IJETCR
Open Access

ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)
Available Online at www.ijetcr.org
Volume 3; Issue 4; July-August-2015; Page No. 96-104

ENHANCING VOICE QUALITY THROUGH IMPROVED E-MODEL: A MATHEMATICAL ANALYSIS

Manish Kumar Giri[1], Girish Tiwari[2]

[1] student, Department of Electronics and Communication Engineering,

Ujjain Engineering College, UJJAIN, M.P., INDIA

manishgiri88@gmail.com

[2]Associate Professor, Department of Electronics and Communication Engineering,

Ujjain Engineering College, UJJAIN, M.P., INDIA

tiwari_girish@yahoo.com

Abstract

The E-model was suggested by ITU-T G.107 standard, which is an ever changing test of calculations to evaluate the network performance whether it is ready to send a VoIP call or not. Several studies have been performed in this field and it highlights the complexity of this model; since a lot of lengthy mathematical calculation is needed because of several parameters which are available in network and it depends upon them. So a need of more simplified version occurs for estimating the voice quality. In this paper, we have assumed simple correction by using standard values to a simplified E-model and calculated the MOS value by taking the default correction coefficients values for 4 basic codecs which are G.711, G.723.1, G.726 and G.729A, and then we show that its predictions mapping with PESQ scores by implementing it with simple calculation. This paper focuses on enhanced VoIP voice call quality evaluation based on simplified E-Model so that the end to end conversation phase identified correctly and organization can be made. In the end, we performed various mathematical calculations to show the variance in MOS with respect to packet loss and better mapping compare to simplified model.

Key Words: VoIP, E-model, PESQ, Voice quality, QOS.

1. INTRODUCTION:

The data transmission between two ends in a network depends on various parameters. So, it can be said that it is not efficient to use a single factor to calculate the voice call quality in a data packet transmission networks. Still in the communication field, the call quality rating is decided by a single number value. Such type value is a basis of call quality estimation and tuning in a network. Voice over Internet Protocol (VoIP) is a technology which supports data transmission in networks on multimedia domain [1]. In past some years, VoIP becomes an important application and it is believed that more data traffic will be generating over the TCP/IP networks. In real-time multimedia applications, the speech quality is affected by the packet loss, jitter, delay and bandwidth [1]. Therefore, it's requirement of VoIP applications to have low delay, low packet loss, low jitter and required bandwidth.

VoIP is based on Internet Protocol network; moreover services provided by the IP networks are good but may not guarantee delay, packet loss, and jitter [2]. So, it's highly required to measure the voice quality under these different environments condition and congestion effects, to prevent the voice quality from these critical problems before they occur. Since voice quality estimation is very important for the end to end users, ITU-T suggested two testing methods subjective and objective testing for evaluating the voice quality. Subjective testing was assumed to be the earliest attempts on this criterion to evaluate the speech quality by giving Mean Opinion Scores (MOS). The MOS test is one of the most popular tests that provide a speech quality rating. The procedure of MOS test was presented by ITU-T Rec. P.800 [3] as users can rate the speech quality from 1 (Poor) to 5 (Excellent) base scales. In general, the listener's numbers are a critical factor in estimating accurate scores. So, subjective testing by using MOS does not support real time estimation which is highly desirable criterion for quality of service these days and it is time consuming, expensive also. Furthermore, in past few years new techniques were developed by estimation in an objective approach (no human view considered) for measuring MOS scores which are: E-model [5] PESQ[4].

PESQ [4] (Perceptual Evaluation of Speech Quality) is speech quality measurement method based on objective analysis. It's an approach which compares two signals;

Corresponding author: Manish Kumar Giri

one is the reference signal and other one is the affected signal. Both signals are transmitted through some test having PESQ algorithm approach and result is PESQ score. Therefore, this method cannot be applied for monitoring real time multimedia calls.

