

# Speech Quality Measurement Methods with Applying PLC Algorithms on Real-time Transmission Control Scheme for VoIP Service

Jinsul Kim

Digital Media Laboratory, Information and Communications University, Seoul, Korea  
Email: jskim72@icu.ac.kr

Hyun-Woo Lee, Won Ryu and Byung Sun Lee

BcN Service Research Group / Broadband convergence Network Research Division,  
Electronics and Telecommunications Research Institute, Daejeon, Korea  
Email: {hwlee, wlyu, bslee}@etri.re.kr

Seung Ho Han, Minsoo Hahn

Speech and Audio Information Laboratory, Information and Communications University, Daejeon, Korea  
Email: {space0128, mshahn}@icu.ac.kr

**Abstract**—According to the discovery of available and suitable audiovisual services over IP, researches on the Quality of Service (QoS) for various multimedia application services are required. This paper presents evaluation of speech quality applying improved Packet Loss Concealment (PLC) algorithms with an end-to-end bidirectional transmission control scheme. We propose real-time transmission control scheme to perform better quality monitoring of VoIP service. In order to evaluate the newly proposed RTCP-XR scheme for VoIP services, we measure in real-time the quality degree of the received speech signal while applying our improved Packet Loss Concealment (PLC) algorithms in G.711 and G.729A codecs using an implemented Session Initiation Protocol (SIP)-based SoftPhone and a 1-port gateway systems. Overall, for the evaluation with scientific exactitude of speech quality applied our PLC algorithms we examine with made phone-calls through the heterogeneous network systems. The experimental results confirm the applied PLC algorithms for improvement of speech quality over IP in real environments with the QoS-factor transmission control mechanism.

**Index Terms**—VoIP, QoS, Speech Quality, RTCP-XR, PLC, PESQ, E-Model, SIP, BcN services

## I. INTRODUCTION

As the interests in supporting IP communication

---

Based on "Multimedia Convergence Service using Multi-functional Systems:3D Virtual Monitoring with Dynamic Voice and Video over IP based on SIP", by Jinsul Kim, Hyun-Woo Lee, Ryu Won, and Byung Sun Lee which appeared in the Proceedings of the IEEE Consumer Communications and Networking Conference 2006, Las Vegas, U.S.A., January 2006. © 2006 IEEE.

tremendously increase, a variety of IP-based real time multimedia application services have been developed rapidly in residential and enterprise communication markets because of their attractive service enhancements and cost savings. Moreover, with the increasing demand for Internet telephony services by using the Voice over IP (VoIP) technology, techniques for improving speech quality over IP networks are one of the hot issues in research area today. Also, QoS-factor monitoring to manage in real-time are required to guarantee the quality of VoIP services. An important factor for VoIP QoS control technique involves the rate control, which is based largely only on network impairments such as the packet loss rate or the delay due to the network congestions [1], [2]. The second important factor is based on the factor that different parts of speech have different perceptual importance and each part of speech does not contribute equally to the overall voice quality [3]. Hence, in our approach, voice packets that are perceptually more important are marked, i.e. acquire priority. If there is any congestion, the packets are less likely to be dropped than packets that are of less perceptual importance. The priority marking based QoS schemes are open loop ones and do not make use of changes in the network impairments. To measure the speech quality we adopt the well-known PESQ (Perceptual Evaluation of Speech Quality) algorithm, which is one of objective perceptual voice quality measurement techniques, based on the ITU-T speech quality evaluation standards [4]. Non-intrusive methods do not need a reference signal and are appropriate for monitoring the real traffic. The ITU-T E-model [5] is the most widely used non-intrusive voice quality measurement

method and may be used to predict the conversational MOS (Mean Opinion Score) directly from the IP network and the terminal parameters [6].

The Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) communications use the RTCP-Receiver Report to feedback IP network conditions from RTP receivers to RTP senders. However, the original RTCP provides overall feedback on the quality of end-to-end networks as a whole [7]. The RTP Control Protocol Extended Reports (RTCP-XR) is a new VoIP management protocol which defines a set of metrics that contains information for assessing the VoIP call quality by the IETF [8]. The RTCP-XR reports the packet loss rate, the packet discard rate (because of jitter) and the distribution of lost/discarded packets. The loss/discard distribution describes the call in terms of bursts (periods during which the loss/discard rate is high enough to cause noticeable quality degradation) and gaps (periods during which lost or discarded packets occur infrequently and hence quality is generally acceptable). To guarantee the call quality, the RTCP-XR can report the call quality directly in terms of the estimated R factor or the mean opinion score (MOS). The R factor is a conversational-quality metric in the range of 0 to 100. And the MOS-LQ is a perceptual listening-quality metric while the MOS-CQ is a conversational-quality metric, both in the range of 1 to 5. The RTCP-XR can be implemented as software integrated into IP phones and gateways inexpensively. Then, the messages containing key call-quality-related metrics are exchanged periodically between IP phones and gateways. However, in order for monitoring the QoS-factor on end-to-end VoIP networks, the RTCP-XR is insufficient to estimate the speech quality because it doesn't have any speech quality monitoring functionality. To solve this problem we propose the update of some components in RTCP-XR scheme.

