

E-모델 기반 통화 품질 분석을 통한 VoIP Planning 및 평가

VoIP Planning and Evaluation through the Analysis of Speech Transmission Quality Based on the E-Model

배성용* 김광훈**
Seong Yong Bae Kwang Hoon Kim

요 약

인터넷의 발달은 통신 기술에 커다란 변화를 요구하고 있다. 특히 공중전화망(PSTN: Public Switched Telephone Network)은 장기적으로 인터넷으로 흡수 통합되고, 현재는 이들간의 연동이 요구되고 있는 실정이다. 이처럼 인터넷에서의 전화 서비스는 인터넷 응용 서비스 중 가장 활용 가치가 높고, 기대를 모으고 있는 서비스이다. 그러나 높은 기대와 적은 서비스 이용료에도 불구하고 인터넷 전화 서비스의 보급이 부진한 이유는 낮은 서비스 품질에 있다. 지금까지 VoIP(Voice over IP) 서비스 품질 평가를 위해 송수신되는 음성에 기반한 다수의 통화 품질 측정 알고리즘들이 사용되어 왔지만, 이러한 알고리즘들은 음성 샘플에 따라 그 결과가 다르고, 일부 알고리즘은 음성 전송 경로의 네트워크 환경 요소를 반영하지 못하는 단점이 있다. E-모델은 이러한 알고리즘들의 문제점을 해결하기 위해 사용될 수 있다. 본 논문에서는 VoIP 단말의 통화 품질은 물론, 음성 전송 경로의 네트워크 품질 손실 요소들을 체계적으로 분석할 수 있는 E-모델의 다양한 분석을 통해 인터넷 전화 사업자나 망 운용자의 VoIP 계획(planning)에 대한 가이드라인과 통화 품질 평가 방법 및 결과를 제시하고자 한다.

Abstract

Voice over Internet Protocol (VoIP) is currently a popular research topic as a real time voice packet transmission method. But current Internet environment do not guarantee the quality of voice when we take a side view of delay, jitter and loss. Up to now, many voice based evaluation algorithms have been used to measure speech quality of VoIP systems. However, these algorithms have the defects that their results are different according to voice samples and some algorithms can not take network environment for speech transmission path. The E-model can be used to solve the problems of these algorithms. In this paper, we introduce VoIP planning guidelines through the various analysis of E-model which can model impairments of network quality as well as VoIP equipment quality systematically. We, also, show the evaluation method and results of speech transmission quality.

Keyword : Internet Telephony, VoIP QoS, VoIP Planning, E-model

1. 서 론

1980년대 후반 Archie, Gopher 등으로 시작된 인터넷은 90년대에 웹(Web)의 출현으로 기존 공중전화망(PSTN: Public Switched Telephone Network)과 같이 전 세계를 연결하는 데이터 통신 네트워크로 자리를 잡게 되었다. 초기 단순한 데이터 검색 및 전송만을

위한 수단이었던 인터넷은 월드와이드웹(WWW: World Wide Web)의 저변 확대로 데이터뿐 만 아니라 음성, 화상 등의 멀티미디어 네트워크로 진화했다. 이 때부터 인터넷은 더 이상 데이터만을 전송하는 수단이 아닌 통합 네트워크 수단으로 태어났다.

이 후, 폭발적인 인터넷 사용자의 증가로 인터넷은 PSTN과의 연동이 너무나도 당연한 결과가 되었으며, 발전을 거듭할수록 음성, 팩스, 문자, 사진, 동영상 등의 멀티미디어 응용이 출현하게 되었다. 즉 기존의 데이터 서비스는 물론, 웹사이트를 통한 데이터베이스 검색, 주문형 비디오(VOD: Video on De-

* 정 회 원 : 한국정보통신기술협회 시험인증연구소 선임연구원
sybac@tta.or.kr(제1저자)

** 정 회 원 : 경기대학교 정보과학부 부교수
kwang@kyonggi.ac.kr(공동저자)

mand) 및 오디오(MOD: Music on Demand), 인터넷을 통한 전화 서비스 등이 가능한 수준으로 발전을 해오고 있다.

여러 인터넷 응용 분야들 가운데 가장 많은 인기와 잠재적인 발전 가능성이 있는 것이 VoIP(Voice over Internet Protocol) 기반의 인터넷 전화 서비스이다. 인터넷 전화 서비스는 기존의 음성 통신 시장에 새로운 변화를 야기하고 있다. 최근 통신 사업자, 인터넷 서비스 제공자, 기업체 및 일반 이용자의 관심이 고조되고 있는 VoIP 기술은 인터넷의 최대 응용 서비스 중 하나로 급부상함과 동시에 시장이 급속히 성장할 수 있는 분야로 주목 받고 있다.

인터넷 전화가 성장할 수 있는 가장 큰 요인은 이동 및 유선 전화보다 상대적으로 저렴한 비용에 있다. 인터넷 인구의 증가와 값싼 서비스 이용료라는 특징으로 볼 때 분명 인터넷 전화는 성장할 수 밖에 없을 것이다. 그러나 초기의 기대와 관심과는 달리 인터넷 전화가 아직까지 대중화되지 못한 데에는 낮은 서비스 품질이라는 커다란 요인이 있다. 기존의 PSTN 전화에서 양질의 서비스를 제공 받던 서비스 이용자들이 적은 서비스 이용료라는 장점에도 불구하고 인터넷 전화를 쉽게 선택하지 않는 이유는 아직까지 인터넷 전화가 PSTN 전화와 동일한 수준의 서비스 품질을 제공하지 못하기 때문이다. 이는 본래 VoIP가 서비스 품질을 보장하지 않은(best effort) IP 망을 기반망으로 음성 데이터 서비스를 제공하기 때문에 전화 서비스 품질 또한 낮을 수밖에 없다.

따라서, VoIP 기술의 인터넷 전화에서 서비스 품질(QoS: Quality of Service)은 기술의 확산 및 이용의 증대를 위해서 반드시 그리고 지속적으로 해결하여야 할 문제이다. VoIP QoS에 영향을 미치는 요소는 매우 다양하다. VoIP 단말(terminal) 내부에서 코덱 지연 및 에코 발생을 비롯하여 IP 망 구간에서 지연(delay), 패킷 손실(packet loss) 및 지터(jitter) 등 다양한 QoS 손실 요인이 있다[1]. 인터넷 전화 사업자 또는 망 운용자는 양질의 인터넷 전화 서비스 품질을 제공하기 위해서 다양한 서비스 품질 손실 요인을 정의하고 분석하여 이들을 최소화할 수 있어야 한다.

