Optimizing Low Speed VoIP Network for Rural Next Generation Network (R-NGN)

Yoanes Bandung, Carmadi Machbub, Armein Z.R. Langi & Suhono H. Supangkat

School of Electrical Engineering and Informatics, Institut Teknologi Bandung
Jl. Ganesha 10 Bandung, 40132 Indonesia
email: bandung@dsp.ee.itb.ac.id, carmadi@lskk.ee.itb.ac.id, langi@lss.ee.itb.ac.id, suhono@lss.ee.itb.ac.id

Abstract. In this paper, we identify some impairments in the Voice over Internet Protocol (VoIP) networks and quantify the levels of voice quality. We propose an optimization model to maximize number of calls in low speed network for Rural Next Generation Network (R-NGN) while maintaining a minimum level of voice quality. The model is based-on Extended E-model to select parameters like voice coder, packet loss level, jitter buffer size, and network utilization. The optimum VoIP access network is achieved by applying G.723.1 5.3 kbps voice coder, packet loss level less than 1%, jitter buffer 50 ms, and network utilization less than 85%.

Keywords: *access network optimization; Extended E-model; voice over IP; voice quality.*

1 Introduction

The growth of Internet Protocol (IP) technology has a significant impact on global telecommunication today. Many people have used this technology for Voice over IP (VoIP) application as an alternative of the expensive traditional telephone usage. Such as in the old circuit switching technology of the Public Switch Telephone Network (PSTN), users can make calls anywhere and anytime as long as they connect to IP connection, but with low cost or sometimes free charge.

In developing countries such as Indonesia, rural communication becomes a strategic importance for people to support economic and social as well as government needs. It is well known that increasing telephone penetration results in increasing economic growth. Research on [1] proposes Rural Next Generation Network (R-NGN) that allows IP-based systems to be used in rural areas.

Figure 1 shows the typical configuration of the R-NGN. Each R-NGN has one community unit (CU) that is acting as its host and is located at the nearest PSTN central office. The CU controls and distributes the traffic to several access units

(AU) located preferably in towns. This configuration also defines community link by applying Ethernet radio communication to connect an access unit (AU) to a community unit (CoU) or nearest CoU from AU to remote CoU that has typical link speed between 128 and 256 kbps [1].
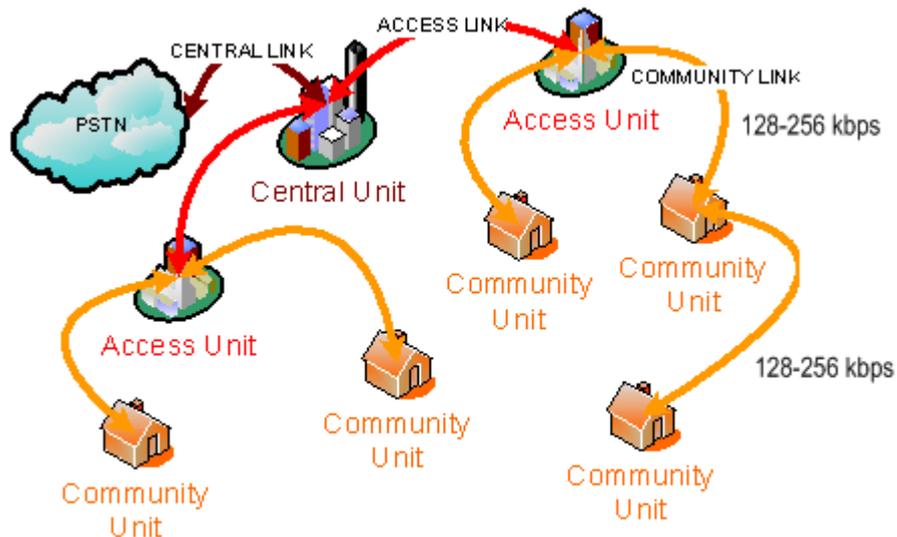
Figure 1 Typical configuration of R-NGN in an area

IP-based network as in R-NGN still lags in providing quality of service (QoS) for real-time applications. VoIP requires more strict quality of service than data applications such as email, World Wide Web (WWW), file transfer, and chat. Like traditional telephone, VoIP requires high availability of resources, good quality of voices, low delay, and low level of packet loss. If VoIP will be used as an alternative of telephone applications or will be integrated to traditional telephone network, then it should provide the same or better quality as in the traditional telephone network [2][3][4].