Packet loss occurs inherently in an IP network due to the network characteristics such as the delay, the jitter, and network congestions. The packet loss stands missing speech samples for a time interval at receiver and brings about the degradation of the received speech quality. Since transmission is a real-time process, it is impossible to re-transmit the missed data when the data is missed. A concealment technique is required for the missed packet reconstruction. Speech codecs have their own PLC mechanisms against the packet loss. At the receiver, a new speech signal is generated to reduce the effect of packet loss by utilizing the PLC mechanism. [9, 10]

In this paper, we proposed an end-to-end bidirectional transmission control mechanism to manage the QoS-factors such as the jitter, the delay, and the packet loss rate. The newly presented QoS-factor transmission control mechanism for the QoS management and monitoring of the speech quality during mutual conversations over IP networks is assessed with the performance analysis of the proposed PLC algorithms in VoIP speech coders. Various types of packet losses are generated by the Error Insertion

Device (EID) and then the PESQ is used to evaluate both the original and the proposed PLC algorithm in G.711 and G.729A. Also, to measure and report the R-factors, our experiments specify the QoS performance with the objective measurement method over the IP network, i.e., E-Model and PESQ.

## II. VOIP NETWORK ARCHITECTURE AND PROCEDURE FOR QoS-FACTOR MANAGEMENT

### A. VoIP System Architecture

Historically, voice has been delivered, either wireless or wired, in a dedicated, one-to-one circuit switched manner. This is because of the unique requirements voice traffic places on a network. It requires the low latency and jitter performance of a voice call, but places high demands on the data capacity at the same time. To guarantee the QoS for VoIP services we proposed a new architecture with a QoS measuring server.

Our proposed architecture, Fig. 1 is designed with a SoftPhone and a 1-port gateway as the main User Agent (UA) which carries QoS measurement components. Our network also includes SIP servers and a QoS-factor monitoring server for the call and QoS control. For the accurate QoS delivering evaluation of the RTCP-XR Scheme on the QoS-factor monitoring server, we applied calls through the PSTN to the SIP-based UA-SoftPhone, the SIP-based UA-SoftPhone to the PSTN, and the SIP-based UA-SoftPhone to the SIP-based UA-SoftPhone.

### B. Real-time QoS-Factor Measurement and Transmission

Because present IP networks are not designed to support the QoS, the quality measurement becomes more important and urgent for more reliable higher quality multimedia services over IP networks. We explore the impact of the individual packet loss, the delay, and the jitter on the perceptual speech quality in VoIP systems. The VoIP service quality evaluation is carried out by firstly encoding the input speech pre-modified with given network parameter values and then decoded to generate degraded output speech signals. In order to obtain an end-to-end (E2E) MOS between the caller-UA and the callee-UA, we apply the PESQ and the E-Model method. In detail, to obtain the R factors for E2E measurement over the IP network we need to get  $I_d$ ,  $I_e$ ,  $I_s$  and  $I_j$ . Here,  $I_j$  is newly defined as in equation (1) to represent the E2E jitter parameter.

$$R\ factor = R0 - I_s - I_d - I_j - I_e + A \quad (1)$$

The ITU-T Recommendation provides most of the values and methods to get parameter values except  $I_e$  for the G.729A codec,  $I_d$  and  $I_j$ . First, we obtain  $I_e$  value after the PESQ algorithm applied. Second, we apply the PESQ values to  $I_e$  value of R-factor as shown in Fig.2. We measure the E2E  $I_d$  and  $I_j$  from our current network

environment. By combining  $I_e$ ,  $I_d$  and  $I_j$ , the final R factor could be computed for the E2E QoS performance results. Finally, obtained R factor is reconverted to MOS by using equation (2), which is redefined by the ITU-T SG12.

$$\begin{cases} R \leq 6.5: MOS = 1 \\ 6.5 \leq R \leq 100: MOS = 1 - \frac{7}{1000}R + \frac{7}{6250}R^2 - \frac{7}{100000}R^3 \\ R \geq 100: MOS = 4.5 \end{cases} \quad (2)$$