A new objective approach suggested by ITU-T G.107 [5] describe the E-model, a mathematical equation based network planning tool that combines and relates all the parameters which are necessary in the estimation of voice quality and it also benefits by combining all parameters in a single factor R, called R value which can be used by researchers for calculating MOS depending upon scale defined. The estimation of network voice quality can be shown accurately by this approach and it is reasonably correct for it. But universal approval has not been given to it for valid measures of networks. Because it is only the estimation of transmission planning purposes not for the actual customer opinion [5]. Not in the case of PESQ [4] which is developed to analyze subjective tests commonly used in assessment of the voice quality by humans. The objective approach of E-model is in demand and widely deployed by users related to research field and quality assessment of voice is successfully analyzed. So suggestion regarding simplification of E-model came [1, 6] to reduce the more complex original E-model [5].

In this paper mapping between the PESQ MOS scores and improved simplified E-model MOS scores are provided by using a simplified version of the E-model corrected for 4 common codecs. The variance of MOS with respect to packet loss should be accurate.

2. RELATED WORK AND LITERATURE SURVEY:

As suggested in Assem et al. [15] we are using the same model for different range of values of parameters. This model has several improvements to the existing E-model [7].

A. Simplified E-model

Due to the complexity of original E-model it was highly required some simplification in that version and this simplification provided by [7]. Since easy evaluation of QoS is needed and suggestion for the simplified E-model came [6], so this model contain only the most important parameters required for calculation of QoS in the system. The simplified E-model has been shown in equation (1) through calculation of rating factor R.

$$R = R_0 - I_{codec} - I_{pl} - I_d \quad \text{.... (1)}$$

Where R_0 is the signal to noise ratio present in the system in the channel region, I_d represents the sender to receiver delays, I_{codec} is the codec relating factor depends on respective codecs and the I_{pl} is the packet loss rate in percentage within a specified time interval. This R value is used to calculate and mapping the MOS score.

B. Improved Simplified E-model

This model comes into picture for calculation of voice quality MOS rating by a simplified approach [8] of the previous E-model. A series of mathematical equations is a part of this version of model and network parameters are also included in it. The calculation itself containing several elements and can be expressed by the following equation (2)

$$R = R_y - I_d + A \quad \text{..... (2)}$$

Where R_y denotes PESQ score corrected second order functions [15] this standard objective method defined by ITU-T recommendation P.862 [4], I_d is the delay impairment factor or average delay time within a specified time interval and A is the advantage factor present in the communication system. The description and method of calculating these parameters is as follows.

1) R_y: A PESQ score corrected second order function representation to obtain better results as shown in [8]. R_y can be expressed by the following equation (3).

$$R_y = aR_x^2 + bR_x + c \quad \text{... (3)}$$

Where R_x itself a part of the simplified E-model which contains same parameters like (1), R_x can be obtained by the following expression (4) and a, b, c are codecs coefficients as shown in Table 1 and derived in section 3.

$$R_x = R_0 - I_e - I_{pl} \quad \text{.... (4)}$$

Table -1 Codec Special Coefficients [8]

Codec used	a	b	c
G.711	0.18	-27.90	1126.62
G.723.1	0.039	-4.2	166.61
G.726	0.046	-4.53	168.09
G.729A	0.063	-8.08	311.72

1.1) R_0: R_0 is the signal to noise ratio present in the system in the channel region. But calculation of R_0 is very difficult considering several noise sources. In regard of noise, ITU-T G.113 [8] provided the common value of R_0. Since, the downfall of voice occurs in the conversion process in network region so it back reduces the theoretical value to 93.2 [5]. So, we are also using the R_0 value equal to 93.2 in this paper.

1.2) I_e: I_e is the equipment impairment (codec quality) factor as defined in [8] and [9]. The distortion generated by different codec represents the codec distortion and further leads to voice distortion and impairments occurrence because of signal conversions. These days, its value is determined by taking the consideration from codec G.113 literature [8] and Table 2 is part of it.