The implementation of VoIP in R-NGN also faces the limitation of resource capacity [5]. As described above, the resources of community link are typically between 128 and 256 kbps. These are low speed networks compared to the demand of users for multi-applications including email, World Wide Web (WWW), file transfer, chat, and VoIP applications. R-NGN must provide high quality multiservices including telephone, facsimile, messaging, email and Internet. The system is required to deliver video conferences and next generation multimedia applications in the future without major infrastructure changes [1].

In order to increase revenue of the implementation of VoIP services, we may have a question of how to maximize the number of calls while maintaining some levels of quality of service. This problem requires appropriate configuration of network access for VoIP applications.

There are three approaches for configuring VoIP network access to maintain the quality of service and to achieve an optimal network. The first method is *trial and error* but it is a very expensive and a time-consuming method. It cannot guarantee to achieve an optimal network. Second, rule of thumb method that sets certain values for quality of service requirements such as delay, jitter, and packet loss. This method may cause over provisioning and inefficient network. Third, analytical method based on E-model [6] is used to predict voice quality from calculation of some parameters. This method will not be valid when network characteristic is changed [7].

The next question is what method provides most efficient quality of VoIP applications? It is needed to understand their characteristics that are affected by voice coder, packet loss level, size of jitter buffer, and network utility. By understanding those characteristics of VoIP, an optimization method can be applied to achieve maximum VoIP calls with minimum level of quality of service.

The rest of this paper is described as follows. Section 2 describes some research works related to this research. In Section 3, we review E-model and Extended E-model. In this section, we describe analytical approach to find transmission rating factor *R* based on those models. In Section 4, we present optimization model of low speed network access and describe the result of the simulation. In Section 5, we describe the validation test in a testbed. Section 6 describes the conclusions.

2   Related Work

Work in [5] identifies two important characteristics of data transmission in IP network. The first characteristic is delay required to transmit data from a source to a destination. Typically, the delay consists of two components: i) fixed delay, which includes processing, transmission, & propagation delay, and ii) variable delay such as queueing delay. The second characteristic is the level of packet loss. Understanding those characteristics is important in order to configure the network properly.

There are other research projects that focus on measurement of quality of service based-on delay and packet loss parameters [3][5][8]. International Telecommunication   Union–Telecommunication   Sector   (ITU-T)

Recommendation G.114 [9] recommends delay (one way delay or latency) less than 150 ms for interactivity communications, and delay more than 400 ms will be unacceptable. Work in [10] describes that delay variation (jitter) less than 20 ms will be highly accepted, but it will be unaccepted if the delay variation is more than 50 ms. Packet loss level less than 10% is still tolerable [11].

Research work in [7] optimizes VoIP access network by selecting some parameters including voice coder, packet loss level, and network utilization. This work shows three scenarios: i) optimizing voice coder for given link capacity, packet loss level, and link utilization, ii) optimizing voice coder and packet loss level for given link capacity and link utilization, and iii) optimizing voice coder and link utilization for given link capacity and packet loss level.

In this research we extend previous work [7]. We want to optimize low speed VoIP access network for R-NGN. The main objective is to maximize the number of calls by selecting some parameters including low bit rate voice coders (G.723.1 5.3/6.3 kbps and G.729 8 kbps), packet loss level, jitter buffer size, and network utilization while providing a minimum level of voice quality.

We use the Extended E-model proposed in [12] for analyzing the voice quality degradation. We describe three scenarios of optimization: i) optimizing voice coder, ii) optimizing voice coder and packet loss level, iii) optimizing voice coder, jitter buffer size, and network utilization. Since the work in [7] focuses on 256 kbps and 1.544 Mbps link speed, we focus on low speed network as the typical link speed of the community link capacity. We are choosing 128 kbps and 256 kbps to represent the typical community link speed as proposed in [1].

