended models that can provide subjective and objective measurements of

 ISBN:978-1-61208-113-7

conversational quality. Section IV analyzes several objective measurement factors that determine quality of service. Section V provides the experimental study via simulations. Section VI concludes our study and briefly discusses our future work.

II. Related Work

There are numerous approaches proposed to objectively measure speech quality in VoIP. Robinson and Yedwab [10] proposed a Voice Performance Management system to monitor call quality in real-time by proactively monitoring, alerting, troubleshooting and reporting network performance problems. Robinson and Yedwab [10] concluded that only packet loss, jitter and latency show the correlations between QoS and QoE.

Gierlich and Kettler [11] provided insight into the impact of different network conditions and the acoustical environment on speech quality. Testing techniques for evaluating speech quality under different conversational aspects were also described. Gierlich and Kettler [11] argued that there is no single number that can objectively indicate speech quality; and pointed out that overall speech quality is a combination of different single values from different speech quality parameters. Wang et. al., [12] designed and implemented a QoS-provisioning system that can be seamlessly integrated into current Cisco VoIP systems. Wang et. al., [12] also described Call Admission Control (CAC) mechanisms (Site-Utilization-based CAC and Link-Utilization-based CAC) to prevent packet loss and over-queuing in VoIP systems.

Myakotnykh and Thompson [13] described an algorithm for adaptive speech quality management in VoIP communications, which can show a real-time change in speech encoding parameters by varying voice packet sizes or compression (encoding) schemes. The algorithm involves the receiver making control decisions based on computational instantaneous quality level (which is calculated per talkspurt using the E-Model) and perceptual metric (which estimates the integral speech quality based on latency, packet loss and the position of quality degradation period in the call). Myakotnykh and Thompson [13] calculated the maximum achievable quality level for a given codec under specific network conditions, packet playout time, packet delay before jitter buffer and degradation in quality caused by traffic burstiness and high network utilization. The algorithm however results in an increase in average quality without increasing individual call quality.

Raja, Azad and Flanagan [14] designed generalized models to predict degradation in speech quality with high accuracy, in which genetic programming is used to perform symbolic regressions to determine Narrow-Band (NB) and Wide-Band (WB) equipment impairment factors for a mixed NB/WB context. Zha and Chan [15] described two algorithms for objective measurement of speech quality: single-ended (needing only to input the degraded speech signal) and double-ended (needing both the original and degraded speech signals). The algorithm developed by Zha and Chan [15] can objectively measure in real-time speech quality using statistical data mining methods.

Several algorithms have also been proposed to optimize some of the existing ITU-T models. The goal of optimization is to enhance existing models by correcting weaknesses that are identified in the models. Gardner, Frost and Petr [16] proposed an algorithm to optimize the E-Model by considering coder selection, packet loss, and link utilization. The authors however stated that the algorithm would have to be enhanced if used in a wide area network involving multiple users. Mazurczyk and Kotulski [17] proposed an audio watermarking method based on the E-Model and the MOS, which provides speech quality control by adjusting speech codec configuration, playout buffer size and amount of Forward Error Correction (FEC) mechanism in VoIP under varying network conditions.

One of the limitations of the E-model is the fact that the model does not consider the dynamic nature of underlying networks that support VoIP. This limitation is addressed by several authors designing adaptive playout buffering to improve voice quality in VoIP. Most of these studies either optimize the E-Model, the PESQ or combine the PESQ and the E-Model to propose a more holistic solution. Mazurczyk and Kotulski [17] highlighted two problems that are associated with adaptive playout buffering: how to estimate current network status and how to transfer network status data to the sending or receiving side. Wu et. al., [18] admitted that VoIP playout buffer size has long been a challenging optimization problem, as buffer size must balance the dynamics of conversational interactivity and VoIP speech quality. Wu et. al., [18] stated that the optimal playout buffer size yields the highest satisfaction in a VoIP call. Wu et. al., [18] investigated the playout buffering dimensions in Skype, Google Talk and MSN Messenger. Wu et. al., [18] concluded that MSN Messenger produces the best performance in terms of adaptive playout buffering, while Skype does not adjust its playout buffering at all. Narbutt and Davis [19] stated that the management of playout buffering is not regulated by any standard and is therefore vendor specific. Narbutt and Davis [19] proposed a scheme that extends the E-Model and provides a direct link to perceived speech quality. Narbutt and Davis [19] evaluated various playout algorithms in order to estimate user satisfaction from time varying transmission impairments including delay, echo, packet loss and encoding scheme.