이는 곧 VoIP 단말의 성능과 전형적인 최선(best effort)의 서비스 품질을 제공하는 IP 기반 인터넷 망 최적화 문제로 국한할 수 있다. IP 기반 인터넷 서비스 품질의 최적화는 IntServ(Integrated Service), Diff-Serv(Differentiated Service), RSVP(Resource Reservation Protocol) 및 MPLS(Multi Protocol Label Switching) 등 기존에 제공되는 다양한 네트워크 QoS 기술들을 적용할 수 있다. 결국, VoIP 인터넷 전화 사업자 또는 망 운용자는 일정 수준의 서비스 품질을 보장하기 위해서는 VoIP 단말과 VoIP 망의 서비스 품질을 체계적으로 분석하고, 시스템 및 망 구축과 운용에 대한 계획(planing)이 필요하다.

본 논문에서는 E-모델이라는 통화 품질 평가 모델의 분석을 통해서 서비스 사업자 및 망 운용자의 VoIP planing에 대한 가이드라인을 제시하고, 실제 통화 품질을 평가하기 위한 방법과 그 결과를 보이고자 한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 기술 연구로서 음성 기반 통화 품질 평가 알고리즘의 종류와 문제점에 대해 알아보고, 3장에서는 VoIP 서비스 품질에 대해 설명한다. 4장에서는 ITU-T(International Telecommunications Union-Telecommunication standardization sector)에서 정의하고 있는 E-모델에 대해 IP 텔리포니 환경에서 VoIP 서비스 품질에 영향을 미치는 손실 요소들을 살펴보고, 이들간의 상관 관계를 알아본다. 그리고 5장에서는 E-모델 분석 결과를 통해 VoIP planing을 하기 위한 가이드라인을 제시하고, 마지막으로 6장에서는 VoIP 통화 품질을 평가하기 위한 방법과 그 결과를 소개한다.

2. 관련 기술 연구

음성 통화 품질 평가에 처음으로 사용된 방법이 MOS(Mean Opinion Score)이다. MOS는 다수의 훈련을 받은 평가원들이 이미 정의된 평가 절차에 의해 통화 품질을 점수로 산정하여 그 평균값을 구하는 방법이다[2]. MOS값은 다른 방법의 측정값보다 직관적으로 품질 등급을 이해하기 쉽기 때문에 일반적인

평가 기준값으로 널리 사용된다. 그러나 이 방법은 근본적으로 주관적인 특성을 가지고 있으며, 평가에 많은 인력과 시간이 소요되는 단점을 가지고 있다. 따라서 적은 비용과 객관적인 품질 평가를 위해 PSQM(Perceptual Speech Quality Measurement)이라는 새로운 평가 알고리즘이 소개되었다. PSQM은 음성 압축 코덱에 의해 영향을 받는 음성의 인지되는 품질을 측정하기 위해 고안되었다[3,4]. 따라서 이 방법은 기본적으로 데이터 네트워크 전송에서 발생하는 패킷 손실, 지연 및 지터와 같은 네트워크 손실 요소들을 반영하지 못한다. 이를 해결하기 위한 방법으로 향상된 PSQM 알고리즘인 PSQM+가 개발되었는데, 이 방법에서는 네트워크 손실을 반영한 평가 품질을 얻을 수 있다[4]. 그러나 이 방법도 여전히 인지되는 음성 기반 평가 결과를 제공한다는 한계를 가지고 있다. PAMS(Perceptual Analysis Measurement System)는 PSQM+와 같이 네트워크의 손실을 반영하여 품질 평가를 할 수 있으면서, PSQM과는 다른 신호 처리 모델을 사용하여 LQS(Listening Quality Score)와 LES(Listening Effort Score)라는 보다 상세한 평가 결과를 제공할 수 있다[4]. 또한 PESQ(Perceptual Evaluation of Speech Quality)는 PSQM+와 PAMS의 장점을 취합하여 개발된 알고리즘으로서, PSQM+나 PAMS와 동일한 시험 결과를 제공한다[4,5].

지금까지, VoIP 통화 품질 평가를 위해 앞에서 설명한 여러 알고리즘이 사용되어 왔지만, 이러한 알고리즘들은 송수신된 음성 기반하여 인지되는 음성의 품질을 평가하기 때문에 음성 샘플에 따라 품질이 달라지고, 일부 알고리즘에 대해서는 음성 전송 경로에 대한 네트워크의 환경 요소를 품질에 반영하지 못하는 문제점이 있다. E-모델은 이러한 음성 기반 평가 알고리즘들의 문제점을 해결하기 위해 사용될 수 있다.

3. VoIP 서비스 품질

일반적으로 음성 서비스 품질은 접속 품질과 통화 품질로 구분된다. 따라서 VoIP 서비스 품질은 다음과 같은 요소로 구분할 수 있다[6].

- 접속 품질(호 설정 시간)
 - ✓ 시스템 내부 셋업 소요 시간
 - ✓ 호 제어 신호 처리를 위한 시스템간(예: 게이트키퍼-게이트웨이) 연결 소요 시간
 - ✓ 인증, 과금과 같은 부가 서비스 처리를 위해 서버 접속에 소요되는 시간
- 통화 품질(호 명료성)
 - ✓ 종단간(End-to-end) 지연: 코덱 지연 + 패킷화 및 버퍼링 + 네트워크 전송 지연
 - ✓ 통화 품질 영향 요소: 코딩 잡음, 에코, 패킷 손실, 대역폭 부족 등

표 1은 VoIP 서비스 품질과 관련한 주요한 품질 파라미터를 보여주고 있다.

<표 1> VoIP 서비스 품질 파라미터

구분	접속품질(접속지연/접속률)	통화품질(호 명료성)
품질 파라미터	DTD	[주관적 파라미터] MOS
	PDD	
	CCD	[객관적 파라미터] PSQM PAMS PESQ E-모델(R-factor)
	Call Cut-Off Ratio	
	Call Completion Ratio	

<참고>

- DTD (Dial Tone Delay: 발신음 지연)
- PDD (Post Dialing Delay: 다이얼 후 접속 지연)
- CCD (Call Clearing Delay: 호 해제 후 통화 단절음 지연)

접속 품질 파라미터에서 DTD는 사용자가 수화기를 들고(off-hook)난 후 다이얼 톤을 듣는데 까지 걸리는 시간을 나타내며, PDD는 사용자가 착신 번호 입력을 마치고 난 후 링백(ring-back)톤을 듣는데 까지 걸리는 시간을 의미한다. CCD는 사용자가 통화를 마치고 호를 끊었을 경우(on-hook) 통화 단절음을 듣는데 까지 걸리는 시간을 의미한다. 호 완료율(Call Completion Ratio)은 사용자가 호를 여러 번 시도하였을 경우 정상적으로 호 통화가 된 비율을 나타내며,

이와는 반대로 정상적으로 호가 연결되지 않은 비율이 호 절단율(Call Cut-Off Ratio)이다.