Table- 2 Coding information and I_e value [8]

Encoder used	codec type	Bit rate (Kbit/sec)	I_e Value
PCM	G.711	64	0
ACELP	G.723.1	5.3	19
ADPCM	G.726	24	25
CS-ACELP	G.729A	8	11

1.3) I_{pl}: I_{pl} is the packet loss rate in percentage within a specified interval evaluated by a fix numbers of data packets. The percentage evaluated is the packet loss when senders' packet is not received by the receiver correctly or in order. It can be expressed by the following formula (5)

$I_{pl} = (1 - n/ds)$ .....(5)

Where n is the number of packets and ds is difference between highest and lowest priority sequences [15]. The calculation of ds can be done from equation (6).

$ds = ls - ss + 1$ .....(6)

Where, ls and ss are the highest priority and lowest priority sequences respectively.

2) I_d: Factors related to delay contributing to I_d taken from ITU-T G.107 [5]. If we considered perfect echo cancellation then the function I_d will be only a function of delay d and can be calculated from some complex calculation of equations as shown with plot of Figure 1. However this model shows good efficiency for one way delay less than 400ms as shown in Figure 1 (referenced from AT&T simplified model). It's found that the reasonable one way end to end delay should not cross 150 ms for having a good R score and hence provide a good call quality [11]. For more fitting values of delay over than 600 ms, a 6th order polynomial is derived as shown in the following equation (7). We have used this formula for calculation of I_d value in our paper.

$I_d = -2.468\times10^{-14}d^6 + 5.062\times10^{-11}d^5 - 3.903\times10^{-8}d^4 + 1.344\times10^{-5}d^3 - 0.001802d^2 + 0.103d - 0.1698$ ...... (7)

Where d is end to end delay or mouth to ear delay in ms. the response of I_d against the value of d is shown in fig 1.

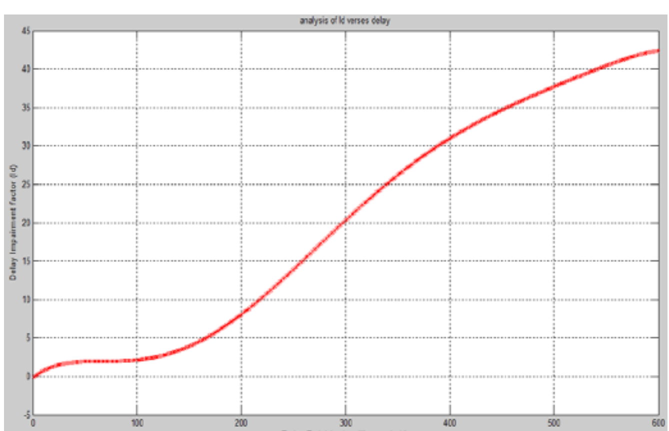

Fig -1: I_d versus one-way delay (d)

3) A: The entity advantage factor, A means advantage in accessing the network parameters, introduced into transmission planning for the E-model (ITU-T G.107) [5]. And further application of this factor can be used as an input value to the E-model. Temporarily A values are listed in [5] as show in Table 3. Assuming our communication system is conventional of wire bound type then we neglect A value.

Table -3: Examples for the Advantage Factor A [5]

Communication system	Maximum value of A
Wire bound	0
Cellular networks	5
Geographical areas	10
Multi hop satellite connections	20

The R value of the E-model is finally converted to MOS score that will further show the quality of service in user opinion as shown in Table 4, theoretical range of rating factor R from 0 to 100. R=0 represents of the lowest quality and R=100 represents the best quality. The R factor value conversion to MOS value can be easily done from the help of equation (8).

For $R<0$: $MOS=1$

For $0< R < 100$: $MOS = 1 + 0.035*R + R*(R-60)\,(100-R)*7*10^{-6}$

For $R>100$: $MOS=4.5$ ……. (8)

The respective mapping of MOS score with R factor is shown in fig 2.