We present an analysis of the delay, jitter, and packet loss parameters. The analysis of delay budget planning is based-on recommendation of Telecommunications Industry Association (TIA) [13], except for the jitter analysis that is caused by queueing delay. In this research, we model the queueing system as an M/G/1 system with Pareto service time to represent the self-similarity of VoIP traffic instead of traditional M/M/1 system used in [7]. We assume that the packet loss is the point on the tail of the waiting time distribution where packets are dropped.

3   Overview of E-model and Extended E-model

This section presents a brief introduction to E-model and Extended E-model. As described in the recommendation of ITU-T G.107, E-model is a computational model that combines all the impairment parameters into a total value. The E-model is not a measurement tool, but an end-to-end transmission planning tool.

The E-model assesses the combined effects of varying transmission parameters that affect the conversation quality of narrow band telephony. The E-model is based on the assumptions that transmission impairments can be transformed into psychological factors and psychological factors on the psychological scale are additive. The output of the E-model is called a transmission rating factor R that is defined in [6] as:

$$R = Ro - Is - Id - Ie + A \tag{1}$$

where Ro represents the effect of background and circuit noise, Is represents the impairments occurring simultaneously with the voice signal (quantization), Id represents the impairments caused by delay, and Ie represents the impairments caused by low bit rate voice coders and packet loss level. The advantage factor A can be used for compensation when there are other advantages of access to the user, e.g., the advantage factor for mobile telephony is assumed to be 10. Because there is no agreement has been reached for VoIP applications, we can drop the advantage factor from the R calculation [2].

In this research, we use a simplified analytic expression for the factor R as proposed in [14]. The factor R is a function of observable transport level quantities [2]. Both Ro (effect of background and circuit noise) and Is (effect of quantization) describe the impairment that has to do with the signal itself. Since none of them depends on the underlying transport network, we rely upon the set of default values as recommended for these parameters [6]. Choosing these default values, the factor R can be written as [2][14]:

$$R = 93.2 - Id - Ie \tag{2}$$

We replace the default values as described in [2][14] from 94.2 to 93.2 due to the revision of the ITU-T G.107 E-model Recommendation (see [6]). According to the recommendation, slight deviation should be considered insignificant for practical planning purposes.

The factor R can be transformed into a Mean Opinion Score (MOS) scale to indicate the subjectivity of voice quality as follows [6]:

$$MOS = \begin{cases} 1 & R < 0 \\ 1 + 0.035R + R(R-60)(100-R) \cdot 7 \cdot 10^{-6}, & 0 < R < 100 \\ 4.5 & R > 100 \end{cases} \tag{3}$$

A new formula is proposed in [12] to quantify the effects of packet loss and delay variation or jitter on speech quality in VoIP applications. The formula is called Extended E-model and incorporated into ITU-T G.107 E-model. It

extends the coverage of the current E-model for MOS prediction as well as network planning.

The Extended E-model includes *Ij* that represents the impairments caused by jitter. Because of its incorporation to the E-model, the output of the Extended E-model is also a transmission rating factor *R* that is defined as [12]:

$$R = Ro - Is - Id - Ie - Ij + A \tag{4}$$

Using the default values as described in the E-model recommendation, the simplified analytic expression for the factor *R* of the Extended E-model can be formulated as follows:

$$R = 93.2 - Id - Ie - Ij \tag{5}$$

In the following sections, we will focus on how to extract the delay impairment *Id,* packet loss impairment *Ie,* and jitter impairment *Ij.* These impairments are caused by some parameters that will be varied depending on the underlying transport network.

3.1   Delay Impairment *Id*

In this section, we will describe the delay of budget planning for calculating the delay impairment *Id.* The analysis is based on TIA recommendation. As shown in Figure 2, end-to-end delay or one-way delay consists of packetization delay, serialization delay, switching delay, queueing delay, propagation delay, decoding delay, and jitter buffer delay.

In this research, we investigate the M/G/1 system to derive the queueing delay instead of the M/D/1 queueing model proposed by TIA or the M/M/1 queueing model used in [7].