통화 품질은 객관적으로 평가하기에 어려움이 있으나 효율적인 네트워크 구축 및 서비스를 위해서는 이에 대한 객관적이고 과학적인 접근이 요구된다. 일반적으로 통화 품질은 MOS라는 5단계의 평가 기준에 의해 품질 분류가 이루어져 왔다[2].

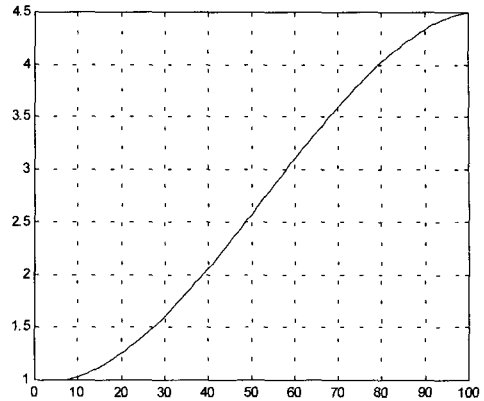
〈표 2〉 MOS 평가표

점수	품질	잡음 정도
5	Excellent	Imperceptible
4	Good	Just Perceptible, not Annoying
3	Fair	Perceptible and slightly Annoying
2	Poor	Annoying but not Objectionable
1	Bad	Very Annoying and Objectionable

그러나 MOS값은 다수의 사람들이 자기 나름대로의 평가를 내리고 평균값을 구하는 방식으로 주관적이며 비효율성이 크다. 따라서 ITU-T에서는 통화 품질에 영향을 미치는 요소들을 체계적으로 분석하기 위한 새로운 방법으로 E-모델을 통한 R값을 제시하고 있다. R값은 MOS값을 비롯하여 PSQM, PAMS, PESQ와 같이 보낸 음성과 받은 음성을 비교하여 품질을 평가하는 방법과는 달리 음성이 지나가는 네트워크 경로의 환경 요인을 분석하여 품질을 평가하는 방법이다. 따라서 음성 샘플에 상관없이 동일한 결과값을 제시한다. 아래의 그림 1은 R값과 MOS값과의 관계를 나타낸 것이다[7].

〈표 3〉 음성 전송 품질의 분류

E-모델 R값	품질	만족도	추정 MOS값
$90 \leq R \leq 100$	최상	매우 만족	4.3 이상
$80 \leq R \leq 90$	높음	만족	4.0 이상
$70 \leq R \leq 80$	중간	소수의 사용자 불만	3.6 이상
$60 \leq R \leq 70$	낮음	다수의 사용자 불만	3.1 이상
$50 \leq R \leq 60$	나쁨	거의 모든 사용자 불만	2.6 이상



〈그림 1〉 R값과 MOS값의 관계

4. E-모델 분석

E-모델은 ITU-T에서 정한 “전송품질평가모델”이라고 할 수 있는데, 여기서 다루는 전송 품질의 의미는 종단간(mouth to ear)에서의 음성 전송 시 품질을 말한다. 즉, 어떤 주어진 네트워크 환경하에서 통화를 하려고 할 때 사용자가 직접 느낄 수 있는 주관적인 품질을 보편적인 인지 측면에서 최대한 객관적으로 평가하고자 착안한 모델이라고 할 수 있다. E-모델은 모든 망에서의 전송 품질을 측정하고 망 계획을 수립하기 위한 도구로서 제안되었으나, IP 전화에서의 손실(impairment)을 적절하게 수용할 수 있어 인터넷 전화 품질 평가에 널리 적용되고 있다. E-모델의 결과값은 R값이라고 불리는 스칼라(scalar)이다. R값이 가질 수 있는 범위는 50부터 100사이의 값으로 50 이하의 값은 수용되지 않는 값이며, 협대역(300에서 3400Hz)의 음성 전송 품질에서 가질 수 있는 최대값은 94.7이다. E-모델은 종단간 전송 품질에 대한 특정 손실 파라미터와 이들간의 상호 작용에 관련된 몇 개의 모델들로 구성되어 있다. 전송 품질 R값을 결정하는 E-모델 수식은 다음과 같다[8].

$$R = R_o - I_s - I_d - I_e, eff + A \quad (1)$$

R_o : 송수신측 신호 크기와 실내 및 회선의 전기적인 소음을 고려한 기본적인 신호

대 소음의 비율(S/N)

Is : 신호 크기 수준, 사이드 톤(side tone), 그리고 PCM(Pulse Code Modulation) 양자화 왜곡(quantizing distortion) 등과 같은 실시간 또는 동시에 발생하는 음성 전송 손실들의 합

Id : 송화자 에코(talker echo), 수화자 에코(listener echo) 그리고 절대 지연(absolute delay) 시간과 같은 음성 신호에 대한 지연 손실들의 합

Ie : 낮은 비트율(low bit-rate) 코딩 방식의 압축 코덱 사용에 따른 특정 장비 내에서 발생하는 장비 손실. Ie,eff는 패킷 손실을 고려한 압축 코덱 사용에 따른 장비 손실을 나타낸다.

A : 이동(mobility)이나 위성(satellite) 전화 같이 새로운 서비스 사용에 따른 이득값.

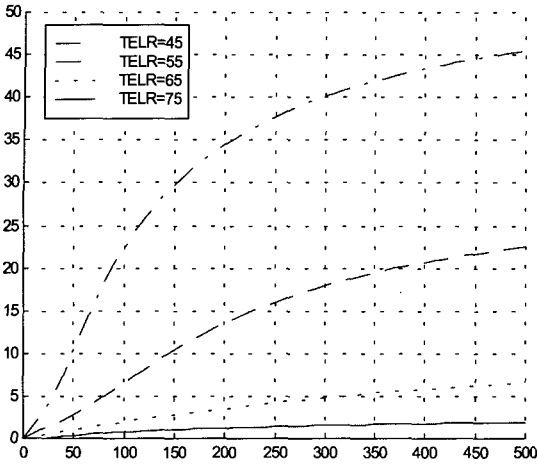
Ro값을 구성하는 식은 G.107[8]에 정의되어 있다. 구성 식 중에서 변경 가능한 주요 파라미터 값은 SLR (Send Loudness Rating)과 RLR (Receive Loudness Rating)이다[9]. SLR은 송신 스피커에 울리는 소리 압력 대 전압 발생의 비를 dB (Decibel)로 나타낸 단위이며, RLR은 수신 스피커에 걸린 전압 대 발생한 소리 압력의 비를 dB로 표현한 것이다. 따라서 송신-수신의 종단간 전체 Loudness Rating인 OLR (Overall Loudness Rating)은 SLR+RLR과 같다. 즉 OLR이 0dB인 경우는 송신측 소리의 크기가 수신측에 동일한 세기로 전달됨을 의미한다. G.108[9]의 그림 9에 의하면 OLR이 6~10dB 사이일 경우 R값이 가장 높다. 참고로 [8]에서 정의하고 있는 OLR의 권고값은 SLR이 8 dB, RLR이 2 dB인 10dB이며, 이 경우 Ro는 94.7의 값을 갖는다.