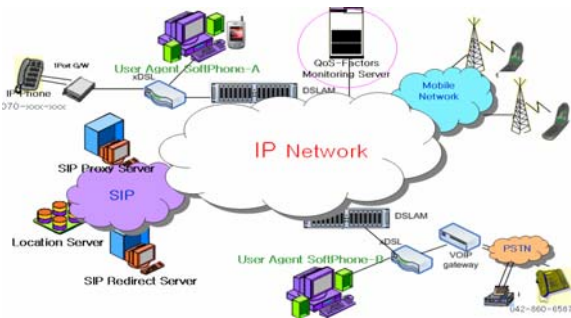


Fig. 1. VoIP network architecture

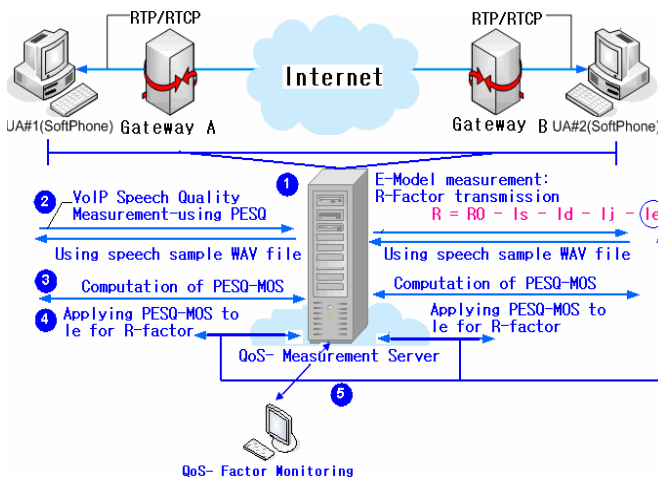


Fig. 2. Procedure of QoS-Factor measurement and transmission

### III. QoS-FACTOR TRANSMISSION CONTROL MECHANISM BASED IN RTCP-XR SCHEME

#### A. Operation Procedure with Voice Packet Transmission

The RTP protocol is used for transmitting real-time data information and the RTCP, for sending the control information. The main function of the RTCP is to provide a detailed representation of the voice packets exchanged during an RTP session. Its structure includes the sender report (SR) and the receiver report (RR) transmitted periodically to all participants in the session. It aims at providing a feedback on the quality of the transmission (e.g., delay jitter, average packet loss, etc.), where transmitters

send “sender reports” and receivers send “receiver reports” using the RTCP-XR. While the SR includes transmission and reception information for active senders in the session, the RR would also contain the reception information for non-active senders. The QoS measurement server receives the QoS-factor and reports them to the QoS-monitoring interface every 2-second. Fig. 2 also shows that the procedure of the RTP/RTCP/RTCP-XR packet transmission with the QoS measurement procedure.

#### B. Conversational Speech Quality Monitoring with Modified RTCP-XR Scheme

For the conversational speech quality we have developed an RTCP-based packet structure to provide an end-to-end bidirectional transmission control that can report the delay, the jitter, and the packet losses in a timely manner. Our packet structure is similar to the RTCP-XR, which is primarily defined to provide more detailed statistics, particularly for multicast applications [8]. In our case, the RTCP-XR scheme was specifically designed to report the delay, the jitter, and the packet losses for every frame of voice signal. Also, the loss and the discard rates are designed to be calculated for each session at the end-receiver in order to get real-time measured values. The original RTCP-XR packet type defined can be used for speech quality monitoring. However, it is not familiar with real-time speech quality reporting for bidirectional conversation through VoIP networks. We modified the RTCP-XR packet scheme in order for conversational speech quality reporting and monitoring because communication services like the VoIP must be recognized as the mutual conversation one.

For the delay monitoring as one of the significant QoS-factor BT-5 is formatted with both sub-block 1 and sub-block 2. SSRC\_1 and SSRC\_2 are for the sender and the receiver numbers which is defined randomly. The DLRR in sub-block 1 reports the one way delay between the sender and the receiver. The DLRR in sub-block 2 reports the round trip delay which is measured using the one way delay information from DLLR in sub-block 1. Fig. 3 shows the message format-II of BT-5 for delay monitoring. In Fig. 4, information of the jitter and the packets is controlled in BT-6. At first, in order to manage the jitter, we categorize it into three levels of min\_jitter, max\_jitter, and mean-jitter. Those are reported as the cumulative effects of jitter values obtained through the jitter buffer in our cases. Second, to get the real-time mutual communication speech quality monitoring field of packet count information of the RTP/RTCP is included in the report block. The Tx/Rx RTP Packets format in BT-6 is designed for monitoring sender/receiver RTP. The Rx RTCP Packets are the RTCP XR packets which are received at the QoS measurement server, and the Tx RTCP Packets are the RTCP XR packets which are sent from the SoftPhone (or Iport Gateway) to the QoS measurement server. The bidirectional QoS-factor is also measured in real-time and reported to the QoS-

Monitoring device. The final result value of the cumulative packet loss is also included in our BT-6 scheme. The format of our BT-7 scheme is similar to the standard RTCP-XR. However, the loss and the discard rates are computed as soon as the call session starts by using the cumulative packets controlled by BT-6. The accuracy of each packet counting in real-time is important. Especially, for time synchronization we set our execution time of both the SoftPhone and the 1-port gateway to the NTP (National Time Protocol). That is, the host system time is synchronized with the national standard time. Finally, by applying the QoS-factor obtained from the bidirectional speech quality measurement MOS-LQ and MOS-CQ is defined separately for the listening quality and the conversational quality. Other factors in BT-7 are just added to the standard format scheme as shown in Fig. 5.

