THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA Vol.28, No.3E 2009. 9. pp. 99~108

Voice Quality Criteria for Heterogenous Network Communication Under Mobile-VoIP Environments

Jae-Hun Choi*, Soonuk Seol**, Joon-Hyuk Chang*
*School of Electronic Engineering, Inha University of Incheon
**Platform Service Department Central R&D Laboratory, KT
(Received August 4 2009; accepted September 10 2009)

Abstract

In this paper, we suggest criteria for objective measurement of speech quality in mobile VoIP (Voice over Internet Protocol) services over wireless mobile internet such as mobile WiMAX networks. This is the case that voice communication service is available under other networks. When mobile VoIP service users in the mobile internet network based on packet call up PSTN and mobile network users, but there have not been relevant quality indexes and quality standards for evaluating speech quality of mobile VoIP. In addition, there are many factors influencing on the speech quality in packet network. Especially, if the degraded speech with packet loss transfers to the other network users through the handover, voice communication quality is significantly deteriorated by the transformation of speech codecs. In this paper, we eventually adopt the Gilbert-Elliot channel model to characterize packet network and assess the voice quality through the objective speech quality method of ITU-T P. 862. 1 MOS-LQO for the various call scenario from mobile VoIP service user to PSTN and mobile network users under various packet loss rates in the transmission channel environments. Our simulation results show that transformation of speech codecs results in the degraded speech quality for different transmission channel environments when mobile VoIP service users call up PSTN and mobile network users.

Keywords: *Heterogenous network, ITU-T P. 862.1 MOS-LQO, Gilbert-Elliot Channel Model, Mobile-VoIP*

I. Introduction

As the needs for high quality communication services have been increased, the convergence of the different networks can be recently seen, e.g. merging conventional mobile communication networks (2G, 3G) and wireless mobile internet networks such as mobile WiMAX or WiBro network. The convergence of the various network causes the possibility of degraded voice quality due to the different speech codec standards and network protocols of each network.

In the recent trends, mobile VoIP services over wireless mobile internet such as WiMAX networks are considered as the most interesting topic. Figure 1 shows the comparison of the mobile VoIP network and conventional mobile communication network. Since the wireless mobile internet network adopts packet-switched transport based on IP, the speech quality through the mobile VoIP network may suffer from time-varying degradations such as the packet loss and jitter. In mobile VoIP and most wired VoIP applications, jitter is not considered as an important factor of the network performance and the voice quality due to jitter buffers. However, packet loss significantly affects the voice distortion and the major reasons of the voice deterioration.

Corresponding author: Joon-Hyuk Chang (changjh@inha.ac.kr)
School of Electronic Engineering, Inha University, 253 Yonghyeon-dong, Nam-gu, Incheon 401-751, Republic of Korea

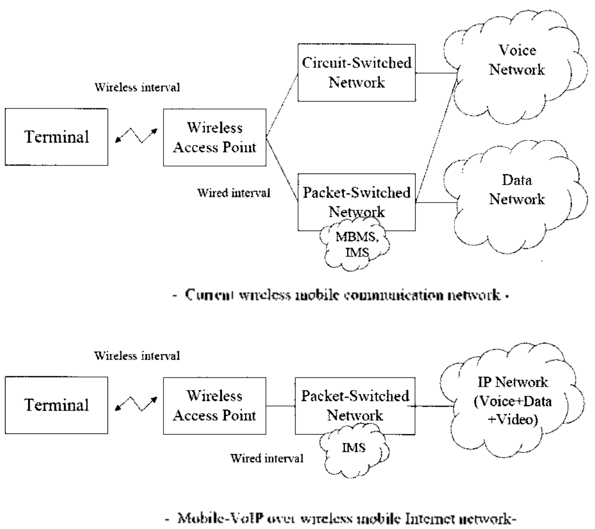


Fig. 1. Classification of the mobile VoIP network and the current wireless mobile communication network by Ji -Hyung Han KT (2008) [1].

Even though there have been many attempts to evaluate the voice quality of the conventional wire-less communication system such as 2G and 3G (WCDMA) and a wired VoIP for the QoS (Quality of Service), there is not a relevant quality criteria and the QoS standards of voice quality for the mobile VoIP [2]-[4]. Furthermore, there are not the objective quality criteria and QoS standards for the various call scenarios of heterogenous network communication such as the cases of mobile VoIP (G.729A)-to-CDMA (2G: EVRC), mobile VoIP (G.729A)-to-CDMA (3G: AMR), and mobile VoIP (G.729A)-to-PSTN (G.711) according to the handover effect [5].

