Mobility Impact on VoIP Quality over Radio Access Network

by

Seyong Park

A Dissertation Submitted

to

The Graduate School in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in

Electrical Engineering

Stony Brook University

May 2009

Stony Brook University

The Graduate School

Seyong Park

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree,

hereby recommend acceptance of this dissertation.

Sangjin Hong, Dissertation Advisor

Associate Professor, Department of Electrical & Computer Engineering

Hang-Sheng Tuan, Chairperson of Defense

Professor, Department of Electrical & Computer Engineering

Milutin Stanacevic

Assistant Professor, Department of Electrical & Computer Engineering

Jungwoo Lee

Associate Professor, School of Electrical Engineering in Seoul National University

This dissertation is accepted by the Graduate School

Lawrence Martin

Dean of the Graduate School

Abstract of the Dissertation

Mobility Impact on VoIP Quality over Radio Access network

by

Seyong Park

Doctor of Philosophy

in

Electrical Engineering

Stony Brook University

2009

Voice over IP (VoIP) has been deployed over the wireline networks for the last decade. It achieved remarkable success of the adaptation as the replacement of circuit switched legacy voice service. It also enabled the service providers to converge data and voice services over a single broadband network. This evolutionary activity has been also initiated in the wireless network as the emergence of broadband wireless network. VoIP has been initially applied to the Wireless Local Area Network (WLAN) over 802.11a/b/g interface. It also has been applied to the Wireless Wide Area Network (WWAN) since 3G networks were deployed. In addition, convergence of wireline and wireless network would accelerate the adoption of VoIP application over the wireless network.

Within several key technical challenges such as transport efficiency and QoS, we selected Voice quality assurance as the dissertation topic since it becomes more critical issue on the mobility network, especially over the handoff period where the additional packet transmission delay, jitter and loss occurs. We presents the enhanced packet based VoIP quality monitoring and measurement method as well as a validation method on a cross-check with the speech quality measurement.

We also presents an enhanced method for the VoIP quality measurement from the existing E-model, and applied the method to the EVDO Radio Access Network (RAN) as an experiment. The EVDO RAN has been deployed as 3G network in US while it typically supports up to approx. 2Mpbs forward link and up to 153 kbps over reverse link to offer data service an the evolution of CDMA2000.

However, EVDO also offers the technical challenge to overcome the performance fluctuation upon the mobility. We address the technical issues on the existing E-model when it is applied to the EVDO, and propose an enhanced method expanding the existing E-model well suited to the VoIP over the EVDO. We also use the Anique+ for the reference to verify that the proposed solution produces the superior accuracy to the existing E-model based measurement.

Once we analyzed the correlations between E-model and Anique+ as a reference, we applied to the mobility impact on EVDO network by measuring the packet loss, jitter and data throughput impact by the EVDO forward link handoff. We also investigate the unvoiced frame effect on the packet loss by detecting the VoIP packet type. Unvoiced frame type is useful to compromise VoIP quality impact by the burst packet loss during the handoff since it takes more than 50% of the VoIP frames.

We investigated the mobility impact on the VoIP service caused by the handoff, and define a validation model with leverage of both E-Model and Anique+. Our experimental analysis over the EVDO network proved the mobility impact on the VoIP quality due to handoff. We achived the optimal size of receiver buffer and data throughput that are recommended for VoIP application configuration to minimize the mobility impact.

All glory to my Lord, Jesus Christ

To my wife, Sohee Kim,

my mother, Soon-Gu Yeo and my father Hee-Hyong Park

Contents

List of Figures

List of Tables

Acronyms

Acronym	Definition
AG	Access Gateway
ANSI	American National Standards Institute
BTS	Base Transceiver Station
CDMA2000	Code Division Multiple Access 2000
CINR	Carrier to Interference-plus-Noise Ratio
DRC	Data Rate Control
EVDO	Enhanced Voice and Data Only
EVRC	Enhanced Variable Rate Codec
FER	Frame Error Rate
GGSN	Gateway GPRS Support Node
GPRS	General packet radio service
GSM	Global System for Mobile communications
IP	Interner Protocol
ITU-T	International Telecommunicaiton Unit
MOS	Mean Opinion Score MS Mobile Station
PDSN	Packet Data Serving Node PER Packet Error Rate
PESQ	Perceptual Evaluation of Speech Quality
PPP	Point-to-Point
QoS	Quality of Service RNC Radio Network Controller
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol
UDP	User Datagram Protocol
VAD	Voice Activation Detection
VoIP	Voice over Internet Protocol
WLAN	Wireless Local Access Network

Acknowledgments

"When pride comes, then comes dishonor, But with the humble is wisdom." Proverb 11:2

I am grateful to my wife, Sohee who's supporting me with her prayer, sacrifice on her professional career to raise kids as well as husband. Her love and support is priceless and appreciated in God's grace. I also thank my son John Jaeyeon and Catherine Soohyun for their cheerful welcoming on coming home. I appreciate my mother for the support and love on me for her entire life.

God is always good. 10 years ago, God handed me over Professor Sangjin Hong who became not only a research advisor, but a dedicated mentor keeping me encouraged to complete the work. As my advisor, he showed me the role model as the advisor in academy with his patience and kindness. His humility and hunger for the knowledge engraved my heart as my attitude for my future career. My sincere gratitude for him, and may God bless him and his family.

I thank Doug Kim and Do-Seok Kim for their collaborations, and other friends and colleagues of Alcatel-Lucent including Peter Kim, Dr. Chang Kim and Dr. Jung Ah Lee for their their comments and collaborations on the dissertation.

My special gratitude to Professor Dae-Joon Hwang in SunKyunKwan University in Korea who motivated and encouraged me at the college to go out of the small cage and learn the world. I also have many thanks to Mr. Andrew Chon for mentoring on

my professional career as well as marriage.

I also have my sincere gratitude for my families in Korea and the ones who've been praying for me.

Chapter 1

Introduction

1.1 Voice over IP over the Wireless Network

VOICE OVER IP (VOIP) has been deployed in wireless network domain mainly over data network. As the early stage, however, the VoIP service has been more prevalent in the wireline network since it provided the consistent transport network performance from the core network to the access network. However, VoIP still faces many technical challenges in the mobile network due to the unpredictable transport network performance as follows:

- Voice quality assurance:As a nature of the packet switched network, the voice quality is not consistent due to the packet loss, delay and jitter while the circuit switched network provides consistent voice quality. VoIP service should meet the equivalent quality of the voice service over the circuit network.

- Transport efficiency enhancement: Since VoIP traffic is transmitted over the packet network, it requires the packet header containing the source and destination address as well as other related control messages. Some VoIP packet without header compression typically carries from 50% to more than 200% packet overhead. For example, if we transmit 20-byte EVRC voice frame over 40-byte UDP/IP/PPP packet format, its packet overhead is 200%. Header compression is a must-have function to carry VoIP traffic over wireless network. However, interoperability within the network systems are still remaining issues.

- Differentiation of the services and security: Since broadband network carries mixture of different service traffic, VoIP traffic should be prioritized at the internetworking systems in order to minimize the performance impact. However, End-to-End QoS architecture has many technical issues from scalability perspective.

In order to assess the voice quality over the packet switched network, ITU-T initiated the voice quality assessment standard, and defined several models supporting various stages including network planning to network maintenance. As a subjective test, Mean Opinion Score (MOS) has been used for the traditional source of the voice quality test. MOS is a standard way to evaluate the voice quality conducted by the subject group testing in the strictly controlled environment. Even though MOS is a popular method to determine the voice quality for the device such as codec, due to the restriction on the testing environment, it may not be feasible to measure the voice quality over real service network systems. Objective tests have been proposed to overcome this constraint, and conduct the voice quality test with relatively less cost.

The objective test is to measure the voice and audio quality employing measurement test systems with the knowledge of the input data.

E-model based model also became another method to monitor and measure the voice quality with account of packet network's performance monitoring. It monitors the packet network by measuring such as packet delay, loss and jitter that are part of key performance index, and provides the computed assessed voice quality over the network.

This document presents a proposal of an E-model based voice quality measurement model employing the mobility impact. Existing E-model was designed on a wireline network. In addition, this document also use Anique+ [16] as the reference to validate the measurement accuracy of the proposed method. Anique+ is a non-intrusive model recently standardized in ANSI. Auditory non-intrusive quality estimation plus (Anique+) [16] is a perceptual model to estimate the speech quality without reference information while it simulates the functional role of human auditory system.

1.2 Overview of Related Researches

1.2.1 Packet Level VoIP quality Measurement

E-Model is a computational model used for the transport network plan. Given various input parameters such as end-to-end delay, echoes, side-tones, loudness, the E-model calculates a scalar quality rating value, R, which corresponds the overall conversational quality. While ITU-T G.107 [4] presents E-model overall parameters used for the R factor computation, companion recommendations including G.113, G.108 and G.109 [5] [6] [7] provide further guidance about specific codec related speech impairments, E-model applications and definitions of transmission quality. Since E-model is defined as framework for the network planning method, there is limited accuracy on the measurement of the voice quality. This leads many proposals on enhancement and extension of the current E-model. Ding and Goubran [11] investigated the effect of packet loss and delay jitter in E-model and proposed the extension of E-model with a new formula. Sun and Ifeachor [10] presented a combined method of E-model and PESQ in order to enhance the accuracy of the voice quality measurement. In [23], Sun and Ifeachor also investigated the impact of the packet loss location on the voice quality. Their experiment showed that the unvoiced speech segment showed less impact on the perceived speech quality. They also analyzed that packet loss of the beginning of the voiced segments have more severe impact than the end of the voiced segments. Even though the location of packet loss affects the voice quality, it is very difficult to distinguish the VoIP packet in terms of voiced versus unvoiced segment. Typical VoIP network carries the encoded voice segments over RTP [18] frames. RTP header [22] contains the codec type information, but it does not provide whether or not it carries voiced segment. However, RTP payload [19] [21] provides an information indicating what kind of information is carried over the RTP frame. RFC-4749 [21]defines the RTP payload structure for the ITU-T G.729.1 codec data. Its Frame Type field specifies the frame type with frame size. If the RTP frame carries non-speech data, its decimal value is set to 15. Using this information, we would analyze how the frame type affects the VoIP quality measurement. We also use Gole and Rosenblush's [9] expand E-model in our work as reference in this paper.

1.3 Main Contribution

The objective of this study is to provide an enhanced VoIP measurement method at the packet level monitoring and measurement. Speech quality measurement should be the optimal course to measure the VoIP quality. However

Three main contributions are summarized in the following:

- Identification of the VoIP quality correlation between packet and speech domains
- Expanded E-Model based VoIP quality measurement with the cross-layer packet inspection
- Employment of the mobility factor in the packet level VoIP quality measurement and monitoring

1.4 Organization of the Proposal

The remainder of this report consists of three sections with the following summary:

- Study of the correlations of the VoIP quality measurement between Packet level based and Speech quality: Chapter 2 presents the study to investigate the discrepancy of existing E-model based VoIP quality measurement method from the emerged speech quality measurement method. Anique+ that became an ANSI standard was used to measure the voice quality over the VoIP service. From this study, I proposed an expanded E-Model based VoIP quality measurement method that would enhance the correlation to the speech quality measurement.

- Expanded E-model for VoIP quality measurement with the employment of packet information: Chapter 3 presents the detailed study on the application layer packet structure that can be exploited to figure out the packet information type that can be either voice or non-voice frame. This chapter presents the several techniques to compose the traffic type upon the codec type, and estimation of the VoIP quality measurement accuracy by employing the traffic type.

- Mobility impact on the expanded E-model: Chapter 4 presents the impact of the handoff on the VoIP quality. Mobility is the key element of wireless network. Mobile station performs the handoff when it approaches the edge of the serving node. As part of the handoff process, there is a short period that transmission of the VoIP packets are intermitted caused by the packet discard, call control transfer as example. This intermission has not been considered as packet loss while it is heavily affected to the speech quality. This chapter investigate the handoff behavior of several wireless technologies and present an enhanced method to assess the mobility impact over the expanded E-model.

In Section 5, I would present the remaining issues and future works that will be pursued as part of the thesis.

Chapter 2

Comparative Analysis on VoIP Quality Measurements

2.1 Introduction

VOICE OVER IP (VOIP) has been deployed in wireless network domain mainly over data network. As the VoIP becomes more prevalent in the main stream of the telephony service, quality measurement became a crucial issue. ITU-T initiated the voice quality measurement standard, and defined several models supporting various stages including network planning to network maintenance. As a subjective test, Mean Opinion Score (MOS) has been used for the traditional source of the voice quality test. The MOS is a standard way to evaluate the voice quality conducted by the subject group testing in the strictly controlled environment. Even though the MOS is the most effective method to determine the voice quality for the device such as codec, it may not be applicable to measure the voice quality over real service network systems due to the restriction on the testing environment. Objective tests have been proposed and standardized for measuring the voice and audio quality by a measurement test systems. According to the knowledge of the input data, the objective tests can be described as intrusive models or non-intrusive models. A recent paper by Antony W. Rix et. al. [1] described details about these models with exemplary standards.

