E-model Modification for Case of Cascade Codecs Arrangement

Miroslav Voznak

***Abstract*—**Speech quality assessment is one of the key matters of voice services and every provider should ensure adequate connection quality to end users. Speech quality has to be measured by a trusted method and results have to correlate with intelligibility and clarity of the speech, as perceived by the listener. It can be achieved by subjective methods but in real life we must rely on objective measurements based on reliable models. One of them is E-model that we can consider as mainly adopted method in IP telephony. This method is based on evaluation of transmission path impairments influencing speech signal, especially delays and packet losses. These parameters which are common in IP network can affect dramatically speech quality. In this article, a new modification of E-model, that takes into consideration the cascade codecs arrangement, is presented. The proposed a correction function improves the current computational non-intrusive approach that is described in recommendation ITU-T G.107, so-called E-model.

***Keywords*—**Cascade codecs arrangement, E-model, MOS, PESQ method, R-factor, Speech quality.

I. INTRODUCTION

THE methodologies to evaluate speech quality can be sub-divided into two groups according to the approach applied, conversational and listening. Conversational tests are based on mutual interactive communication between two subjects through the whole transmission chain of the tested communication system. These tests provide the most realistic testing environment but are they are very time consuming. Listening tests do not provide such plausibility as conversation tests but are recommended more frequently [1]. According to methods of assessment, speech quality evaluation methodologies can be subdivided as subjective methods and objective methods. To evaluate speech quality, MOS (Mean Opinion Score) scale as defined by the ITU-T recommendation P.800 is applied [7].

II. MOS SCALE

The basic scale as prescribed by the recommendation is depicted on Fig. 1. In order to avoid misunderstanding and incorrect interpretation of MOS values, ITU-T published ITU-T recommendation P.800.1 in 2003 [8].

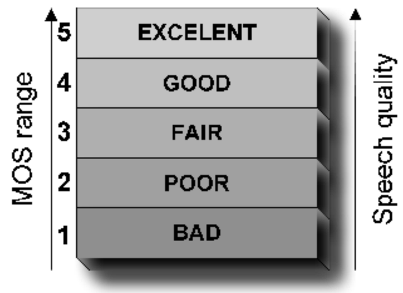

Fig. 1. MOS Scale.

This recommendation defines scales both for subjective and objective methods as well as for individual conversational and listening tests [7], [8].

A. Subjective Evaluation Methods

These methods are based on evaluation by human beings (listeners), i.e. subjects. During the testing, samples are played to a sufficient number of subjects, and their results are subsequently analysed statistically. Subjects can evaluate the speech quality on a five-degree scale in accordance with the MOS model as defined by ITU-T.

The best known representatives of these measurements include methods such as ACR (Absolute Category Rating) or DCR (Degradation Category Rating). Major disadvantages of these methods are high requirements on time, final evaluation being influenced by listener's subjective opinion and most of all impossibility to use them for testing in real time [7], [10].

B. Objective Evaluation Methods

The use of objective methods substitutes the necessity to involve humans in the testing by mathematical computational models or algorithms. Their output is again a MOS value or, depending on the algorithm applied, a different value which can be transferred to a MOS value using a suitable mapping function. The aim of objective methods is to estimate, as precisely as possible, the MOS value which would be obtained by a subjective evaluation involving sufficient number of evaluating subjects. Objective testing's exactness and efficiency is therefore a correlation of results from both subjective and objective measurements [1], [10]. Objective methods can be sub-divided into two groups, Intrusive and Non-intrusive.

Manuscript received August 21, 2011. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 218086.

Miroslav Voznak is an associate professor with the Department of Telecommunications, VSB –Technical University of Ostrava, Ostrava, Czech Republic (phone: +420 597321699; e-mail: miroslav.voznak@vsb.cz).

C. Intrusive Approach

The core of intrusive (also referred to as input-to-output) measurements is the comparison of the original sample before releasing it into a transmission chain of a communication system with the output degraded sample, transmitted through the system [1], [10].

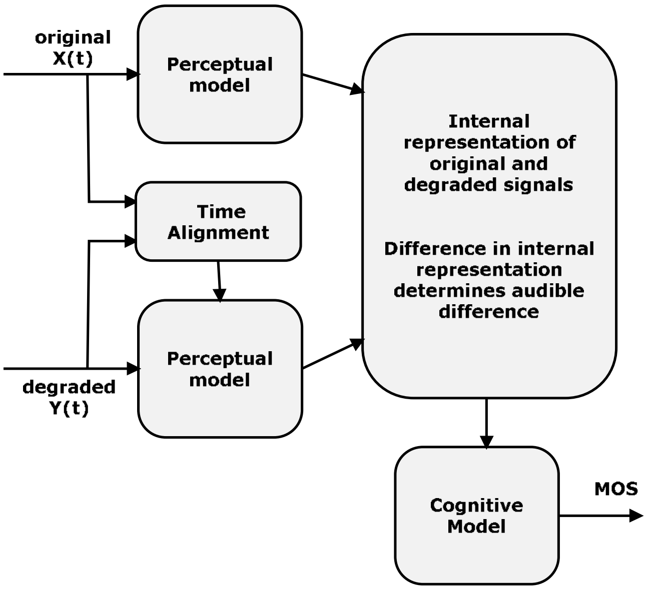

Fig. 2. The basic philosophy used in PESQ.