Internet traffic as well as VoIP traffic is widely known to have a self-similarity. The Pareto distribution is a suitable model for such traffic [12]. Research in [15] aims to find adequate stochastic models for the arrival process and holding time in VoIP network. The results show that the call arrivals are well modeled by Poisson process while the call holding times by generalized Pareto distribution. For such models, M/G/1 queueing model is established to analyze the queueing performance.
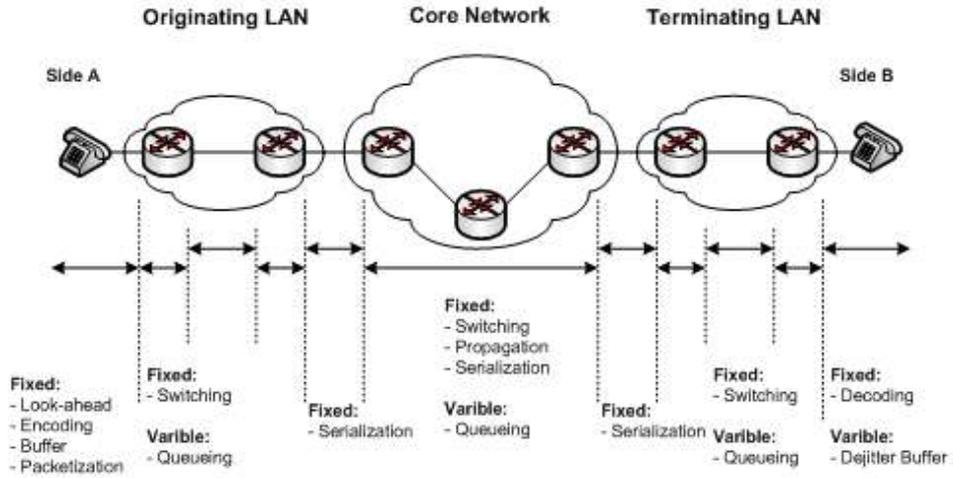
Figure 2 End-to-end delay components in VoIP networks.

We use one parameter (shape only) of Pareto distribution function defined over the nonnegative real numbers as [16]:

$$F(x) = 1 - 1/[(1+x)^{\beta}], x \geq 0 \tag{6}$$

where β is the shape parameter. The corresponding density function is:

$$f(x) = \beta/[(1+x)^{\beta+1}], x \geq 0 \tag{7}$$

The queueing delay is obtained from the average waiting time of a customer in the system $E[\tilde{w}]$ as follows [17]:

$$E[\tilde{w}] = \frac{\rho}{1-\rho} \frac{C_x^2 + 1}{2} E[\tilde{x}] \tag{8}$$

where ρ is traffic intensity, C_x^2 is squared coefficient of variation (SCV), and $E[\tilde{x}]$ is the first moment of the distribution.

The delay impairment factor, *Id,* represents all impairments due to delay of voice signals and consists of impairments due to Listener Echo, Talker Echo, and Absolute Delay.

$$d = Idte + Idle + Idd \tag{9}$$

Factor *Idd* represents the impairment caused by absolute one-way delay, *Ta*. Factor *Idle* relates to the average round trip delay in four wire-loop, *Tr*, while *Idte* relates to the average one-way delay from the receiver side to the point in the end-to-end path where a signal coupling occurs as a source of echo, *T*. The one-way delay *d* can be expressed as follows [18]:

$$d = Ta = T = \frac{Tr}{2} \tag{10}$$

Instead of using the complicated computation equations [6], a simplified equation is provided in order to obtain *Id* [18]:

$$Id = 0.024d + 0.11(d - 177.3)H(d - 177.3) \tag{11}$$

$$where \begin{cases} H(x) = 0, x < 0 \\ H(x) = 1, x \geq 0 \end{cases} \tag{12}$$

The work in [18] proposes a more accurate fit to the curve from [6] as follows:

$$Id = 1.618.10^{-12}d^6 - 1.765.10^{-10}d^5 + 6.447.10^{-8}d^4 - 8.221.10^6 d^3$$
$$- 0.0002315\, d^2 + 0.0352\, d - 0.0243 \tag{13}$$

3.2   Equipment Impairment *Ie*

ITU-T G.113 [19] provides guidance related to transmission impairments caused by voice compression and packet loss level. *Ie* values can be determined trough subjective tests according to ITU-T P.833 (2001). Network congestion will be a major problem that affects the level of packet loss [2].