E-Model is a computational model used for the transport network plan. Given various input parameters such as end-to-end delay, echoes, side-tones, loudness, the E-model calculates a scalar quality rating value, R, which corresponds the overall conversational quality. While ITU-T G.107 [4] presents E-model overall parameters used for the R factor computation, companion recommendations including G.113, G.108 and G.109 [5] [6] [7] provide further guidance about specific codec related speech impairments, E-model applications and definitions of transmission quality. Since E-model is defined as framework for the network planning method, there is limited accuracy on the measurement of the voice quality. This leads many proposals on enhancement and extension of the current E-model. Ding and Goubran [11] investigated the effect of packet loss and delay jitter in E-model and proposed the extension of E-model with a new formula. Sun and Ifeachor [10] presented a combined method of E-model and PESQ in order to enhance the accuracy of the voice quality measurement. Gole and Rosenblush [9] expand the use of E-model to the VoIP performance monitoring tool by exploiting transport-level metrics for the purpose of monitoring conversational voice quality. This method was used for several papers as reference model to determine the VoIP quality. Their method is used in our work as reference for the comparative analysis with ANIQUE+ model.

This paper presents the measurement of the VoIP quality in the packet and the speech domains, and provides the analysis illustrating the correlation between packet level impairments and speech level distortion. In this paper, we use ANIQUE+ [16] that is a non-intrusive model recently standardized in ANSI. Auditory non-intrusive quality estimation plus (ANIQUE+) is a perceptual model to estimate the speech quality without reference information while it simulates the functional role of human auditory system.

2.2 RELATED WORKS

2.2.1 E-model based VoIP quality measurements

Many researches have used metrics based on the E-model in order to predict the VoIP quality. As defined in ITU-T G.107 [4], R value is derived as follows:

$$R = R_0 - I_s - I_d - I_e + A, \tag{2.1}$$

where R_0 is the original voice quality including noise source; I_s is a combined impairment related to the voice signal; I_d represents the impairments related to the delay of the voice signals including echoes and one-way transmission delay; I_e is equipment related impairments such as random packet-loss that is typically caused by the low-bit rate codec. G.113 [7] provides codec specific I_e values. The advantage factor A can be applied if there are other advantages of access to the user. G.107 [4] provides several default values upon the access transport types as recommended upper limit for the network planning engineer. Cole and Rosenblush [9] extended the use of E-model from the network planning tool to network monitoring tool with simplified R value computation within the transport level measurement including delay, jitter and packet loss as

$$R \sim 94.2 - 0.024d + 0.11(d - 177.3)H(d - 177.3) - \gamma_1 - \gamma_2 ln(1 + \gamma_3 e) \tag{2.2}$$

where γ_1, γ_2, γ_3 are specified in [9] upon the codec type and H(x) is Heavy-side function, for all x $\geq$ 0, H(x) yields 1 and for all x<0, it yields 0. They used the default value of I_s defined in [4] and reduced the expression of the R factor.

As defined in G.107, R factor can be translated to the estimated MOS_E as follows:

$$UserR = 94 - I_d - I_e^*, \tag{2.3}$$

$$I_e^n = I_e + (k(I_d - I_e))e^{-y/t_3} \tag{2.4}$$

Where K is constant with a nominal value of 0.7, t_3 is a time constant (typically 30 - 60 seconds), and y is the time delay since the last burst.

As defined in G.107, R factor can be translated to the estimated MOS (i.e. MOS_{CQE}) as follows:

$$
\begin{aligned}
MOS_{CQE} &= 1, R < 0, \\
&= 1 + 0.035R + R(R-60)(100-R)7 \cdot 10^{-6}, 0 < R \leq 100, \\
&= 4.5, R > 100,
\end{aligned}
$$

2.2.2 Anique+ based VoIP quality measurement

Anique+ consists of three models that are Frame distortion D_F, Mute impact D_M and Non-speech impact D_N models. Anique+ computes the overall objective distortion D_X by sum of the three distortions as follows:

$$D_X = D_F + D_M + D_N \tag{2.5}$$

D_X is translated to the estimated MOS to yield objective speech quality Q_X:

$$Q_x = -3.5min(D_x, 1) + 4.5 \tag{2.6}$$

The D_F is modeled from the distortions of speech and audible background noise depicted as follows:

$$D_F = D_S + D_M \tag{2.7}$$

D_S is the distortion in speech that is obtained by accumulating frame distortions for active speech over time and normalized by the total number of active speech frame

T_s:

$$D_S = \frac{1}{T_s} \sum_{m \in S} X(m) \tag{2.8}$$

where $X(m)$ is the output of frame distortion model ranging from 0 to 1 at the $m_t h$ frame. D_B is the distortion in the audible background noise and is estimated as

$$D_B = \frac{1}{T_B} \sum_{m \in \beta} \{\alpha_F (P_{env}(m) - P_{th} + \beta_F)\} X(m) \tag{2.9}$$

where $P_{env}(m)$ is the DC-value of modulation power spectrum at the m_{th} frame, P_{th} is the threshold for audible background noise. T_B is the number of frames for background noise, α_F and β_F are weighting factors. X(m) is the output of the frame distortion model that Anique+ [16] uses. D_M is a mute impact model to detect the unnatural mutes in speech signaling and estimate the impact on the perceived quality. Unnatural mutes become the most common distortion caused by the poor voice activity detection (VAD) in digital transmission network such as burst packet drop. Unnatural start and stop are detected on the mute impact model. D_M is modeled as follows:

$$D_M(t) = \sum_{i=0}^{k} exp\left[\frac{1(t - t_i)}{\tau}\right] \cdot u(t - t_i) \tag{2.10}$$

Where u(x) is a unit step function, which is 1 for $x \geq 0$ and 0 for $x < 0$, and v_i is the instantaneous distortion of the i-th mute at time t_i. For each mute, perceived distortion is raised by the amount of v_i at the end of the mute event and decays over time with the time constant τ. The instantaneous distortion of the i-th mute is estimated by

$$V_i = p_1 log_2(L_i) + P_2 \tag{2.11}$$

where L_i is the length of $i_t h$ mute and p_1 and p_2 are constants. D_N is a non-speech model detecting and estimating the impact of the annoying non-speech activities such as a distortion of bit information in the voice packets. Speech decoder may generate annoying non-speech sounds upon receipt of the distorted non-speech activities. D_N is the impact of non-speech estimated to be proportional to the accumulated frame log-power (p_{acc}) in the non-speech activity region as

$$D_N = q_1 P_{acc} + q_2 \tag{2.12}$$

where q_1 and q_2 are constants. With these distortion models, Anique+ should perform the training procedure so that the model can estimate the distortion level on speech signals.

2.3 COMPARATIVE ANALYSIS OF VOIP QUALITY

Even though the E-Model and the ANIQUE+ estimate voice quality in different domains, we expect that a similarity might exist to some extent in the attributes of these methods as packet loss and jitter impairment affect the resulting speech waveform as well as the R factor of E-Model. Specifically in speech waveform domain, jitter impairments can be realized as unnatural speech, which can be reflected as the frame distortion, D_F, in the ANIQUE+. In addition, bursty packet loss can result in unnatural mutes in speech waveforms which in turn can be explained by the mute distortion, D_M. We investigate the correlations between the E-model and the ANIQUE+ in the following aspects:

- Jitter and frame distortion
- Packet loss vs. mute impact

We use the E-models that Cole and Rosenblush [9] specified as primary references in order to perform the comparative analysis with the ANIQUE+. We further extend their models with employment of packet error factors.

2.3.1 Jitter and speech distortion

Excessive packet transmission delay may affect the conversational voice quality caused by the loss of the packet synchronization in the codec. Packet transmission jitter is another cause of the irregular voice quality. In Anique+, speech distortion model is used to detect similar occasion to the transmission jitter by monitoring active speech signal as well as audible background noises.

Packet transmission jitter may occur by packet buffering at the transmission node or communication host such as I/O buffer and application memory. In [8], Packet transmission delay due to jitter buffering was added to the one-way delay. Further, [9] presents the probability of packet loss due to the jitter as follows:

$$e_{jitter} \sim P\{l > bg\} \tag{2.13}$$

where l is jitter length, bg is the jitter buffer size. The impact of jitter on voice quality can be reflected in different manner by the far-end listeners depending on different jitter buffer management algorithms and packet sizes. In the decoding of speech packets, the impact of jitter can be implemented as deletion, insertion, and contraction of speech packets. And the resulting speech waveform may or may not suffer from voice quality degradation depending on the location and degree of jitter distortions. The e_{jitter} will be compared with the D_F specified in (7). Using Pearson product-moment correlation coefficient [17], Correlation function between e_{jitter} and D_F is defined as follows:

$$C_F = \frac{n\sum e_{jitter_i} D_{F_i} - \sum e_{jitter_i} \sum D_{F_i}}{\sqrt{n\sum e_{jitter_i}^2 - (\sum e_{jitter_i})^2}\sqrt{n\sum D_{F_i}^2 - (\sum D_{F_i})^2}} \tag{2.14}$$

where C_F is correlation function between 0 and 1, and are expected values. This study also imposes the difference in distortion source between two methods. Coding distortions as well as distortions caused by jitter are blended into D_F, whereas e_{jitter} considers jitter distortions only. However, our test assumes that there is no coding distortion.

2.3.2 Packet loss vs. mute impact

Over the VoIP network, consecutive packet losses become the most common cause of unnatural mute from voice quality perspective. Random packet loss effect was measured and presented in [7], for example on G.729a codec, with 49 as largest I_e value for the case of an average packet loss of 16%. The measured I_{ef} sequentially increases upon the increase of the average packet loss. Cole and Rosenblush [9] defined the packet loss effect I_{ef} as follows:

$$I_{ef} \sim \gamma_1 + \gamma_2 ln(l + \gamma_3 e) \tag{2.15}$$

where e is the total loss probability and the γ_i's are fitting parameters and variable upon the codec type. They also analyzed difference between random and burst packet loss. From the measured data using G.711 codec, they observed that the I_{ef} will significantly increase when the burst packet loss is greater than 4%. In our experiment using G.711 codec of the following section, we use their model employing the random and burst error in order to compare with mute impact on Anique+, however, we refine the e to e+ as defined below:

$$I_{ef} \sim 30ln(1 + 15e^{+})H(0.04 - e^{+}) + 19ln(1 + 70e^{+})H(e^{+} - 0.04) \tag{2.16}$$

$$e^{+} = \frac{1}{N}\left(\sum_{N} C_i\right) \cdot 0.01 \tag{2.17}$$

where e+ is average packet loss ratio over the testing period, and $H(x)$ is Heavyside function, for all $x \geq 0$, $H(x)$ yields 1 and for all $x < 0$, it yields 0. $H(x)$. We also apply the I_{ef} to the receny effect as depicted in (2-4). Correlation function between I_{ef} and D_M is defined as follows:

$$C_M = \frac{n \sum I_{ef_i} D_{M_i} - \sum I_{ef_i} \sum D_{M_i}}{\sqrt{n \sum I_{ef_i}^2 - (\sum I_{ef_i})^2} \sqrt{n \sum D_{M_i}^2 - (\sum D_{M_i})^2}} \tag{2.18}$$

where C_M is between 0 and 1, and are expected values.

2.4 SIMULATION EXPERIMENT

From the comparative analysis, we defined correlation functions C_F and C_M. The C_F is correlation between packet delay/jitter and speech distortion, and the C_M is correlation between packet loss and mute distortion. Simulation experiment has been done on the packet loss impairments as follows:

- Measurement of packet loss (e.g. random loss and burst loss)
- Measurement of MOS from E-model and ANIQUE+ with D_F and D_M

As illustrated in Figure 2-1, we performed the simulation to compare the voice qualities upon the impairments on voice frames from speech and packet level measurement. As the analysis tools, we used the ANIQUE+ for speech quality measurement and G.107 E-Model for the packet level voice quality measurement.

We also used a free software [18] [19]to generate the impairments on the sample voice frames. Mute impairments were implemented to 10 speech files from 5 female and 5 male talkers. The length of each speech file is about 8 seconds long and

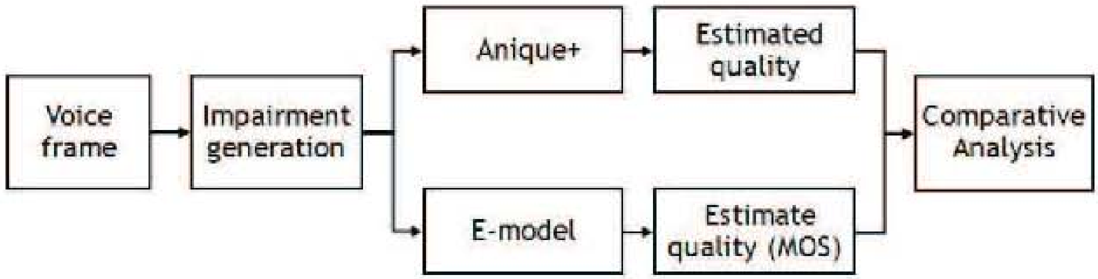

Figure 2-1: VoIP quality test configuration

contains a sentence pair. For simplicity, we assumed that the payload size is 60 msec and replaced the impacted parts of speech with silence in waveform domain.