III. The Analysis of Quality of Experience Measurement Models

In this section we briefly review three commonly adopted ITU-T recommended QoE measurement methods: Mean Opinion Score (MOS), Perceptual Evaluation of Speech Quality (PESQ) and E-Model.

ISBN:978-1-61208-113-7

A. Mean Opinion Score

Mean Opinion Score or MOS has been endorsed by ITU-T as a subjective method to evaluate voice transmission quality. The MOS test involves using a group of testers (listeners) to assign a rating to a voice call. The quality is rated on a scale of 1 to 5, with 1 = bad, 2 = poor, 3 = fair, 4 = good and 5 = excellent [2]. The arithmetic mean of the scores provided by all listeners becomes the final MOS value of the voice call. Assessment ratings can also be obtained by clustering the test results as "Good or Better" or as "Poor or Worse", and further calculating the relative ratio or percentage of each type of results. For a given voice call, these results are expressed as "Percentage Good or Better" (%GoB) and "Percentage Poor or Worse" (%PoW) [3]. Table I shows the MOS rating, %GoB, %PoW and the correlation between each rating [4].

Table I: Subjective Ratings for Measuring QoE

User Satisfaction	**MOS** (5)	**%GoB** (100)	**%PoW** (0)
Very Satisfied	4.3-4.4	97.0-98.4	0.2-0.1
Satisfied	4.0-4.29	89.5-96.9	1.4-0.19
Some Dissatisfied	3.6-3.9	73.6-89.5	5.9-1.39
Many Dissatisfied	3.1-3.59	50.1-73.59	17.4-5.89
Nearly All Dissatisfied	2.6-3.09	26.59-50.1	37.7-17.39
Not Recommended	1.0-2.59	0-26.59	99.8-37.69

The advantage of the MOS is that it can provide an offline analysis of end-user opinions. However, MOS tests cannot provide an absolute reference for the evaluations; that is, MOS ratings are dependent on the expertise of listeners [1]. Moreover, MOS tests cannot be used in large scale experiments that involve a large number of users because of the involved overhead (e.g., test setup). Moreover, MOS tests are unrepeatable by nature.

B. Perceptual Evaluation of Speech Quality

ITU-T P.862 (PESQ) involves the comparison between the original or reference speech samples and the ones traversed through a test network channel. The more similar the output signal is to the reference signal, the higher the score assigned to the quality of the transmission channel. The comparison result, referred to as the PESQ, is in the range of -0.5 to 4.5 which can be linearly projected to the corresponding MOS score. The PESQ, however, cannot comprehensively represent conversational voice quality due to its exclusion of several network and system parameters including end-to-end delay, echo, listening level, sidetone, loudness loss, Enhanced Variable Rate Codec (EVRC) [5].

C. E-Model

The E-Model is designed to measure the instant user perceived quality instead of the cumulative effect during an entire conversation. The E-Model assumes that individual impairment factors are additive on a psychological scale and combines the cumulative effects of these factors into the Transmission Rating Factor, R, which can be transformed into other quality measures like the MOS, Percentage Good or Better (%GoB) or the Percentage Poor or Worse (%PoW). The R-rating is on a scale of 0 to 100, with high values of R between 90 and 100 interpreted as excellent quality, while lower values of R indicate a lower quality. Values of R below 50 are considered unacceptable and values above 94.15 are assumed to be unobtainable in narrowband telephony. The E-Model measures individual impairment factors at different points in time to compute the R-rating. The value of the R-rating is consequently associated with measurements taken at a given time point and does not reflect the dynamic nature of quality during the entire length of a conversation. The following formula shows the computation of the R-rating:

$$R = R_0 - I_s - I_d - I_e + A \quad (1)$$

R_O represents the basic signal-to-noise ratio, including noise sources such as circuit noise and room noise. The factor I_s is a combination of all impairments which occur simultaneously with the voice signal. The factor I_d represents the impairments caused by delay, and the effective equipment impairment factor I_e represents impairments caused by low bit-rate codecs and packet-losses of random distribution. The advantage factor A corresponds to the user allowance due to the convenience when using a given technology.

The E-Model not only takes in account the transmission statistics (transport delay and network packet loss), but it also considers the voice application characteristics, like the codec quality, codec robustness against packet loss and the late packets discard. However, the impairment due to playout buffer size is simply excluded, which consequently results to the ignorance on how the dynamic varying of the playout buffer size affects QoE throughout an entire conversation. In this paper, we conduct various simulations and experiments to demonstrate that the exclusion of the affect of the playout buffer in the E-model lessen the accuracy of measurements of the conversational quality as perceived by end-users.

IV. The Analysis of Objective Quality Measurement Factors

Quality of Service is determined primarily by latency, packet loss and jitter presented in transmission networks which eventually impose their impacts on user perceived QoE. The ITU-T recommends that as long as the value of latency in a network running VoIP is lower than 150ms, users are expected to be highly satisfied if other factors that may affect network performance are negligible. When latency is between 150ms and 200ms, perceivable performance degradation in quality is expected and becomes more severe with the increase of latency.

VoIP is also sensitive to packet loss, which may result in unintelligible conversation if the packet loss rate is high.

 ISBN:978-1-61208-113-7

The ITU-T recommends that as long as the packet loss rate is less than 1%, users are expected to be highly satisfied if other factors that may affect network performance are negligible. Our simulation via OPNET also shows the same observation. We discuss and show the experimental setup later in Section V. As shown in Figure 1, with increasing packet loss rate, the end user perception of quality gradually decreased.

While gradually increasing the value of packet loss in the network beyond the recommended 1% level, the researcher assigned a score based on the MOS scale. All other impairments were kept at levels where there effects will not be obvious so as not to interfere with the effect of packet loss. Figure 1 shows that as packet loss increased in the transmission network, the end user perception of quality gradually decreased. It is expected that if packet loss decrease in the transmission network and other factors are kept constant, the end user perception of quality will improve. Figure 1 also shows the packet loss in this OPNET simulation was not evident for all three IP phones until after 25 sec.

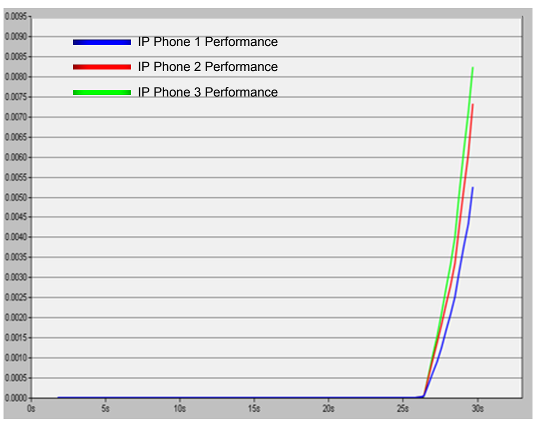


Figure 1: The Effect of Packet Loss Rate on QoE.

V. Experimental Study

This section discusses some of the results that were obtained during experiments that we conducted.

A. Experiment Setup

Figure 2 shows the experimental setup in OPNET that was used to simulate 200 audio conference calls, each with 30s duration. The simulation model included three different participants in the conference. Each of the IP phone nodes had DHCP assigned IP addresses on the same network segment with the two client computer nodes. The experiment simulated a typical packet switched network scenario and used a fixed jitter buffer size that does not dynamically adjust with varying jitter in the network. MOS scores from users associated with the three different IP phone nodes were also simulated and recorded in the experiment.