Ding & Goubran apply one frame per packet and use random distribution to find *Ie* [12]. This impairment increases logarithmically with packet loss rate as follows:

$$Ie = Ie\_opt + C1.\ln(1 + C2.loss\_rate) \tag{14}$$

where *Ie_opt* is the optimum (without packet loss) *Ie* as defined in [19], *loss_rate* is the amount of packet loss in percent, and factors *C1, C2* are constants for adjusting the shape of the curve. Figure 3 shows the relation of factor *Ie* vs. packet loss level.
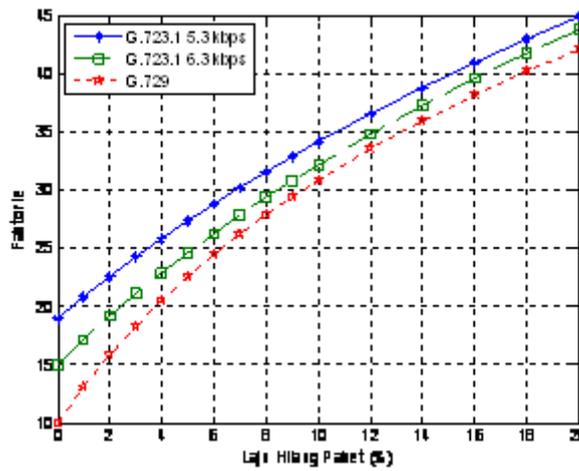
Figure 3 Factor *Ie* vs. packet loss level.

3.3   Jitter Impairment Ij

Ding & Goubran use Pareto distribution to model network delay that is caused by jitter [12]. The jitter impairment *Ij* is defined as:

$$Ij = C1 \cdot H^2 + C2 \cdot H + C3 + C4 \cdot e^{-Tj/K} \tag{15}$$

where   *C1, C2, C3, C4* are coefficients, *Tj* is a fixed buffer and *K* is a time constant. Figure 4 describes the relation of factor *Ij* vs. Hurst parameter when *Tj* = 60 ms.

4   Optimization of Low Speed VoIP Access Network

In this section, we will present an optimization model for maximizing the number of calls in low speed network capacity (128 and 256 kbps) while maintaining a minimum level of voice quality.

As stated above, our optimization model is based on selection of some parameters such as low bit rate voice coder (G.723.1 5.3 kbps and 6.3 kbps, and G.729), packet loss level, the size of jitter buffer, and network utilization. The formulation of the optimization model is defined as follows:

Maximize: Number of calls in the IP conection;
Subject to: R (coder, packet loss, jitter buffer, network utilization) >= 55;

The objective function of the optimization problem is to maximize the number of calls in the low speed network capacity. For the constraint, the minimum voice quality of the factor R is 55 or 2.8 in MOS measurement. We choose this value because we work with critical parameters of low speed networks that will produce low voice quality. By analyzing the voice quality in a testbed, we ensure that the minimum voice quality with factor $R=55$ is still satisfying for users in rural area. In this problem, the factor R is characterized by some parameters like voice coder, packet loss level, jitter buffer size, and network utilization.
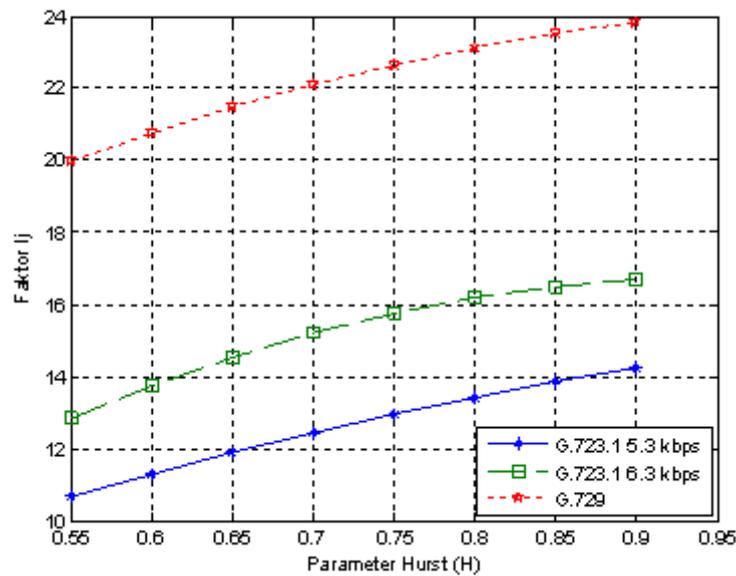
Figure 4   Factor Ij vs. Hurst parameter when $Tj = 60$ ms.

