# VoIP 품질 측량 도구 및 품질 기반의 요금 부과 방안 연구

정회원 정 윤 찬*, Ibanez Al Ann*

# VoIP Quality Metric and Quality-based Accounting Scheme

Younchan Jung*, Ibanez Al Ann* *Regular Members*


## 요 약

인터넷 음성통신은 이용영역이 무선통신 환경으로 이동함에 따라 유선 통신을 이용할 때와 비교하여 전송 중 더 높은 패킷 손실을 직면하게 된다. 따라서 적정 수준의 통화품질을 보장해줄 수 없을 상황도 고려할 수 밖에 없다. 이런 상황에서 이동 VoIP 서비스의 실시간 품질 추적이 중요한 기술 요소로 대두되었다. 이 논문에서는 두 가지 요인 즉, 평균적인 패킷 분실 정도와 패킷 손실의 뭉침 정도가 인간이 인지하는 통화 품질에 미치는 영향을 연구한다. 또 이 두 요인을 실시간으로 측정할 수 있는 '이동 평균' 방식을 제안한다. 이 이동평균 방식에 따라 실시간으로 측정된 두 요소로 통화품질을 추정할 경우에 얼마나 정확하게 추정가능한지를 확인하기 위하여, 이 논문에서는 기존에 나와 있는 비 실시간 품질 측정 도구를 이용하여 측정한 품질 추정 값과 이동평균 방식으로 측정한 두 요소에 의한 통화품질 예측 값을 비교한다. 이 비교 분석을 통하여 이동평균 방식으로 측정한 두 요소를 품질 측량 도구로 사용할 수 있음을 입증한다. 마지막으로 품질과 요금 부과 관계를 명확히 연관시켜줄 수 있는 품질 기반 요금 부과 체계를 제안한다.


**key Words :** Quality of Service, VoIP, Quality Monitoring, Quality Metric, Quality-based accounting


## ABSTRACT

As VoIP systems move to wireless environments with much higher average packet loss rates than wired networks, it becomes less possible for the network to assure a reasonable QoS. So, real-time quality monitoring for mobile VoIP applications is an important issue to be explored. In this paper, we explore perceptual quality dependency on two parameters: the burst loss rate and average burst length. Also, we propose a simple 'moving average' approach with α aiming to measure those parameters on real-time basis. In order to find how accurately the two parameters measured estimate the real perceptual quality, we compare actual measured PESQ scores with estimated value by matching the measured quality metric to the trained MOS table. Finally, we propose the quality-based accounting system, which can set obvious continuities between quality and billing.


## Ⅰ. Introduction

Typically, real-time applications imposes very strict quality-of service (QoS) requirements on the Internet Protocol (IP)-based networks[1]-[3]. In both wired and wireless networks, the primary challenge in providing QoS is network congestion, which can cause unacceptable packet loss and delay[4]. In wireless networks, there is an additional challenge of QoS[5], [6]. Though packet

※ This study was also supported by the Research Fund, 2008 of The Catholic University of Korea.
* 가톨릭대학교 정보통신전자공학과 통신네트워크 연구실(ycjung@catholic.ac.kr, alannibanez@yahoo.com.ph)
논문번호 : KICS2009-11-570, 접수일자 : 2009년 11월 13일, 최종논문접수일자 : 2009년 12월 17일

loss in wired networks is typically caused by excessive congestion, wireless and mobile networks typically suffers much more loss due to the sources of disruption come from the physical layer (e.g., handoff from one cell to another), or from the link layer (e.g., collision or congestion in CSMA/CA operation). Then, it becomes possible for wireless QoS to be degraded to a lower level of services.

Users expect a certain level of VoIP (Voice over Internet Protocol) services in either wired or wireless networks[7]-[9]. For now, there is no universal solution for VoIP services to satisfy users who are accustomed to dedicated circuit call service[8]. In order to manage QoS, there has been essentially two complementary approaches aiming to estimate the user-perception of speech quality: subjective testing and objective testing[3]. Subjective tests are usually time consuming to perform[9], 10]. On the order hand, objective testing techniques measure physical properties of a system. These properties can then be used to predict perceived performance[2] ,[9], [11]. However, it is necessary to have customers use the VoIP application at varying level of QoS and ask for feedback on their experience. A key consideration in this paper is the fact that from users viewpoints, there has been no quantitative way to measure the service quality on real-time basis.