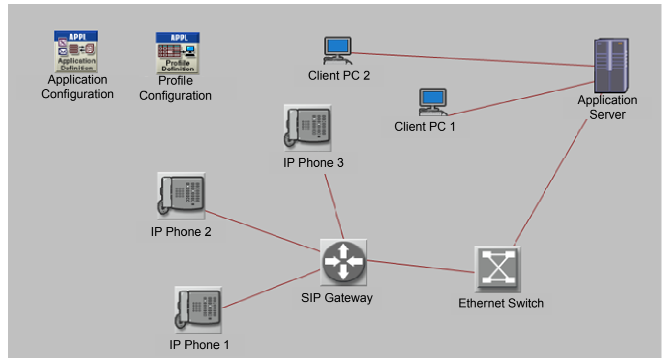


Figure 2: Experiments Setup.

B. Experiment Results

Figure 3 shows the main observations that were recorded during the simulations. The results include MOS values associated with jitter delay, jitter loss and end-to-end delay for the 200 voice calls measured at each IP Phone.

Figure 3(a) shows how static jitter buffer sizes affect the MOS of users. The MOS is taken at every 0.2s during the simulation. The result shows that users' opinion fluctuated throughout the duration of the conversation as a result of the inability of the jitter buffer size to adapt to varying values of jitter in the network.

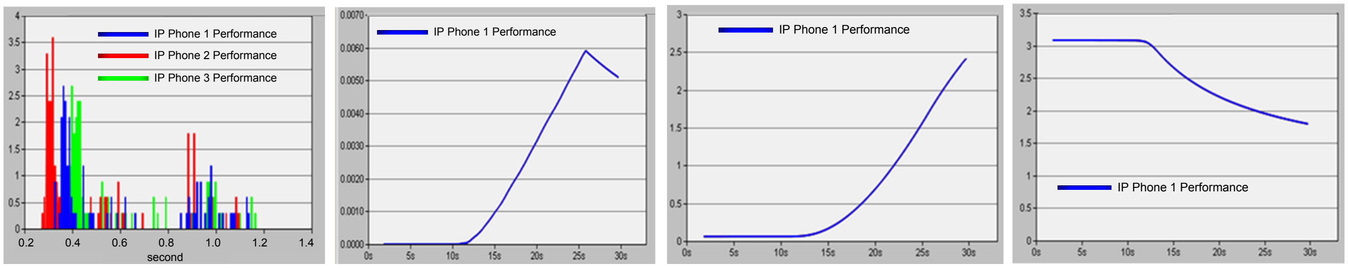
Since the similar performance results were observed for all three IP phones in our experiments, we only show the experiments result from IP phone 1. Figure 3(b) shows the jitter measured at IP phone 1 during the simulation, which demonstrates that after 10s jitter was continuously increasing for about 15s. In the experiment, the jitter varying greatly with the change of various conditions (e.g., available bandwidth, packet loss rate). Figure 3(c) shows the end-to-end delay presented in the network was negligible at the beginning while the network condition was good; however, started increasing significantly after about 15s while the network condition getting worse. Correspondingly, Figure 3(d) shows the MOS varying and reflecting the change of the network condition.

C. Playout Buffering

Mitigating the impact of jitter involves collecting packets in a jitter buffer and playing the packets out relative to the size of the jitter buffer. The size of the jitter buffer affects the end-to-end delay on the network and also the packet loss rate. The size of the jitter buffer may either be kept fixed (static playout buffering) or dynamically adjusted (adaptive playout buffering) with respect to the variations in jitter presented in the network.