From the simulation tests, as depicted in Table 1 and Table 2, we obtained average numbers of the measurement attributes as the quality metrics including D_F, D_M, MOS_A, R and MOS_E, where D_F, D_M and MOS_A is given by the ANIQUE+ and R and MOS_E are given by the E-model based computation.

During the simulation, we could observe that both the E-model and the ANIQUE+ produced erroneous results upon the packet loss. However, normalized results from 10 speech files seem to be reasonable to see the similarity between the ANIQUE+ and the E-model based voice quality measurement.

2.4.1 Analysis on Packet loss - burst loss

On the burst packet loss simulation, we generated a single burst packet loss within the voice frame, and measured the voice quality impacts. Since we used clean speech material that doesn't contain audible background noise, it is not necessary to consider D_B in equation (2.9). We used the E-model with burst mode derived from the

	0 ms	60ms	120ms	180ms	240ms
D_F	0.252	0.261	0.270	0.291	0.273
D_M	0.008	0.078	0.101	0.123	0.130
MOS_A	3.54	3.30	3.15	3.00	3.07
R	94.2	86.0	80.3	76.0	72.4
MOS_E	4.5	4.2	4.0	3.9	3.7

Figure 2-2: Voice quality measurements upon the burst packet loss

equation (2.16).

As shown in Figure 2-2, the burst packet loss increased the speech distortion as well as mute impact. We also observed high correlations between the voice quality degradation and the mute impact from the ANIQUE+ test. As presented in Figure 2-2, both ANIQUE+ and E-model provide lower voice quality scores as we have more impairments. Also, MOS_A and MOS_E are highly correlated each other. This provides the validation of the E-model method to measure the voice quality on the real network.

It has to be noted that there's a gap between MOS_A and MOS_E scores, and MOS_A doesn't reach its maximum value of 4.5 when there's no impairment. This is due to the variability of talker and speech utterances for which ANIQUE+ has not been validated. This problem is common to all speech waveform-based objective models up to now.

	0 ms	60ms	120ms	180ms	240ms
D_F	0.252	0.248	0.241	0.249	0.254
D_M	0.008	0.051	0.085	0.151	0.197
MOS_s	3.54	3.40	3.34	3.08	2.90
R	94.2	86.0	80.3	76.0	72.4
MOS_E	4.5	4.36	4.29	4.21	4.13

Figure 2-3: Voice quality measurements upon the random packet loss

2.4.2 Analysis on random packet loss

We performed another simulation for the random packet loss. We used the E-model with random packet loss model derived from equation (2.16). From the test result, we could also observe a similarity in results between the E-model and the ANIQUE+.

As presented in Figure 2-3, the mute effect is a primary source of the voice quality degradation. We also observed similar results to the first simulation test from the MOS scores from ANIQUE+ and E-model method. We also observed that the random packet loss provides more consistent voice quality degradation than the burst packet loss that showed a flip result at 240 msec burst packet lost. It was observed that the speech distortion at 240 msec burst packet loss decreased compared to small packet losses. We left the root cause analysis for the future study.

2.4.3 Comprehensive Analysis

Combined with burst andrandom packet losses, we found that there are strong correlations between ANIQUE+ and E-model based voice quality measurements. As illustrated in Figure 2-4 and depicted in Table 2-5, we found there are over 95%

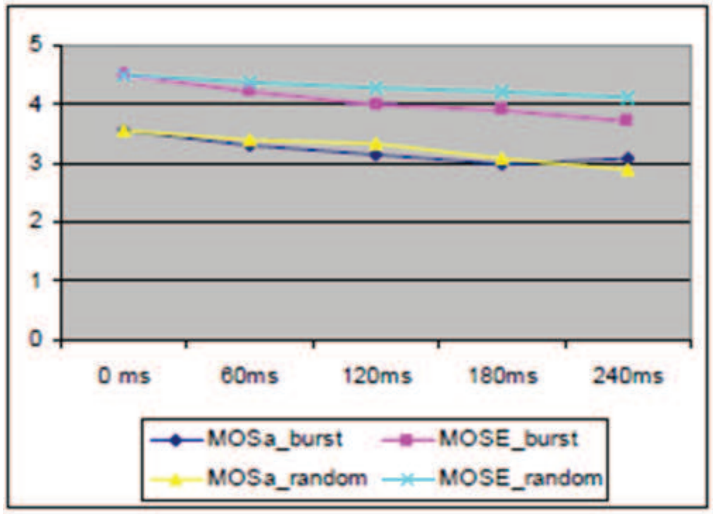

Figure 2-4: MOS of ANIQUE+ and E-model voice quality measurements

	D_M	D_F	MOS
Burst Packet Loss	0.98	0.1	0.96
Random Packet Loss	0.97	0.79	0.95

Figure 2-5: Correlations of D_M and D_F from MOS

correlations in quality measurements between ANIQUE+ and E-model.

From the results of Figure 2-5, we also confirmed that mute impact is the primary source of the quality distortion in the pack loss impairments from the simulation test while the speech distortion was limited to the quality degradation at the random packet loss. Speech distortion might be affected by the number of packet-loss locations where the random packet loss typically incurs.

2.5 Conclusion

We conducted the comparative analysis between the expanded E-model and the ANIQUE+. From the exhaustive simulation tests, we could achieve reasonable results to show the high correlation between the E-model and the ANIQUE+. However, we also realized that there are many challenging areas that we should further investigate including the unvoiced frame effect, jitter modeling and validation. We also want to do further investigation over the live wireless network for the future work.

Chapter 3

VoIP Quality measurement with structured frame information

3.1 Introduction

VOICE OVER IP (VOIP) has been deployed in wireless network domain mainly over data network. VoIP has become a critical application for the wireless network to fulfill the flexible and cost-effective voice service over packet network. One of the technical challenges in VoIP over wireless network is to maintain the voice quality due to the dynamic traffic characteristics. As the VoIP becomes more prevalent in the main stream of the telephony service, quality measurement became a crucial issue. ITU-T initiated the voice quality measurement standard, and defined several models supporting various stages including network planning to network maintenance. As a subjective test, Mean Opinion Score (MOS) has been used for the traditional source of the voice quality test. MOS is a standard way to evaluate the voice quality conducted by the subject group testing in the strictly controlled environment. Even though MOS is the most effective method to determine the voice quality for the device such as codec, due to the restriction on the testing environment, it may not be applicable to measure the voice quality over real service network systems. Objective tests have been proposed and standardized. As a substitute, objective test is to measure the voice and audio quality by a measurement test systems. According to the knowledge of the input data, objective tests can be described as intrusive models or non-intrusive models. A recent paper by Antony W. Rix et. al. [1] described details about these models with exemplary standards. In this paper, we will use Anique+ [16] that is a non-intrusive model recently standardized in ANSI. Auditory non-intrusive quality estimation plus (Anique+) [16] is a perceptual model to estimate the speech quality without reference information while it simulates the functional role of human auditory system.

E-Model is a computational model used for the transport network plan. Given various input parameters such as end-to-end delay, echoes, side-tones, loudness, the E-model calculates a scalar quality rating value, R, which corresponds the overall conversational quality. While ITU-T G.107 [4] presents E-model overall parameters used for the R factor computation, companion recommendations including G.113, G.108 and G.109 [5] [6] [7] provide further guidance about specific codec related speech impairments, E-model applications and definitions of transmission quality. Since E-model is defined as framework for the network planning method, there is limited accuracy on the measurement of the voice quality. This leads many proposals on enhancement and extension of the current E-model. Ding and Goubran [11] investigated the effect of packet loss and delay jitter in E-model and proposed the extension of E-model with a new formula. Sun and Ifeachor [10] presented a combined method of E-model and PESQ in order to enhance the accuracy of the voice quality measurement. In [23], Sun and Ifeachor also investigated the impact of the packet loss location on the voice quality. Their experiment showed that the unvoiced speech segment showed less impact on the perceived speech quality. They also analyzed that packet loss of the beginning of the voiced segments have more severe impact than the end of the voiced segments. Even though the location of packet loss affects the voice quality, it is very difficult to distinguish the VoIP packet in terms of voiced versus unvoiced segment. Typical VoIP network carries the encoded voice segments over RTP [18] frames. RTP header [22] contains the codec type information, but it does not provide whether or not it carries voiced segment. However, RTP payload [19] [21] provides an information indicating what kind of information is carried over the RTP frame. RFC-4749 [21]defines the RTP payload structure for the ITU-T G.729.1 codec data. Its Frame Type field specifies the frame type with frame size. If the RTP frame carries non-speech data, its decimal value is set to 15. Using this information, we would analyze how the frame type affects the VoIP quality measurement. We also use Gole and Rosenblush's [9] expand E-model in our work as reference in this paper.

This paper presents an expanded E-model employing the RTP packet header information to analyze the impact of the packet type on the VoIP quality impairment. We identified the performance impact by the unoviced frame type that normally occupies more than 50% of the entire voice frames as positive attribute on the VoIP packet loss.

3.2 RELATED WORKS

3.2.1 VoIP over Real-time Transport Protocol

Real-time Transport Protocol (RTP) [18] provides end-to-end transport function for the applications transmitting real-time data such as voice, audio and video. It has been widely used to transfer the encoded voice traffic. In RTP header, there are several fields such as version, padding enclosure, payload type, sequence number and timestamp. Payload Type (PT) field is to identify the medium type to be carried over the RTP frame. In VoIP application, it is used to determine the encoder type. Sequence number is used by the receiver to detect the pack loss and restore the packet sequence. Timestamp field reflects the sampling instant of the first byte in the RTP data packet. It is used for the synchronization and jitter calculation. It is dependent on the payload format that is normally determined by the payload type (e.g. encoder type). RTP Control Protocol (RTCP) provides monitoring and report functions based on the periodic transmission of the control packets to the participants in the RTP session. Its primary function is to report the quality of the RTP transmissions. Upon the reports, the application may optimize its encoding controls. The RTP payload format is also defined upon the codec type that is specified in the PT field since each codec generates its own packet format. For example, The RTP payload format for G.729.1 audio codec is specified in RFC 4749 [21]. RTP payload format for EVRC codec is also specified in RFC3558 [24]. Each standard also defined the frame type that depicts the encoding rate and frame size of the payload. Four-bit FT field of G.729 payload header over RTP [21] is defined in Figure 3-1. The FT value 15

FT	Encoding rate	Frame size
0	8 kbps	20 bytes
1	12 kbps	30 bytes
2	14 kbps	35 bytes
3	16 kbps	40 bytes
4	18 kbps	45 bytes
5	20 kbps	50 bytes
6	22 kbps	55 bytes
7	24 kbps	60 bytes
8	26 kbps	65 bytes
9	28 kbps	70 bytes
10	30 kbps	75 bytes
11	32 kbps	80 bytes
12-14	Reserved	
15	NO_DATA	0

Figure 3-1: Payload header of G.729.1: Field Type

indicates that there is no audio data in the payload. It may be used to perform the flow control by updating the MBS value where the maximum receivable bit rate is defined. RFC3558 also defined the 4-bit frame type in the Table of Content entries. It describes the different encoding rates of EVRC and SVM codecs with variable frame sizes. This FT information should be used what type of frame is lost, and determine how much the packet loss will affect the VoIP quality in RTP level. We facilitate the frame type information to analyze the difference in the VoIP quality impact upon the packet loss from the transport level.

FT	Rate	Frame size
0	Blank	0 byte
1	1/8	2 bytes
2	1/4	5 bytes
3	1/2	10 bytes
4	1	22 bytes (171 bits with 5 padding bits end with zeros)
5	Erasure	0 (will be dropped)

Figure 3-2: Frame Type for EVRC and SMV codecs

3.2.2 Structured Information from the Encoded packet

Some codec type also provides Encoded embedded control information on the encoded payload. In addition to the frame type field in the RTP payload header, embedded control information will enhance the performance of the packet level VoIP quality measurement. However, this approach also requires more cost that the VoIP measurement tool performing the deep packet inspection. As illustrated in Figure 3-3, G.729.1 codec provides bit stream ranged from 20 bytes to 80 bytes upon the encoding rate from 8 kbps to 32 kbps. In L2 payload on 12 kbps encoding rate, there is class information bits defined in order to enhance frame erasure concealment. Class information bits present if the payload contains voiced segments or unvoiced segments with the corresponding signal information that are either onset or transition. Onset comprises all voiced frames with stable characteristics, and transition presents relatively weak characteristic so that the class might be changed. For example, voiced transition seems to move to the unvoiced onset class, and unvoiced transition would move to voiced onset class. However, this class information is not available for L1 8 kbps encoding rate. This may limit us to apply the class field to overall G.729.1

SYNC	NBIT	L1 (160 bits)	L2 (80)	L3 .... L12 (40)..........(40)

Figure 3-3: G.729.1 Payload structure

codec. However, we decided to use this field as part of this study to analyze how the embedded control information enhances the VoIP quality measurement.