It is suggested that there should be an easy way to link the intangible user-perceived quality with the tangible quality factors such as packet loss and delay, which can be observed at the edge-device. The main focus in this paper is to investigate how to use the quality factors aiming to find the quality metric which would lead to estimate the subjective quality with accuracy. The quality metric should be measured in the terminal itself[8], [12], [13]. VoIP terminals typically have jitter buffers, which cause the terminal to interpret all delays exceeding a given delay threshold as discarded packet losses. For the reason that packet loss at the edge-device incorporates both pure lost and discarded packets, we focused on the use of packet losses for the purpose of finding the quality metric.

The basic packet loss model has been based on Gilbert-Elliot (GE) model, which distinguishes at first a good and secondly a bad state with different loss rates[14], [15]. The model is fit to the typical VoIP call, which have properties of low or zero packet loss for most of a call (say a three minute call) and high loss rates (say 30%) for short periods (say two 3 seconds) of the call. The burst defined in the GE model is a period of time during which the packet loss rate is high enough to be problematic (say bad). In this paper, quality monitoring activities are implemented repeatedly in every short period of time (say 22 seconds) within the whole call. Then, it is difficult to measure the GE model-based bursts in such a short period. So, in order to measure the characteristics of the burst loss appearances with simplicity in every short period, we use a new type of the burst loss model having two parameters: the burst loss rate and average burst length where the burst loss rate is defined as the frequency of the loss event of one or consecutive lost packets and the average burst length as average length of a loss event, respectively. Existing studies of packet loss distribution suggest that once bad states with loss rates of typically 30% occur, the loss events are predominantly one packet in length and to a lesser extent of two or three packets[16], [17]. Our model covers a run length of up to 8 packet loss consecutively. Using the model, we derive the trained MOS (mean opinion score) table where each element represents the perceptual quality level for a certain burst loss condition. For the purpose of finding the reliable quality metric, we analyze quality estimation errors, which depends on how to measure the burst loss rate and average burst length. In the section V, we use them as accounting information for the quality-based accounting system, which is one of challenging study areas towards all IP ages[18] ,[19], [20].

The quality metric is described in section II, followed by description of the trained MOS table in section III. Section IV describes how to use the burst loss parameters to estimate the VoIP quality level. Finally, we propose the simple quality-based accounting system in section V, followed by our conclusions in section VI.

## Ⅱ. Quality metric by using burst loss rate and burst length

To determine listening quality in the subjective quality assessment, users are asked to rate 'the quality of the speech' on a five-point scale. However, subjective tests are unsuitable for real-time applications. On the other hand, the objective approach, which estimates the perceptual quality by measuring the physical factors which affects the quality degradation, has more recently been applied to real-time assessment. As described in the previous section, packet loss impairments can be seen as the primary determinant of perceived quality for future IP terminals. Specially in mobile environment, the sources of packet-level disruption can come from the handoff from one cell to another or from the link layer collision or congestion in CSMA/CA operation. Differently from Gilbert-Elliot model, we interpret a burst loss event as packet losses in a sequence of length $B$ where $B \geq 1$. As shown in Fig. 1, our burst loss model alternates a good state with packet loss rate = 0 and a burst state with packet loss rate = 1.