This type of testing includes, among other, the following methods: PSQM (Perceptual Speech Quality Measurement), PAMS (Perceptual Analysis Measurement System) developed by British Telecommunications and PESQ (Perceptual Evaluation of Speech Quality) [1], [10]. PESQ is the most common and most elaborate objective intrusive method. Computational technique applied by this method combines PAMS' robust temporal alignment techniques and the PSQM' exact sensual perception model. Its final version is contained in ITU-T recommendation P.862 [1], [9], [10], [15]. The basic philosophy of PESQ approach is depicted on Fig.2, as was stated above, the princip of this intrusive test is based on comparison of original and degraded signals, their mathematical analysis using FFT and interpretation in the cognitive model.

D. Non-Intrusive Approach

Contrary to intrusive methods which need both the output (degraded) sample and the original sample, non-intrusive methods do not require the original sample. This is why they are more suitable to be applied in real time. Yet, since the original sample is not included, these methods frequently contain far more complex computation models. Examples of these types of measurements frequently use INMD (in-service nonintrusive measurement device) that has access to transmission channels and can collate objective information about calls in progress without disrupting them. These data are further processed using a particular method, with a MOS value as the output [1], [10]. The method defined by ITU-T recommendation P.563 or a more recent computation method E-model defined by ITU-T recommendation G.107 are examples of such measurements [3].

III. E-MODEL

Complexity of modern networks requires that individual parameters of the transmission path are not assessed separately but rather that all their possible combinations and their interaction are considered. This can partially be achieved by an expert estimate based on the parameters of the transmission path, yet using a computation model is a more systematic approach. The E-model is a computation model which takes into account all the links between transmission parameters. Its output is a scalar labelled R which is a function of total expected call quality. The E-model is based on the "equipment impairment factor" method. The original structure of this model was developed by Swedish expert Nils-Olof Johannesson, member of the Voice Transmission Quality from Mouth to Ear group under ETSI. This model was further developed by the SG12 group under ITU-T and it was published in ITU-T recommendation G.107 as the E-model [1], [3], [13]. The structure of the connection reference model is depicted on Fig. 3.

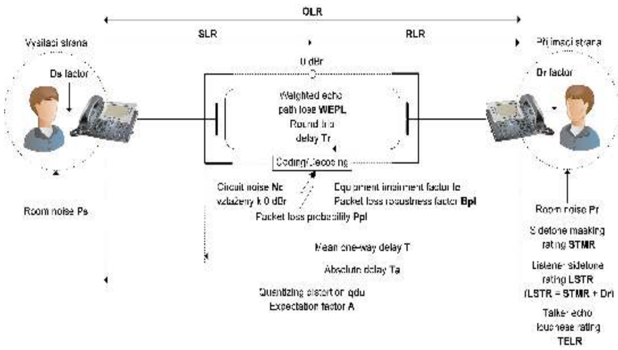

Fig. 3. Reference model for E-model computation.

The computational model consists of various mathematical operations over all parameters of the transmission system. The computation itself can be split into several elements and can be expressed by the following equation (1).

$$R = R_o - I_s - I_d - I_{e-eff} + A \tag{1}$$

Ro represents the basic delay of the signal from the noise which includes all types of noise, such noises caused by the device's electrical circuit and noises arisen on the wiring. *Is* comprises all possible impairments combinations that appear more or less simultaneously with a useful voice signal. Factor *Id* represents all impairments which are caused by different combinations of delays. This impairment factor is expressed by relation (2) where is divided into the three factors.

$$I_d = I_{dte} + I_{dle} + I_{dd} \quad (2)$$

The factor *Idte* gives an estimate for the impairments due to talker echo and the factor *Idle* represents impairments due to listener echo. Echo-cancellation is solved in recommendation ITU-T G.168 and can be effective suppressed. The factor *Idd* represents the impairment caused by too-long absolute delay *Ta*, which occurs even with perfect echo cancelling.

For Ta ≤ 100 ms we can assume *Idd = 0* because a negligeable influence appears in the R-factor but with increasing delay the overall R-factor is affected. This situation is depicted on Fig. 4.

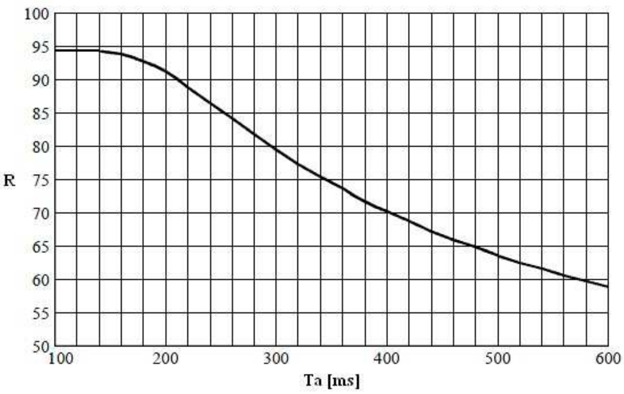

Fig. 4. Dependency of R-factor on absolute delay.