In general, there are two type methods for the voice quality measurement. The mean opinion score (MOS) as the subjective method is mostly used in the voice quality evaluations recommended by the ITU-T P. 800 [6]. The ITU-T P. 800 MOS tech-nique is a reliable approach to assess the voice quality but it is difficult to be conducted due to time con-suming and expensive costs. In order to address these problems, the objective methods have been proposed to provide the estimation of the objective MOS such as the intrusive objective measurement of Perceptual Analysis/Measurement System (PAMS) [7], Perceptual Evaluation of Speech Quality (PESQ), ITU -T P. 862 [8] or nonintrusive techniques of ITU-T E-model [9] and ITU-T P. 563 [10]. The ITU-T E-model is widely used to predict the voice quality through the system parameters. Indeed, the E-model as the nonintrusive method has been adopted to evaluate the voice quality for wired VoIP applications in the current internet network. But, the E-model method is restricted to the cases which communicate with heterogenous network communication such as public switched telephone systems or current mobile co-mmunication networks by the various call scenarios. Especially, the E-model can not continuously measure the network performance or the end-to-end voice quality when mobile VoIP service users call up the 2G network (EVRC), 3G network (AMR) and PSTN (G.711) users under the mobile VoIP environments.

In this paper, we propose the objective speech quality criteria and QoS standards to provide mobile VoIP services over wireless mobile internet network under heterogenous communication networks. For the various types of call scenarios, we evaluate the voice quality according to the transformation of the speech codecs under the heterogenous network co-mmunication. Various packet losses are generated by the Gilbert-Elliot channel model which models the packet network based on IP network and the ITU-T P. 862. 1 MOS-LQO (Mean Opinion Score-Listening Quality Objective) and PESQ are used to evaluate the degraded voice quality due to the packet loss [14][15].

II. The subjective and objective mea-surements for voice quality in tele-communication networks

In telecommunication networks, the quality of communicated speech is the most important criterion of QoS which could be categorized with three com-ponents including the speech or voice communication quality, the service performance, and the terminal equipment performance [11]. Among the three com-ponents of QoS, the voice communication quality mainly contributes to the whole communication quality perceived by a user. The evaluation measurement of the voice quality has been divided into the subjective and the objective methods.

The subjective method of voice communication quality is representatively the traditional mean opinion score (MOS) defined in ITU-T P. 800 which ranks the subjective rating of user satisfaction, where each score ranges from 1 to 5. The MOS scores are averaged as a mean opinion to represent the five point scales with a definition of voice quality: Excellent (5), Good (4), Fair (3), Poor (2), and Bad (1). Even though the subjective MOS method is broadly used to evaluate the voice transmission quality, the MOS method has many defects which are lacks of statistical reliability and time consuming, expensive costs. Also, it is not possible to implement the subjective MOS method for the real-world network connections.

In order to address these disadvantages, the various objective evaluation techniques such as the ITU-T P. 862 PESQ, ITU-T G. 107 E-model, ITU-T P. 563, and PSQM have been proposed [8][10]. Among these objective measurements for voice quality evaluations, the PESQ and E-model are dominantly used for mobile wireless network, planning public switched telephone networks (PSTNs), and current VoIP systems.

In particular, the E-model is the most popularly used in non-intrusive objective measurement which predicts conversational MOS by directly computing the system parameters such as the packet loss, delay, jitter and speech codec obtained from the IP network [9]. Since the E-model depends on the system parameters extracted from the IP networks

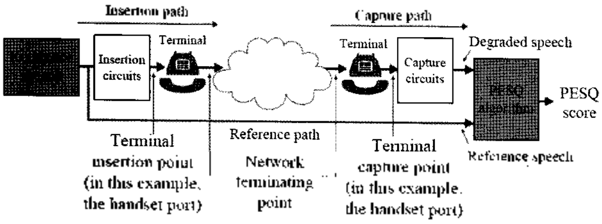


Fig. 2. The block diagram of PESQ evaluation for speech quality.

for evaluating the voice quality, it is hard to continuously measure the communicated voice quality such as the call scenarios for CDMA (2G or 3G)-to-PSTN and PSTN-to-CDMA (2G or 3G).

Specifically, if the handover occurs under mobile VoIP network environments, the users in mobile VoIP network could call up the mobile wireless network users. However, the E-model could only predicts the conversational MOS before the occurrence of the handover so the E-model is restricted for the heterogenous network communication which has the different speech codecs and protocols.