The alternate states are controlled by two parameters: the burst loss rate ($B_{rate}$) and average burst length ($B_{leng}$). Dwell times in the good state are geometrically distributed with mean $\frac{1}{B_{rate}}$ and those in the burst state are distributed based

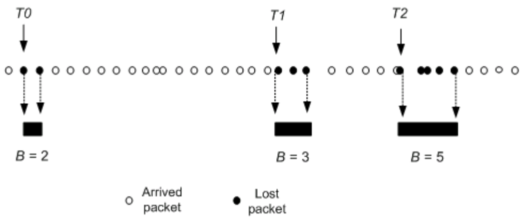


Fig. 1. Burst loss model

on the following probability density function with mean $B_{leng}$.

$$P^{B_{leng}}(x) = \begin{cases} \{1.0 \quad \text{if } x = 1 & \text{for } B_{leng} = 1 \\ \begin{cases} 0.25 & \text{if } x = 1 \\ 0.5 & \text{if } x = 2 \\ 0.25 & \text{if } x = 3 \end{cases} & \text{for } B_{leng} = 2 \\ \begin{cases} 0.1 & \text{if } x = B_{leng} - 2 \\ 0.2 & \text{if } x = B_{leng} - 1 \\ 0.4 & \text{if } x = B_{leng} \\ 0.2 & \text{if } x = B_{leng} + 1 \\ 0.1 & \text{if } x = B_{leng} = 2 \end{cases} & \text{for } B_{leng} \geq 3. \end{cases} \tag{1}$$

Using the good and burst states enables a simple and useful fit for measurement processes looking for the quality metric. As a result, difficulties in distinguishing the different burst loss appearances can be reduced. At first, the distinguishing process involves to measure the burst loss rate, which represents how frequently a burst loss event occurs, and secondly to measure the average burst length, which corresponds to the average number of packets lost in a bundle when a burst loss event occurs.

In the impairment scenario, the $B_{rate}$ and $B_{leng}$ range from 1% to 10% and from 1 to 8, respectively. We aim to obtain a trained MOS table in the scope of the scenario, that is, $E^{MOS}_{(i,j)}$, $i$ = 1, 2, ⋯, 10 and $j$ = 1, 2, ⋯, 8.

## Ⅲ. Obtaining trained MOS table

The voice sample considered for our test was retrieved from OPTICOM CO. This sample lasts around 22 seconds involving 10 different languages. It was encoded using the G.711 codec with a sampling rate of 8 KHz and a bit rate of 64 Kbps. The G.711 speech data is grouped into 20 ms packets, results in an original stream (ORIG) including 1081 packets of with each length of 160 bytes. Next, the ORIG is converted into lots of distorted streams (DSTDs) suffered from artificial burst loss events with parameters of

$B_{rate}$ and $B_{leng}$. A burst loss event occurs randomly with the frequency of $B_{rate}$ and the number of packets in a burst loss is $B_{leng}$ in average. Here, $B_{rate}$ ranges from 0% to 10% and $B_{leng}$ spans from 1 to 8 by using (1). We computed 1200 distorted streams, the DSTD($B_{rate}$, $B_{leng}$, $k$) where $k$ = 1, 2, ⋯, 15. That is, for a given pair of ($B_{rate}$, $B_{leng}$), 15 different distorted streams are generated.

In order to predict the subjective quality, we used an objective method for speech quality assessment: Perceptual evaluation of speech quality (PESQ) based on ITU-T Recommendation P.862. We used the PESQ tool made by OPTICOM (WWW.OPTICOM.de). PESQ score ($M^{MOS}_{(i,j,k)}$) was produced associating with the DSTD($i,j,k$) where $i$ = 1, 2, ⋯, 10, $j$ = 1, 2, ⋯, 8, and $k$ = 1, 2, ⋯, 15. That is, the value of $M^{MOS}_{(i,j,k)}$ is computed after comparing the ORIG with the DSTD($i,j,k$). Then, we can train the MOS table, $E^{MOS}_{(i,j)}$, $i$ = 1, 2, ⋯, 10 and $j$ = 1, 2, ⋯, 8 using

$$E^{MOS}_{(i,j)} = \frac{\sum_{k=1}^{15} M^{MOS}_{(i,j,k)}}{15}. \qquad (2)$$

That is, the trained value of $E^{MOS}(i,j)$ is used to predict the perceptual quality for the degraded speech data suffering from the condition of $B_{rate}$ = $i$ and $B_{leng}$ = $j$.