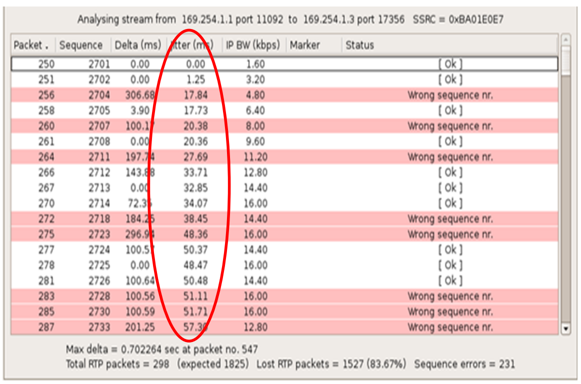
Figure 4 shows two different Wireshark interfaces for the same VoIP session. Wireshark allows the playback of different segments of the entire conversation stream. Playing back voice segments and matching the values of jitter present

 ISBN:978-1-61208-113-7



(a) MOS jitter Delay (b) Average Jitter (sec) (c) Average Delay (sec) (d) Average MOS

Figure 3: Voice Application Performance



Analysing stream from 169.254.1.1 port 11092 to 169.254.1.3 port 17356 SSRC = 0xBA01E0E7

Packet .	Sequence	Delta (ms)	Jitter (ms)	IP BW (kbps)	Marker	Status
250	2701	0.00	0.00	1.60		[Ok]
251	2702	0.00	1.25	3.20		[Ok]
256	2704	306.68	17.84	4.80		Wrong sequence nr.
258	2705	3.90	17.73	6.40		[Ok]
260	2707	100.1	20.38	8.00		Wrong sequence nr.
261	2708	0.00	20.36	9.60		[Ok]
264	2711	197.74	27.69	11.20		Wrong sequence nr.
266	2712	143.88	33.71	12.80		[Ok]
267	2713	0.00	32.85	14.40		[Ok]
270	2714	72.3	34.07	16.00		[Ok]
272	2718	184.25	38.45	14.40		Wrong sequence nr.
275	2723	296.94	48.36	16.00		Wrong sequence nr.
277	2724	100.5	50.37	14.40		[Ok]
278	2725	0.00	48.47	16.00		[Ok]
281	2726	100.64	50.48	14.40		[Ok]
283	2728	100.56	51.11	16.00		Wrong sequence nr.
285	2730	100.59	51.71	16.00		Wrong sequence nr.
287	2733	201.25	57.3	12.80		Wrong sequence nr.

Max delta = 0.702264 sec at packet no. 547
Total RTP packets = 298 (expected 1825) Lost RTP packets = 1527 (83.67%) Sequence errors = 231

Figure 4: The Wireshark Screenshot Showing the Presence of Jitter.

in each segment confirmed that jitter affects the end user perception of quality.

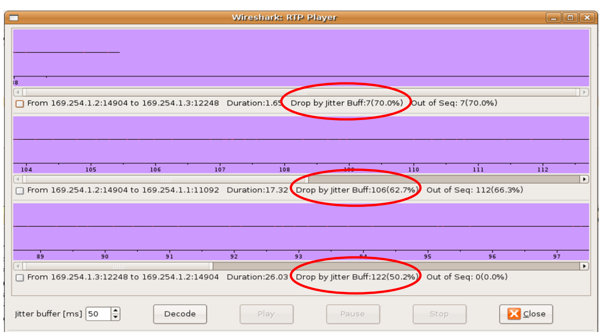


Figure 5: The Wireshark Screenshot Showing the packets dropped by jitter buffer.

Figure 5 shows the results of a typical conversation analyzed using the experimental design shown in Figure 2. Different speech segments of the VoIP session, the source and destination address of the speech segment, together with the number of packets that are out of sequence and those dropped by jitter buffer are shown in Figure 5. Several VoIP sessions were tested using different values of jitter buffer, packet loss, available bandwidth, QoS mechanism, noise reduction, echo cancellation, and different choice of codec.

Static playout buffer size is easier to implement when compared to adaptive playout buffering. However, static playout buffering might sometimes be either too small or too large to handle the fluctuating values of jitter in the network. When the buffer size is too small, slower packets are not played out and dropped. When the buffer size is too large, packets are held longer than necessary and consequently reduce conversational quality.