3.2.3 GSM over RTP using frame format on the payload

When the GSM frames are carried over the RTP frames, it provides different frame formats upon the voiced and unvoiced frames. As part of the Internet Protocol Harmonization Over Networks (TIPHON) project, ETSI TS 101 318 [25] defines the GSM codec frame formats upon the encoding rate as well as frame type in terms of voice and unvoiced information. RTP encodings are defined in terms of the GSM codecs, which are full-rate [26], half-rate [27] and enhanced full-rate [28] GSM codecs. The bits of the codec parameters are located in the same way as they are defined in the respective codec specifications [26] [27] [28]. For example, RTP encoding for full-rate GSM codec [26] defines 264 bits beginning with 4 bit of signature in addition to 260 bits of CSM codes as depicted in Table 3. For the unvoiced frame that is defined as silence indication frame (SID), it is identified by particular bits of SID code word consisting of 95 bits Xn(i) parameters in the error protection class I, which all shall be 0 (see figure 3-4 in ETS 300 909 [11]). These bits are:

- the most significant bit (bit number 2) of all Xn (0..12) parameters in every sub-frame
- the middle bit (bit number 1) of the Xn(0..12) parameters in sub-frames 1, 2 and 3
- the middle bit (bit number 1) of the Xn(0..3) parameters in sub-frame 4.

In addition to the full-rate GSM codec, RTP encoding also supports half-rate (GSM 60.20 [27]) and enhanced full-rate (GSM 60.60 [28]) codec. Similar to the full-rate GSM codec RTP encoding, both GSM codecs of RTP encoding support the SID frame. Half-rate GSM codec [27] of RTP encoding sets the SID codeword consisting of 79 bits with '1'. Enhanced full-rate GSM codec of RTP encoding sets SID codeword consisting of 95 bits with '1'.

It should be more effective to estimate the VoIP quality if the measurement exploits the SID frame format to detect the unvoiced frames.

3.2.4 E-model based VoIP quality measurements

Since wireless network transmits the VoIP traffic over multiple network nodes, its quality may be inconsistent upon the topology (e.g. the number of nodes), RF conditions (e.g. signal interference or transmission power) and mobility (e.g. handoffs). A capacity study [12] over Ad Hoc wireless Network showed that the each node may get the end-to-end throughput equivalent to $O()$, where n is the number of hops. De-

Parameter	Bit orders	
Signature	1^{st} - 4^{th}	Every frame is started with 0xD as magic number
$LAR_c(1)$ - $LAR_c(7)$	5^{th} - 40^{th}	Short-term filter parameters
N_1	41^{st} - 47^{th}	Each $X_c(n)$ is 3 bit long. Four sub-frames follow after filter parameters. Each sub-frame consists of Long-Term Prediction (LTP) parameters that are $N_{1..4}$ and $b_{1..4}$ and ecoding parameters RPE parameters
b_1	48^{th} - 49^{th}	
M_1	50^{th} - 51^{st}	
X_{max1}	52^{nd} - 57^{th}	
$X_1(0)$-$X_1(12)$	58^{th} - 96^{th}	
. .	. .	
N_4	209^{th} - 215^{th}	
B_4	216^{th} - 217^{th}	
M_4	218^{th} - 219^{th}	
X_{max4}	220^{th} - 225^{th}	
$X_4(0)$-$X_4(12)$	226^{th} - 264^{th}	

Figure 3-4: The order of GSM 06.10 full rate speech codec parameters in the canonical RTP encoding

crease of the end-to-end link capacity will cause the voice quality degradation directly or indirectly upon the traffic flow control.

An experimental test [14] also showed that the VoIP capacity significantly decreased upon increase of the number of hops. Many researches used metrics based on the E-model in order to predict the VoIP quality. As defined in ITU-T G.107 [4], R value is defined in Equation (2.2) Gole and Rosenblush [9] extended the use of E-model from the network planning tool to network monitoring tool with simplified R value computation within the transport level measurement including delay, jitter and packet loss as defined in Equation (2.3). We applied the equation proposed by Gole and Rosenblush [9] to the initial experiment to estimate the R value of VoIP quality over the wireless network.

3.3 UNVOICED FRAME IMPACT ON VOIP QUALITY

In the former study [30], we investigated the correlation between the packet level VoIP quality measurement using modified E-Model and the speech level VoIP quality measurement using Anique+ from the following perspectives: " Packet transmission delay/jitter and speech distortion " Packet loss vs. mute impact In this study, we further expand the previous VoIP quality measurement [30] with the exploit of unvoiced frame impact by detecting the VoIP frames as either voice or unvoiced packet. There are two factors that we considered for the primary attributes in terms of the unvoiced frame impact on the VoIP quality measurement as follows: " The number of unvoiced frames in the impairments " Imbedded control message effect on the unvoiced frames

3.3.1 Number of the unvoiced frames in the impairments

The loss of the unvoiced frames may affect VoIP quality much less than the one of the voce frames. From the previous study regarding the VoIP QoS [33], it was assessed that more than 50% of entire conversation may carry the unvoiced frames due to the voice activity. We investigated the unvoiced frame impact on the VoIP quality with the range of the occupancy from none to 60%, and applied to the Ief and R factor based on the equation (3). We revisit the equation (2.15) and revise it with the unvoiced frames loss factor as follows:

$$I_e^* \sim \gamma_1 + (\gamma_2 * \lambda) ln(l + (\gamma_3 * \lambda)e) \qquad (3.1)$$

,where the λ is the voice activity factor.

3.3.2 Imbedded control message on the unvoiced frames

Depending on the codec implementation and applications, even though the packet carries unvoiced frame, it may also contain embedded control messages such as flow control, power control messages. For example, over the wireless network, if the unvoiced packet carrying the power control message is discarded, the mobile would assume that it should increase transmission power that might cause interference to other mobiles. In addition, the VoIP application protocol may also be affected to maintain the connection state by the impairment of the unvoiced packets. Typical wireless VoIP applications maintain the connectivity verification not only in the transport layer such as using the TCP timeout, but also in application layer using embedded counter. Even though the TCP timeout is typically used in the VoIP application over the wired Internet, it may be too long to be used in a wireless environment especially where the packet loss may cause the performance degradation to other mobiles due to the interference result from the power control corruption. In out study, we assume that the VoIP application may use the application layer embedded control messages, and consecutive packet loss of the control message will cause the call session termination. We also ranged the number of the consecutive packet loss from 3 to 60 that is equivalent to 60 msec to 300 msec. We add this factor as an additional attribute to the equation (3.1), and define the following:

$$I_e^* \sim \gamma_1 + (\gamma_2 * (\lambda + \delta) ln(l + (\gamma_3 * (\lambda + \delta))e) \tag{3.2}$$

We also investigate how the timer of the consecutive packet loss reflects the Ie* as well as R factor.

3.3.3 Case Study: EVRC on CDMA RAN

We performed the packet loss simulation to analyze the unvoiced frame effect. As shown in Figure 3-5, EVRC Rate Set 1 provides four different frame types. We only uses three frame types including full rate with 22 bytes, half rate with 10 bytes and 1/8 rate with 2 bytes.

Full rate and half rate of voice frame carries the voice information while 1/8 rate carries unvoiced frame that is due to the voice activity effect. We also collected the field data that are distributed as follows:

- 40% of Full rate frame
- 10% of Half rate frame
- 50% of 1/8 rate frame

We only considered the bearer traffic since it is carried over UDP/IP that does not support for the retransmission while the call control messages over TCP/IP supports for the retransmission upon the packet loss. Hence we excluded the amount of the control traffic allocation from the VoIP quality effect estimation. In addition, we also allocated 10% of the entire traffic to call control messages such as call processing, OAM messages. Since 1/8 rate unvoiced frame does not affect the voice quality, we exploit the loss of 1/8 rate frame as unaffected packet loss. As depicted in Figure

EVRC (Rate set 1 8.55Kbps)									
traffic type	Traffic allocation (%)	EVRC	VoIP frame	Packet overhead		Total	Occupancy (%)	Packet loss Impact weigh	Actual VoIP quality impact
				UDP/IP	L2				
bearer traffic	90%	full rate	22	28	6	56	36%	65%	23%
		half rate	10	28	6	44	9%	29%	3%
		1/4 rate	4	28	6	38	0%	0%	0%
		1/8 rate	2	28	6	36	45%	6%	3%
Total			38	112	24	174	90%	100%	29%
Call Control traffic									
			Signaling frame	TCP/IP	L2				
control message	10%		64	40	12	116	10%	N/A	N/A

Figure 3-5: CDMA RAN traffic profile

3-5, we computed the actual packet loss effect with consideration of unvoiced frame effect as 29%.

3.4 Conclusion

In this Chapter, we studied how the unvoice frame impact the VoIP quality measurement data. If we do not consider the unvoiced frame as null parameter of the VoIP quality impact, the packet monitoring tool may produce too exaggerated VoIP quality degradation upon the packet loss. With the compensation of the unvoiced frame effect on the packet loss, the packet monitoring tool should provide more realistic VoIP quality measurement upon the packet loss statistics.

As consequence, the E-model based VoIP quality measurement tool should consider to integrate the unvoiced frame effect as a compensated parameter on a particular VoIP quality measurement.

Chapter 4

Mobility Impact on VoIP Quality over wireless network

4.1 Introduction

VOICE OVER IP (VOIP) has been deployed in wireless network domain mainly over data network. As the VoIP becomes more prevalent in the main stream of the telephony service, quality measurement became a crucial issue. ITU-T initiated the voice quality measurement standard, and defined several models supporting various stages including network planning to network maintenance. As a subjective test, Mean Opinion Score (MOS) has been used for the traditional source of the voice quality test. The MOS is a standard way to evaluate the voice quality conducted by the subject group testing in the strictly controlled environment. Even though the MOS is the most effective method to determine the voice quality for the device such as codec, it may not be applicable to measure the voice quality over real service network systems due to the restriction on the testing environment. Objective tests have been proposed and standardized for measuring the voice and audio quality by a measurement test systems. According to the knowledge of the input data, the objective tests can be described as intrusive models or non-intrusive models. A recent paper by Antony W. Rix et. al. described details about these models with exemplary standards [1].

E-Model is a computational model used for the transport network plan. Given various input parameters such as end-to-end delay, echoes, side-tones, loudness, the E-model calculates a scalar quality rating value, R, which corresponds the overall conversational quality. While ITU-T G.107 [4] presents E-model overall parameters used for the R factor computation, companion recommendations including G.113, G.108 and G.109[5][6][7] provide further guidance about specific codec related speech impairments, E-model applications and definitions of transmission quality. Since E-model is defined as framework for the network planning method, there is limited accuracy on the measurement of the voice quality. This leads many proposals on enhancement and extension of the current E-model.

We applied the E-model to the wireless network domain, especially on the mobility environment. On the mobile network, user traffic may redirect upon the handoff that switches the serving cell site when the mobile user moves out of the cell boundary. When the user traffic switches its traffic path, the mobile network system may discard the user traffic. However, it is hard to be known due to the different network protocol between the core network and access network that will be discussed in details on the following section. We investigated how the wireless network introduces a hidden impairment on the VoIP quality due to the mobility. This paper presents the analysis on the variety of transport layer including packet and frame layers. In addition, the analysis is also compared to the voice quality measurement using ANIQUE+ tool. Gole and Rosenblush [9] expand the use of E-model to the VoIP performance monitoring tool by exploiting transport-level metrics for the purpose of monitoring conversational voice quality. This method was used for several papers as reference model to determine the VoIP quality. Their method is used in our work as reference for the comparative analysis with ANIQUE+ model.

This paper presents the measurement of the VoIP quality in the packet and the speech domains, and provides the analysis illustrating the correlation between packet level impairments and speech level distortion. In this paper, we use ANIQUE+[16] that is a non-intrusive model recently standardized in ANSI. Auditory non-intrusive quality estimation plus ANIQUE+ is a perceptual model to estimate the speech quality without reference information while it simulates the functional role of human auditory system.