Is값은 OLR이 너무 작아 발생하는 음질 감소, 통화 시 발생하는 사이드 톤에 의한 품질 손실 그리고 음성의 디지털화 과정에서 발생하는 양자화(quantization) 잡음에 의한 손실 등을 고려한 것으로, G.107[8]에 그 구성식과 입력 파라미터들의 권고값

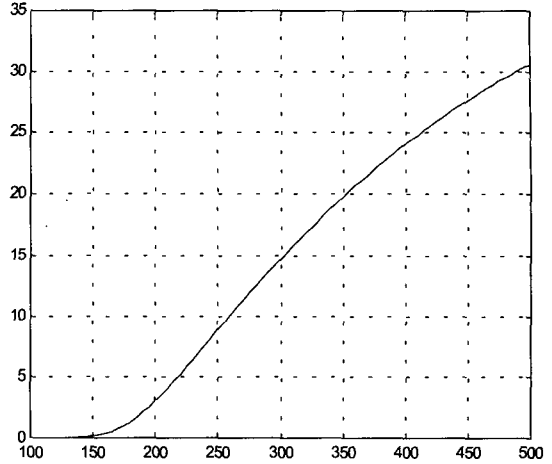
을 제시하고 있다. 여기서 사이드 톤이란 특별한 형태의 에코로서 송화자의 음성이 다시 송화자의 귀에 들리는 잡음을 말하며, 발생하는 원인은 전화기 회선의 밸런스 저항(impedance)과 PSTN 종단 회선의 입력 저항과의 불일치에서 비롯된다. [8]에서 제시하는 파라미터의 기본값을 적용하면, Is는 1.43이란 값을 얻을 수 있다.

Id값은 음성 신호 지연(delay)에 따른 모든 손실들을 나타내며, 송화자 에코(talker echo) 지연과 수화자 에코(listener echo) 지연에 따른 잡음과 직접적인 음성 지연에 따른 품질 저하 요소를 포함하고 있다[8]. 일반적으로 수화자 에코는 송화자 에코에 영향을 미치는 네트워크의 요소 및 연결 특성에 의존적이다. 따라서 송화자 에코에 대한 충분한 제어가 이루어진다면, 수화자 에코는 거의 무시할 수 있다. 먼저 에코 지연(echo delay)에 따른 잡음의 경우, 송화자 에코 지연 손실에서 TELR(Talker Echo Loudness Rating)이 중요한 파라미터로 작용하는데, 이 값은 앞서 설명한 OLR과 EL(Echo Loss)의 합으로 표현된다. 즉, 에코 신호가 발생하는 경로의 전체 손실도를 dB로 나타낸 것으로 G.107[8]에서는 65dB를 권고하고 있다. TELR이 65dB보다 작을 경우 주어진 시간 지연에 대하여 TELR이 65dB이상인 경우보다 R값이 작아짐을 알 수 있다. 그림 2는 45dB에서 75dB 사이의 TELR값에 대하여 송화자 에코가 R값에 미치는 영향을 보여주고 있다. TELR이 65dB일 경우 지연이 300ms이내에서는 5이하의 R값 저하가 나타난다.

한편, 음성 시간 지연에 의한 직접적인 품질 저하 요소는 그림 3에 나타나 있다. 150ms 미만의 지연에서는 일반적으로 시간 지연을 거의 감지할 수 없으므로 R값에 영향이 없으나, 그 이상 지연이 증가하면 급격하게 R값이 감소한다. 또한 R값 90이상을 만족하려면 시간 지연에 따른 R값 손실이 대략 4.7 이하이어야 하는데, 이 때 시간 지연의 최대값은 200ms이다. 이는 두 사람이 대화할 때, 한 사람이 말을 마치고 다른 사람이 말을 이어갈 경우(turn talking), 약 200ms의 시간 중단이 필요한 것과 일치한다. 즉, turn talking 지연이 200ms으로, 시간 지연이 그 이상이면



〈그림 2〉 예코 잡음에 따른 품질 저하 요소의 분석



〈그림 3〉 시간 지연에 따른 품질 저하 요소의 분석

자연스러운 대화가 불가능하고 대화 의미의 왜곡 등 품질 저하가 발생할 수 있음을 의미한다[10].

I_e 는 낮은 비트율(low bit-rate) 코딩 방식에 의한 압축 코덱 사용으로 장비 내에서 발생하는 음성 품질 손실을 나타내는 값으로, 음성 신호의 압축 코덱 종류에 따라 그 값이 달라진다. G.113 Appendix I[11]에서는 다양한 음성 코덱에 따라 알맞은 I_e 값을 권고하고 있는데, 패킷 손실이 발생하는 경우도 함께 제시하고 있다. 식 (1)에서 $I_{e,eff}$ 가 의미하는 것은 I_e 값과 패킷 손실율을 함께 고려했을 때의 실질적인 I_e 값을 의미한다.

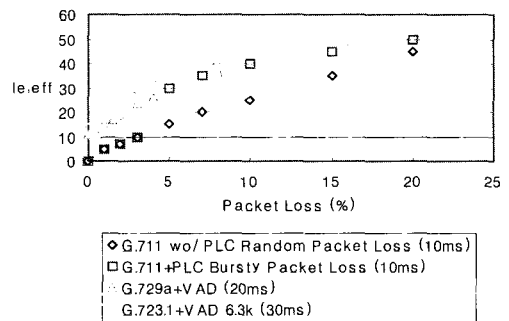
표 4에서는 코덱의 종류에 따른 I_e 값을 보여주고 있다. $I_{e,eff}$ 의 경우에는 ITU-T에서 G.729a, G.723.1, G.711 코덱의 패킷 손실 발생시 $I_{e,eff}$ 값을 일부 실험 데이터를 근거로 제안하고 있으나, 이는 한정된 샘플

시험 결과에 기반을 둔 값이므로 절대적이라고 할 수 없다. 또한, 연속적인 음성 패킷 여러 개를 한꺼번에 잃어버리는 burst packet loss의 경우, 현재 ITU-T Study Group 12에서 연구가 진행중이며, 현재의 $I_{e,eff}$ 값에는 반영되어 있지 않다. Random packet loss의 경우, Bpl이라는 Packet Loss Robustness 값을 정의하여 패킷 손실 정도에 따른 $I_{e,eff}$ 값을 수식을 통해 구할 수 있다[8]. 그림 4는 G.711, G.729a, G.723.1 6.3k 코덱에 대하여 패킷 손실율을 증가시키면서 계산한 $I_{e,eff}$ 값을 보여주고 있다. 그림에서, 약 3% 미만 패킷 손실의 경우 G.711 코덱은 burst나 random 구분 없이 동일한 $I_{e,eff}$ 값을 갖지만, 5% 이상 패킷 손실이 발생하면 burst packet loss의 결과가 random packet loss보다

〈표 4〉 음성 코덱에 따른 I_e 값

코덱 종류	I_e
G.711	0
G.729a+VAD	11
G.723.1 6.3k	15
G.723.1 5.3k	19

* VAD: Voice Activity Detection



〈그림 4〉 패킷 손실율에 따른 코덱별 장비 손실

MOS값이 1등급 정도 저하됨을 알 수 있다. 여기에서, $I_{e,eff}$ 값은 수많은 샘플 시험을 통해 실험적으로 유추한 값을 알아두어야 한다.