In contrast to using fixed buffer size, adaptive playout buffering continuously estimates the instant jitter in a network, and further dynamically adjusts the playout buffer to match the varying jitter during a whole conversation. This ensures that packets associated with each conversation segment are played out with the same playout delay. In order to ensure the optimal QoE for end-users, adaptive playout buffering will guarantee that conversational quality is maintained at consistent levels, compared to fluctuating levels associated with static playout buffering.

The output of the E-Model is expected to be a reflection of the quality associated with the entire length of a conversation. However, the computation uses snapshots of impairment values taken at different points within an entire conversation and does not include the affect of the playout buffer. Figure 3(a) shows that the MOS measured at different points during the conversation varied from the beginning of the conversations to the end of the conversations. The measured MOS shown in Figure 3(d) is corresponding to the jitter variations shown in Figure 3(b). At the beginning of the conversations when jitter was zero or almost zero, the highest MOS value was recorded. As jitter increased in the network, the MOS value recorded decreased. Towards the end of the simulation, the MOS value recorded started to increase again.

Adaptive playout buffering would ensure that the variations in jitter in the network are matched by a corresponding value of playout buffer size. This would also ensure that the end-to-end delay in the network is kept at optimal levels and rate of packet loss in the network is kept below the recommended 1%. We will report on the inclusion of the playout buffer in the E-Model in a subsequent paper.

The results of our simulations show that the relevance

 ISBN:978-1-61208-113-7

of adaptive playout buffering should not be ignored in the evaluation of end user perception of conversational quality. Including measurements of adaptive playout buffering into the computations of the E-Model will either reduce the value of the R-rating or keep the value the same, depending on calculated effects of jitter buffer. A modification to the value of the R-rating will provide a more accurate mapping to the MOS. This will in turn give a more accurate prediction of the end user perception of conversational quality.

VI. Conclusion

Through the series of experiments conducted in this study and data analyzed, it was obvious that the quality of experience of end users is a subject that should never be overlooked in the design and development of VoIP solutions. Quality of experience ultimately determines the migration of users from circuit-switched networks to VoIP and the next generation of services. Quality of service related to network configuration and performance was shown in this study as ineffective in ensuring the desired user experience, if other factors that contribute to the quality of experience of end users are not brought to cognizance.

In our future work, we will pursue a longitudinal study tracking the effects of cost of VoIP service, security provision, usability, human behavior, call session and reliability on quality of experience. The goal would be to determine if there are long term effects of these factors on quality of experience. Development of an approach to capture the opinions of end-users about these factors in different parts of the world might be informative. Developing a methodology to conduct subjective tests to indentify how human behavior varies with time during a conversation depending on the quality impairments present would be a future goal.

References

[1] Sat, B., and Wah, B. W. (2009). Analyzing Voice Quality in Popular VoIP Applications. IEEE MultiMedia, vol. 16, no. 1, pp. 46-59. doi:10.1109/MMUL.2009.2.

[2] Zwar, E. J., and Munch, B. (2006). Voice Quality and Network Capacity Planning for VoIP. Retrieved from http://my.gartner.com/portal/server.pt?open=512&objID=260&mode=2&PageID=3460702&resId=493265&ref=QuickSearch&sthkw=G00140936. Last accessed: 5/6/2010.

[3] Narbutt, M., and Davis, M. (2005). Assessing the Quality of VoIP Transmission Affected by Playout Buffer Scheme. 4th International Conference on Measurement of Speech and Audio Quality, Prague, Czech Republic.

[4] International Telecommunications Union (1996). ITU-T P.800. Methods for Subjective Determination of Transmission Quality. Retrieved from http://www.itu.int/rec/T-REC-P.800-199608-I/en. Last accessed: 5/6/2010.