As shown in Fig. 2, the ITU-T P. 862 Perceptual Evaluation of Speech Quality (PESQ) as the intrusive objective method is commonly used to measure the end-to-end QoS of the networks since the PESQ compares the distortion between a original reference signal and a degraded signal. Also, the PESQ method has been used for evaluating the QoS of the current mobile wireless networks as call quality evaluation models [12][13].

Considering the evaluation measurement of call quality for the end-to-end user under mobile VoIP networks which communicate with heterogenous network, the PESQ is more suitable than the conventional E-model technique. In addition, it is possible to continuously measure the end-to-end call quality according to the occurrence of handover. Since the PESQ scores range from -0.5 to 4.5, it is not possible to linearly compare the PESQ scores with ITU-T P. 800 subjective MOS scores. For this reason, the ITU-T P. 862. 1 MOS-LQO (Mean Opinion Score-Listening Quality Objective) of the estimated MOS average scores as the transformation of the PESQ raw data is provided. Therefore, in order to establish QoS criteria and indexes for mobile VoIP users communicating with the PSTN and current mobile wireless network users by the handover, the ITU-T P. 862. 1 MOS-LQO methods should be introduced for the heterogenous network communication [14][15].

II. The call quality measurement based on ITU-TP. 862. 1 MOS-LQO for the heterogenous network communication

2.1. The call scenarios of the Mobile-VoIP

With the trends of merging the conventional wireless communication networks and the mobile wireless internet networks, the effort of improving the speech or voice communication quality between the conventional networks such as public switched telephone network (PSTN) and the wireless networks (2G, 3G) and mobile VoIP network is one of the hot issues. According to the convergence network environments, there are many call scenarios from the PSTN and wireless network to mobile VoIP network or reversely. Based on the classifications of the network environments, the various call scenarios are summarized in Table 1 mobile VoIP users are defined as the consumers which utilize the telephony services through the wireless mobile internet such as WiBro.

Considering the classifications of the networks, call scenarios for mobile VoIP users experiencing different network communication could be divided into two cases. One is the case that mobile VoIP users communicate with the homogeneous network consisting of the same speech codecs and network protocols and the other is the scenario that mobile VoIP users call up the heterogeneous network using the different speech codec standards and network protocols. When the handover from mobile VoIP network to the heterogeneous network occurs, we should consider the call scenarios according to handover effects. In Table 2, the call scenario based on the change of the network is shown as a representative case. Also, the effect of call quality according to handover is given in Table 3.

Table 1. Classification of communication networks and relationship with mobile VoIP.

networks	note
PSTN	Circuit networks Target network which communicates PSTN users with mobile VoIP users
Wired internet	Packet networks, wired VoIP Target network which communicates PSTN users with mobile VoIP users
WLAN	Packet networks, wireless VoIP WLAN is classified as mobile VoIP according to the mobility
WiBro	Packet networks. mobile VoIP user networks
CDMA	Circuit networks (Voice) Target network which communicates PSTN users with mobile VoIP users
WCDMA	Circuit networks (Voice) Target network which communicates PSTN users with mobile VoIP users
Satellite	Satellite communication networks Applying the Backhaul networks

Table 2. Classification of call scenario under mobile VoIP environments.

Network change of mobile VoIP users according to handover (HO)			Network of the counterpart callers	Change of communication network
Type	prior to HO	following HO		
Homogeneous network	WiBro	WiBro	PSTN	(2) Homogeneous network is reserved.
	WiBro	WiBro	Wired internet	(1) Homogeneous network is reserved.
	WiBro	WiBro	WiBro	(1) Homogeneous network is reserved.
	WiBro	WiBro	CDMA	(2) Heterogeneous network is reserved.
Heterogeneous network	WiBro	CDMA	PSTN	(5) Heterogeneous network →Heterogeneous network
	WiBro	CDMA	Wired internet	(3) Homogeneous network →Heterogeneous network
	WiBro	CDMA	WiBro	(3) Homogeneous network →Heterogeneous network
	WiBro	CDMA	CDMA	(4) Heterogeneous network ›Homogeneous network

※ The wired internet network is classified with the homogeneous network because the wired internet network is based on packet switched network such as WiBro and so there is not the transformation of the protocol and the speech codecs

Table 3. The effects of speech quality according to call scenario.

The transformation of the call scenario according to HO	Effect of call quality
(1) Homogeneous network is reserved.	There are no changes.
(2) Heterogeneous network is reserved.	There are no changes. (Initial level is reserved)
(3) Homogeneous network →Heterogeneous network	There is a possibility to reduce the voice quality due to the transformations of speech codecs and the handover effect
(4) Heterogeneous network →Homogeneous network	There is a possibility of improving the voice quality according to the enhancement of delay, jitter.
(5) Heterogeneous network →Heterogeneous network	There is a possibility of the improvement or the degradation of voice quality according to the type of the heterogeneous networks.