식(1)의 A는 이동이나 위성 전화와 같이 새로운 서비스 사용에 따른 이득값을 나타낸다. 일반적으로 서비스 사용자는 새로운 형태의 서비스가 초기에 낮은 서비스 품질을 제공하더라도 어느 정도의 수준의 품질은 수용할 수 있는데, 이는 이득값으로 산정되어 R 값에 증가 요인으로 반영된다. 그러나 이 값은 시간이 지남에 따라 서비스 품질이 개선되고 사용자가 서비스 사용에 익숙해지면서 점차 감소한다. 인터넷 전화는 새로운 서비스라기 보다는 기존의 유선 전화 서비스를 대체한 것으로 보아 그 값을 0로 권고하고 있다. 참고로 [8]에서 정의하고 있는 A의 권고값은 표 5와 같다.

〈표 5〉 통신 시스템 별 이득값 A

통신 시스템 예	A 최대값
Legacy(유선)	0
건물 내에서의 무선	5
지리적 또는 차량 이동중의 무선	10
위성(다중 홉 연결)	20

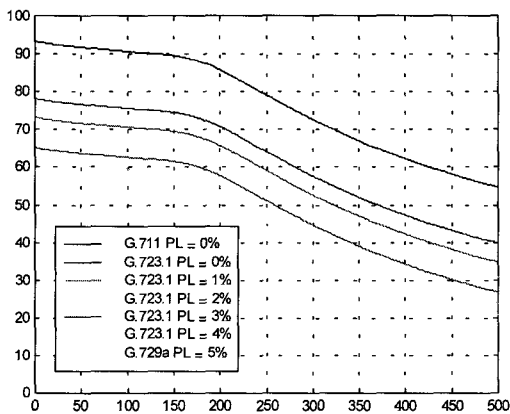
5. R값 분석을 통한 VoIP Planning

표 3의 음성 품질에 대한 분류에서 R값이 70 이상이면 인터넷 전화 서비스가 가능한 기준으로 분류할 수 있다. 식 (1)에서 R_0 , I_s , 그리고 A 값을 G.107[8]에서 제안하는 기본값으로 적용하면 다음과 같이 간단한 식으로 나타낼 수 있다.

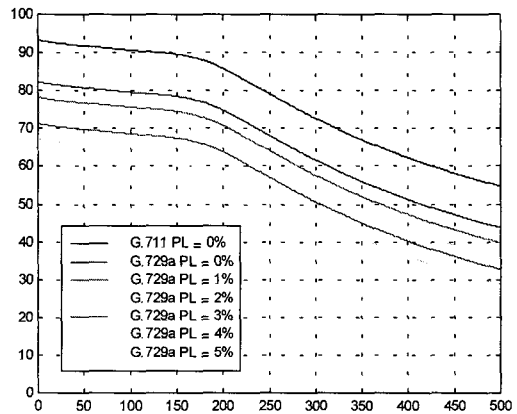
$$R = 93 - I_d - I_{e,eff} \quad (2)$$

따라서 원하는 통화 품질을 얻기 위해서는 I_d 와 $I_{e,eff}$ 에 의한 R값 손실이 23을 초과하여서는 안된다. 앞에서 설명한대로 I_d 는 지연에 의한 손실이고, $I_{e,eff}$ 는 장비에서 패킷 손실을 고려한 코덱에 의한 손실을 의미하므로 지연, 코덱, 패킷 손실율의 세가지 변수를 변화시키면서 다양한 상황을 가정해 볼 수 있다.

그림 5에서는 G.729a 및 G.723.1 6.3k 코덱을 적용했을 때 패킷 손실율 및 지연에 따른 R값을 보여주고 있다. 한 예로, G.729a 8k 코덱 사용에서 패킷 손실율이 2%인 경우, 종단간 지연이 150ms 미만이면 "Medium" 품질을 만족할 수 있음을 알 수 있다. 이때, 코덱 자체의 지연은 $T_F(N+1) + T_L(T_F)$: 프레임 길이, T_L : look ahead 시간, N : 패킷당 프레임 수)를 통해 구할 수 있는데[12], G.729a는 최소 25ms의 코덱



(a) G.729 코덱



(b) G.723.1 코덱

〈그림 5〉 코덱별 패킷 손실율과 지연에 따른 R값

〈표 6〉 R≥70(MOS: 3.6) 기준을 만족하는 VoIP 네트워크 환경의 예

지연/손실 코덱	코덱 지연 시간	패킷당 프레임 수	지터 버퍼 크기	패킷 손실	네트워크 지연 시간
G.729a+VAD	45ms	3(30ms)	4(120ms)	1.0%	41ms
G.729a+VAD	65ms	5(50ms)	2(100ms)	1.0%	41ms
G.729a+VAD	45ms	3(30ms)	3(90ms)	2.0%	27ms
G.729a+VAD	25ms	1(10ms)	3(30ms)	2.5%	28ms
G.723.1 6.3k	97.5ms	2(60ms)	1(60ms)	0.5%	30.5ms
G.723.1 5.3k	67.5ms	1(30ms)	1(30ms)	0.2%	34.5ms
G.723.1 6.3k	67.5ms	1(30ms)	1(30ms)	1.0%	34.5ms

〈표 7〉 R≥75(MOS: 3.8) 기준을 만족하는 VoIP 네트워크 환경의 예

지연/손실 코덱	코덱 지연 시간	패킷당 프레임 수	지터 버퍼 크기	패킷 손실	네트워크 지연 시간
G.729a+VAD	45ms	3(30ms)	3(90ms)	0.5%	42ms
G.729a+VAD	35ms	2(20ms)	3(60ms)	1.0%	37ms
G.729a+VAD	45ms	3(30ms)	2(60ms)	1.0%	27ms
G.729a+VAD	25ms	1(10ms)	2(20ms)	1.3%	38ms
G.723.1 6.3k	67.5ms	1(30ms)	1(30ms)	0.1%	34.5ms

지연을 가지며, G.723.1은 67.5ms의 코덱 지연을 가지고 있다. 여기서, T_L look ahead 시간은 코덱 알고리즘 수행시간으로 G.729는 5ms, G.723.1은 7.5ms의 지연 시간을 가진다. R값이 70(MOS: 3.6)과 75(MOS: 3.8) 이상의 경우를 기준으로 다양한 VoIP 네트워크 환경을 분석한 결과를 표 6과 표 7에서 보다 자세하게 설명하고 있다.