Fig. 2 shows the PESQ MOS scores for $B_{leng}$ = 1, 3, 5, and 8. It can clearly be seen that as $B_{leng}$ increases, MOS score decreases. However, there is a upper bound range where the difference in quality is meaningless. The severe burst loss region, which corresponds to the burst loss conditions with $B_{rate} \geq 7\%$ and $B_{leng} \geq 5$, is likely to impose a limit on accuracy of the trained MOS table. Recall that we defined the burst loss rate as the frequency of the loss event of one or consecutive lost packets and the average burst length as average length of a loss event, respectively. Then, $B_{rate} \times B_{leng}$ corresponds to the packet loss rate. Our study results can be considered to be effective if the network impairments condition belongs to the region of packet loss rate $\leq 35\%$.

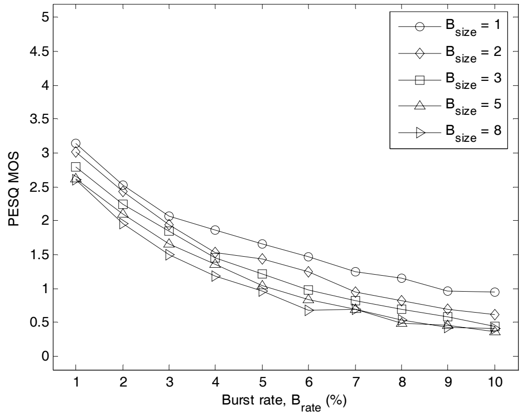


Fig. 2. VoIP quality dependency on different burst packet loss conditions

## Ⅳ. VoIP quality estimation using the quality metric

The objective assessment approach, which estimates the perceptual quality by measuring the network quality parameters, can be applied to real-time assessment. For the purpose of estimating VoIP quality on real-time basis, the important point is related to the quality metric which leads to estimating the subjective quality accurately and easily. In order to estimate perceptual quality accurately, the quality metric must be well chosen. In the previous section, we proved that the burst loss rate and average burst length can be used effectively as the quality metric except the condition of $B_{rate} \geq 7\%$ and $B_{leng} \geq 5$. Also these two parameters need to be measured on a real time basis. Fortunately, the RTP header contains a packet sequence number field. Monitoring this field at the edge-device enables to measure the quality metric. Fig. 3 shows a simple approach we propose that aims to measure the quality metric: the burst loss rate and average burst length. Motivated by the work on TCP round trip time estimation, a weighting

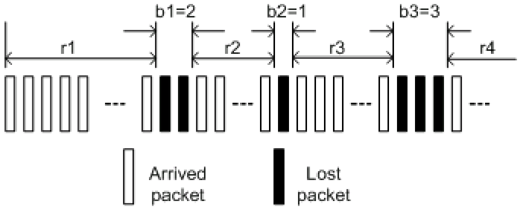


Fig. 3. Measuring the quality metric

factor $\alpha$ can be used to obtain the burst loss rate and average burst length[7].

The moving average for the burst loss rate ($R_i$) can be measured using

$$R_i = (1-\alpha)\times R_{i-1} + \alpha \times \frac{1}{r_i}, i=1,2,\cdots,K \quad (3)$$

where $r_i$ is the number of packets arrived between the ($i$-1)th burst loss event and $i$th burst loss event and $R_0$ = 0. The $K$th burst loss is the final burst loss event in the unit period of quality monitoring. Here we assume that one cycle for quality monitoring is fixed to 22 seconds which corresponds to the length of our voice sample. This means that VoIP quality monitoring repeats every unit of 22 seconds. Also the moving average for the burst length ($B_i$) can be measured using

$$B_i = (1-\alpha)\times B_{i-1} + \alpha \times b_i, i=1,2,\cdots,K \quad (4)$$

where $b_i$ is the number of packets lost in sequence when the $i$th burst loss event occurs and $B_0$ = 1. In order to investigate the effect of $\alpha$ on accuracy of our quality estimation, we used 10 different $\alpha$ values: 0.002, 0.004, 0.006, 0.008, 0.01, 0.02, 0.04, 0.06, 0.08, and 0.1.