2.2. The voice quality measurement based on the ITU-T P. 862. 1 MOS-LQO for the QoS criteria of Mobile VoIP

As we mentioned before, since mobile VoIP service offers the voice service in the packet-switched networks such as mobile WiMAX networks based on IP, the packet loss affects the voice quality of end-to-end users. Also, mobile VoIP users experience the heterogenous network communication according to the handover where heterogenous networks have different speech codec standards and protocols. Therefore, the E-model method could not evaluate the voice quality of the heterogenous network communication due to the handover. For this reason, the PESQ method is more suitable than the E-model method, because the PESQ measurement is possible to continuously evaluate the actual communicated call quality for end-to-end users under the heterogenous network communication occurring to the handover.

Since the PESQ-MOS scores range from -0.5 to 4.5 as mentioned in Section II, we adopt the mapping function of the ITU-T P. 862. 1 MOS-LQO to linearly compare the raw data from the PESQ with the subjective MOS of ITU-T P. 800. Through the transformation of the raw data from the PESQ into MOS-LQO based on ITU-T P. 862. 1, we could

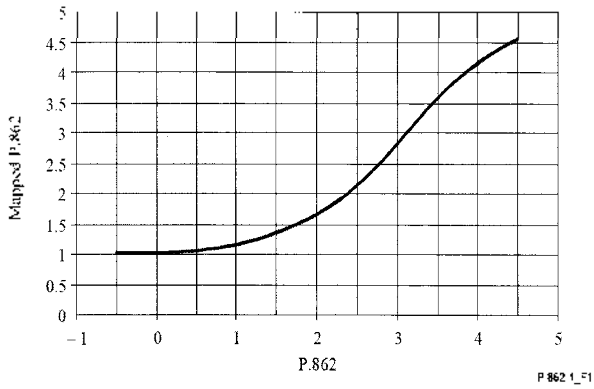


Fig. 3. ITU-T P. 862.1-P. 862 algorithm's mapping graph.

establish the QoS criteria for the heterogenous network communication under mobile VoIP environments.

The ITU-T P. 862. 1 Recommendation offers a single mapping function to convert the P. 862 raw scores into the MOS-LQO scores. The mapping function could be given as follows [14]:

$$y = 0.999 + \frac{4.999 - 0.999}{1 + e^{-1.4945 \cdot x + 4.6607}} \qquad (1)$$

where y depicts the mapped MOS-LQO score which should be within ±0.01 absolute of the curve according to the graph of the mapping function and x denotes the raw data obtained from the PESQ algorithm. Also, the ITU-T P. 862. 1 recommendation provides the mapping graph for technical engineers to conveniently compare the PESQ-MOS with ITU-T P.862. 1 MOS-LQO. In this regards, Figure 3 shows the mapping graph of ITU-T P. 862. 1 Rec [14].

IV. Experimental results

The voice service in mobile VoIP of wireless mobile internet based on the packet-switching is transmitted with IP. For modeling the packet network environments, we substantially employ the Gilbert-Elliot channel model based on the two state Markov model where the Gilbert-Elliot channel model has two states: Good (G) and Bad (B) [16][17]. Fig. 4 represents the Gilbert packet loss model. As shown

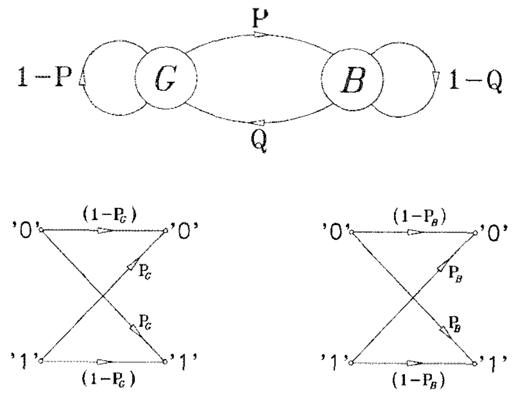


Fig. 4. Gilbert-Elliot channel model.

in Fig. 4, state G depicts a packet received without a packet loss and state B denotes the packet loss respectively. Also, we represent the probabilities associated with the transition between G and B as the symbols P and Q where P is the probability of transition from G state to B state and Q is the probability of transition from B state to G state, respectively. Based on the Gilbert-Elliot channel model, we first introduce the mean bit error probability the BER generated by this packet loss channel model. The BER is given as follows:

$$BER = \frac{P}{1-\gamma} P_B + \frac{Q}{1-\gamma} P_G \tag{2}$$

where

$$\gamma = 1-(P+Q) \tag{3}$$

is a correlation coefficient of the packet errors and of the burst or random characteristic of the communication channel. In (3), $\gamma \approx 0$ means a nearly random error channel environment, while $\gamma \approx 1$ denotes a totally bursty channel environment [16].