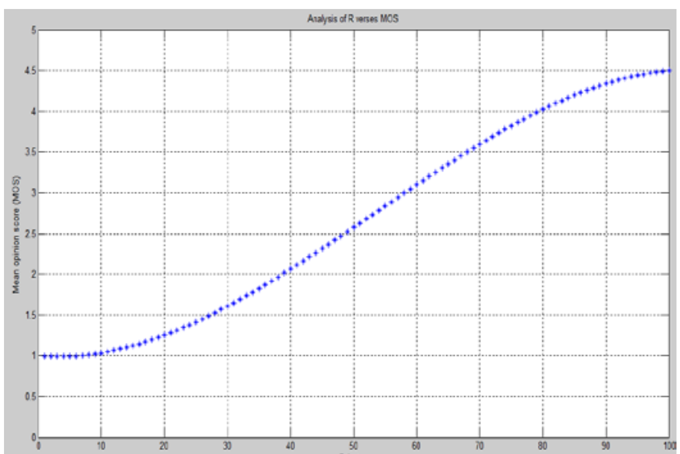

Fig-2: R factor vs. MOS plot

Table -4: R and MOS value ranges

R- value	Satisfaction level (user opinion)	MOS
90 to 100	Every user satisfied	4.5 plus
80 to 90	Lots of user Satisfied	4.0 to 4.5
70 to 80	Some user not satisfied	3.6 to 4.0
60 to 70	Many users not satisfied	3.1 to 3.5
50 to 60	Almost every user not satisfied	2.6 to 3.1
0 to 50	Not recommended	1.0 to 2.5

3. MATHEMATICAL CALCULATIONS

In this section we have calculated the values of R_x (which is part of simplified E-model scores) and R_y (which are corrected PESQ scores) for 4 different codecs as specified in the above section. For calculating the value of R_y we have taken the standard values of a, b, and c from table 1, moreover the factor R_x depends upon R_0, I_e and I_{pl}. The value of I_e is taken from standard defined values which are available in [8]. I_{pl} depends upon n and ds values which are number of packets and the difference between highest and lowest priority sequences. In our calculation we have assumed the value of n = 10, the ls value range considered from 11 to 15 and the ss value from 0 to 4 for simplification. This further gives the ds value range from 8 to 16. We have calculated the R_x and R_y scores for this 9 set of values and there results and plots are given below.

Table -5: R_x and R_y values for different codecs

	FOR n = 10 ds = ls-ss+1		G.711 CODEC		G.723.1 CODEC		G.726 CODEC		G.729A CODEC	
S.N.	DS	I_{pl}	R_x	R_y	R_x	R_y	R_x	R_y	R_x	R_y
1	8	-25	118.2	343.6632	99.2	133.7550	93.2	145.4610	107.2	169.5299
2	9	-11.11	104.311	174.8854	85.31	92.1448	79.31	98.1623	93.31	106.3049
3	10	0	93.2	89.8632	74.2	69.6900	68.2	73.1010	82.2	73.2249
4	11	9.0909	84.1091	53.3574	65.1091	58.4804	59.1091	61.0445	73.1091	57.7297
5	12	16.666	76.534	45.6632	57.534	54.0633	51.534	56.8055	65.534	52.7716
6	13	23.0769	70.1231	55.2904	51.1231	53.8233	45.1231	57.3427	59.1231	54.2244
7	14	28.571	64.629	75.3163	45.629	56.1667	39.629	60.8121	53.629	59.5906
8	15	33.333	59.867	101.4632	40.867	60.1033	34.867	66.0655	48.867	67.3183
9	16	37.5	55.7	131.038	36.7	64.9987	30.7	72.3735	44.7	76.4237

The conversion of MOS to R value of PESQ scores can be done by using a lengthy Candono Formula as in [12] or by the simplified 3rd-order polynomial [13] as shown in (9).

$$R = 3.026\times MOS^3 + 25.314\times MOS^2 + 87.060\times MOS - 57.336 \qquad \ldots\ldots (9)$$

The values of R_x and R_y can be mapped with each other and these plots are shown here.