In the scenario 1, we select voice coder that is allowed in order to maximize the number of calls while maintaining guaranteed level of voice quality. By modifying the model as proposed in [7], an optimization model can be presented as follows:

Set :
       Coder
Parameters :
       Id{Coder}, Ie{Coder}, Ij{Coder}
Variables :
       Portion{Coder}, Feas{Coder} binary
Objective :

Maximize Calls : sum {i in Coder}
(Feas[i]*Portion[i]*Link*Util/Rate[i])
Subject to : #1
93.2 – (Id[i] + Ie[i] + Ij[i]) * Feas[i] ≥ 55
Subject to : #2
Sum{i in Coder} Portion[i] = 1

This problem is considered an "assignment" type optimization due to the number of calls that will be maximized with one of the *Set* combinations [7]. The variable *Portion* is used to assign the calls to a particular combination in the *Set*. Strict assignment would require the *Portion* to be a binary integer (1 or 0). In order to avoid this non-linearity, the program is allowed to make fractional assignments. Nevertheles, the assignment theorem ensures that the solution produced will always exhibit an assignment of 1 or 0 for every *Portion.*

The binary variable *Feas* is used to penalize coders that do not meet the constraints and limit the working set to $R \geq 55$ [7]. The program will simply switch the variable from 1 to 0, which eliminates voice coder from participating in the objective function. This is known as a penalty function that is non-linear. Being non-linear, the program will find the first coder that meets the constraints and will not look for others that could produce a better objective function. This problem can be solved by setting all variables to 1 during program initialization. The program will look at all variables and reverse them if they do not meet the constraints. For the model, we write equations as follows:

Maximize $a_1x_1x_4 + a_2x_2x_5 + a_3x_3x_6$
Subject to $c_1x_1 + c_2x_2 + c_3x_3 <= 38.2$
$x_4 + x_5 + x_6 = 1$
$x_1, x_2, x_3 \geq 0 \; ; \; 1 \geq x_4, x_5, x_6 \geq 0$

where a_i for $i = 1, 2, 3$ are the numbers of calls for G.723.1 5.3 kbps, G.723.1 6.3 kbps, and G.729 8 kbps. While c_i for $i = 1, 2, 3$ are the totals of impairment factors like *Id, Ie,* and *Ij*. The variable x_1, x_2, x_3 are *Feas* and x_4, x_5, x_6 are *Portion.*

The number of *calls* is obtained from this equation:

$$number\_calls = speed \cdot \frac{util}{coder\_rate} \qquad (16)$$

where *speed* represents the speed of network (kbps), *util* represents the utilization or traffic intensity (percent), and *coder_rate* represents the bit rate of

voice coder. We assumed that *coder_rate* of G.723.1 5.3 kbps was 20.8 kbps, G.723.1 6.3 kbps was 21.9 kbps, and G.729 8 kbps was 31.2 kbps.

The optimum solution is achieved by applying G.723.1 5.3 kbps voice coder. The maximum number of calls is 4.92 in 128 kbps link speed, and 9.86 in 256 kbps link speed.

In the scenario 2, we select voice coder and packet loss level allowed. The optimum solution is achieved by applying G.723.1 5.3 kbps voice coder and packet loss level being allowed is 1%.