At end of the quality monitoring unit, the measured burst loss rate and burst length come to ($R_K = br$) and ($B_K = bl$), respectively. We obtained $(br^{\alpha}_{(i,j,k)}, bl^{\alpha}_{(i,j,k)})$ for the DSTD($i$, $j$, $k$), $i$ = 1, 2, ⋯, 10, $j$ = 1, 2, ⋯, 8, and $k$ = 1, 2, ⋯, 15. The data volume correspond to the amount of 1200 records where each record contains 10 pairs of ($br$, $bl$) for 10 different $\alpha$ values. In order to find how accurately our quality metric estimates the perceptual quality, we calculated the difference between $M^{MOS}_{(i,j,k)}$ (actually measured PESQ scores) and the estimated value by matching the measured quality metric of $(br^{\alpha}_{(i,j,k)}, bl^{\alpha}_{(i,j,k)})$ to the trained MOS table using

$$\begin{aligned} D^{\alpha}_{(i,j,k)} &= \left| M^{MOS}_{(i,j,k)} - E^{MOS}_{(ir^{\alpha}_{(i,j,k)}, il^{\alpha}_{(i,j,k)})} \right| \\ ir^{\alpha}_{(i,j,k)} &= nt^{10}(br^{\alpha}_{(i,j,k)}) \\ il^{\alpha}_{(i,j,k)} &= nt^{8}(bl^{\alpha}_{(i,j,k)}) \end{aligned} \quad (5)$$

where the function $nt^n(x)$ converts $x$ and rounds the result to the nearest integer among from 1 to n. Then, the estimation error is defined as

$$E^{\alpha} = \frac{\sum_{i=1}^{10}\sum_{j=1}^{8}\sum_{k=1}^{15} D^{\alpha}_{(i,j,k)}}{1200}. \quad (6)$$

In Fig. 4, first 10 bars show the effect on estimation errors as a function of $\alpha$. The approach of ‘moving-average’ can be implemented simply because it needs memory-less property. On the other hand, Last bar in Fig. 4 shows the estimation error when we use the approach of ‘time-average’ to obtain average values relating to the burst loss rate and average burst length. It can be seen that ‘time-average’ case shows best on estimation error performance. However, the

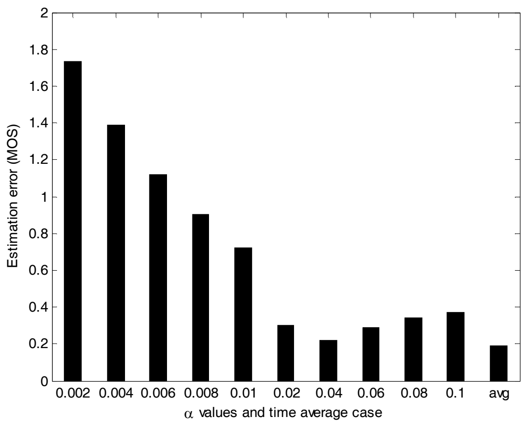


Fig. 4. Estimation errors for moving-average and time-average approaches

condition of $\alpha$ = 0.04 in 'moving-average' case shows similar performance to 'time-average' case. Then, considering benefits from implementation viewpoints, we argue that the 'moving-average' method with $\alpha$ = 0.04 can be used in VoIP quality monitoring restricting the estimation error within 0.02 in MOS when the quality monitoring repeats every unit of around 22 seconds.

## V. Quality-based accounting application

As VoIP systems move to wireless environments, with much higher average packet loss rates than wired networks, the end user is less likely to receive what he/she has contracted for, and is more likely to be dissatisfied with the service provided. A vital factor in ensuring customer satisfaction in service quality limitations, is to set obvious continuities between quality and billing. This has motivated us in the past to propose 'quality-based billing'[8], where the amount charged to the end user for a particular conversation depends on the QoS that was actually delivered. We extend the previous work using the quality metric explored in this paper.