중단간 지연 시간은 크게 코덱에 의한 지연, 네트워크 지연, 지터 버퍼(jitter buffer)에 의한 지연으로 구분할 수 있다. 지터 버퍼의 경우는 네트워크에서 발생하는 지터를 보정해주기 위한 것으로 패킷 손실 과도 밀접한 관련이 있다. 만약 네트워크에서 발생하는 지터가 지터 버퍼보다 크다면 필연적으로 지터 버퍼에 의한 패킷 손실이 발생하기 때문이다. 하지만 지터 버퍼를 너무 크게 설정하면 그에 비례해서 전체 지연이 증가하므로 네트워크 환경을 고려하여 적절한 지터 버퍼를 설정하는 것이 중요하다. 실제 VoIP 장비에서는 동적(dynamic) 지터 버퍼를 적용하고 있어서 상황에 알맞도록 자동적으로 조정해 주고 있다.

표 6에서 G.729a 코덱의 경우, 패킷당 프레임수가 3(30ms)이고 지터 버퍼 크기가 4(120ms)이면 최대 패킷 손실율은 1%, 최대 네트워크 지연 시간은 41ms이다. 이를 다르게 해석하면 일반적으로 서비스 사업자 네트워크 지연 시간이 최대 30ms를 넘지 않는 것으로 볼 때, G.729a 코덱은 패킷 손실율 1%대에서 패킷당 프레임 수 3, 지터 버퍼 크기 4(120ms)로 설정하여 서비스가 가능함을 의미한다. 한편, 패킷 손실율이 1%에서 패킷당 프레임수가 5(50ms)이면 최대 지터 버퍼 크기가 2(100ms)이다. 또한, 패킷 손실율이 각각 2%와 2.5%일 경우는 패킷당 프레임수 3(30ms), 지터 버퍼 크기 3(90ms)과 패킷당 프레임수 1(10ms), 지터 버퍼 크기 3(30ms)에서 서비스가 가능함을 나타낸다. G.729a 코덱과 달리 G.723.1 코덱은 보다 제한된 네트워크 환경에서 사용이 가능하다. 한 예로, 6.3k의 경우, 패킷 손실율은 최대 1%를 초과해서는 안된다. 즉, 패킷 손실율이 1%의 경우 패킷당 프레임수가 1(30ms)이고 지터 버퍼 크기도 1(30ms)인 조건에서 허용 가능한 네트워크 지연 시간은 34.5ms이다.

따라서 G.723.1 6.3k 코덱은 사용 가능한 최대 패킷 손실율이 1%이다. 이에 반해, G.723.1 5.3k 코덱은 더 제한된 조건에서 사용이 가능하다. G.723.1 5.3k 코덱의 최대 패킷 손실율은 0.2%이다.

표 7은 R 값이 75 이상인 네트워크의 다양한 조건을 나타낸다. R 75 조건에서 G.729a 코덱의 최대 패킷 손실율은 1.3%이다. 그러나 1.3%의 미만에서는 패킷당 프레임수와 지터 버퍼 크기를 다양하게 설정하여 사용할 수 있다. 반면, G.723.1 코덱의 최대 패킷 손실율은 0.1%이다.

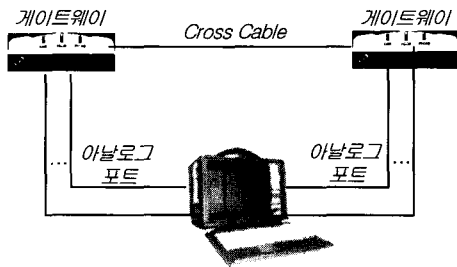
앞에서 살펴본 바와 같이 G.729a 코덱보다 G.723.1 코덱과 같이 압축율이 커지면 품질은 저하되지만 대역 용량면에서는 이득을 얻을 수 있다. 따라서 VoIP 품질을 일정 수준 이상 보장해 주기 위해서는 네트워크의 품질을 정확히 파악하여 그에 알맞은 코덱을 선택하는 것이 매우 중요하다.

6. VoIP 통화 품질 평가

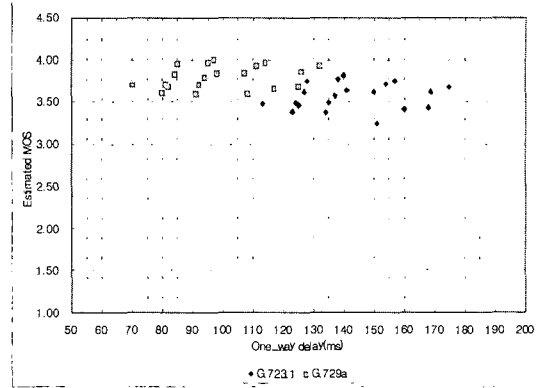
6.1 VoIP 단말 통화 품질 평가

그림 6은 VoIP 게이트웨이의 통화 품질 측정을 위한 시험 구성도이다.

시험에서, 모든 시험 대상 장비는 G.723.1 6.3k와 G.729a 코덱을 사용하고, 최소 지터 버퍼 크기는 30ms로 설정한다. 계측 장비는 Empirix사의 HammerIT 장비를 사용하고 사용한 측정 알고리즘은 PSQM이며, 이 값을 MOS값으로 치환하여 결과를 표시한다. HammerIT 장비를 시험 대상 게이트웨이의 FXS