$Ie-eff$ comprises impairments caused by using a particular voice codec, occurrence of packet loss and its resistance against losses. Specific impairment factor values for codec operation under random packet-loss have formerly been treated using tabulated, packet-loss dependent *Ie*-values. Now, the packet-loss robustness Factor *Bpl* is defined as a codec-specific value. The packet-loss dependent effective equipment impairment factor *Ie-eff* is derived using the codec-specific value for the equipment impairment factor at zero packet-loss *Ie* and the packet-loss robustness factor *Bpl*, both listed in Appendix I of ITU-T G.113 for several codecs. With the packet-loss probability *Ppl, Ie-eff* is calculated using the equation (3).

$$I_{e-eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}} \quad (3)$$

BurstR is the so-called burst ratio, which is defined as ratio between „Average length of observed bursts in an arrival sequence“ and „Average length of bursts expected for the network under random loss“. When packet loss is random *BurstR* = 1 and when packet loss is bursty *BurstR* > 1. For packet loss distributions corresponding to a 2-state Markov model with transition probabilities *p* between a found and a loss state, and *q* between the loss and the found state, the burst ratio can be calculated as *1/(p+q)*. The packet-loss *Ppl* is expressed in relation (4) then *BurstR* can be calculated as (5).

$$P_{pl} = 100 \cdot \frac{p}{p+q} \quad (4)$$

$$BurstR = \frac{1 - P_{pl}}{q} \cdot 100 \quad (5)$$

As can be seen from equation (3), the effective equipment impairment factor in case of *Ppl* = 0 (no packet-loss) is equal to the *Ie* value defined in Appendix I of [ITU-T G.113].

Last, parameter *A* slightly adjusts the final quality depending user's concentration [4], [5]. The value of conventional (wirebound) communication system is A=0, mobility by cellular networks in a building A=5, Mobility in a geographical area or moving in a vehicle A=10 and access to hard-to-reach locations, e.g., via multi-hop satellite connections A=20. It should be noted that the values above are only provisional. The use of the factor A and its selected value in a specific application is up to the planner's decision. Additional background information on the advantage factor A can be found in Appendix II to ITU-T G.113.

For all input parameters used in the algorithm of the E-model, the default values are listed in recommendation ITU-T G.107 and it is strongly recommended to use these default values for all parameters which are not varied during planning calculation. If all parameters are set to the default values, the calculation results in a very high quality with a rating factor of R = 93.2.

IV. Experiment with Codecs Tandeming

Two kinds of experimental measurements were carried out – simulation using the E-model and, to enable comparison, measurement using the objective intrusive method PESQ.

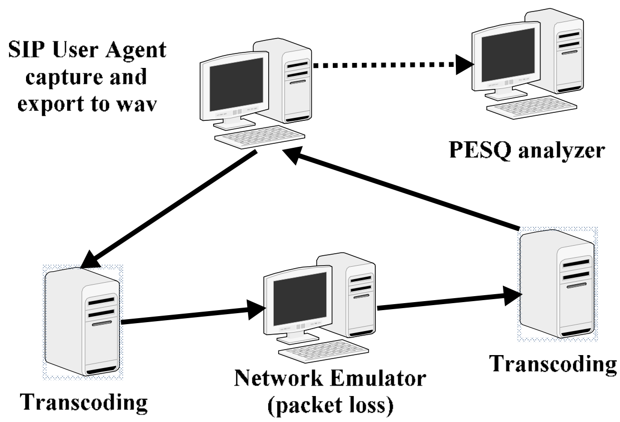

Fig. 5. Measuring testbed.

Measurements reflected the impact of codec tandeming, i.e. use of several voice codecs consecutively along the transmission path. In addition, the impact of IP telephony packet losses in one of the segments on the transmission path on the overall call quality was also taken into account. Fig. 5 shows the structure of the measuring testbed in which measurements using the PESQ method were carried out [14].

Voice codec G.711 A-law was applied between the SIP User Agent's and transcoding gateway. Between the two transcoding gateways, codecs G.711 µ-law, G.729, G.726 and G.723.1 (ACELP) were applied in sequence. In addition, packet loss ranging from 0 – 10% with a 0.5 % step was simulated at Network emulator.

A. Simulation in E-model

In order to simulate quality using the E-model, the author of the article developed an application in Java in accordance with ITU-T recommendation G.107 (04/2009) [3]. In addition to parameters defined by the recommendation, the application can also determine the final quality depending on codec tandeming. The chapter dealing with the E-model indicates that the impact of codecs is covered by parameter $Ie-eff$. In the E-model, the overall impact of codecs in cascade is therefore defined as a sum of partial impacts $Ie-eff$ for individual parts of the transmission path. This can be expressed by equation (6). By applying the prior equation, we can determine the impact for each part of the transmission path separately in equation (7).

$$I_{e-ef}f = I_{e-eff\,1} + I_{e-eff\,2} + \ldots + I_{e-eff\,n} \qquad (6)$$

$$I_{e-eff_n} = I_{e_n} + \left(95 - I_{e_n}\right) \cdot \frac{P_{pl_n}}{\dfrac{P_{pl_n}}{BurstR} + B_{pl_n}} \qquad (7)$$

The computational application using the E-model developed by author is currently in version 1.3.