As shown in Fig. 5, the place where the quality parameters can be measured is in the terminal itself. The endpoint measures the burst loss rate and average burst length using 'moving-average' method with $\alpha$ = 0.04 every 22 seconds. Then the burst loss rate and average burst length measured corresponds to accounting information, which needs to be sent to the gatekeeper. After the AAA (Authentication, Authorization, and Accounting) server collects the accounting information during call authorized period, it can compute call accounting using the cumulative data based on the quality prediction method suggested in section IV.

The gatekeeper and AAA server would involve into accounting information transfer to implement the above 'quality-based accounting system'. As shown in Fig. 6, the quality reporting messages flow during an RTP session [6]. The two terminals (T1 and T2) are responsible for measuring accounting metric values. An accounting message can be sent to the gatekeeper every after a certain interval (say 22 seconds). Then, the gatekeeper goes between the device and the accounting server (say AAA server).

## VI. Conclusion

In this paper, we used the burst loss model which has two parameters: the burst loss rate and average burst length. They have benefits in a simple and useful fit for measurement processes looking for the quality metric. We experimented PESQ assessments based on ITU-T Recommendation P.862 and obtained the trained MOS table. We proposed a simple 'moving average' approach with $\alpha$ aiming to measure the burst loss rate and average burst length. And in order to

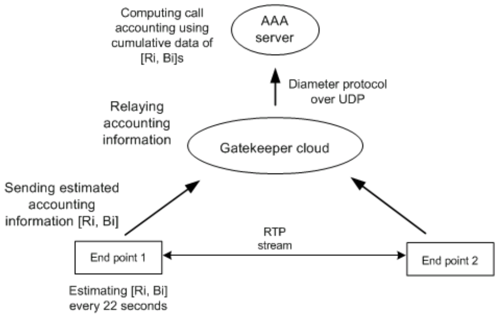


Fig. 5. Quality-based accounting system

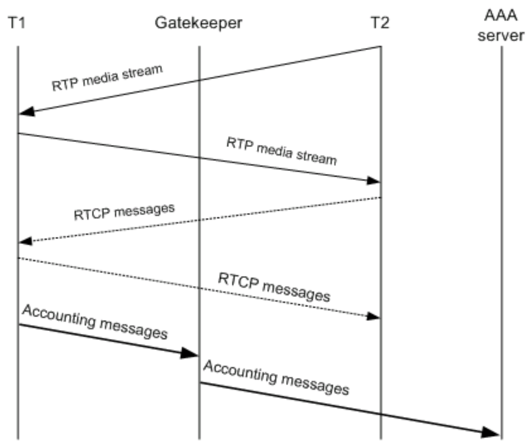


Fig. 6. Accounting information, gatekeeper and AAA server

find how accurately the two parameters measured estimates the perceptual quality, we calculated the differences between actual measured PESQ scores and estimated value by matching the measured quality metric to the trained MOS table. We found that using the burst loss rate and average burst length as the quality metric, the 'moving-average' method with $\alpha = 0.04$ can be suitable in VoIP quality monitoring restricting the estimation error within 0.02 in MOS when the quality monitoring repeats every unit of around 22 seconds. As one of applications, we proposed the quality-based accounting system. This is a step forward to the future AAA area because a vital factor in ensuring customer satisfaction in VoIP service quality limitations is to set obvious continuities between quality and billing.

## References

〔1〕 R. Ramjee, J. Kurose, D. Towsley and H. Schulzrinne: 'Adaptive playout mechanisms for packetized audio applications in wide-area networks', In Proc. IEEE Infocom, 1994, pp. 680-688.

〔2〕 N. Egi, H. Aoki and A. Takahashi: 'Objective quality evaluation method for noise-reduced speech', MESAQIN 2007.

〔3〕 A. Takahashi, H. Yoshino, and N. Kitawaki: 'Perceptual QoS Assessment Technologies for VoIP,' IEEE Communications Magazine, July 2004, pp.28-34.