〈그림 6〉 VoIP 게이트웨이 통화 품질 측정 시험 구성도



〈그림 7〉 아날로그 VoIP 게이트웨이 통화 품질 측정 결과

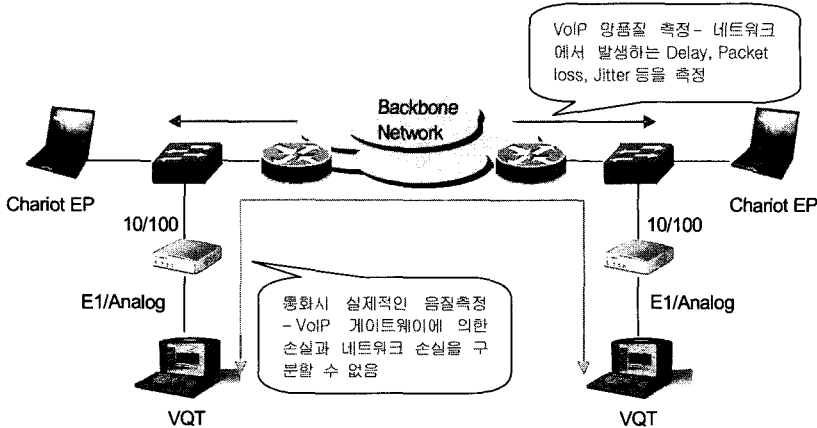
(Foreign Exchange Station) 포트에 연결하여 동시에 최대 8개까지의 콜을 시도하여 약 2분간 통화를 유지하면서 음질을 측정하며 같은 과정을 20분에 거쳐 반복한다. 그림 7은 이와 같은 방법으로 구해진 20종의 아날로그 전화기용 인터넷 전화 게이트웨이의 통화 품질 시험 결과이다.

측정한 시험 구성과 동일한 조건에서 이론적인 R 값과 그에 해당하는 MOS값을 구해보면 다음과 같다.

- G.723.1 6.3k
 - ✓ 지연: 코덱 67.5ms + 지터 버퍼 30ms = 97.5ms
 - ✓ 패킷 손실: 0%
 - ✓ E-모델 R값: 75
 - ✓ MOS값: 3.8
- G.729a
 - ✓ 지연: 코덱 25ms + 지터 버퍼 30ms = 55ms
 - ✓ 패킷 손실: 0%
 - ✓ E-모델 R값: 80
 - ✓ MOS값: 4.0

그러나, 실제 시험에서 얻은 결과값은 다음과 같다.

- G.723.1 6.3k
 - ✓ 지연: 110ms~175ms
 - ✓ MOS값: 3.2 ~ 3.8
- G.729a



<그림 8> 네트워크 환경을 고려한 VoIP 통화 품질 측정 방법

- ✓ 지연: 70ms~132ms
- ✓ MOS값 : 3.6 ~ 4.0

시험 결과값에서 MOS값을 R값으로 치환해 보면 G.723.1 6.3k 코덱은 62~75, 그리고 G.729a 코덱은 70~80 사이가 되므로, 각 코덱의 결과값은 이론적인 값과 최대 13과 10 정도의 차이를 보여주고 있다. 이러한 차이는 장비에서 신호 이득비, 지터 버퍼 특성, 코덱 성능 차이, 그리고 프로세싱 지연 등에 기인한 것이다. 그리고, 시험에서 사용 가능한 FXS 포트도 동시에 여러 개의 통화를 시도하므로 장비 프로세스의 성능에 따라 그 결과가 다르게 나타난다. E-모델에서는 장비에서 발생할 수 있는 손실 중 코덱과 관련한 부분만을 반영하므로 반드시 VoIP 장비에 대한 성능 검증을 통해 추가적으로 발생할 수 있는 품질 손실을 최소화해야 한다.

6.2 E-모델 통화 품질 평가

VoIP 게이트웨이에서 발생하는 품질 손실 요소를 통화 품질 측정을 통해 파악하였다면, 다음은 E-모델에서 다루고 있는 네트워크 환경 요소 중 지연, 지터 및 패킷 손실을 측정하는 방법에 대해 설명한다.

NetIQ사의 Chariot VoIP 프로그램을 이용하면 G.711, G.729 또는 G.723.1 음성 코덱에 각각 알맞은

UDP(User Datagram Protocol)/RTP(Real Time Protocol) 데이터그램을 전송하여 네트워크에서 발생하는 지연, 지터, 패킷 손실과 같은 품질 요소를 측정할 수 있다. 그림 8에서는 Chariot EP(Endpoint)간의 경로로 음성 패킷을 보내어 품질 요소를 측정하는 간단한 예를 보여주고 있다. Chariot은 R값 도출을 위해 *Id*, *le,eff*의 경우, 네트워크에서 측정된 결과값과 측정 시 설정한 코덱을 적용하고, 나머지 변수인 *Ro*, *Is*는 ITU-T G.107[8]의 기본값을 적용하여 R값 및 MOS값으로 그 결과를 알려준다. 반면, Agilent사의 VQT 장비는 PAMS 알고리즘 등을 이용하여 네트워크의 품질을 반영한 VoIP 게이트웨이의 실질적인 통화 품질을 측정하는데 효과적이지만, 게이트웨이의 품질 손실과 네트워크의 품질 손실을 구별할 수는 없다. 이러한 2가지 방법으로 실제 통화 품질 측정을 하면, Chariot에 의해 도출된 MOS값과 VQT에 의해 구해진 MOS값에는 차이가 있다. 이 차이는 Chariot에서 사용하는 E-모델은 VoIP 장비에서 압축 코덱 사용에 의한 지연 및 잡음 손실만을 고려하고 있지만, VQT는 게이트웨이 내에서 발생하는 모든 손실 요인을 반영하여 통화 품질을 측정하기 때문이다. 따라서 이러한 경우, 반드시 앞에서 설명한 VoIP 게이트웨이의 자체 통화 품질 측정을 통해 장비에서 코덱에 의한 품질 손실 외에 추가적으로 발생하는 다른 손실 요소들을 줄여나가야 한다.

또한, VoIP 게이트웨이의 성능은 시간에 따라 거의 변함이 없으나 네트워크 성능의 경우에는 시간에 따라 변화가 심한 특성을 가지고 있으며, 이는 사람들의 인터넷 접속 형태에 따라 달라진다. 따라서 올바른 VoIP 망 품질 측정을 위해서는 적용할 코덱 종류, 지터 버퍼, 음성 패킷 길이, 동시 통화량을 결정한 후 이 시나리오를 바탕으로 최소 24시간에 걸쳐 지속적으로 품질을 측정하여야 한다.

표 8은 그림 8과 같이 Chariot EP를 이용하여 주요 인터넷 전화 사업자의 E-모델 R값을 측정된 결과이다. 측정 조건과 방법은 다음과 같다.