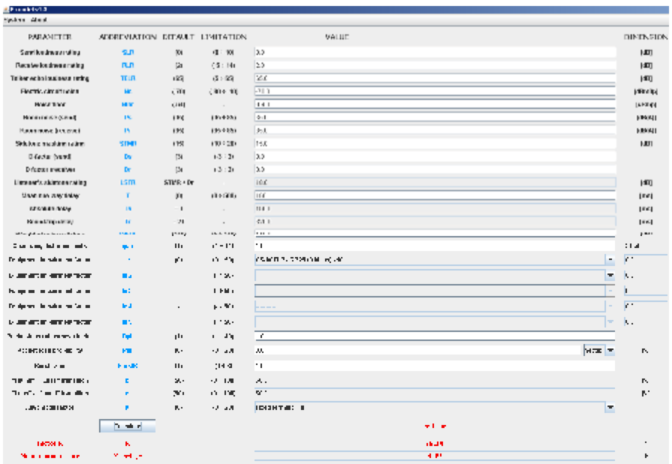

Fig. 6. E-model v1.3 application's main window.

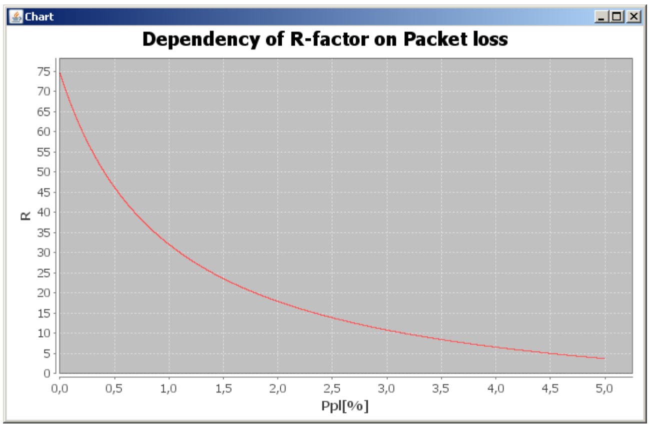

Fig. 7. Dependency of R-factor on PL.

It also contains a function visualising the relations between R and MOS value and various parameters in a chart. The chart can be further amended as required. Fig. 6 is an example of this application's main window.

The example of codec tandeming for three codecs G.711 A-law, G.729 and again G711 A-law is depicted on Fig. 7. The dependency of R-factor on packet loss is expressed for one-way delay 80 ms and packet loss range 0-5 %, this chart function is included in the menitoned application.

B. Evaluation using PESQ

The experimental testbed for evaluation using PESQ method was established in the IP telephony laboratory in Ostrava and was set up as shown on Fig. 5. The SIP User Agent comprised both transmitting and receiving section. The User Agent consisted of SPITFILE software system, SPITFILE was developed at our department [11], [12]. This application ran under Linux (Ubuntu 9.10) on a virtual machine created by VMware. Transcoding Gateways were Cisco Systems routers equipped with Voice IOS in which transcoding was carried out. The Network was simulated by the NE1000 component from US company Simena which enables to emulate different network parameters. In our case, it emulated loss of transmitted packets.

C. E-model's Approach to Codecs Tandeming

Having studied ITU-T resources, we can assume that the SG12 working group which developed the E-model, did not concern itself with codec tandeming. This was the impetus to carry out this experiment. To establish to what extent the E-model in this particular case converges to reality, it was necessary to adopt a suitable method for comparison. The standard objective method is the evaluation according to PESQ defined by ITU-T recommendation P.862 [9].

V. Results and Correction Function Draft

The measurements were carried out for about one month in the IP telephony laboratory at the Department of Telecommunications in Ostrava. Measurements were carried out six times for each codec and each loss value in order to eliminate possible statistical variances and errors. Thus, more than 1 200 values were obtained. Tables below (1–5) provide an overview of results obtained using the E-model and PESQ for individual codecs in individual loss ranges.

Table 1. Results for codec G.711 A-law.

Packet loss [%]	**G.711 A-law**		
	E-model [MOS]	**PESQ [MOS]**	**Difference [MOS]**
0	4.41	4.10	0.31
0.5	4.14	3.95	0.19
1	3.83	3.82	0.01
1.5	3.53	3.69	0.16
2	3.26	3.57	0.31
2.5	3.01	3.46	0.45
3	2.79	3.35	0.56
3.5	2.61	3.24	0.63
4	2.44	3.15	0.71
4.5	2.30	3.06	0.76
5	2.17	2.98	0.81
5.5	2.06	2.90	0.84
6	1.96	2.83	0.87
6.5	1.87	2.76	0.89
7	1.80	2.71	0.91
7.5	1.73	2.66	0.93
8	1.67	2.61	0.94
8.5	1.61	2.57	0.96
9	1.56	2.54	0.98
9.5	1.52	2.52	1.00
10	1.48	2.50	1.02

Table 2. Results for codec G.711 μ-law

Packet loss [%]	**G.711 μ-law**		
	E-model [MOS]	**PESQ [MOS]**	**Difference [MOS]**
0	4.41	4.10	0.31
0.5	4.14	3.98	0.16
1	3.83	3.85	0.02
1.5	3.53	3.73	0.20
2	3.26	3.61	0.35
2.5	3.01	3.48	0.47
3	2.79	3.36	0.57
3.5	2.61	3.23	0.62
4	2.44	3.10	0.66
4.5	2.30	2.97	0.67
5	2.17	2.84	0.67
5.5	2.06	2.71	0.65
6	1.96	2.58	0.62
6.5	1.87	2.45	0.58
7	1.80	2.31	0.51
7.5	1.73	2.17	0.44
8	1.67	2.04	0.37
8.5	1.61	1.90	0.29
9	1.56	1.76	0.20
9.5	1.52	1.62	0.10
10	1.48	1.48	0.00