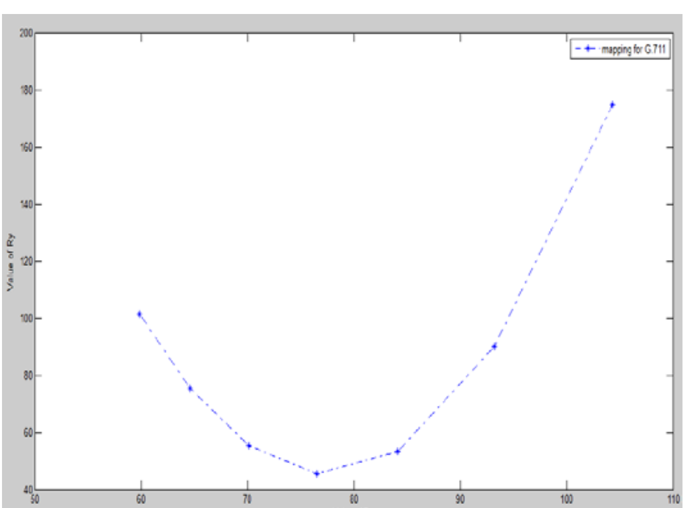

Fig-3: Relationship between R_x and R_y (G.711)

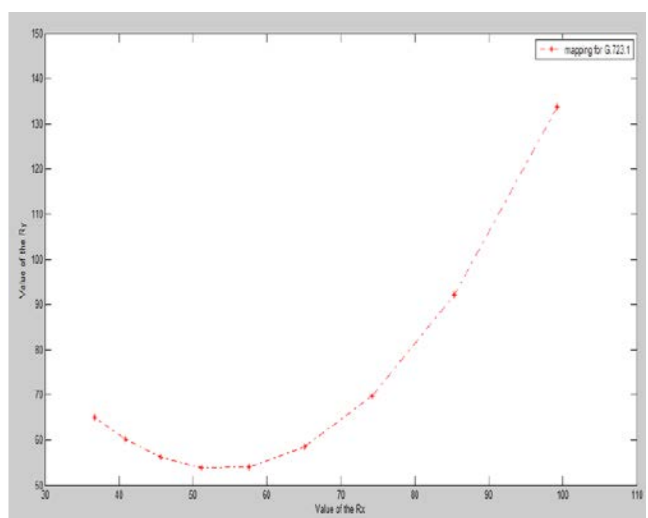

Fig-4: Relationship between R_x and R_y (G.723.1)

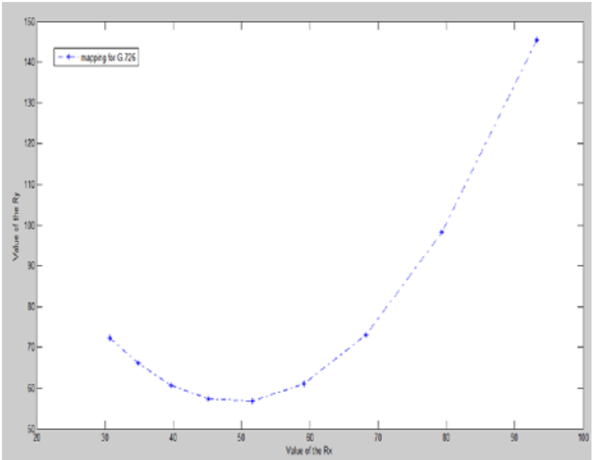

Fig-5: Relationship between R_x and R_y (G.726)