- 측정 시간: 48시간 연속 측정
- 측정 주기: 2분 30초
- 측정 보고 주기: 3초
- 데이터그램: 30ms
- 압축 코덱: G723.1 6.3K
- Silence Suppression: On (50% Active)
- 지터 버퍼: 60ms (2개 데이터그램)

패킷화(packetization) 지연 및 패킷의 전송 지연 등을 포함하는 프로세싱 지연 5ms, 그리고 지터 버퍼 60ms로 최소 128ms의 지연이 필요하다. E-모델 분석을 통해 알아본 바와 같이 R값 70 이상을 만족하려면 종단간 지연은 최소한 200ms 이하여야 한다. 측정 결과, 종단간 지연은 대부분의 사업자가 네트워크에서의 지연이 10~30ms 이하로 발생하여 대략 150ms이하를 만족한다. 지터는 바로 이전 패킷과의 평균적인 지터를 나타내는 RFC1889 지터값과 최대 지터값으로 구분하여 측정할 수 있다. 측정 결과, 최대 지터값은 B사를 제외하고는 모두 60ms를 초과하였으며, 이 경우 VoIP 단말에서 필연적으로 패킷 손실이 발생하며, 패킷 손실은 R값에 반영된다. 만약 지터값이 기준치를 초과하면, 사업자는 네트워크의 QoS 관리를 통해 지터에 의한 패킷 손실을 최소화하여 R값을 향상시켜야 한다. 패킷 손실율은 네트워크 구간에서의 평균 패킷 손실율을 나타내며, A사의 경우 약 0.2%의 패킷 손실로 인해 대략 1 정도의 R값 저하를 나타낸다. 반면, 나머지 사업자의 패킷 손실율은 통화 품질에 큰 영향을 미치지 않는다.

〈표 8〉 E-모델 통화 품질 측정 결과

	A사	B사	C사	D사
MOS값	3.73	3.79	3.78	3.80
R값	73.1	74.0	74.0	74.4
종단지연(ms)	151	151	149	136
최대지터값(ms)	153	19	94	80
RFC1889 지터(ms)	0.33	0.07	2.36	0.73
패킷손실율(%)	0.24	0.01	0.04	0.09

상기 표에서 각 사업자의 결과값은 무작위로 추출된 최소 10개 사이트에서 측정된 값의 평균값을 나타낸다. G.723.1 6.3K 코덱을 사용할 경우, 최대 R값과 MOS값은 각각 75.1과 3.83이다. 측정 결과, 평균 R값과 MOS값은 각각 'medium' 품질 70과 3.6 이상을 만족하나, 인터넷 트래픽의 특성상 통화 구간 및 시간에 따른 품질 변화가 발생한다. 종단간 지연은 G.723.1 6.3K 코덱을 사용할 경우, 코덱 지연 67.5ms,

7. 결 어

인터넷 전화의 통화 품질은 지속적으로 해결하여야 할 과제중의 하나이다. 그 동안 이를 해결하기 위한 방법은 음성 기반의 VoIP 통화 품질 측정 알고리즘을 사용하여 VoIP 단말의 성능을 파악하거나, VoIP 망의 최적화를 통해 부분적이고 단계적으로 이루어졌다. 이로써 아무리 좋은 성능의 시스템일지라도 양질의 네트워크의 품질이 제공되지 않으면 원하는 통화 품질을 보장할 수 없다. 따라서 인터넷 전화 사업자나 망 운용자는 VoIP 기술의 인터넷 전화 서비스에서 일정 수준의 통화 품질을 제공하려면 VoIP 단말의 성능은 물론, 네트워크의 품질을 일정 수준으로 유지할 수 있어야 한다. 이를 위해서 사업자나 망 운용자는 VoIP 단말 및 네트워크의 품질 손실 요소를 체계적으로 정의하고 분석하여, 손실 요인을 최소화할 수 있어야 한다.

본 논문에서는 음성 기반 통화 품질 측정 알고리즘의 단점을 보완하고, 음성 전송 경로에 대한 네트워크의 품질 요소를 반영할 수 있는 E-모델의 다양한 분석을 통해 VoIP planning을 하기 위한 가이드라인을 제시하고, 실제적으로 VoIP 단말 및 네트워크의 품질을 측정하기 위한 방법 및 그 결과를 소개하였다.

본 논문의 결과는 실제 인터넷 전화 서비스를 제공하는 사업자나, VoIP 망 운용자에게 있어서, VoIP 시스템 및 망을 구축하거나 운용하는데 필요한 planing 정보나 가이드라인으로 적극 활용될 것으로 예상된다.

참고 문헌

- [1] DaeHo Kim, SeongGon Choi, Jun Kyun Choi, "Performance Evaluation for QoS Guaranteed VoIP service with OPNET and Related QoS Technologies", ICACT'2001, Feb., 2001.
- [2] ITU-T P.800, 'Methods for subjective determination of transmission quality', Aug., 1996.
- [3] ITU-T P.861, 'Objective quality measurement of telephone-band(300-3400Hz) speech codecs', Feb., 1998.
- [4] Stefan Pracht, Dennis Hardman, 'Voice Quality in Converging Telephony and IP Networks', Agilent Technologies, 2000.
- [5] ITU-T 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.
- [6] 이인섭, "음성 서비스 품질 표준화 동향", TTA IT Standard Weekly, 2003-40호, 2003.10.
- [7] ITU-T G.109, 'Definition of categories of speech transmission quality', Sep., 1999.
- [8] ITU-T G.107, 'The E-model, a computational model for use in transmission planning', July, 2002.
- [9] ITU-T G.108, 'Application of the E-model: A planning guide', Sep., 1999.
- [10] TIA/EIA/TSB116, 'IP Telephony Equipment Voice quality recommendations for IP telephony', Mar., 2001.
- [11] ITU-T G.113 Appendix I, 'Provisional planning values for the equipment impairment factor Ie and packet-loss robustness Bpl', May, 2002.
- [12] ITU-T G.114, 'One-way transmission time', May, 2000.

◎ 저자 소개 ◎



배 성 용 (Seong Yong Bae)

1992년 충남대학교 전산학과 졸업(학사)
1994년 충남대학교 대학원 전산학과 졸업(석사)
2002년 경기대학교 대학원 전산학과 수료(박사)
1994년 2월~2001년 11월 한국전자통신연구원
2001년 12월~현재 한국정보통신기술협회 시험인증연구소 선임연구원
관심분야 : VoIP, BcN, Open API, Workflow, BPM
E-mail : sybae@tta.or.kr



김 광 훈 (Kwang Hoon Kim)

1984년 경기대학교 전산학과 졸업(학사)
1986년 중앙대학교 대학원 전산학과 졸업(석사)
1994년 University of Colorado at Boulder, Computer Science, MS
1998년 University of Colorado at Boulder, Computer Science, Ph.D
1986년 2월~1991년 8월 한국전자통신연구원
1993년 5월~1994년 8월 American Educational Products, Inc., Professional DB Consultant
1994년 9월~1995년 8월 Colorado Advanced Software Institute, Research Assistant
1995년 9월~1997년 2월 Aztek Engineering, Inc., Software Engineer
1998년 3월~현재 경기대학교 정보과학부 조교수, 부교수
관심분야 : Workflow, BPM, Groupware, Database, VoIP, BcN
E-mail : kwang@kyonggi.ac.kr