Combining (2) and (3), the transition probabilities of P and Q are given by:

$$P = (1-\gamma)(1-\frac{P_B - BER}{P_B - P_G}) \tag{4}$$

$$Q = (1-\gamma)\frac{P_B - BER}{P_B - P_G} \tag{5}$$

The channel environments could be either in good state G or in the bad state B, where the mean error probabilities of P_G and P_B are nearly $P_G=0$ and $P_B=0.5$, respectively. In the our simulation environments, we set the $P_G=0$ and $P_B=0.5$. Applying the $P_G=0$ and $P_B=0.5$ for (4) and (5), the transition probabilities of (4) and (5) are represented such that

$$P = 2(1-\gamma)BER \tag{6}$$

$$Q = (1-\gamma)(1-2BER) \tag{7}$$

Based on the Gilbert-Elliot channel model, we evaluate the call quality of the end-to-end in terms of the variable packet loss rates using ITU-T G. 191 STL (Software Tool Library) for the various call scenarios, where the call scenarios are divided into three cases of mobile VoIP (G.729A)-to-CDMA (2G: EVRC), mobile VoIP (G.729A)-to-CDMA (3G: AMR)) and mobile VoIP (G.729A)-to-PSTN (G.711) [18] [21]. First, the PESQ method of objective measurement is conducted and the PESQ results are transformed into MOS-LQO by the ITU-T P. 862. 1 mapping function for the linear comparison with subjective MOS scores.

In our simulation environments, the channels are assumed to be the random packet losses and burst packet losses. The random packet loss implies that the packet loss is generated uniformly under random loss channel characteristic, while the burst packet loss happens continuously for the consecutive frames. In order to conduct the experiments in terms of the variable packet loss rates and the call scenarios, the packet loss rates are set with the values from 1% to 50% and the call scenarios are assumed to be G.729A-to-EVRC, G.729A-to-AMR, and G.729A-to-G.711, respectively. For the test implementation, thirty test phrases, spoken by four male and four female speakers extracted from the NTT database, were used. Each phrase consisted of two different 8- second meaningful sentences.

To assess the call quality of the heterogenous network communication for the variable packet loss

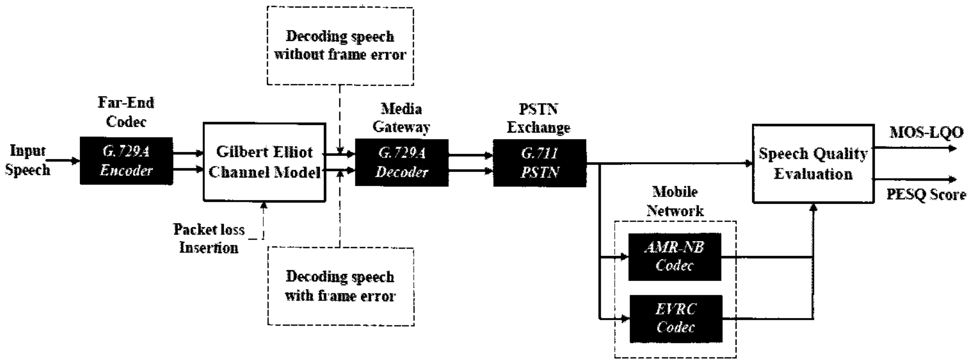


Fig. 5. Block diagram of experimental setup.

Table 4. Relationship between the subjective MOS and user's satisfaction.

MOS	User Satisfaction
4.3	Satisfied
4.0	Very Satisfied
3.6	Some Users Dissatisfied
3.1	Many Users Dissatisfied
2.6	Nearly All Users Dissatisfied
1.0	Not Recommended

rates and the call scenarios under the Gilbert–Elliot channel model, we establish the experimental environments as shown in Fig. 5. In Fig. 5, the PESQ scores were obtained and converted into the ITU–T P. 862.1 MOS–LQO to linearly compare with subjective MOS scores.

At first, we assessed the end–to–end call quality for the three types of call scenarios of G.729A→G.711→AMR (3G), G.729A→G.711→EVRC (2G), and G.729A→G.711 (PSTN) for the packet losses from 1% to 50% under the random and burst channel environments.