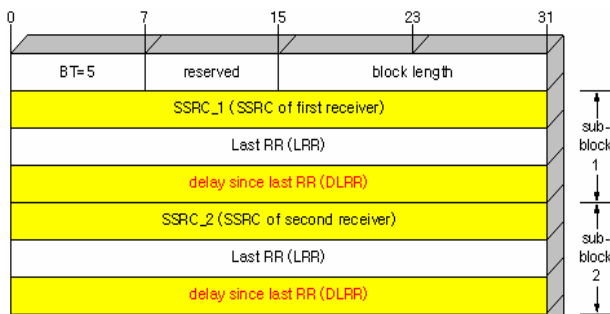


Fig. 3. BT=5 delay monitoring for transmission control

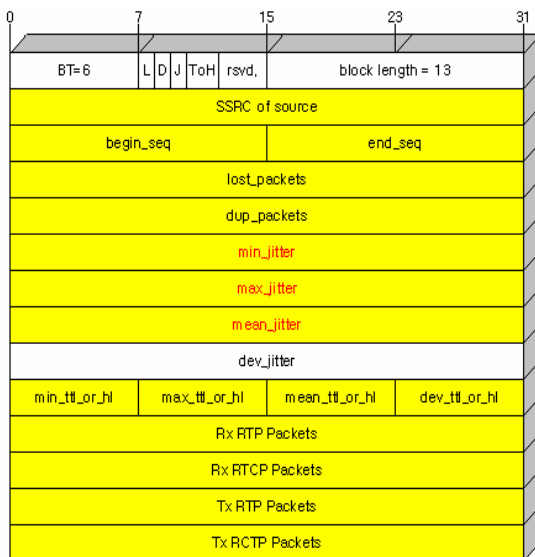


Fig. 4. BT=6 jitter, packets for transmission control

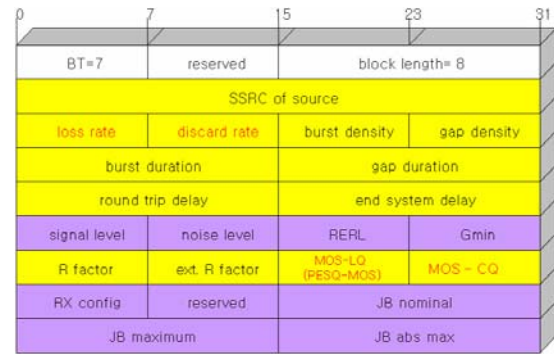


Fig. 5. BT=7 loss & discard rate, MOS-LQ, & -CQ for transmission control

#### IV. PROPOSED PLC ALGORITHMS

When packet losses occur over IP networks, the PLC algorithms employed in speech codecs reconstruct lost speech frames based on the previously received speech information. The PLC algorithm employed in G.711 Appendix I repeatedly inserts pitch period detected from the previous speech in history buffer [9] which is called the pitch period replication method. The PLC algorithm employed in G.729A estimates an excitation signal and synthesis filter parameters from last good frame [10].

In this paper, two techniques are proposed to construct new PLC algorithms for G.711 and G.729A codecs, respectively. The proposed PLC algorithms adaptively utilize both-side information for lost frames based on the class classification and adaptively control gains according to the average magnitude variation across the frames. We apply the proposed algorithm to our system to prove packets loss compensation in both off-line test environment in which the decoder is modified by our PLC algorithm and real-time environment, by the transmission control mechanism based on the RTCP-XR scheme with given QoS-factors.

##### A. Applying G.711

First of all, the proposed PLC algorithm uses adaptively both-side information for lost frames with the buffer to store a succeeding frame. When both-side information for lost frame is available, the missing speech is reconstructed by adaptive both-side information utilization. The last good frame is classified into either the class A or the class B by the decision tree. To make the decision tree, we use the CART (Classification And Regression Tree) [11] and analyze 1000 Korean dialogue speech. The ZCR (Zero Crossing Rate) [12], the ratio between the average magnitude [12] of the current frame and that of the adjacent frame, the average magnitude ratio between the subframes in current frame, and ratio between the first value and the second value of the autocorrelation function [12] are evaluated as the feature parameters.

If the last good frame is classified into the class A, the

proposed PLC algorithm uses only succeeding frame information to reconstruct missing speech by using the pitch period replication method which has same procedures with that employed in G.711 Appendix I. On the other hand, if the last good frame is classified into the class B, the proposed PLC algorithm uses both previous frame information and succeeding frame information to reconstruct missing speech. Reconstructing speeches by using previous information and succeeding information are averaged and the averaged signal becomes the final reconstructing speech for missing speech.