[5] International Telecommunications Union (2007). ITU-T P.862 Corrigendum. Retrieved from http://www.itu.int/rec/T-REC-P.862-200710-I!Cor1/en. Last accessed: 5/6/2010.

[6] Telecommunications Industry Association (2005). Telecommunications, IP Telephony Equipment and Voice Quality Recommendations for IP Telephony. Retrieved from http://ftp.tiaonline.org/TR-41/tr4112inactive/Public/Latest_Revision_of_PN-4689/PN4689LB.pdf. Last accessed: 5/6/2010.

[7] Morris, M. G., Venkatesh, V., Davis, G. B., and Davis, F.D. (2003). User Acceptance of Information Technology: Toward a Unified View. Retrieved from http://proquest.umi.com/pqdlink?vinst=PROD&fmt=6&startpage=-1&ver=1&vname=PQD&RQT=309&did=420086661&exp=10-30-2014&scaling=FULL&vtype=PQD&rqt=309&TS=1257038425&clientId=43838&cfc=1. Last accessed: 5/6/2010.

[8] Becvar, Z., Mach. P., and Bestak, R. (2009). Impact of Handover on VoIP Speech Quality in WiMAX Networks. Eighth International Conference on Networks, icn, pp.281-286, Gosier, Guadeloupe, France.

[9] International Telecommunications Union. (2008). The E-Model. Retrieved from http://www.itu.int/ITU-T/studygroups/com12/emodelv1/tut.htm. Last accessed: 5/6/2010.

[10] Robinson, P. and Yedwab, D. (2009). Voice and Video Application Performance Management in UC Deployments. Retrieved from http://www.ucstrategies.com/uploadedFiles/UC_Information/White_Papers/Psytechnics/Psytechnics UCS_Final_042109%5B1%5D.pdf. Last accessed: 5/6/2010.

[11] Gierlich, H. W. and Kettler, F. (2006). Advanced speech quality testing of modern telecommunication equipment: an overview. Signal Processing, 86(6), 1327 - 1340.

[12] Wang, S., Mai, Z., Xuan, D., and Zhao, W. (2006). Design and Implementation of QoS-Provisioning System for Voice over IP. IEEE Transactions on Parallel and Distributed Systems, vol. 17, no. 3, pp. 276-288.

[13] Myakotnykh, E. S. and Thompson, R. A. (2009). Adaptive Speech Quality Management in Voice-over-IP Communications. Fifth Advanced International Conference on Telecommunications, aict, pp.64-71, Venice/Mestre, Italy.

[14] Raja, A., Azad, R. M. A., and Flanagan, C. (2008). VoIP Speech Quality Estimation in a Mixed Context with Genetic Programming. 10th Annual Conference on Genetic and evolutionary computation, Atlanta, Georgia, United States.

[15] Zha, W. and Chan, W. (2005). Objective Speech Quality Measurement Using Statistical Data Mining. EURASIP Journal on Applied Signal Processing, no. 9, 1410-1424.

[16] Gardner, M., Frost, V.S. and Petr, D.W. (2003). Using optimization to achieve efficient quality of service in Voice over IP networks. IEEE International Performance, Computing, and Communications Conference, Phoenix, Arizona, United States.

[17] Mazurczyk, W. and Kotulski, Z. (2007). Adaptive VoIP with Audio Watermarking for Improved Call Quality and Security. Journal of Information Assurance and Security 2, 226-234. Retrieved from http://www.mirlabs.org/jias/mazurczyk.pdf. Last accessed: 5/6/2010.

[18] Wu, C., Chen, K., Huang, C., and Lei, C. (2009). An Empirical Evaluation of VoIP Playout Buffer Dimensioning in Skype, Google Talk and MSN Messenger. Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital and Video, Williamsburg, VA, United States.

[19] Narbutt, M. and Davis, M. (2005). Assessing the Quality of VoIP Transmission Affected by Playout Buffer Scheme. 4th International Conference on Measurement of Speech and Audio Quality, Prague, Czech Republic.

 ISBN:978-1-61208-113-7