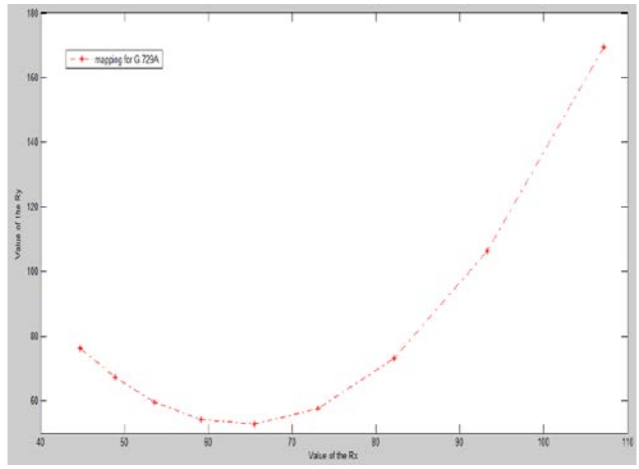

Fig-6: Relationship between R_x and R_y (G.729A)

4. SIMULATION AND RESULTS

In this section we have presented simulation plot between the PESQ MOS scores, simplified E-model and the improved simplified E-model which shows the mapping between these MOS values. The comparison between these scores shows that the improved model is better mapped with the standard PESQ MOS scores. For our understanding we have taken the packet loss percentage range between 0 to 20 % and plots between packet loss rate and respective MOS scores are given. The calculated values from our analysis are given below in tabular form for respective four types of codecs, which we have got through simulation.

Table -6: Comparison in MOS Values for G.711 Codec

S.N.	PACKET LOSS (%)	PESQ MOS	IMPROVED SIMPLIFIED E-MODEL MOS	SIMPLIFIED E-MODEL MOS
1	0	4.3356	4.2280	4.2952
2	2	3.9965	3.8381	4.2382
3	4	3.6033	3.4159	4.1757
4	5	3.4093	3.2134	4.1424
5	6	3.2250	3.0237	4.1079
6	10	2.6456	2.4414	3.9582
7	20	2.5133	2.3112	3.5163

Table -7: Comparison in MOS Values for G.723.1codec

S.N.	PACKET LOSS (%)	PESQ MOS	IMPROVED SIMPLIFIED E-MODEL MOS	SIMPLIFIED E-MODEL MOS
1	0	3.5824	3.3939	3.6111
2	2	3.4372	3.2423	3.5163
3	4	3.3027	3.1033	3.4191
4	5	3.2401	3.0391	3.3696
5	6	3.1809	2.9786	3.3196
6	10	2.9809	2.7759	3.1156
7	15	2.8200	2.6145	2.8542
8	20	2.7613	2.5560	2.5908

Table -8: Comparison in MOS Values for G.726 Codec

S.N.	PACKET LOSS (%)	PESQ MOS	IMPROVED SIMPLIFIED E-MODEL MOS	SIMPLIFIED E-MODEL MOS
1	0	3.7389	3.5594	3.3196
2	2	3.5874	3.3992	3.2184
3	4	3.4465	3.2519	3.1156
4	5	3.3809	3.1840	3.0637
5	6	3.3191	3.1203	3.0116
6	10	3.1133	2.9098	2.80115
7	15	2.9585	2.7533	2.5383
8	20	2.9232	2.7179	2.2794

Table -9: Comparison in MOS Values for G.729A Codec

S.N.	PACKET LOSS (%)	PESQ MOS	IMPROVED SIMPLIFIED E MODEL MOS	SIMPLIFIED E-MODEL MOS
1	0	3.7444	3.5652	3.9582
2	2	3.5461	3.3558	3.8769
3	4	3.3608	3.1632	3.7917
4	5	3.2749	3.0748	3.7477
5	6	3.1943	2.9923	3.7030
6	10	2.9306	2.7252	3.5163
7	15	2.7457	2.5405	3.2692
8	20	2.7267	2.5216	3.0116