Second, the proposed PLC algorithm controls adaptively gains according to the average magnitude variation across the frames. The variation is evaluated by  $G$ .  $G_I$  only in increasing case and  $G_D$  only in decreasing case are respectively calculated by

$$G_I = \frac{2E_{i-1} - E_{i-2}}{E_{i-1}}, \quad G_D = 2 - \frac{2E_{i-2} - E_{i-1}}{E_{i-1}} \quad (3)$$

where  $E$  is the average magnitude, and  $i$  is the frame index. Evaluated  $G_I$  or  $G_D$  is set as  $G$ . The attenuation factor is increased or decreased up to  $G$  for first loss period of 10 msec. From after that to the period of 60 msec, the attenuation factor is linearly decreased from the  $G$  to 0. In other words, G.711 Appendix I PLC attenuates the reconstructing speech with fixed factor [9] while the proposed algorithm has varying attenuation factor according to evaluated  $G$ .

#### B. Applying G.729A

For G.729A codec when both-side information is available, the proposed PLC algorithm modifies the LSP (Line Spectrum Pairs) interpolation coefficients for the first subframe based on the class of the last good frame which is classified into either the class A or the class B by the decision tree. It is same with that used in the proposed G.711 PLC algorithm.

In the class A, the modified LSP interpolation is given as

$$q_i = 0.48q_i^{(\text{previous})} + 0.52q_i^{(\text{future})}, \quad i = 1, \dots, 10 \quad (4)$$

where  $q_i$  is the  $i$ -th LSP coefficient for the current lost frame,  $q_i^{(\text{previous})}$ , for the previous frame, and  $q_i^{(\text{future})}$ , for the future frame. And, in the class B, the modified LSP interpolation is given as

$$q_i = 0.52q_i^{(\text{previous})} + 0.48q_i^{(\text{future})}, \quad i = 1, \dots, 10. \quad (5)$$

Second, the proposed PLC controls adaptively gains according to the average magnitude variation across the frames. The PLC algorithm employed in G.729A attenuates

the adaptive and fixed codebook gains as follows.

$$g_c^m = 0.98g_c^{(m-1)}, \quad g_p^m = 0.9g_p^{(m-1)} \quad (6)$$

where  $m$  is the subframe index,  $g_c$  is the fixed codebook gain, and  $g_p$  is the adaptive codebook gain.

The proposed PLC algorithm modifies the attenuation factors based on the variation of average magnitudes between previous two frames. The variation is evaluated by  $G$  with same way proposed in G.711 PLC algorithm. With this  $G$ , the adaptive and fixed codebook gains are attenuated by the new factors as follows.

$$\begin{cases} G > 1: g_c^m = 0.988g_c^{(m-1)}, g_p^m = 0.92g_p^{(m-1)} \\ 0.7 \leq G \leq 1: g_c^m = 0.98g_c^{(m-1)}, g_p^m = 0.9g_p^{(m-1)} \\ G < 0.7: g_c^m = 0.977g_c^{(m-1)}, g_p^m = 0.88g_p^{(m-1)} \end{cases} \quad (7)$$

## V. SIMULATION SURROUNDINGS

This paper evaluates the performance of the QoS-factor transmission control mechanism with the proposed PLC algorithm under different packet loss simulating conditions. The PLC algorithm is tested in both in off-line environment and real-time environment with G.711 and G.729A codecs.

In the off-line environment, Fig. 6 shows the experimental procedure. First, speech data are encoded in the encoder. Second, packet losses are generated by the EID. Third, encoding files are decoded with and without packet losses. Finally, decoding files with packet losses are evaluated with the reference speech decoded without a packet loss in the speech quality evaluation module. The performance of the proposed PLC algorithms for G.711 and G.729A are compared to that of the G.711 Appendix I PLC algorithm and that of the PLC algorithm employed in G.729A, respectively.

In real-time environment, our VoIP system applies the QoS-factor transmission control mechanism based on the RTCP-XR scheme for the QoS-factor management and monitoring. Fig. 7 outlines the configuration of simulation environment applying proposed QoS-factor scheme and improved PLC algorithms under tests. The PLC algorithms are tested each other interactively using the SoftPhone and 1port gateway. The performance of the proposed PLC algorithm for G.729A is compared to that of the PLC algorithm employed in G.729A. And the reference system for G.711 has no PLC mechanism and simply inserts silence for a missing period. Thus, the proposed PLC algorithm for G.711 is employed in G.711. To show the effect of proposed PLC algorithm, the speech quality of the reference system is compared to that of the system with the proposed PLC algorithm when packet losses occur.