Table 3. Results for codec G.729

Packet loss [%]	**G.729**		
	E-model [MOS]	**PESQ [MOS]**	**Difference [MOS]**
0	4.14	4.07	0.07
0.5	4.06	3.99	0.07
1	3.98	3.90	0.08
1.5	3.89	3.82	0.07
2	3.81	3.74	0.07
2.5	3.73	3.66	0.07
3	3.65	3.58	0.07
3.5	3.57	3.50	0.07
4	3.49	3.43	0.06
4.5	3.41	3.36	0.05
5	3.34	3.29	0.05
5.5	3.27	3.22	0.05
6	3.20	3.15	0.05
6.5	3.13	3.09	0.04
7	3.07	3.03	0.04
7.5	3.01	2.96	0.05
8	2.95	2.91	0.04
8.5	2.89	2.85	0.04
9	2.83	2.79	0.04
9.5	2.78	2.74	0.04
10	2.72	2.69	0.03

Table 4. Results for codec G.726

Packet loss [%]	**G.726**		
	E-model [MOS]	**PESQ [MOS]**	**Difference [MOS]**
0	4.24	4.27	0.03
0.5	4.16	4.20	0.04
1	4.09	4.14	0.05
1.5	4.02	4.07	0.05
2	3.94	4.01	0.07
2.5	3.86	3.94	0.08
3	3.79	3.87	0.08
3.5	3.71	3.80	0.09
4	3.64	3.73	0.09
4.5	3.57	3.66	0.09
5	3.50	3.59	0.09
5.5	3.43	3.52	0.09
6	3.36	3.44	0.08
6.5	3.29	3.37	0.08
7	3.23	3.29	0.06
7.5	3.17	3.22	0.05
8	3.11	3.14	0.03
8.5	3.05	3.06	0.01
9	2.99	2.98	0.01
9.5	2.94	2.90	0.04

10	2.88	2.82	0.06

Table 5. Results for codec G.723.1

Packet loss [%]	G.723.1		
	E-model [MOS]	PESQ [MOS]	Difference [MOS]
0	3.79	3.78	0.01
0.5	3.72	3.77	0.05
1	3.65	3.75	0.10
1.5	3.58	3.72	0.14
2	3.52	3.69	0.17
2.5	3.46	3.66	0.20
3	3.39	3.62	0.23
3.5	3.33	3.58	0.25
4	3.27	3.53	0.26
4.5	3.21	3.48	0.27
5	3.16	3.42	0.26
5.5	3.10	3.36	0.26
6	3.05	3.29	0.24
6.5	3.00	3.22	0.22
7	2.95	3.14	0.19
7.5	2.90	3.06	0.16
8	2.85	2.98	0.13
8.5	2.80	2.89	0.09
9	2.76	2.79	0.03
9.5	2.71	2.69	0.02
10	2.67	2.58	0.09

Results as shown in tables 1–5 clearly indicate significant heterogeneity between data obtained by means of the E-model and by means of PESQ [9].

If we take MOS values obtained by means of PESQ as correct, MOS values simulated by the E-model need to be adjusted using a correction function. This function is based on the second order polynomical regression (8) that was applied for individual combination of codecs and the relevant coefficients were claculated, they are listed in Table 6.

$$MOS(Ppl) = x + y \cdot P_{pl} + z \cdot P_{pl}^{\ 2} \tag{8}$$

The codec G.729 is chosen for a demonstration (9) and regression function with coefficients is expressed as:

$$MOS(Ppl) = 4.07378 + 0.17635 \cdot P_{pl} + 0.00380 \cdot P_{pl}^{\ 2} \tag{9}$$

Let $MOS(Ppl)'$ stand for values after correction and $MOS(Ppl)$ stand for original values. Values depend on the current Ppl size, i.e. on loss rate of packets transmitted. The resulting function is described by equation (10).

$$MOS(Ppl)' = MOS(Ppl) - \left\{ \begin{array}{l} \underbrace{a}_{A} + \\ + \underbrace{\left[b \cdot (Ppl - c)^2 - d\right]}_{B} + \\ + \underbrace{(Ppl \cdot e)}_{C} \end{array} \right\} \tag{10}$$

Fig. 8 explains individual correction function's elements. Parameter A in equation (10) determines the shift in all values irrespective on loss rate Ppl by a up or down, provided the difference between the original and estimated MOS values is constant. Parameter B adjusts the original values where they differ from the referential ones hyperbolically, i.e. the loss rate Ppl is such that the difference between original MOS values is the highest. Coefficient c stands for the top of the hyperbola where the difference is the biggest; coefficient d stands for its size and coefficient b determines the twist of the hyperbola depending on the sign and its overall "average". Lastly, parameter C adjusts the output data where the difference between the original and estimated MOS values grows linearly with the increasing loss rate. Table 6 provides an overview of the proposed values of correction function's coefficients for individual codecs.