〔4〕 K.C. Claffy: 'Measuring the Internet', IEEE Internet Computing, 2000, 1, (4), pp. 73--75.

〔5〕 Adam Petcher: 'QoS in Wireless Data Networks', http://www.cs.wustl.edu/~jain/cse574-06/ftp/wireless_qos/index.html

〔6〕 A. J. Saliba, M. A. Beresford, M. Ivanovish, and P. Fitzpatrick: 'User-perceived quality of service in wireless data networks', Personal and Ubiquitous Computing, 2005, Vol.9, (5), pp. 413-422.

〔7〕 B. Moon and A.H. Aghvami: 'Quality-of service mechanisms in All-IP wireless access networks', *IEEE J. on Seleceted Areas in Comm.Proc.-Commun.*, Vol.22, No.5, pp. 873-888, June 2004

〔8〕 R.B. Simon: 'VoIP quality assessment: taking account of the edge-device', IEEE Trans. Audio, Speech and Language Processing, 2006, 11, (14), pp.1977-1983.

〔9〕 A.W. Rix, J. G. Beerends, D.S. Kim, P. Kroon and O. Ghiza: 'Objective assessment of speech and audio quality-technology and applications', IEEE Trans. Audio, Speech and Language Processing, 2006, 11, (14), pp.1890-1901.

〔10〕 A. Lakaniemi, J. Rost and V.I. Raisanen: 'Subject VoIP speech quality evaluation based on network measurements', IEEE Internet Computing, 2001, (4), pp.73-75.

〔11〕 Younchan Jung and J.W Atwood: 'Dynamic adaptive playout algorithm using interarrival jitter and dual use of $\alpha$', IEE Proc.-Commun., 2006, 4, (153), pp.279--287.

〔12〕 'Packet-based multemedia communications systems ', ITU-T Recommendation H.323, July 2003.

〔13〕 'RTP Control protocol Extended Reports (RTCP XR) ', ITU-T, RFC 3611, Nov. 2003.

〔14〕 L. Carvalho, J. Angeja and A. Navarro: 'A new packet loss model of the IEEE 802.11g wireless network for multimedia communications', {\em IEEE Trans. on Consumer Electronics}, 51, (3), pp.809--814, Aug. 2005.

〔15〕 L. Carvalho, J. Angeja and A. Navarro: 'A new packet loss model of the IEEE 802.11g wireless network for multimedia communications', IEEE Trans. on Consumer Electronics, 51, (3), pp.809-814, Aug. 2005

〔16〕 'Packet loss distributions and packet loss model', ITU-T, COM 12-D97-E, Jan. 2003.

〔17〕 'Modelling burst packet loss within the E-Model ', ITU-T, COM 12-D104-E, Jan. 2003.

〔18〕 C. Metz: 'AAA protocols: authentication, authorization, and accounting for the Internet', *IEEE Internet Computing*, pp.75-79, November 1999.

〔19〕 F. Ghys, A. Vaaraniemi: 'Component-based charging in a next-generation multimedia network', *IEEE Communications Magazine*, 41,

(1), pp.99-102, Jan. 2003
〔20〕 Younchan Jung and J. William Atwood: 'QoS-Related accounting architectures for VoIP services,' in Proceedings of The 20th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2005), Jeju, Korea, July 4-7, 2005, Proceedings Vol.4, pp. 1475-1476.

**정 윤 찬** (Younchan Jung) 정회원

1980년 2월 경북대학교 전자공학과 졸업
1990년 2월 KAIST 전기 및 전자공학과 석사
1994년 8월 KAIST 전기 및 전자공학과 박사
1996년 3월~현재 가톨릭대학교 정보통신전자공학부 교수
<관심분야> 네트워크 QoS, VoIP, 네트워크 보안

**Ibanez Al Ann** 정회원

2009년 2월 Mapua Institute of Technology, Information Technology 학사
2009년 9월~현재 가톨릭대학교 정보통신전자공학과 석사과정
<관심분야> 컴퓨터 네트워크, Quality of Service, 멀티미디어 통신