4.2 Related works

4.2.1 Layered transport architecture in mobile network

Existing mobile networks consist of layered network architecture. When the mobile station connects to the server over the IP core network, the user traffic (e.g. bearer traffic) is transmitted over multi-layered transport protocols. As an example, Figure

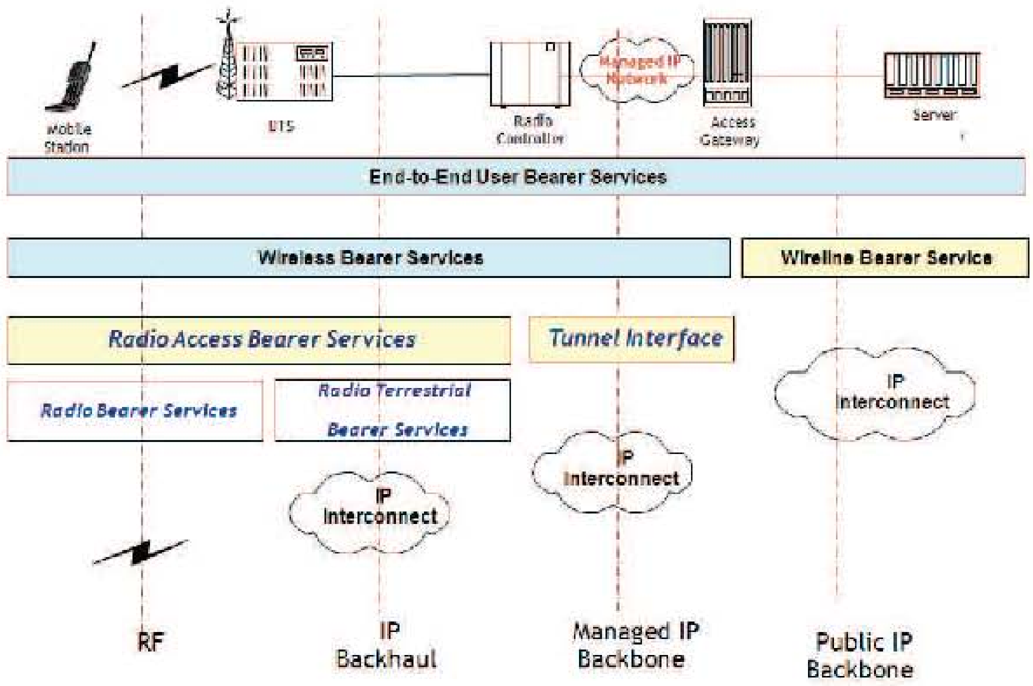

Figure 4-1: Layered mobile network architecture

4-1 illustrates a logical hierarchy of the end-to-end bearer services in typical mobile network systems. When the mobile station application software connects to the remote server over the internet, the user traffic is typically segmented and encapsulated into multiple layer of the transport protocol at each network interface of the following:

- The Access Gateway (AG) (e.g. PDSN, GGSN)
- Radio Network Controller (RNC)
- Base Transceiver Station (BTS)

The AG interfaces between public IP network and managed IP core network that is owned by the operator. When it receives the IP packets from the remote server, it will segment the IP packets and encapsulates over the wireless bearer service transport protocol. QoS is enforced between the MS and Access Gateway as defined by the

Service Level Agreement (SLA). RNC is another network node interfacing between managed network and Radio Access network. When the RNC receives the packet from the AG, it should encapsulate them over another transport protocol between BTS and RNC. On a typical mobile network, RNC contains the frame selector that is a functional unit to handle the handoff. The RNC serves a cluster of BTS geographically located in neighbor areas. When the BTS receives the traffic packets, it terminates the packet overhead and transmits the radio frames to the mobile stations. This layered network architecture causes a complicated network structure for the VoIP quality measurement exercise. Quality measurement using packet network monitoring scheme presented by Gole and Rosenblush [9] may be hard to be applied on this architecture since packet delay and discard at each mobile station might be too costly to be measured. Even though we apply Gole and Rosenblush's method to the BTS and RNC over the bearer service transport network, it may not be accurate. When the mobile traffic path changes due to the handoff, they are carried over another BTS-RNC network. Since the mobile station address is not embedded in the BTS-RNC transport packet header, it is not easy to measure the entire VoIP quality over the mobility.

4.2.2 Handoff over the mobile network

Handoff is a key function to support for the continuity of the mobile service as the mobile users moves from the serving cell site to the neighbor cell sites. As the mobile user moves to the boundary of the serving cell site, the RF signal gets weaker so

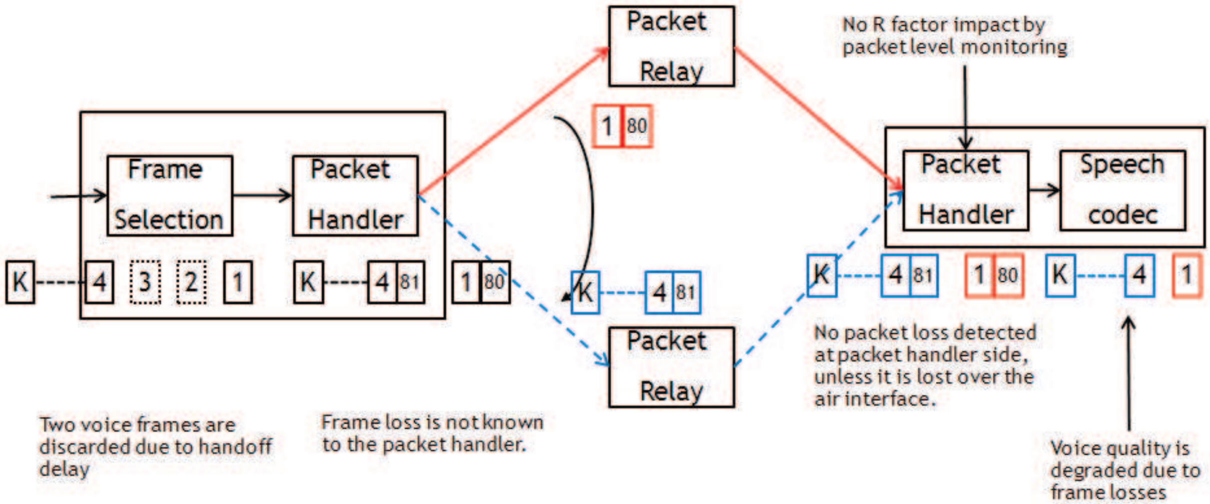

Figure 4-2: Handoff effect over the radio access network

that the mobile cannot receive the message over the air due to the excessive frame errors. Hence mobile should be able to monitor the RF signal strength if it becomes lower or higher. Based upon the RF signal strength, the mobile informs the serving BTS that it should initiate the handoff to release the current call resource of the serving BTS, and relocate the call resource to the new serving BTS. In principle, there are two strategies conducting the handoff depending upon the radio resource and handoff strategy. Hard handoff is to transfer the call resource to the new cell site after releasing the call resource from the current serving. Soft handoff allows the mobile can use multiple call resource during the handoff procedure, and release the call resource after completing the handoff process.

Prior researches [41,42,44,46,47] showed that the vertical handoff currently takes too long to meet the existing carrier grade services. Pagtzis [41] showed the latency of vertical handoff between WLAN and GPRS take more than 3 seconds from WLAN to GPRS and 5 seconds from GPRS to WLAN. Due to the excessive handoff, VoIP quality over vertical handoff is not measurable from the existing voice quality metrics [9, 10] perspective.

There are two areas, Air interface and Backhaul interface where the VoIP quality may be impacted due to the handoff. As illustrated in Figure 4-2, there are three functional unites that are packet handler function at the RNC, packet relay at the BTS and packet handler at the MS. The following describes the functions on each unit:

- Packet Handler (PH) at the RNC: PH performs the packet processing when it receives the user traffic from the Frame Selector or BTS. The PH functions are dependent on the backhaul transport network bandwidth and protocol in addition to the QoS requirement on the user service. Frame Selector (FS) in the RNC manages traffic resource for the traffic path and switches the traffic path upon the BTS request of the handoff. It exchanges the control messages with the serving BTS and the target BTS when it switches the traffic path.

- Packet Relay (PR) at BTS: PR is a counterpart of the PH conducting the transport protocol handling function for the backhaul network. It also provides the packet processing function for the air interface with the MS. When it receives the packets over the backhaul interface, it forwards the packets to the target MS.

- Packet Handler (PH) at MS: PH at MS terminates the packet format of the over-the-air interface and forwards the user traffic (e.g. VoIP frame) to the application layer (e.g. voice codec) unit.

The following describes how those functional units perform at the handoff process over hypothetic CDMA radio network:

- When the MS detects the pilot channel of its neighbor cell sites, it informs the serving cell site that it starts receiving the neighbor cell site's pilot signals. When the mobile decides the handoff, it sends the handoff request message to the target neighbor cell site via the serving cell site.

- The serving cell site forwards handoff request message to the target cell site via RNC. When the target cell site responds the handoff request message, the RNC and the new serving cell site establishes the new bearer traffic connections in order to transfer the user traffic over the backhaul network.

- Since this connection is different from the previous one between the RNC and the old serving cell site, there is no prolongation to trace the packet level quality measurement using the packet header information unless the RNC keeps tracking the transition of the bearer traffic connections. However, it becomes difficult due to the limited memory space, processing power and number of cell sites connected by a single RNC.

- When the MS completes the handoff, it sends a request message to the old serving cell site to release the traffic resource. Upon reception of the request message, the old serving cell site releases the traffic resource for the MS and terminates the bearer traffic connection with the RNC.

4.2.3 RNC performance analysis and estimation

As presented above, it is not practical for an RNC to trace all bearer traffic connections on the high-mobility environment. We can derive the memory space requirement of the RNC to perform the VoIP quality monitoring for the handoff as follows:

$$D = BHCA * Th * Oh \tag{4.1}$$

where, D is the required memory space, $BHCA$ is Busy Hour Call Attempt, Th is data throughput over the air and Oh is the overhead including packet overhead control plane message overhead.

For example, the number of handoff per RNC is determined by the Busy Hour Call Attempts (BHCA) and the average number of handoff per call. If the MS is serviced under high-mobility environment, 5 $\sim$ 7 handoffs typically occur on a single call. If BHCA of a RNC is 500K, then the number of handoff on a RNC is estimated 2.5M - 3.5M. Since each handoff requires a new transport network session, the system should trace entire handoff session including backhaul session establishment that is typically 15% of the user traffic where 10% is for control message overhead and 5% is for packet overhead. If we assume that the handoff process requires 5 second period to complete the resource transition, the RNC requires the following amount of the memory space is estimated 14.375 $\sim$ 20.125 GB per hour that is assessed by the following: Total memory space with 500K BHCA* 5 $\sim$ 7 second per handoff * 8Kbps average data throughput * 1.15 overhead

This will require 1.4 $\sim$ 2 GB data access on every handoff period that is 5 - 7 seconds. It also requires that the RNC processor conduct the packet header analysis to trace the end-to-end VoIP quality through the packet level monitoring.

4.2.4 Mobility impact parameters on VoIP quality

In order to determine the Mobility impact on VoIP quality from the packet level monitoring, the following are identified as the mobility impact parameters that will be employed as extended E-model:

1. Handoff ratio per call (HOR): Average handoff occurrence per call
2. Handoff latency(HOL): Average handoff period from start to completion
3. Average Frame Error Rate (FER): average FER over the air interface per call
4. Intermittence ratio (IR): Percentage of intermittence of packet transmission over the entire handoff period.Traffic intermission is normally dependent on mobile call processing software implementation. Mobile can either stop sending the bearer traffic over the air, or keep sending the traffic.
5. Signaling Traffic Overhead (STO): when the handoff is triggered, MS and BS and source BS and target base station exchanges the control messages to proceed the HO function. This control message overhead will cause either traffic congestion or processing latency to complete the HO process.

The handoff ratio is dependent on the degree of mobility. Higher mobility causes more handoffs per call. Handoff latency is normally determined by the handoff strategy and call resource management function of the RNC. For example, if the operator wants to release the call resource in conservative way after the new call resource is established, then the latency will be relatively longer than normal handoff procedure. Average FER is typically configured as a constant value in our study since it is determined by RF engineering design before the network deployment. Intermittence ratio is an empirical parameter presenting how many packets are to be discarded during the handoff upon the traffic path switching. As illustrated in Figure 2, when the handoff occurs, the bearer traffic path switches from the serving cell site to the neighbor cell site where the packet relay forwards the traffic to the mobile station or vice versa. Being different from the conventional voice call, VoIP is a data service and it is not simultaneously serviced by the multiple cell sites due to the resource as well as complexity of traffic processing.

4.2.5 Extended E-model based on mobility impact parameter

Many researches have used metrics based on the E-model in order to predict the VoIP quality. As defined in ITU-T G.107 [4], R value is derived as follows:

$$R = R_0 - I_i - I_d - I_e + A \tag{4.2}$$

where R0 is the original voice quality including noise source; Is is a combined impairment related to the voice signal; Id represents the impairments related to the delay of the voice signals including echoes and one-way transmission delay; Ie is equipment related impairments such as random packet-loss that is typically caused by the low-bit rate codec. G.113 [7] provides codec specific Ie values. The advantage factor A can be applied if there are other advantages of access to the user. G.107 [4] provides several default values upon the access transport types as recommended upper limit for the network planning engineer.

Cole and Rosenblush [9] extended the use of E-model from the network planning tool to network monitoring tool with simplified R value computation within the transport level measurement including delay, jitter and packet loss as

$$R \sim 94.2 - 0.024d + 0.11(d - 177.3)H(d - 177.3) - \gamma_1 - \gamma_2 \ln(1 + \gamma_3 e) \tag{4.3}$$

where γ_1, γ_2, γ_3 are specified in [9] upon the codec type and H(x) is Heavy-side function, for all $x \geq 0$, $H(x)$ yields 1 and for all $x < 0$, it yields 0. They used the default value of I_s defined in [4] and reduced the expression of the R factor.

Using the handoff parameters, R factor is refined as follows:

$$R = R0 - Is - Id - Ie + A \tag{4.4}$$

$$Rm = R - Im \tag{4.5}$$

$$I_m = F_H OR(HOL + STO + I_R) + FER \tag{4.6}$$

where I_m is the impairment due to the mobility impact, $F_H OR$ is estimation function of handoff impacts including HOL: handoff latency STO: Switching Traffic Overhead; IR: Intermittence of the packet transmission; FER: Frame Error Rate on the Over-The-Air traffic. HOL and IR can also be assessed in terms of number of packet discards as follows:

$$Number\ of\ Packet\ discarded = (HOL + IR) * frame\ transmission\ rate \quad (4.7)$$

The number of packet discards are counted as the packet loss at (2). HOL is an explicit delay that is known as part of the handoff protocol, and IR is implicit delay that is dependent on the implementation. Each radio access network has different HOL and IR values since handoff strategies and protocols are different in each radio access network. This paper covers EVDO as our experimental radio access network.