The PESQ results and the converted MOS–LQO scores are summarized in Table 5 and 6. In general, the MOS scores which get equal to 4.0 or higher than 4.0 mean the acceptable satisfaction and toll quality. However, the MOS scores which might get higher than 3.1 and smaller than 3.6 represent that some users are slightly satisfied but certain users could be

Table 5. The PESQ and MOS-LQO results of various speech codecs based on the Gilbert-Elliot channel model under random channel environments.

	G.729A→G.711 →AMR (3G)		G.729A→G.711 →EVRC (2G)		G.729A→G.711 (PSTN)	
error rates	PESQ	MOS-LQO	PESQ	MOS-LQO	PESQ	MOS-LQO
0%	3.391	3.396	3.314	3.285	3.531	3.584
1%	3.251	3.194	3.206	3.129	3.326	3.314
3%	3.128	3.013	3.097	2.967	3.184	3.102
5%	3.000	2.824	2.986	2.804	3.039	2.896
10%	2.742	2.459	2.750	2.467	2.766	2.509
15%	2.523	2.180	2.554	2.216	2.554	2.221
20%	2.329	1.958	2.360	1.990	2.352	1.988
25%	2.222	1.846	2.257	1.878	2.245	1.870
30%	2.085	1.720	2.137	1.764	2.098	1.736
35%	1.954	1.613	2.001	1.646	1.967	1.624
40%	1.879	1.560	1.934	1.600	1.883	1.567
45%	1.754	1.474	1.809	1.509	1.770	1.490
50%	1.627	1.400	1.678	1.428	1.582	1.410

Table 6. The PESQ and MOS-LQO results of various speech codecs based on the Gilbert-Elliot channel model under burst channel environments.

	G.729A→G.711 →AMR (3G)		G.729A→G.711 →EVRC (2G)		G.729A→G.711 (PSTN)	
error rates	PESQ	MOS-LQO	PESQ	MOS-LQO	PESQ	MOS-LQO
0%	3.391	3.396	3.314	3.285	3.531	3.584
1%	3.391	3.396	3.314	3.285	3.517	3.566
3%	3.127	3.011	3.084	2.948	3.209	3.130
5%	2.935	1.732	2.898	2.678	2.996	2.820
10%	2.616	2.293	2.603	2.274	2.658	2.349
15%	2.335	1.961	2.340	1.965	2.370	2.001
20%	2.038	1.677	2.059	1.693	2.059	1.697
25%	1.980	1.629	2.005	1.647	2.003	1.650
30%	1.954	1.613	1.985	1.634	1.983	1.635
35%	1.874	1.552	1.906	1.572	1.891	1.566
40%	1.694	1.439	1.740	1.464	1.712	1.449
45%	1.697	1.443	1.748	1.471	1.751	1.454
50%	1.625	1.400	1.686	1.432	1.639	1.408

Table 7. MOS-LQO results of the male versus female for AMR speech codec based on the Gilbert-Elliot channel model under random and burst channel environments.

	G.729A→G.711→AMR (3G)			
	Random channel environments		Burst channel environments	
error rates	Male	Female	Male	Female
0%	3.597	3.194	3.597	3.194
1%	3.367	3.021	3.597	3.194
3%	3.225	2.801	3.258	2.765
5%	3.033	2.615	2.917	2.548
10%	2.625	2.282	2.450	2.135
15%	2.374	1.986	2.083	1.838
20%	2.129	1.787	1.769	1.584
25%	2.007	1.685	1.718	1.540
30%	1.870	1.570	1.708	1.518
35%	1.738	1.488	1.641	1.463
40%	1.685	1.435	1.533	1.346
45%	1.573	1.375	1.540	1.346
50%	1.483	1.316	1.485	1.315

dissatisfied. In Table 4, the relationship between the subjective MOS and user's satisfaction is summarized. Based on Table 4, if mobile VoIP users call up the conventional mobile network (2G and 3G) and PSTN users, the AMR, EVRC, and G.711 must get smaller or equal to 3% packet losses to guarantee the QoS criteria under the random channel characteristic environments. Also, we can see that the MOS-LQO scores for burst channel environment are similar as shown in Table 6. From Table 5 and 6, we can conclude that the MOS-LQO scores for AMR, EVRC, and G.729A should be at least equal or lower than 3% packet loss rates in terms of the QoS for the call scenarios of mobile-VoIP environments.

The relationship between the voice quality for the male and female is analyzed as shown from Table 7 to Table 9. As shown in Table 7, Table 8, and Table 9, the MOS-LQO scores of the voice for the male get higher scores from 0.1 to 0.4 than the MOS-LQO scores of female. Through the MOS-LQO results from Table 7 to Table 9, mobile VoIP services must consider the fact that the MOS-LQO scores of the male is more higher than the MOS-LOQ results of the female.