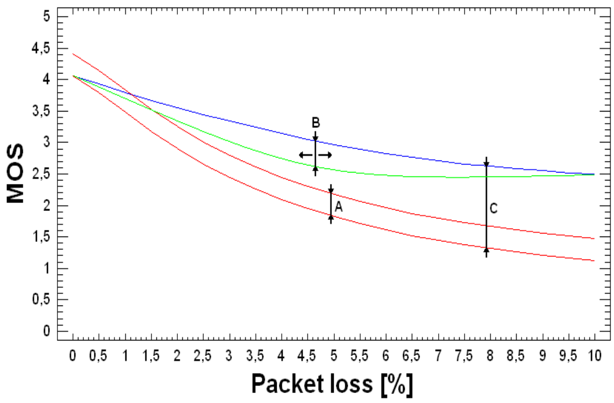

Fig. 8. Correction function parametres explanation.

Table 6. Coefficients of correction function

Codec	Coefficient				
	a	***b***	***c***	***d***	***e***
G.711 A-law	0.34	0.019	5	0.5	-0.14
G.711 µ-law	0.31	0.038	5	0.98	0
G.729	0.07	0	0	0	-0.0035
G.726	-0.06	0.0033	6	0.09	0.015
G.723.1	0.01	0.0115	5	0.315	0.009

Values obtained by means of the E-model, which were adjusted using the correction function with coefficients as defined in Table 6, confirmed improvements against original values (Tables 1–5).

Table 7. E-Model with correction function results for codec G.711 A-law, G.711 μ-law and G.729

PL loss %	G.711 A-law		G.711 μ-law		G.729	
	Model with corr. MOS	Diff. MOS	Model with corr. MOS	Diff. MOS	Model with corr. MOS	Diff. MOS
0	4.10	0.00	4.13	0.03	4.07	0.00
0.5	3.99	0.04	4.04	0.06	3.99	0.00
1	3.83	0.01	3.89	0.04	3.91	0.01
1.5	3.67	0.02	3.73	0.00	3.83	0.01
2	3.53	0.04	3.59	0.02	3.75	0.01
2.5	3.40	0.06	3.44	0.04	3.67	0.01
3	3.29	0.06	3.31	0.05	3.59	0.01
3.5	3.22	0.02	3.19	0.04	3.51	0.01
4	3.14	0.01	3.07	0.03	3.43	0.00
4.5	3.09	0.03	2.96	0.01	3.36	0.00
5	2.98	0.00	2.84	0.00	3.29	0.00
5.5	2.88	0.02	2.72	0.01	3.22	0.00
6	2.82	0.01	2.59	0.01	3.15	0.00
6.5	2.77	0.01	2.45	0.00	3.08	0.01
7	2.72	0.01	2.32	0.01	3.02	0.01
7.5	2.67	0.01	2.16	0.01	2.97	0.01
8	2.62	0.01	2.00	0.04	2.91	0.00
8.5	2.56	0.01	1.81	0.09	2.85	0.00
9	2.54	0.00	1.62	0.14	2.79	0.00
9.5	2.53	0.01	1.42	0.20	2.74	0.00
10	2.52	0.02	1.20	0.28	2.69	0.00

Table 8. E-model with correction function results for codecs G.726 and G.723.1

Packet loss [%]	G.726		G.723.1	
	E-Model with corr.[MOS]	Diff. [MOS]	E-model with corr. [MOS]	Diff. [MOS]
0	4.27	0.00	3.81	0.03
0.5	4.20	0.00	3.79	0.02
1	4.14	0.00	3.76	0.01
1.5	4.08	0.01	3.73	0.01
2	4.01	0.00	3.70	0.01
2.5	3.93	0.01	3.67	0.01
3	3.87	0.00	3.62	0.00
3.5	3.79	0.01	3.58	0.00
4	3.72	0.01	3.53	0.00
4.5	3.65	0.01	3.47	0.01
5	3.57	0.02	3.42	0.00
5.5	3.50	0.02	3.35	0.01
6	3.42	0.02	3.29	0.00
6.5	3.34	0.03	3.22	0.00
7	3.27	0.02	3.15	0.01
7.5	3.20	0.02	3.07	0.01
8	3.13	0.01	2.98	0.00
8.5	3.05	0.01	2.89	0.00
9	2.98	0.00	2.80	0.01
9.5	2.91	0.01	2.70	0.01
10	2.83	0.01	2.60	0.02

Initially, the differences reached up to one MOS grade, and were reduced to hundredth of a MOS grade after the correction, an example is depicted on Fig. 9-13.

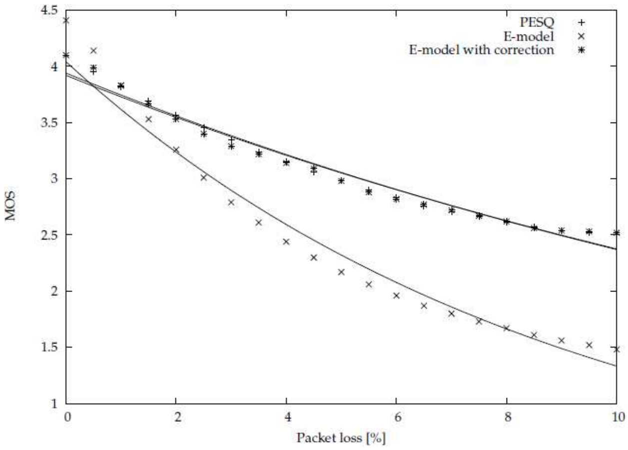

Fig. 9. Graphical results for codec G.711 A-law.