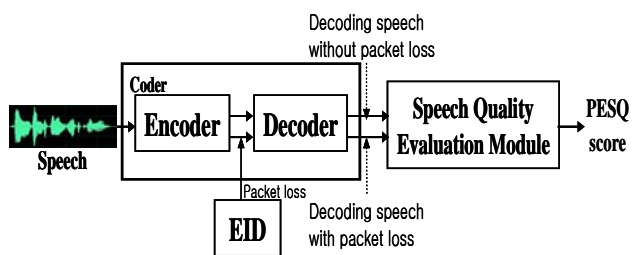


Fig. 6. Experimental procedure

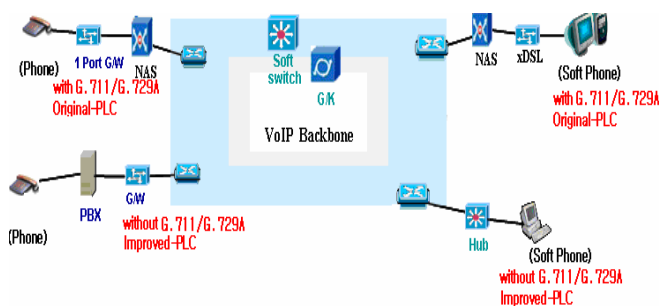


Fig. 7. Configuration of simulation environment applying proposed QoS-factor scheme and improved PLC algorithm

A. Packet Loss Design

A communication channel can be modeled by the Gilbert-Elliot model [13] or the Bellcore model [14]. In the simulation, we use the STL (Software Tool Library) included in ITU-T Recommendation G.191 [15] to model packet losses. The STL provides the EID which generates losses based on the Gilbert-Elliot model and the Bellcore model. Various patterns of losses can be generated by the modification of parameters in the EID.

To model various packet loss environments, we design burst and random packet losses with 1%, 3%, 5%, 10%, and 15% loss rates and it is considered that a packet contains 1 speech frame, 2 speech frames, or 3 speech frames.

B. Performance Evaluation

For PLC algorithm performance evaluation, we use the systemic evaluation method which is PESQ defined by ITU-T Recommendation P.862 for objective assessment of quality. After comparing an original signal with a degraded one, the output of PESQ provides a score from -0.5 to 4.5 as a MOS-like score. The reference speech for the real-time environment simulation and the off-line environment simulation is the original speech and the decoded speech without any packet loss, respectively. In the real-time environment, after getting the PESQ, it is used for E-Model evaluation method as we mention above. Finally, the MOS is acquired.

We use 200 Korean dialogue speech utterances from 2 male and 2 female speakers as test data. The duration of each utterance is about 10 seconds.

VI. SIMULATION RESULTS

The simulation shows that the proposed PLC algorithms with G.711 and G.729A codecs over IP in real environments applying our QoS-factor transmission control mechanism based on the modified RTCP-XR scheme achieve much better result for speech quality than other VoIP systems which use standard scheme and no PLC algorithm during mutual conversation in IP network environment. In the following result tables, ‘fpp’ refers to frame per packet.

A. PESQ Measurement in Off-line Test Environment

Table I shows the results of the PLC performance evaluation for G.711. At 1fpp, the degree of improvement for burst losses is higher than that for random losses. PESQ scores are increased by 0.016 for random losses and 0.103 for burst losses on the average, respectively. On the other hands, at 2fpp and 3fpp, the degree of improvement for random losses is higher than that for burst losses because of the packetization effect. At 2fpp, the degree of the improvement for random losses is 0.112 while that for burst losses is 0.047 on the average. At 3fpp, the degree of the improvement is 0.119 for random losses and 0.022 for burst losses, respectively.

Table II shows the results of the PLC performance evaluation for G.729A. At 1fpp, the difference of the degree of improvement score between random losses and burst losses is not large. PESQ scores are increased by 0.074 on the average. On the other hands, at 2fpp and 3fpp, the degree of improvement for random losses is higher than that for burst losses because of the packetization effect. At 2fpp, the degree of the improvement for random losses is 0.169 while that for burst losses is 0.042 on the average. At 3fpp, the degree of the improvement is 0.124 for random losses and 0.0389 for burst losses, respectively.

Thus, the superiority of the proposed PLC algorithm is proved in the off-line test environment.

TABLE I  
PESQ RESULT FOR OFF-LINE ENVIRONMENT WITH G.711

Loss type \ PLC	PLC	Original PLC			Proposed PLC		
		1fpp	2fpp	3fpp	1fpp	2fpp	3fpp
Random	1%	3.983	3.987	4.018	3.984	4.025	4.066
	3%	3.644	3.630	3.621	3.641	3.717	3.714
	5%	3.440	3.412	3.390	3.451	3.524	3.504
	10%	3.124	3.038	2.965	3.149	3.186	3.123
	15%	2.893	2.784	2.667	2.940	2.959	2.845
Burst	1%	3.975	3.991	3.983	4.021	4.029	4.009

	3%	3.558	3.589	3.600	3.637	3.621	3.600
	5%	3.317	3.262	3.305	3.412	3.297	3.323
	10%	2.839	2.758	2.755	2.972	2.817	2.783
	15%	2.504	2.381	2.407	2.664	2.455	2.446