Table 8. MOS-LQO results of the male versus female for EVRC speech codec based on the Gilbert-Elliot channel model under random and burst channel environments.

	G.729A→G.711→EVRC (2G)			
	Random channel environments		Burst channel environments	
error rates	Male	Female	Male	Female
0%	3.436	3.191	3.436	3.191
1%	3.271	2.987	3.464	3.107
3%	3.147	2.786	3.165	2.731
5%	2.981	2.628	2.838	2.518
10%	2.596	2.338	2.403	2.145
15%	2.369	2.062	2.078	1.852
20%	2.138	1.843	1.778	1.608
25%	2.019	1.737	1.713	1.580
30%	1.896	1.632	1.713	1.554
35%	1.757	1.536	1.649	1.495
40%	1.724	1.475	1.547	1.382
45%	1.614	1.404	1.562	1.380
50%	1.512	1.343	1.512	1.352

Table 9. MOS-LQO results of the male versus female for G.729A speech codec based on the Gilbert-Elliot channel model under random and burst channel environments.

	G.729A→G.711 (PSTN)			
	Random channel environments		Burst channel environments	
error rates	Male	Female	Male	Female
0%	3.728	3.335	3.728	3.335
1%	3.530	3.099	3.817	3.315
3%	3.352	2.852	3.432	2.828
5%	3.141	2.652	3.050	2.590
10%	2.706	2.311	2.530	2.168
15%	2.439	2.003	2.146	1.855
20%	2.170	1.805	1.806	1.587
25%	2.040	1.700	1.752	1.547
30%	1.891	1.581	1.736	1.534
35%	1.756	1.492	1.667	1.465
40%	1.698	1.435	1.542	1.356
45%	1.607	1.372	1.556	1.351
50%	1.499	1.321	1.495	1.321

IV. Conclusions

In this paper, we have proposed the criteria for the QoS standards of mobile VoIP services over wireless mobile Internet such as mobile WiMAX networks in terms of the variable packet loss rates under the environments of the heterogenous network communication. In order to model the mobile VoIP network based on IP, we eventually adopt the Gilbert –Elliot channel model and also set the three types of call scenarios to evaluate the end–to–end call quality where the call scenarios are divided into three cases such as mobile VoIP (G.729A)–to–CDMA (2G: EVRC), mobile VoIP (G.729A)–to–CDMA (3G: AMR), and mobile VoIP (G.729A)–to–PSTN (G.711), respectively. We assess the end–to–end call quality for a criteria and the QoS by introducing the PESQ of the ITU–T P. 862 and MOS–LQO of the ITU–T P. 862.1.

Through the experimental results, we can see that at least higher to 3% packet loss under the random and burst channel characteristics allow mobile VoIP service users to accept the reliable quality. Also, as the MOS–LQO of the male is higher than the female's MOS–LQO, techniques to improve the female's voice quality should be introduced at the network providing the mobile VoIP services for the user's satisfaction. The PESQ scores and MOS–LQO shown in this paper could be used broadly as the criteria and QoS standards for the mobile VoIP services.

Acknowledgement

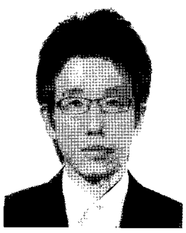
This work was supported by the IT R&D program of MKE/KEIT. [2008–F–045–02] and the research was financially supported by the Ministry of Knowledge Economy(MKE) and Korea Industrial Technology Foundation (KOTEF) through the Human Resource Training Project for Strategic Technology.