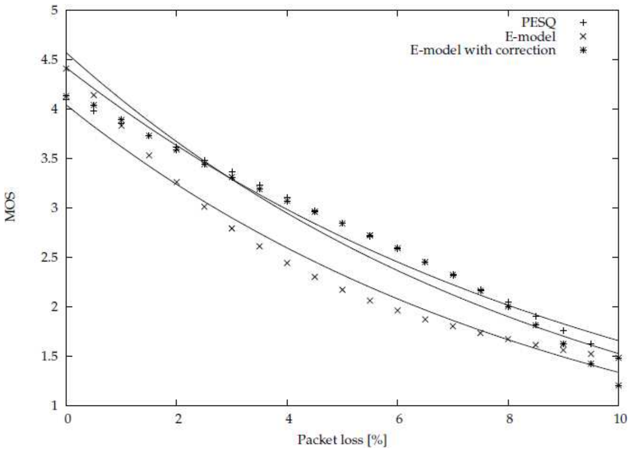

Fig. 10. Graphical results for codec G.711 μ-law.

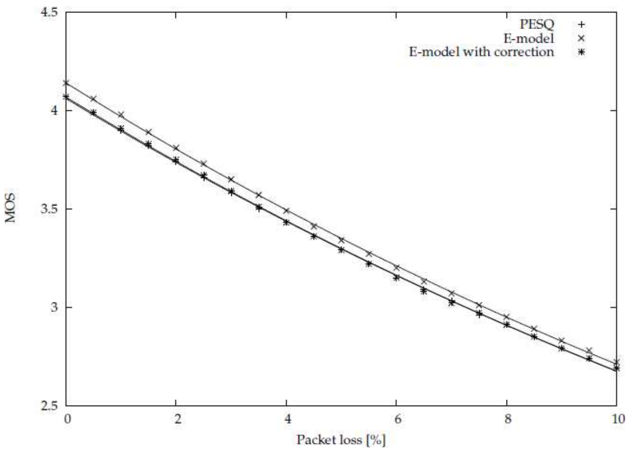

Fig. 11. Graphical results for codec G.729.

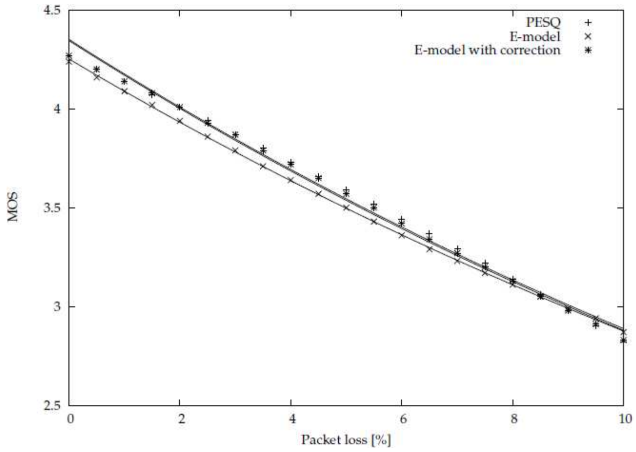

Fig. 12. Graphical results for codec G.726.

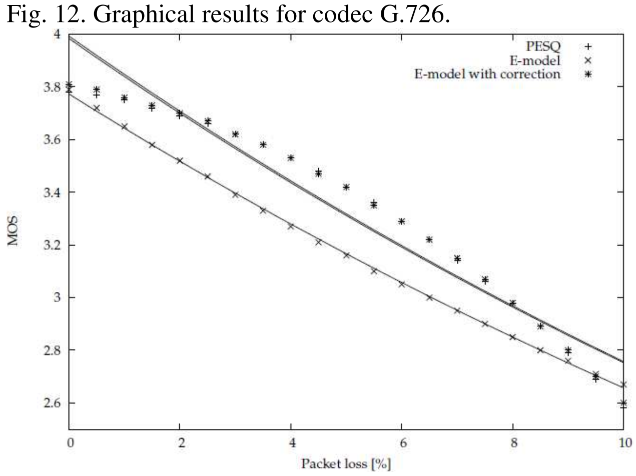

Fig. 13. Graphical results for codec G.723.1 (ACELP).

VI. Conclusion

The E-model brings a modern approach to the computation of estimated quality, allowing for easy implementation. One of its advantages is that it can be applied in real time. The method is based on a mathematical computation model and can be applied as early as the planning stage a new communication system. The E-model is classified as objective non-intrusive methods and is applied primarily in the Voice over IP technology. The latest version of ITU-T recommendation G.107 was drafted in 2009, but the development of the E-model is by no means finished [3]. The second part of the paper aimed at showing and proving in practice that the E-model in its current version does not reflect the reality and that the SG12 group failed to address certain significant influences (such as codec tandeming). To enable results comparison, PESQ was applied in accordance with ITU recommendation P.862 which these days is considered de-facto as a standard for determining MOS. The author would like to stress two significant benefits of their paper. The first being a statement, confirmed by the experiment, that the E-model in tandeming structure does not correspond with results obtained through measurements. The second being the design of a correction function and the way to enhance the E-model. Results depend strongly on choosing suitable coefficients. This is why it is necessary to study the issue further and more deeply and to design its adjustments for other codec structures. Another output of the work is an application in JAVA which enables computations using the E-model for end users.

.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 218086.