Once R_m is estimated, as defined in G.107 [4], R factor should be transformed to the estimated MOS_E using E model as follows:

$$MOS_E = \begin{cases} 0 & \text{if } R_m \leq 0 \\ 1 + 0.035R_m + R_m(R_m - 60)(100 - R_m)7 \cdot 10^{-7} & \text{if } 0 \leq R_m \leq 100 \\ 4.5 & \text{if } x \geq 100 \end{cases}$$

4.3 COMPARATIVE EXPERIMENT ON EVDO NETWORK

We performed experimental tests on the EVDO network using the EVDO handset with data measurement tool that can collect the Over-The-Air (OTA) performance data such as average FER.

4.3.1 EVDO Handoff

EV-DO is a variation of CDMA network to provide high data rate for the CDMA mobile users. Similar to the CDMA technology, it provides the handoff for the mobile services. However, it supports different handoff strategy over downlink and uplink. Same as other typical data service, EV-DO also provide asymmetric data transmission where the downlink is much more loaded than uplink in terms of data traffic amount. The following are EVDO handoff procedure:

1. The entire handoff period is typically between n x 20 msec and sub sec seconds.

2. When MS sees a stronger pilot in the Active Neighbor Set. (The new pilot needs to be stronger than by certain dB called hysteresis to prevent pingpong).

3. MS starts send Null DRC. (DRC is used to indicate the rate that the mobile wants to be served at. E.g. 2.4 Mbps)

4. MS sends Null DRC for integer multiple of a value called "DRCLength" normally 2 slots. This is to make sure that the current serving sector sends off the packets in the buffer to the MS and not overlap the transmission with new serving sector. Also, the MS uses for "Soft/Softer Handoff Delay parameters" before start pointing DRC to new sector. These delays usually set at between 512 that is equivalent to 0.85 seconds and 64 slots that is equivalent to 100 msec. This is estimated time that it takes for AN to switch the packet path from the old sector to the new sector.

5. After Null cover duration, the MS starts sending DRC with the Walsh code

cover for the new sector.

As described above, the handoff delay parameters are variable upon the implementation. Hence, the VoIP quality impact is dependent on the length of the handoff delay. Using equation (4.7), we assessed the number of packet discard ratio for the handoff as follows:

$$Nunber\ of\ packet\ discards\ per\ second\ = (0.85 + 0.1) * 50 = 47.5\ packet/sec \quad (4.8)$$

Since the entire handoff period is variable depending upon the traffic condition, the packet loss ratio over the handoff is defined as above.

4.4 ANALYTICAL MODEL

We used the commercial Data Measurement (DM) tool that can collect the OTA frames and collect the number of frame transmission and reception as well as frame errors. We also used another packet sniffing tool that can measure the user traffic performance at the packet level.

The DM tool is equipped in the laptop, and captures OTA frames between mobile and base station as illustrated in Figure 3. We use Iperf as packet level measurement tool. Iperf that is a freeware of a packet sniffing tool was run on both laptop and server, and it collected the performance data such as data throughput and packet discards. The DM tool collects the OTA frames between the mobile station and base station while the Iperf collects the packet transmission between laptop and server.

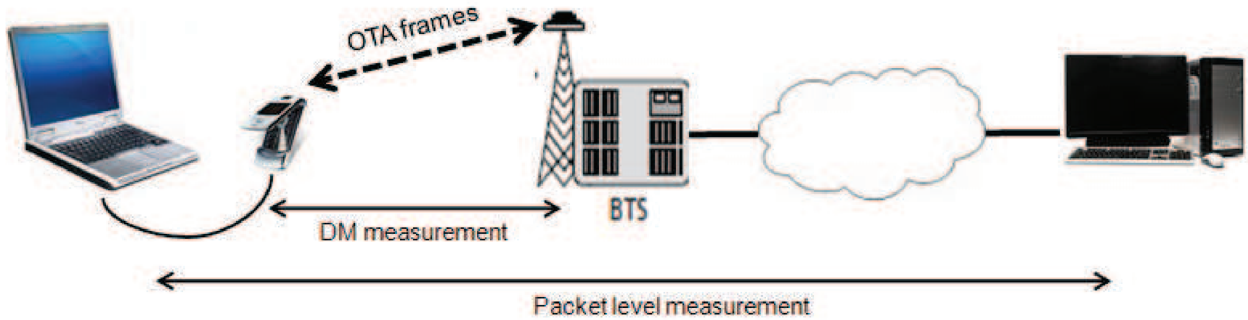

Figure 4-3: Testing configuration

We analyze how these two tools present different performance measurement for the handoff period between user and access network level.

The experiment consists of two phases as follows:

- 1st Phase: HO factor modeling
- 2nd phase: HO factor validation

HO factor modeling is a process to assess the quality impact due to the handoff and determine the HO factor upon. Once the HO factor is determined, it should validated by the comparison with the speech quality test.

4.4.1 Phase I: HO factor modeling

Handoff factor modeling is based on the experimental analysis as illustrated in Figure 4. We conducted the drive test to collect the performance data. From the drive test, we collected the Over-The-Air (OTA) at the DM tool as well as End-to-End packet level data measurement. From the data measurement, we perform the comparative analysis in terms of the following:

- Frame Error Rate (FER) vs. Packet Error Rate (PER)
- Carrier to Interference-plus-Noise Ratio (CINR) measurement for the down link of the OTA
- OTA vs. E2E delay measurement

FER is an average number of the OTA frames discarded due to the frame errors. It's measured at the mobile station and base station. In our paper, we only present down link FER. PER is an average number of IP packets discarded at the user traffic level where the user equipment (e.g. mobile station, server) measures. CINR is measured by the DM tool presenting the down link signal effectiveness. Delay is measured in two layers, which are both OTA interface between base station and mobile station, and E2E between mobile station and server as illustrated in Figure 4-3.

As illustrated in Figure 4, from the comparative analysis, we assessed the Rm and the MOS. We exploit Cole and Rosenblush's model [9] to assess the Rm and equivalent MOS value. HO factor is determined using the handoff estimation function with the E-2-E and OTA performance measurement data.

We also estimated the MOS using the current E model. HO factor is determined with the difference between Estimated MOS and Assessed Quality (i.e. MOSE) derived by Rm using the above equation (4).

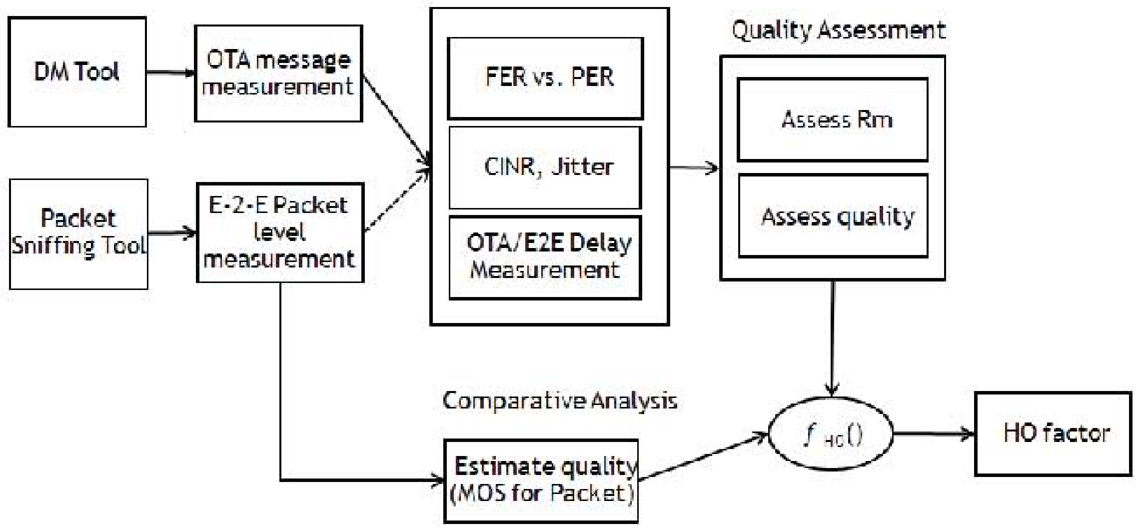

Figure 4-4: Handoff factor Modeling

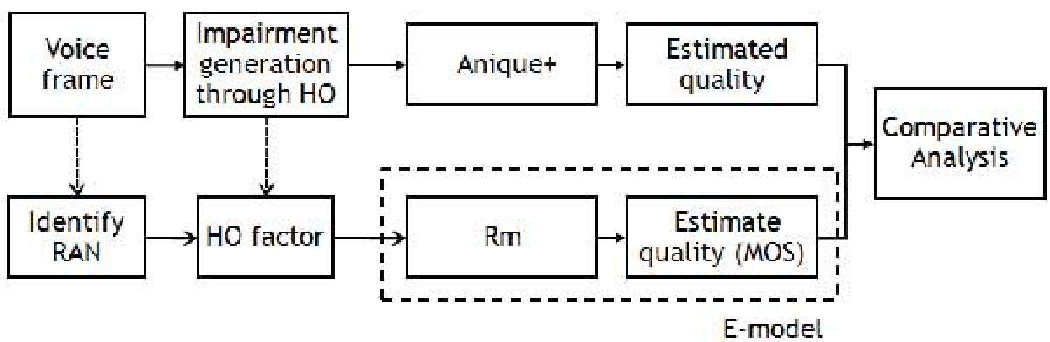

Figure 4-5: HO factor validation process

4.4.2 Phase II: HO factor validation

Once HO factor is determined, it should be validated if it can be applied as a generic parameter.

As illustrated in Figure 5, we conducted the validation process:

1. Voice frames have been impaired through handoff . In order to retrieve better result with less variation, we collected variety of voice frames that are impaired via handoff.

2. Estimate the impaired voice quality using Anique+ tool.

3. Using E model defined in (4.4), we estimate the MOS_E

4. Compare the quality scores between Anique+ and E model.

Comparison of the VoIP quality using Anique+ and E model was presented in Chapter 2.

4.5 EXPERIMENT THROUGH TEST DRIVE

We performed the test drive on the route as depicted in Figure 6. Due to the limited storage space, we performed the data collection over two hours with 20 min period.

We performed the data processing when we collected the performance measurement at iperf [34] as E2E test. The following are measured performance data:

- PER
- Actual transmission throughput
- Jitter

Table 1 describes the testing configurations with the data throughput, buffer size and packet size. results upon the testing configurations.

As described in Table 1, we emulated the VoIP applications over the test. We defined the packet type as UDP and the packet size equivalent to VoIP application. We determined the transmission bandwidth from 100 kbps to 56 kbps. Since VoIP is a real-time application, we limited the buffer size small between 4K and 12K.

Table 4.1: Performance measurements at mobile station

Test	Case I	Case II	Case III	Case IV
Trans. bandwidth (Kbps)	100	56	56	56
Receiver Buffer size (Byte)	8K	4K	8K	12K
Packet size (Byte)	138	138	138	138
Testing period(second)	1200	1200	1200	1200
Data type	UDP	UDP	UDP	UDP

4.5.1 Test case I

Test case 1 provides packet transmission with 100Kbps bandwidth with 8KB buffer. As summary, we observed the handoff impact was mild since the bandwidth and buffer size is large enough to tolerate the packet intermission during the handoff. Figure 4-6 presents the measured packet error rate in the test Case I. We observed that the packet error occurred on handoff occurrence. It was noticed that no handoff was occurred while the vehicle goes through a tunnel where the leaky cable was deployed. While the vehicle moves on high speed, we observed handoff s occurred every 45 - 50 seconds. The packet error rates are between 1 and 11%.

Figure 7 presents transmission jitter measured with the accordance of the packet errors. We observed transmission jitter over the handoff period. It was observed between 7.5 and 45 msec. Most handoff caused greater jitter between 20 and 30 msec. Sometime, we observed that the jitter reached 45 msec. As described in EVDO handoff procedure in the previous section, there are buffering effect during the EVDO handoff due to the Null cover duration as well as data rate change upon mobility.