References

1. J. –H. Choi, S. U. Seol and J. –H. Chang, "A Study on Voice Communication Quality Criteria Under Mobile–VoIP Environments", *The Journal Of The Acoustical Society Of Korea*, vol. 28, no. 2E, pp. 1–8, 2009.
2. L. Ding et al., "Performance study of objective voice quality measures in VoIP," in *Proc. IEEE Symp. on Comput. Commun.*, pp. 197–202, July 2007.
3. A. Kuriltu et al., "Application and verification of the objective quality assessment method according to ITU recommendation series ITU–T P.862," J. Audio Eng. Soc., vol. 54, no. 12, pp. 1189–1202, Dec. 2006.
4. L. Sun and E. C. Ifeachor, "Performance study of objective voice quality measures in VoIP," *IEEE Trans. on Multimedia*, vol. 8, no. 4, pp. 809–820, Aug. 2006.
5. A. F. Duran, E. C. del Pliego, and J. I. Alonso, "Effects of handover on Voice quality in wireless convergent networks," In *Proceedings of the IEEE Radio and Wireless Symposium*, pp. 23–26, Jan. 2007.
6. ITU–T Rec. P. 800, "Telephone transmission quality. Methods for objective and subjective assessment of quality; methods for subjective determination of transmission quality," *International Telecommunication Union, Telecommunications Standardization Sector (ITU–T)*, pp. 24–25 Geneva, Switzerland, 1996.
7. A. W. Rix and M. P. Hollier, "The perceptual analysis measurement system for robust end–to–end speech quality assessment," in *Proc. IEEE ICASSP*, vol. 3, pp. 1515–1518, June 2000.
8. ITU–T Rec. P. 862, "Perceptual evaluation of speech quality (PESQ), an objective method for end–to–end speech quality assessment of narrowband telephone networks and speech codecs," Feb. 2001.
9. ITU–T Rec. G. 107, "The E–model, a computational model for use in transmission planing," March 2005.
10. ITU–T Rec. P.563, "Single–ended method for objective speech quality assessment in narrow–band telephony applications," May 2004.
11. S. Moller, *Assessment and prediction of Speech Quality in Tele–communications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
12. M. J. Ho and A. Mostafa, "AMR call quality measurement based on ITU–T P.862.1 PESQ–LQO," in *Proc. 64th IEEE Vehicular Technology Conference*, pp. 1–5, Sep. 2006.
13. Z. Qiao, L. Sun, and E. C. Ifeachor, "Case study of PESQ performance in live wireless mobile VoIP environments," *in Personal, Indoor and Mobile Radio Communications 19th International Symposium on*, pp. 15–18, Sep. 2008.
14. ITU–T Rec. P. 862.1, "Mapping function for transforming P.862 raw result scores to MOS–LQO", Nov. 2003.
15. ITU–T Rec. P. 800. 1, "Mean opinion score (MOS) terminology," July 2006.
16. M. Mushkin and I. Bar–David, "Capacity and coding for the Gilbert-Elliot channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1277–1290, Nov. 1989.
17. ITU–T Rec. G. 191, "Software tools for speech and audio coding standardization," Nov. 2000.
18. R. Salami, C. Lallamme, B. Bessette, and J. P. Adoul, "ITU–T G.729 Annex A: reduced complexity 8kb/s CS–ACELP codec for digital simultaneous voice and data," *IEEE Commun. Mag.*, vol. 35 pp. 56–63, Sept. 1997.
19. 3GPP TS 26.090 V5.0.0. *AMR speech codec; Transcoding functions*, June 2002.
20. TIA/EIA/IS–127, *Enhanced variable rate codec, speech service option3 for wideband spread spectrum digital systems*, 1997.
21. ITU–T Recommendation G.711 Appendix I, "A high quality low –complexity algorithm for packet loss concealment with G.711," Sept. 1999.

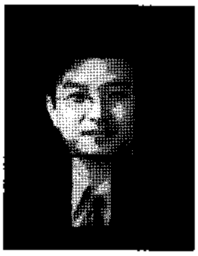
[Profile]

• Jae-Hun Choi



received the BS degrees in electronic engineering from Inha University, Incheon, Korea, in 2007 and he is working toward the MS degree in electronic engineering, at Inha University, Incheon, Korea. His research interests include speech reinforcement, speech enhancement processing, and audio coding.

• Soonuk Seol



received his B.S. in Information and Communication Engineering from the Korea University of Technology and Education in 1998, and M.S. and Ph.D. in Computer Science from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2000 and 2004, respectively. From 1998 he was also a visiting researcher at the Protocol Engineering Center in Electronics and Telecommunications Research Institute (ETRI) until October 1999. Currently, he is with the Central R&D Laboratory at KT, as a senior researcher. His research interests include mobile VoIP/IPTV, mobile computing, QoS network, and protocol engineering.

• Joon-Hyuk Chang

received the BS degree in electronics engineering from Kyungpook National University, Daegu, Korea in 1998 and the MS and PhD degrees in electrical engineering from Seoul National University, Korea, in 2000 and 2004, respectively. From March 2000 to April 2005, he was with Netdus Corp., Seoul, as a chief engineer. From May 2004 to April 2005, he was with the University of California, Santa Barbara, in a postdoctoral position to work on adaptive signal processing and audio coding. In May 2005, he joined Korea Institute of Science and Technology, Seoul, as a Research Scientist to work on speech recognition. Currently, he is an assistant professor in the School of Electronic Engineering at Inha University, Incheon, Korea. His research interests are in speech coding, speech enhancement, speech recognition, audio coding, and adaptive signal processing.