In the scenario 3, we select voice coder and jitter buffer size. The optimal solution is achieved by applying G.723.1 5.3 kbps voice coder, jitter buffer size 50 ms, and network utilization 0.85.

5   Validation Test

Validation tests were run in a testbed to validate the optimization result based-on the Extended E-model. Figure 5 shows the configuration of the testbed that is part of campus network in Institut Teknologi Bandung (ITB). Two client terminals for running SIPp were set up in "IC Design Lab" at PAU Building as originating LAN and "Signal and System Lab" at Labtek VIII Building as terminating LAN. In these LANs, we run Firewall to emulate our network design. The emulation includes link capacity, network delay & jitter, and packet loss level. Between LANs is a high-speed core network that represents a backbone network in the VoIP system.
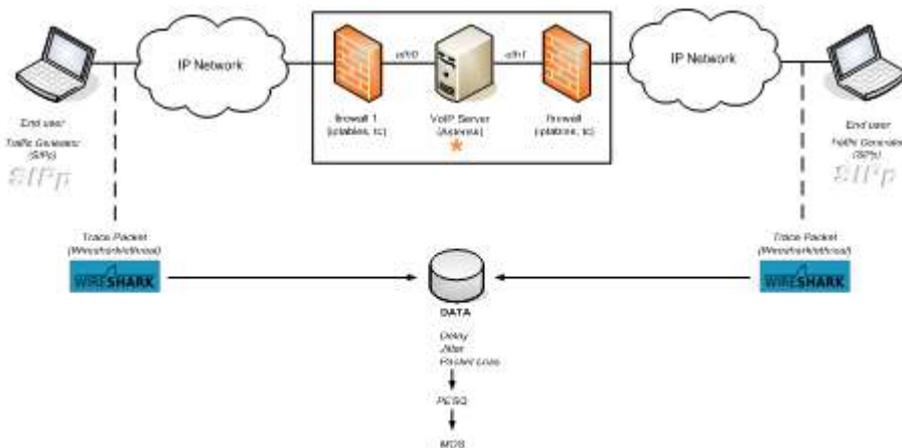
Figure 5 Testbed configuration for the validation test.

In the experiment, we measure the voice quality by applying G.729 8 kbps voice coder. Figure 6 shows the errors between the MOS value from simulation and those from experiment in 128 and 256 kbps link capacity. In 128 kbps link capacity, the mean error was 0.38 MOS and in 256 kbps link capacity, the mean error was 0.74 MOS. These results indicate that the simulation provides good accuracy.
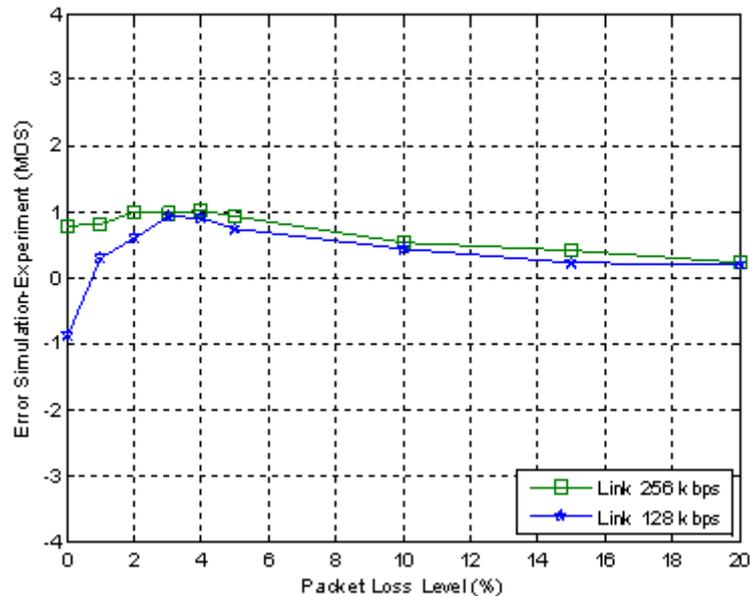
Figure 6 Error between simulation and experiment result.