Similar to Figure 7, we also observed that transmission data throughput was also

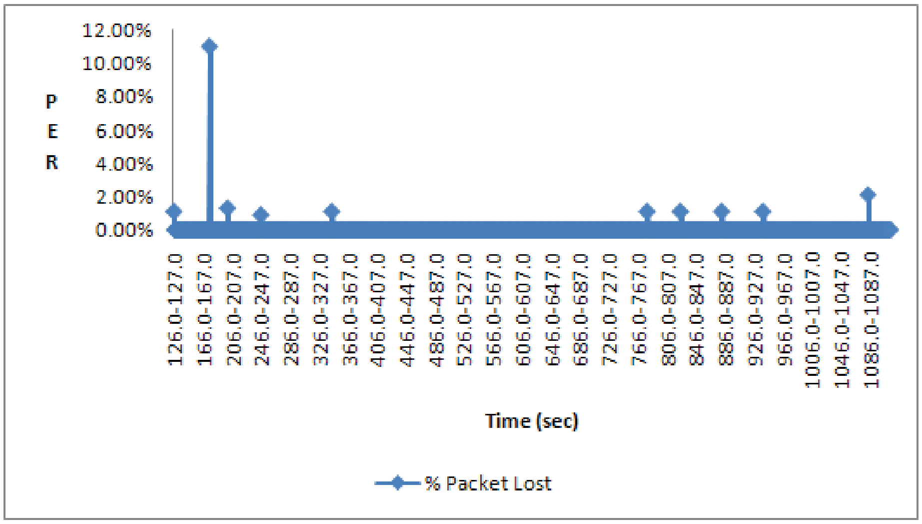

Figure 4-6: Measurement of the Packet Error Rate (Case I)

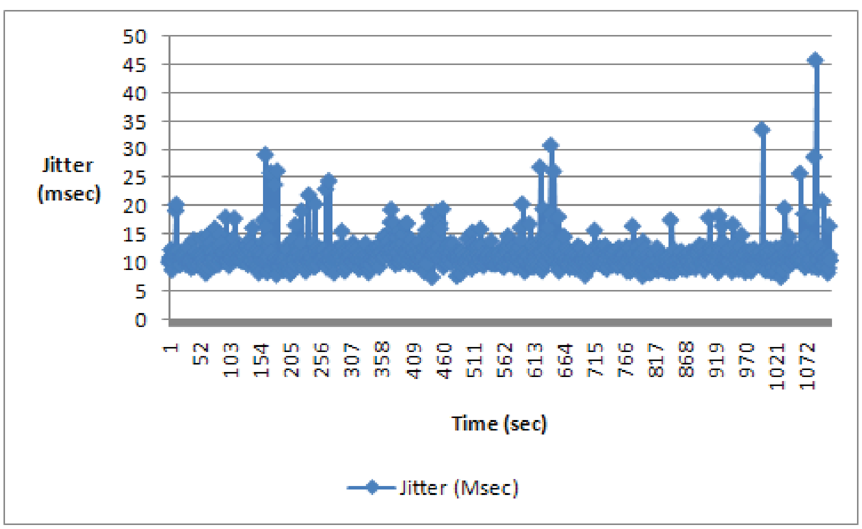

Figure 4-7: Measurement of transmission jitter (Case I)

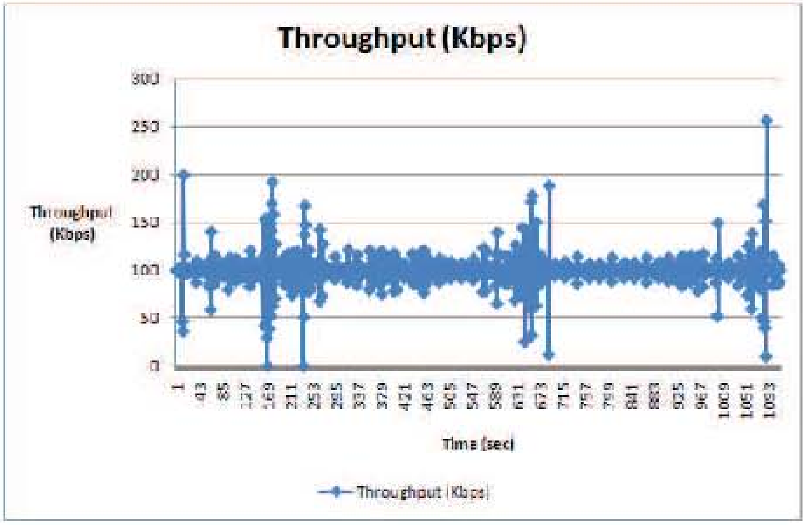

Figure 4-8: Measurement of Actual data throughput (Case I)

fluctuated at the handoff. Due to the jitter effect, the transmission rate was summed up after jitter was dissolved.

4.5.2 Test Case II

Test case II is configured to be closer to the real networking environment that provides narrow bandwidth as well as small buffer size disguise the mobile handset. We observed that the packet discards were hiked at the handoff.

As we assessed in the previous section, we confirmed that the packet loss rate is between 40 and 68 %. As shown in Figure 9, majority of handoff was over 50% percent of the packet loss ratio.

As depicted in Figure 10, the transmission jitter is also twice longer than the test case I. This is also aligned with the packet loss rate.

Actual data rate reflects the handoff as shown in Figure 11. We observed that the number of transmitted data was fluctuated upon the jitter caused by handoff.

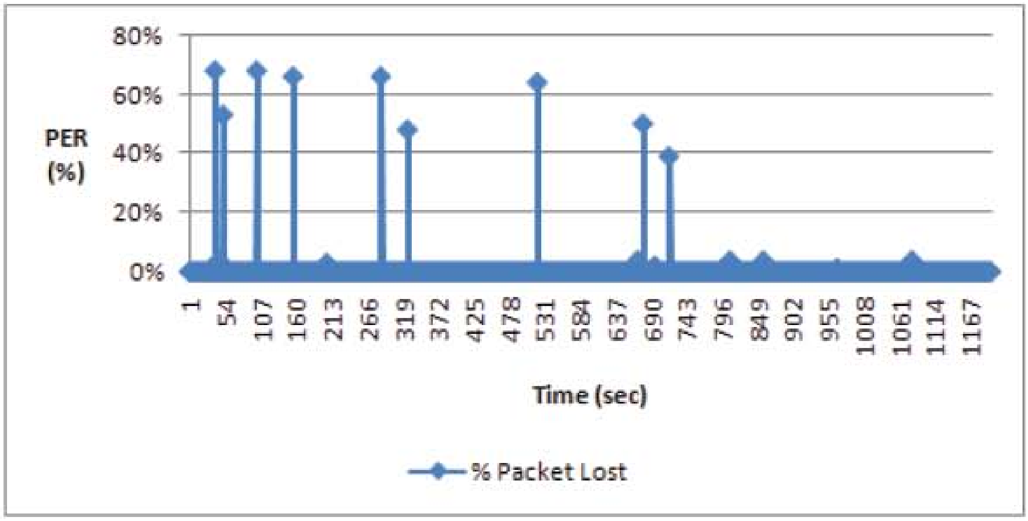

Figure 4-9: Measurement of the Packet Error Rate (Case II)

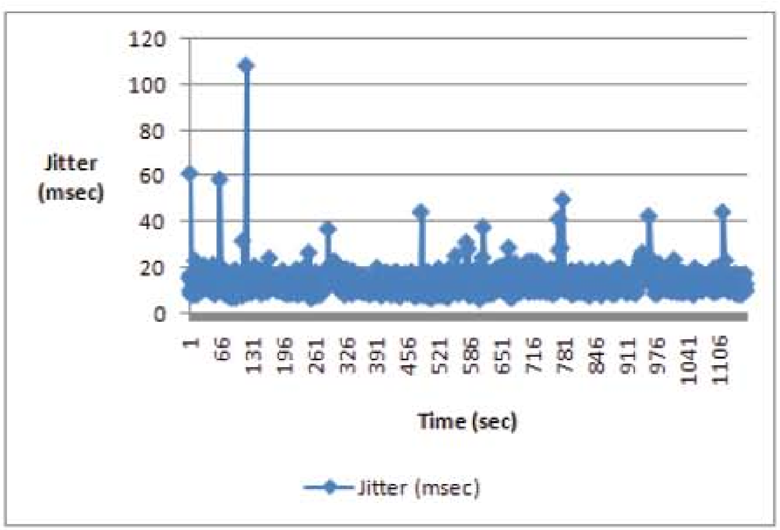

Figure 4-10: Measurement of transmission jitter (Case II)

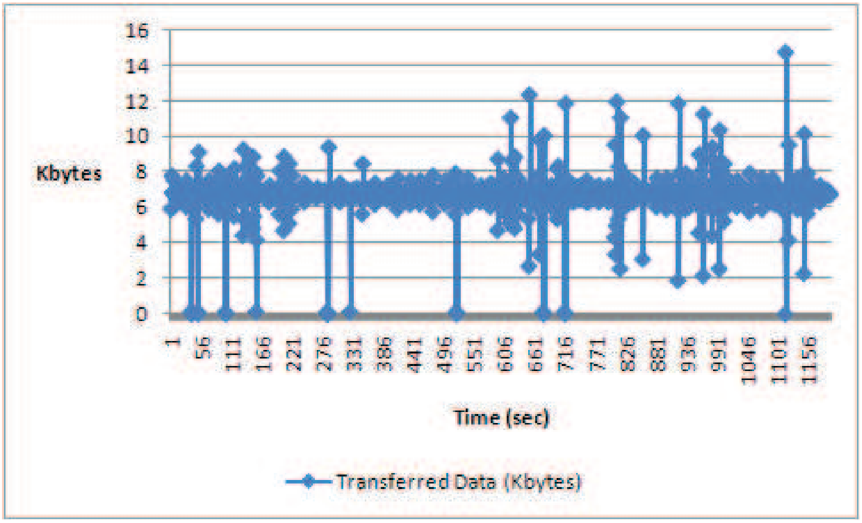

Figure 4-11: Measurement of Actual transmitted data (Case II)

4.5.3 Comprehensive Analysis

As we observed the Test case II that EVDO handoff generated the temporal burst packet loss around 50% percent of transmitted packets. This is clearly not acceptable from the previous study [7][8][9][26] rate as burst error the current voice quality. However, Test case I provides moderate packet loss rate upon the handoff, and it is within the boundary that most voice quality requirements provide.

We also run other two test cases with increase of the receiver buffer size. We observed that increase of the receiver buffer size contribute the less packet loss rate. However, it significantly increased the transmission jitter in several handoffs. As shown in Figure 12 and 13, even though the packet loss decreased, the transmission jitter increased compared to Figure 9 and 10. This is caused by the increase of the packet back log latency. when the buffer gets emptied due to the degradation of the OTA data rate as signal strength gets lower, the VoIP packets in the receiver buffer

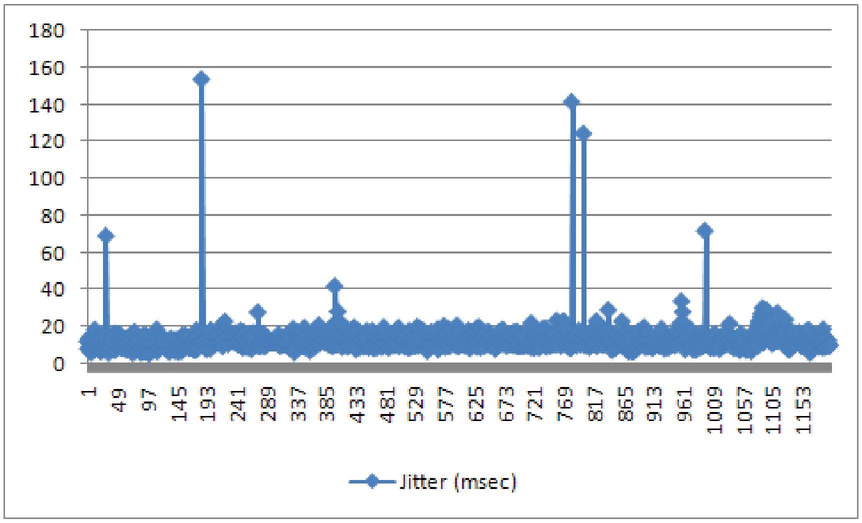

Figure 4-12: Measurement of transmission jitter (Case III)

should be consumed. When the VoIP packets are received, it should be stored into the buffer as the buffer can start forwarding the VoIP packets. If the buffer size is too large, it causes another jitter from the receiver side. Difference of the jitter between Figure 4-10 and 4-13 is due to the difference in buffer sizes in Case II and III.

This showed that the increase of receiver buffer is not a good answer for the VoIP quality enhancement because of its realtime nature.

4.6 Conclusion

From the drive test, we assessed the handoff impact on the VoIP over EVDO access network. We also validated our HO impact was aligned with other previous simulation [26] as well as analyses with experimental test over 802.11 network.

We also found an interesting difference between Saleh [42]and our test. Saleh [42] conducted the test over CDMA2000 system that supports the soft handoff and show

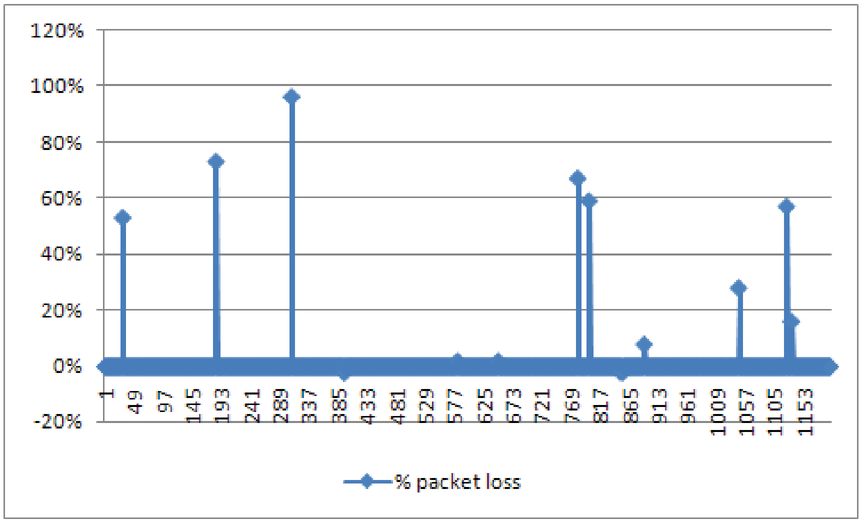

Figure 4-13: Measurement of the Packet Error Rate (Case III)

much less packet discard and jitter issue with even vertical handoff while we observed a burst packet loss and high jitter during the handoff. EVDO network supports a hard handoff for the down link where we measured while it also supports the soft handoff that is same as CDMA2000 RAN architecture.