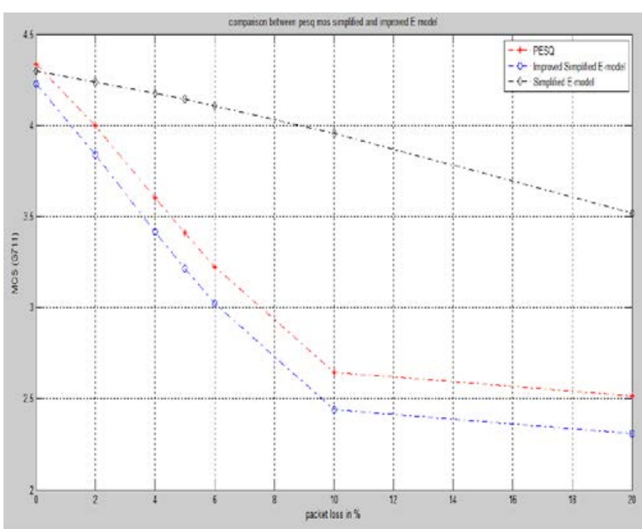

Fig-7: Comparison between different MOS scores (G.711)

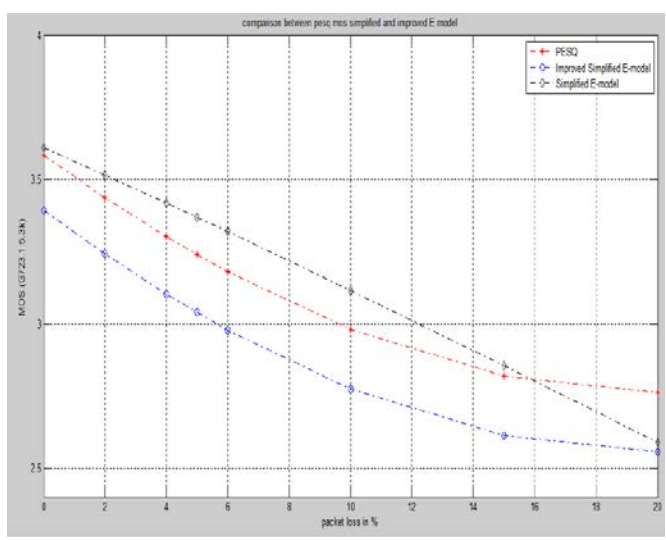

Fig-8: Comparison between different MOS scores (G.723)

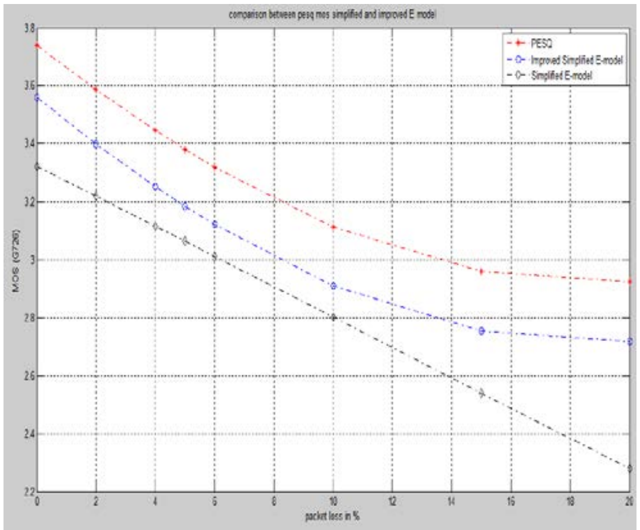

Fig-9: Comparison between different MOS scores (G.726)

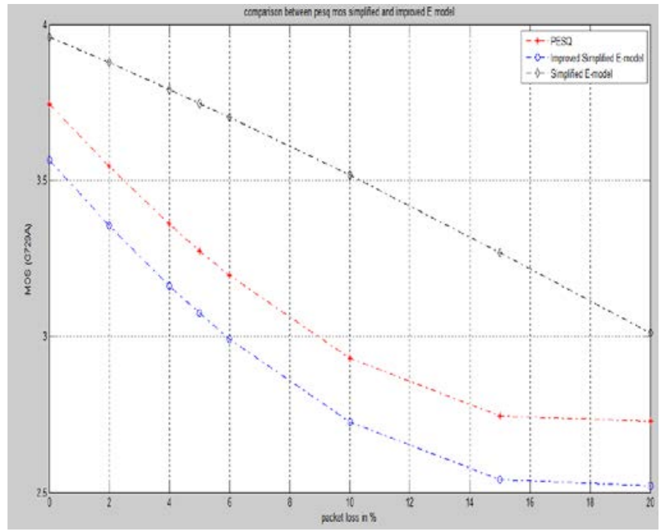

Fig-10: Comparison between different MOS scores (G729A)