6   Conclusion

This research presents an analysis of Voice over IP (VoIP) implementation parameters such as voice coder & jitter buffer size, and presents an analysis of network parameters such as delay, delay variation or jitter, packet loss level, and network utilization. Those parameters are used to optimize the VoIP access network for R-NGN.

The research aims to propose an optimization model of VoIP network access for achieving maximum calls especially in low network capacity, i.e. 128 and 256 kbps, while maintaining a guaranteed voice quality of the conversation. The optimization model uses Extended E-model to select VoIP implementation parameters like voice coder & jitter buffer size and network parameters like packet loss level and network utilization.

In general, the research shows that the optimum solution for the low speed network of the R-NGN will be achieved by applying G.723.1 ACELP 5.3 kbps voice coder rather than G.723.1 MP-MLQ 6.3 kbps and G.729 CS-ACELP 8 kbps, the packet loss level of less than 1%, jitter buffer around 50 ms, and network utilization of less than 0.85.

Acknowledgment

References

[1]   Langi, A., *A Rural Next Generation Network (R-NGN) and Its Testbed*, ITB Journal on ICT, **1 C**(1), pp. 1-15, 2007.
[2]   Karam, M.J. & Tobagi, F.A., *Analysis of the delay and jitter of voice traffic over the internet*, in Proc. of INFOCOM, pp. 824-833, 2001.
[3]   Markopoulou, A.P., Tobagi F.A. & Karam M.J., *Assessing the quality of voice communications over internet backbones*, IEEE/ACM Transactions on Networking, **11**(5), Oct. 2003.
[4]   Agnihotri, S., Aravindhan, Jamadagni, H.S. & Pawate, B.I., *A new tecnique for improving quality of speech in voice over IP using Time-Scale Modification*, in Proc. of ICASSP, **2**, pp. 2085-2088, 2002.
[5]   Bandung, Y. & Langi, A., *Softswitch-based telephony system for Rural Next Generation Network (R-NGN)*, in Proc. of TSSA, Bandung, 2004.
[6]   ITU-T Recommendation G.107, *The E-model, a computational model for use in transmission planning*, Mar. 2005.
[7]   Gardner, M.T., Frost, V.S. & Petr, D.W., *Using optimization to achieve efficient quality of service in voice over IP networks*, The 22nd International Performance, Computing, and Communications Conference (IPCCC), Phoenix, Arizona, 2003.
[8]   Jiang, W. & Schulzrinne, H., *Assessment of VoIP service availability in the current internet*, in Proceedings of the Passive and Active Measurement Workshop, La Jolla, CA, Apr. 2003.
[9]   ITU-T Recommendation G.114, *One way transmission time*, May 2003.

[10] Calyam, P., *Measurement challenges for VoIP infrastructures*, www.internet2.edu, 2002.

[11] Liang, Y.J., Farber, N. & Girod, B., *Adaptive playout scheduling and loss concealment for voice communication over IP networks*, IEEE Transaction on Multimedia, 2001.

[12] Ding, L. & Goubran, R., *Speech quality prediction in VoIP using the extended E-model*, Proc. of IEEE GLOBECOM, 2003.

[13] TIA-116, *Telecommunication Industries Association (TIA)*, Voice Quality Recommendations for IP Telephony, EIA/TIA/TSB-116, 2001.

[14] Cole, R. & Rosenbluth, J., *Voice over IP performance monitoring*, ACM Computer Communication Review, Apr. 2001.

[15] Dang, T.D., Sonkoly, B. & Molnar, S., *Fractal analysis and modeling of VoIP traffic*, in Proceedings of Networks 2004, pp. 217-222, 2004.

[16] Fischer, M.J. & Harris, C.M., *A method for analyzing congestion in pareto and related queues*, The Telecommunication Review, 1999.

[17] Daigle, J.N., *Queueing theory with applications to packet telecommunications,* Boston, Springer, 2005.

[18] Sun, L., *Speech quality prediction for voice over Internet Protocol networks,* Dissertation, University of Plymouth, 2004.

[19] ITU-T Recommendation G.113, *Transmission impairments due to speech processing,* Feb. 2001.