Our analysis and test result showed that VoIP application over the existing EVDO Rev0 network should require at least 1.5 times bandwidth to provide the equivalent voice quality to the existing legacy circuit wireless voice service. However, the buffer size should be limited to twice than data throughput in order to maintain lower jitter due to the packet backlog problem as we observed in Figure 4-13.

Chapter 5

Future Research and Conclusion

5.1 Conclusion

This dissertation addressed the mobility impact on the VoIP service from the voice quality management perspective.

In chapter 2, we performed that comparative analysis between E-model that is a computational model used for the transport network plan and Anique+ that is a perceptual model to estimate the speech quality without reference source. We exploit the model that Gole and Rosenbluth [9] developed as a network VoIP quality monitoring tool. The analytical comparison showed that there are a strong correlation between E-moded based VoIP quality model and Anique+.

In chapter 3, we also investigated if the packet overhead can contribute an accurate VoIP quality monitoring. With leverage of the packet overhead information indicating the VoIP payload type that is either voice, mute or voice+signal, we refined the VoIP quality parameters. We defined the unvoiced packet effect that normally takes more than 50% of voice frame sequence. We employ the unvoiced frame effect as an attribute to revise Gole and Rosenbuth [9] E-model based VoIP quality monitoring tool.

In chapter 4, we measured the handoff impact on the VoIP quality due to the switching latency as part of handoff procedure. From the experimental analysis, we observed that EVDO handoff over the downlink contributes at least additional 1% burst packet loss. We also concluded that the data rate and receiver buffer should be configured twice more than the required bandwidth in order to handle the data throughput and jitter impact due to the handoff.

From the entire research, we concluded that mobility causes VoIP quality degradation from the following impacts:

- Packet loss
- Transmission bandwidth decrease
- Jitter

We also assessed that the unvoiced frame overhead should provide compromise on the VoIP packet loss by checking the VoIP frame type up to 50%.

5.2 Future research

Our research was limited in EVDO network due to the availability of the network testing environment.We also want to pursue the following area to investigate the VoIP quality impact on different handoff types:

- Veritical Handoff between different RAN technologies such as WLAN and WCDMA: Vertical handoff test [45–47] showed that the handoff latency is more than 3 seconds that is not acceptable from 3G network perspective. While the technologies continue enhancing the latency time, other potential enhancement should also be studied.

- Soft Handoff is expected that the handoff latency should be negligible. However, it should be confirmed by the technical experiment.

- Handoff over 4G network: 4G network advocates that it provides broadband access over the air. However, it is not clear how 4G can support for the VoIP with mobility. In order to offer VoIP application over Mobile Wimax that is one of 4G technologies, it is critical to assess the VoIP quality over the mobility environment.

We also want to expand our research to Quality of Experience (QoE) [52–54] level. We learned that the current speech quality model is not applicable to the live network. Typical sample speech source is less than 10 second period, however, it is not applicable to cumulate the VoIP quality measurement. QoE should be employed in order to assess the user experience over the entire VoIP session.

Bibliography

[1] Antony W. Rix, John G.Beerends, Doh-Suk Kim, Peter Kroon and Oded Ghitza, "Objective Assessment of Speech and Audio Quality - Technology and Applications", IEEE Trans. Audio, Speech and Language Processing Vol. 14, NO 6, November 2006.

[2] Jun Tian and Abdelsalam (Sumi) Helal, "Rapid Mobility of Mobile IP over WLAN" 3rd International Conference, ICCNMC 2005, August , 2005. Proceedings pp 259-267

[3] Doh-Suk Kim, "ANIQUE: An auditory model for single-ended speech quality estimation", IEEE Trans. Speech Audio Process., vol 13, no. 5, pp. 821-831, Sep. 2005.

[4] ITU-T G.107 ITU-T Recommendation "The E-model, a computational model for use in transmission planning.", 2000

[5] ITU-T G.108 ITU-T Recommendation "Application of the E-Model: A planning guide", 1999.

[6] ITU-T G.109 ITU-T Recommendation, "Definition of Categories of Speech Transmission Quality", 1999

[7] ITU-T G.113 ITU-T Recommendation G.113 "General Characteristics of General Telephone Connections and Telephone Circuits - Transmission Impairments", Feb, 1996

[8] ITU-T G.114 ITU-T Recommendation, "One-Way Transmission Time", 1996.

[9] R. G. Cole and J. H. Rosenbluth. Voice over IP Performance Monitoring. ACM SIGCOMM, 2001.

[10] Lingfen Sun and Emmanuel C. Ifeachor, "Voice Quality Prediction Models and their Applications in VoIP Networks", IEEE Transactions on Multimedia, Vol. 8, No. 4, August 2006, pp. 809- 820.

[11] L.Ding, R.A. Goubran, "Speech Quality Prediction in VoIP Using the Extended E-Model", in Proc. of the IEEE Globecom 2003 Conference, San Francisco, December 2003, Vol.7, pp. 3974-3978.

[12] J. Li, C. Blake, D. S. J. De Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks", in Proceedings of the 7th ACM International Conference on Mobile Computing and Networking, (Rome, Italy), pp. 61- 69, July 2001.

[13] D. Hole and F. Tobagi, Capacity of an ieee 802.11b wireless lan supporting voip, in In Proceedings of ICC 2004.

[14] Ganguly, S.; Navda, V.; Kim, K.; Kashyap, A.; Niculescu, D.; Izmailov, R.; Hong, S.; Das, S.R., "Performance Optimizations for Deploying VoIP Services in Mesh Networks", Selected Areas in Communications, IEEE Journal on Volume 24, Issue 11, Nov. 2006 Page(s):2147 - 2158

[15] A. D. Clark, "Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality", IP-Telephony Workshop, Columbia University, 2001

[16] Doh-Suk Kim, Ahmed Tarfaf, "Enhanced Perceptual Model For Non-Intrusive Speech Quality Assessment", Proc. IEEE Int. Conf. Acous., Speech, Signal Processing, vol. 1, pp. 829 - 832, 2006.

[17] http://en.wikipedia.org/wiki/Correlation, Pearson product-moment correlation coefficient

[18] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", IETF RFC 3550, 2003

[19] R. Zopf, "Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)", IETF RFC 3389, 2002

[20] M. Handley, V. Jacobson, C. Perkins, "SDP: Session Description Protocol", IETF RFC4566, 2006

[21] A. Sollaud, "RTP Payload Format for the G.729.1 Audio Codec", IETF RFC 4749, 2006

[22] H. Schulzrinne, S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", IETF RFC3551, 2003

[23] Sun, L., Wade, G., Lines, B., and Ifeachor, E. Impact of packet loss location on perceived speech quality. In Internet Telephony Workshop 2001 (New York, Apr. 2001).

[24] A. Li , "RTP Payload Format for Enhanced Variable Rate Codecs (EVRC) and Selectable Mode Vocoders (SMV)" IETF RFC3558 June 2003

[25] ETSI TS 101 318 Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON); Using GSM speech codecs within ITU-T Recommendation H.323, 1998

[26] ETS 300 961: "Digital cellular telecommunications system (Phase 2+); Full rate speech; Transcoding (GSM 06.10)"

[27] ETS 300 969: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech transcoding (GSM 06.20)".

[28] ETS 300 726: "Digital cellular telecommunications system; Enhanced Full Rate (EFR) speech transcoding (GSM 06.60)".

[29] ETS 300 909: "Digital cellular telecommunications system (Phase 2+); Channel coding (GSM 05.03)".

[30] S.Park, D.Kim, S.Hong, "Comparative Analysis on VoIP Quality Measurements over Wireless Mesh Network", MSDL lab Internal report, 2007

[31] K.Medepalli, M.Chuah, S.Park, S.Jajoo, "MPLS-mux: an efficient protocol for providing QoS in IP based wireless networks" VTC 2003-Fall. 2003 IEEE 58th

[32] M. Chuah, K. Medepalli, S. Park and J. Wang, "End to end QoS in 3G IP based radio access networks,", Bell Labs Tech. Journal, July 2002.

[33] Q.Bi, P.Chen, Y.Yang, Q.Zhang, "An Analysis of VoIP Service Using 1 EV-DO Revision A System" IEEE JSAC in Communications, Vol 24, No.1, Jan 2006

[34] Iperf, http://en.wikipedia.org/wiki/Iperf

[35] Seyong Park Kyungtae Kim Kim, D.C. Sunghyun Choi Sangjin Hong, "Collaborative QoS architecture between DiffServ and 802.11e wireless LAN", IEEE VTC, 2003, Vol 2.

[36] Ian Marsh and Bjorn Gronall, "Performance evaluation of voice handovers in real 802.11 networks", The Second International Workshop On Wireless Network Measurement, 2006

[37] P.Galiotos, T.Dagiuklas, D. Arkadianos, "QoS Management for an enhanced VoIP platform using R-factor and Network Load Estimation functionality", High Speed Networks and Multimedia Communications 5th IEEE International Conference, 2002

[38] N.Savkoor, RSankar, and S.Azad, "Handoff Performance Issues in PCS Networks" 1998 IEEE

[39] H.Zhang, L.Xie, J.Byun, P.Flynn and Choon Shim, "Packet Loss Burstiness and Enhancement to the E-Model" IEEE, Proceedings of the Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks (SNPD/SAWN'05) 2005

[40] M. Ylianttila, R. Pichna', J. Vallstrom', J. Makela , A. Zuhedi', P. Krishnamurthy, and K. Pahlavan, "HANDOFF PROCEDURE FOR HETEROGENEOUS WIRELESS NETWORKS", Global Telecommunications Conference, 1999.

[41] Theo Pagtzis, Rajiv Chakravorty, Jon Crowcroft, "Proactive IP Mobility Management for Context-aware all-IP Wireless Access Networks", Proceedings of 6th IASTED/IEEE International Conference on Wireless and Optical Communications (WOC), Banff, Canada, July 2002

[42] Aladdin Saleh, "Mobile IP Performance and Interworking Architecture in 802.11 WLAN/CDMA2000 Networks", IEEE Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR04), 2004

[43] M. Broitman and S. Matantsev, "Determination of Optimal Size of the De-jitter Buffer of a Receiving Router", Automatic Control and Computer Sciences, 2008

[44] Arunesh Mishra, Minho Shin, William Arbaugh, "An Empirical Analysis of the IEEE 802.11 MAC Layer Handoff Process",ACM SIGCOMM Computer Communications Review (CCR) Volume 33, Issue 2 April 2003

[45] Christian Hoene, "Internet Telephony over Wireless Links", Ph.D thesis of Technischen Universitat Berlin, 2006

[46] MIKA YLIANTTILA, "VERTICAL HANDOFF AND MOBILITY SYSTEM ARCHITECTURE AND TRANSITION ANALYSIS", Ph.D thesis, University of Oulu 2005

[47] Enrique Stevens-Navarro and Vincent W.S. Wong, "Comparison between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks", IEEE VTC Spring 2006: 947-951

[48] Akira Takahashi, Atsuko Kurashima, Hideaki Yoshino, "Objective Assessment Methodology for Estimating Conversational Quality in VoIP", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 14, NO. 6, NOVEMBER 2006

[49] A. Takahashi, M. Masuda, and A. Kurashima, Methods of improving the accuracy and reproducibility of objective quality assessment of VoIP speech, IEICE Trans. Commun., vol. E87-B, no. 6, pp. 16601669, Jun. 2004.

[50] Ludovic Malfait, Jens Berger, and Martin Kastner, "P.563The ITU-T Standard for Single-Ended Speech Quality Assessment", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 14, NO. 6, NOVEMBER 2006

[51] Simon R. Broom, "VoIP Quality Assessment: Taking Account of the Edge-Device", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 14, NO. 6, NOVEMBER 2006

[52] Andrew Perkis, Solveig Munkeby and Odd Inge Hillestad, "A model for measuring Quality of Experience", NORSIG 2006

[53] Marcio Nieblas Zapater, Graa Bressan, "A Proposed Approach for Quality of Experience Assurance for IPTV", Proceedings of the First International Conference on the Digital Society ICDS, 2007

[54] LIU Li-yuan, ZHOU Wen-an, SONG Jun-de, "The Research of Quality of Experience Evaluation Method in Pervasive Computing Environment", 1st International Symposium on Pervasive Computing and Applications, 2006

[55] Bin Hong Lee Guan Yan Cai Yu Ge, Winston K.G. Seah, "VoIP Capacity over Wireless Mesh Networks", IEEE Conference on Local Computer Networks, Tampa, Florida, USA, 14-16 November 2006. IEEE Computer Society 2006