Burst	0%	4.2	93	4.4	4.2	93	4.4
	3%	3.2	64	3.3	3.2	65	3.4
	5%	3.0	57	3	3.3	64	3.3
	10%	2.7	50	2.6	2.8	53	2.7

**TABLE II**  
PESQ RESULT FOR OFF-LINE ENVIRONMENT WITH G.729A

Loss type \ PLC	PLC	Original PLC			Proposed PLC		
		1fpp	2fpp	3fpp	1fpp	2fpp	3fpp
Random	1%	3.952	3.912	3.939	3.976	4.003	4.005
	3%	3.569	3.488	3.500	3.615	3.637	3.605
	5%	3.319	3.239	3.234	3.386	3.417	3.359
	10%	2.936	2.818	2.776	3.038	3.030	2.932
	15%	2.658	2.537	2.475	2.787	2.751	2.642
Burst	1%	3.953	4.031	4.019	3.991	4.049	4.039
	3%	3.531	3.634	3.698	3.587	3.663	3.717
	5%	3.275	3.347	3.406	3.352	3.393	3.434
	10%	2.808	2.896	2.936	2.901	2.953	2.974
	15%	2.478	2.543	2.585	2.586	2.604	2.674

**TABLE IV**  
RESULT FOR REAL-TIME ENVIRONMENT WITH G.729A

Loss type \ PLC	PLC	Original PLC			Proposed PLC		
		PESQ	R	MOS	PESQ	R	MOS
Random	0%	3.5	72	3.7	3.5	72	3.7
	3%	3.2	63	3.4	3.3	68	3.5
	5%	3.1	62	3.2	3.2	65	3.4
	10%	2.9	57	2.9	3	58	3.0
	Burst	0%	3.5	72	3.7	3.5	72
Burst	3%	3.2	64	3.3	3.3	67	3.5
	5%	3.1	61	3.2	3.2	64	3.2
	10%	2.9	56	2.9	3.0	57	2.9

**B. PESQ, R-factor and MOS Measurement in the Real-time Environment**

In Table III, the PLC performance evaluation results for G.711 for the real-time environment are summarized as the PESQ, the R-factor, and the MOS. The proposed PLC achieves the PESQ gains of from 0.1 to 0.3 as to loss types. The corresponding gains for the R-factors and the MOSs are also achieved by these PESQ gains. High gains are achieved at random losses with high loss rates.

In Table IV, the PLC performance evaluation results with G.729A for the real-time environment are summarized as the PESQ, the R-factor, and the MOS. The proposed PLC achieves the PESQ gain of 0.1 for all types of loss. The corresponding gains for the R-factors and the MOSs are achieved by these PESQ gains. Finally, even though the MOS is not improved for burst losses at high loss rates, the MOS gain of 0.1 is generally achieved.

**TABLE III**  
RESULT FOR REAL-TIME ENVIRONMENT WITH G.711

Loss type \ PLC	PLC	No PLC			Proposed PLC		
		PESQ	R	MOS	PESQ	R	MOS
Random	0%	4.2	93	4.4	4.2	93	4.4
	3%	3.5	71	3.6	3.8	78	3.9
	5%	3.0	59	3	3.4	68	3.5
	10%	2.5	48	2.5	3.0	58	3

**VII. CONCLUSION**

In this paper, the performance evaluation of speech quality confirms that our proposed PLC algorithm outperforms more efficiently than the original algorithm in the G.711 and G.729A speech codecs. Also, we have proposed an end-to-end bidirectional QoS-factor transmission control mechanism to manage the speech quality in real-time on modified RTCP-XR scheme for QoS monitoring of mutual conversations over IP networks. The newly presented QoS-factor transmission control mechanism is assessed with the performance analysis of the proposed PLC algorithms in off-line and real-time environment. Finally, we have shown that, based on the simulation results the speech quality for VoIP service can be significantly improved by using the proposed PLC algorithm with newly presented transmission control mechanism.

**REFERENCES**

[1] R. Eejaie, M. Handley and D. Estrin, "RAP: An End-to-end Rate-based Congestion Control Mechanism for Realtime Streams in the Internet," *Proc. IEEE INFOCOM*, USA, 21-25 March 1999.

[2] F. Beritelli, G. Ruggeri and G. Schembra, "TCP-Friendly Transmission of Voice over IP," *Proceedings of IEEE International Conference on Communications*, New York, USA, April 2002.