4. CONCLUSION

In this paper some consideration has been used from us for evaluating the E-model parameters which further applied to estimation of speech quality assessment. The primary benefits of deploying E-model that it is based on objective method which applicable in real time domain. While different approach came from ITU-Recommendation, simplified versions of E-model have been introduced by researchers and industry to be used for voice calculation and prediction of call quality. Consequently, we have taken help from an improved simplified E-model which already used in past works and show how we mapped the R factor and MOS values while using help of standard default values in the model for 4 common codecs we used are G.711, G.723.1, G.726 and G.729A. We highlighted its results by implementing mathematical equations and calculated for several values of packet loss rate and analyze the impact of voice quality enhancement factors under several network conditions and use the simplified improved E-model to assess voice quality. The main important advantage of the improved simplified version that, it is less complex than the original E-model model and it is more accurate than the simplified versions used; we considered some benefits from our calculations; the first as confirmed by the calculation results, that the accuracy found by simplified version of E-model is less than PESQ based scores. The second, the correction coefficients value which are taken from standard values can be also derived to further improve the simplified E-model to predict the call quality. The third, implementation of the results (MOS scores) using the improved simplified E-model for 4 common codecs.

REFERENCES

1. John Q. Walker, “Assessing VoIP Call Quality Using the E-model”, NetIQ Corporation.
2. Junsheng Zhang and Xiaohua Sun, “The VoIP phone QoS protection in the wide-are network”, Computer learning .2006 No.6 .17-18.
3. ITU-T Recommendation P.800, “Methods for subjective determination of transmission quality”, Geneva, 08/1996.
4. ITU-T Recommendation P.862, “Perceptual evaluation of speech quality(PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs”, Febrauary 2001.
5. ITU-T Recommendation G.107, “The E-model: A computational model for use in transmission planning”, Geneva, 04/2009.
6. Chunlei Jiang and Peng Huang, “Research of Monitoring VoIP Voice QoS”, International Conference on Internet Computing and Information Services, 2011.
7. Pystechnics Limited, “The E-Model, R Factor and MOS”, 23 Museum street Ipswitch, Suffolk United Kingdom. December 2003. 10-41.
8. ITU-T Recommendation G.113, “Transmission impairments due to speech processing”, 2001.
9. ITU-T Recommendation P.833, “Methodology for derivation of equpment impairment factors from subjective listening-only tests”, 2001.
10. R.G.Cole and J.Rosenbluth, “Voice over IP performance monitoring”, ACM comput. Commun. Rev., vol. 31, no. 2,pp. 9-24, April 2001.
11. S.Pracht and D.Hardman: “Voice Quality in Converging Telephony and IP Networks, Agilent

Technologies", White Paper, http://literature.agilent.com/litweb/pdf/5980-0989E.pdf

12. C. Hoene, H. Karl, and A. Wolisz, "A perceptual quality model for adaptive VoIP applications", Int. Symp. Performance Evaluation of Computer and Telecommunication Systems(SPECTS'04), SanJose, CA.
13. L. Sun, "Speech Quality Prediction for voice Over Internet Protocol Networks", Ph.D dissertation, Univ. Plymouth, UK., Jan 2004.
14. Marta Carbona and Luigi Rizzo, "Dummynet Revisited", ACM SIGCOMM Computer Communication Review Volume 40 Issue 2, April 2010,12-20.
15. Assem et al., "Monitoring VoIP call quality using improved simplified E-model', 2013 International Conference on Computing, Networking and Communications, Communication QoS and System Modeling Symposium.