References

[1] I. Pravda and J. Vodrazka, “Voice quality planning for NGN including mobile networks,” in *Proc. 12th International Conference on Personal Wireless Communications (PWC 2007)*, Prague, 2007.

[2] M. Kavacky, E. Chromy, L. Krulikovska and P. Pavlovic, “Quality of Service Issues for Multiservice IP Networks, “ *In Proc. SIGMAP 2009 International Conference on Signal Processing and Multimedia Applications*, Milan, Italy, July 2009, pp. 185 – 188.

[3] *The E-model: A computational model for use in transmission planning*, ITU-T Recommendation G.107, Geneva, 04/2009.

[4] *Application of the E-model: A planning guide*, ITU-T Recommendation G.108, Geneva, 09/1999.

[5] *Definition of categories of speech transmission quality*, ITU-T Recommendation G.109, Geneva, 09/1999.

[6] *Transmission impairments due to speech processing,* ITU-T Recommendation G.113, Geneva, 11/2007.

[7] *Methods for subjective determination of transmission quality,* ITU-T Recommendation P.800, Geneva, 08/1996.

[8] *Mean Opinion Score (MOS) terminology,* ITU-T Recommendation P.800.1, Geneva, 07/2006.

[9] *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, ITU-T Recommendation P.862, Geneva, 02/2001.

[10] M. Voznak, F. Rezac and M. Halas, ”Speech Quality Evaluation in IPsec Environment, ” *In Proc. 12th International Conference on Networking, VLSI and Signal Processing (ICNVS '10)*, Cambridge, England, February 2010, pp. 49-53.

[11] M. Voznak and F. Rezac, ”SPAM over Internet telephony, ” *In Proc. 11th International Conference on Research in Telecommunication Technology (RTT 2009),* Srby, Czech Republic, September 2009, pp. 119-121.

[12] M. Voznak, “Speech bandwith requirements in IPsec and TLS environment”, *In Proc. 13th WSEAS International Conference on Computers,* Rhodes, Greece, July 2009, pp.217-220.

[13] A. E. Mahdi and D. Picovici, ”Advances in voice quality measurement in modern telecommunications, ” *Digital Signal Processing, Volume 19, Issue 1,* January 2009, pp.79-103.

[14] M. Voznak, M. Tomes, Z. Vaclavikova and M. Halas, ”E-model Improvement for Speech Quality Evaluation Including Codecs Tandeming, ” *In Advances in Data Networks, Communications, Computers,* Faro, Portugal, November 2010, pp. 119-124.

[15] J. Kouril and H. Atassi, "Objective Speech Quality Evaluation. A primarily Experiments on a Various Age and Gender Speakers Corpus," *In Proc. 8th WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing,* December 2009 Puerto de la Cruz, SPAIN, pp. 333-336.

[16] M. Voznak,M. and J. Rozhon, ”SIP Back to Back User Benchmarking, ” *In Proc. The Sixth International Conference on Wireless and Mobile Communications ICWMC 2010*, IEEE Computer Society, September 2010, Valencia, Spain, pp. 92-96.

[17] A. Kovac, M. Halas and M. Voznak, ”Impact of Jitter on Speech Quality Estimation in Modified E-model, ” *In Proc. Research in Telecommunications Technology RTT 2010*, Velké Losiny, September 2010, pp. 209-215.

[18] H. Farsi, M. Mozaffarian and H. Rahmani, "Improving Voice Activity Detection Used in ITU-T G.729.B, "*In Proc. 3rd WSEAS International*

Conference on Circuits, Systems, Signal and Telecommunications, January 2009 Ningbo, China, pp. 11-15.
[19] F. Alvarez-Vaquero and J. Sanz-Gonzalez,"Network VoIP for corporative environment design*," In Proc. 7th WSEAS International Conference on Telecommunications and Informatics,* May 2008, Istanbul, Turkey, pp. 194-198.
[20] M. Voznak and J. Rozhon, "Performance testing and benchmarking of B2BUA and SIP Proxy, " *In Proc. 33rd International Conference on Telecommunications and Signal Processing TSP 2010,* August 2010, Baden near Vienna, Austria, p.497-503.
[21] M. Voznak and J. Rozhon, "SIP Infrastructure Performance Testing, " *In Proc. The 9th International Conference on Telecommunications and Informatics,* Catania, Italy, May 2010, p. 153-158.
[22] A. Maraj and S. Rugova, "The role of softswitch in NGN network," *In Proc. 8th WSEAS International Conference on Data Networks, Communications, Computers,* November 2009, Baltimore, pp. 179-184.
[23] K. Molnar and V. Vrba, "DiffServ-Based User-Manageable Quality of Service Control System,"*In Proc. 7th International Conference on Networking (ICN 2008),* April 2008, Cancun, Mexico, pp. 485-490.
[24] I. Baronak and M. Halas, "Mathematical representation of VoIP connection delay," *RADIOENGINEERING Volume: 16 Issue: 3,* September 2007, pp. 77-85.

Miroslav Voznak holds position as an associate professor with Department of Telecommunications, Technical University of Ostrava. He received his M.S. and Ph.D. degrees in telecommunications, dissertation thesis “Voice traffic optimization with regard to speech quality in network with VoIP technology” from the Technical University of Ostrava, in 1995 and 2002, respectively. Topics of his research interests are Next Generation Networks, IP telephony, speech quality and network security.