- [3] J. C. De Martin, "Source-driven Packet Marking for Speech Transmission over Differentiated-Services Networks," *Proceedings of IEEE ICASSP 2001*, Salt Lake City, USA, May 2001.
- [4] ITU-T Recommendation 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.
- [5] ITU-T Recommendation G.107, "The E-Model, a computational model for use in transmission planning," 2000.
- [6] R.G. Cole and J. H. Rosenbluth, "VoIP over IP Performance Monitoring," *Journal on Computer Communications Review*, vol 31, no.2, April 2001.
- [7] H. Schulzrinne, S. Casner, R. Frederick and V. Jacobson, "RTP: A Transport Protocol for Real-Time Application," *IETF RFC 3550*, July 2002.
- [8] T. Friedman, R. Caceres, A. Clark, "RTP Control Protocol Extended Reports," *IETF RFC 3611*, Nov. 2003.
- [9] ITU-T Recommendation G.711 Appendix I, "A high quality low-complexity algorithm for packet loss concealment with G.711," Sept. 1999
- [10] ITU-T Recommendation G.729 Annex A, "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited-linear-prediction (CS-ACELP)," Nov. 1996.
- [11] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, JOHN WILEY&SON, 2001, pp. 396-412.
- [12] L. R. Rabiner, R. W. Schafer, *Digital Processing of Speech Signals*, Prentice Hall, 1978.
- [13] Mushkin, M, Bar-David, I, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Transactions on Information Theory*, Vol. 35, Nov. 1989, pp 1277-1290.
- [14] Vijay K. Varma, "Testing Speech Coders for Usage in Wireless Communications Systems," *Proceedings of IEEE workshop on Speech Coding for Telecommunication*, Oct. 1993, pp 94-94.
- [15] ITU-T Recommendation G.191, "Software tools for speech and audio coding standardization," Nov. 2000.

**Jinsul Kim** was born in Namwon, Chonbuk, South Korea, in 1972. He received the B.S. degree in computer science from University of Utah, Salt Lake City, Utah, USA, in 2002, and the M.S. degree in digital media engineering from Information and Communications University (ICU), Daejeon, South Korea, in 2005. Since 2005, he has been working as a researcher in Broadband convergence Network Interworking Technology Laboratory, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. His research interests include Mobility, NGN, Multimedia Communication and Voice/Video over IP.

**Hyun-Woo Lee** was born in Seoul, South Korea, in 1970. He received the B.S. degree in electronics engineering from Hankuk Aviation University, Kyonggi, South Korea, in 1993, and his M.S. and Ph.D. degree in communication and information engineering from Hankuk Aviation University, Kyonggi, South Korea, in 1995 and 2005. Since 1995, he has been working as senior researcher, in Broadband convergence Network Interworking Laboratory, Electronics and Telecommunications Research

Institute (ETRI), Daejeon, Korea. His research interests include Mobility, Wireless LAN, Interworking Technology and VoIP.

**Won Ryu** was born in Busan, South Korea, in 1958. He received the B.S. degree in the department of computer science & statistics from Pusan National University, Busan, South Korea, in 1983, and the M.S. degree in the department of computer science & statistics from Seoul National University, Seoul, South Korea, in 1988, respectively, and his Ph.D. degree in information engineering from SungKyunKwan University, Kyonggi, South Korea, in 2002. Since 1989, he has been joining as a researcher and currently he is a project leader in the Broadband convergence Network Interworking Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. His research interests include All-IP based Convergence Service & Control Plane, Mobility, NGN, Wire/Wireless Interworking Technology, and VoIP.

**Byung Sun Lee** was born in Seoul, South Korea, in 1956. He received the B.S. degree in the department of mathematics from SungKyunKwan University, Kyonggi, South Korea, in 1980, and the M.S. degree in the department of computer science from Dongkuk University, Seoul, South Korea, in 1982, respectively, and his Ph.D. degree in the department of computer science from Korea Advanced Institute of Science and Technology University, Daejeon, South Korea, in 2003. Since 1982, he has been joining as a researcher and currently he is a project leader in the Broadband convergence Network Service Research Group, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. His research interests include All-IP based Convergence Service & Control Plane, Mobility, NGN, IMS Technology, OpenAPI Technology and VoIP.

**Seung Ho Han** was born in Seoul, South Korea, in 1981. He received the B.S. degree in avionics engineering from Hankuk Aviation University, Kyonggi, South Korea, in 2003, and the M.S. degree in electronics engineering from Information and Communications University (ICU), Daejeon, South Korea, in 2005. He is currently working toward Ph.D. degree in school of engineering at ICU. His research interests include speech coding, and VoIP.

**Minsoo Hahn** was born in Seoul, South Korea, in 1956. He received the B.S. and the M.S. degrees in electrical engineering from Seoul National University, Seoul, South Korea, in 1979 and 1981, respectively, and the Ph.D. degree in electrical and electronics engineering from University of Florida, Florida, USA, in 1989. From 1982 to 1985, he was with Korea Research Institute of Standards and Science (KRISS), Deajeon, South Korea. From 1990 to 1997, and he was with Electronics and Telecommunications Research Institute (ETRI), Deajeon, South Korea. In 1998, he has been a faculty member of the School of Engineering, Information and Communications University (ICU). Currently, he is a Full Professor in School of Engineering, ICU and a Director in Digital Media Laboratory, ICU. His research interests include speech and audio coding, speech synthesis, noise reduction and VoIP. Recently, he is located general chair for Speech Engineering Community, South Korea.