Regular paper

Simple analysis of the impact of packet loss and delay on voice transmission quality

Krzysztof Perlicki

Abstract — In this paper a simple analytical expression based on E-model is applied to analyse the impact of packet loss and delay on voice transmission quality. The relationship between overall transmission quality rating, packet loss and delay G.711, G.723 and G.729 codecs is presented.

Keywords — E-model, packet network, packet loss, delay, transmission quality rating.

1. Introduction

The IP network will become an ubiquitous infrastructure, which is used by numerous applications with various requirements and with different characteristics of generated traffic. In particular, it concerns to data and voice applications. In the future, the voice communication is expected to migrate from the public switched telephone network (PSTN) to the packet (IP) transmission. Due to constant improvements within the years, traditional voice communications over the PSTN is characterized today by low delay, high availability and adequate voice quality. For the packet transmission to compete with the PSTN, it should provide the same level of quality, which implies stringent delay, packet loss and also packet delay variation (jitter). The packet delay variation is bounded by the dejitter buffer, which removes small amounts of variation by increasing delay and/or by the increasing packet loss [1].

The E-model [2, 3] is a network planning tool used in the design of networks for carrying voice applications. The model estimates the relative impairments to voice quality. It combines the impairments caused by transmission parameters into R factor (overall transmission quality rating). The factor R is given by [3]:

$$R = R_o - I_s - I_d - I_e + A\,. \qquad (1)$$

The factor R_o expresses the basic signal-to-noise ratio. The factor I_s represents all impairments which occur more or less simultaneously with the voice signal, such as: too loud speech level, non-optimum sidetone, quantization noise, etc. The delay impairment factor (I_d) sums all impairments due to delay and echo effects. The equipment impairment factor (I_e), represents impairments which are caused by low bit-rate codecs used in special equipment; this factor can also be used to take into account the influence of packet loss. The advantage factor A represents an advantage of access which certain systems may provide in comparison to conventional systems (for wirebound communication system factor A is equal zero). Typically, the values of the R factor are categorized as shown in Table 1. This table also contains equivalent transformed values of R into *MoS* (mean opinion score).

Table 1
R factor, *MoS* and the user satisfaction measure [3]

R	*MoS*	User satisfaction
$90 < R < 100$	$4.34 < MoS < 4.5$	Very satisfied
$80 < R < 90$	$4.03 < MoS < 4.34$	Satisfied
$70 < R < 80$	$3.60 < MoS < 4.03$	Some users dissatisfied
$60 < R < 70$	$3.10 < MoS < 3.60$	Many users dissatisfied
$50 < R < 60$	$2.58 < MoS < 3.10$	Nearly all users dissatisfied

An interesting aspect of the E-model is that terms R_o, I_s, I_d, I_e and A are additive and that delay and packet loss contributions are isolated into I_d and I_e, respectively. The others factors, like noise (R_o) and a connection that is too loud (I_s) of a packetized voice call are not fundamentally different from traditional switched telephone network.

2. Delay in packet network – I_d factor

The delay for packetized voice call, where the most important contributions are encoding, packetization, propagation, queuing, service, dejittering and decoding delay is larger than for traditional switched telephone network where the end-to-end delay is equal to propagation delay and switching delay. Figure 1 shows the relationship between the I_d factor and the absolute end-to-end (mouth-to-ear) delay (T_a). The results are derived from E-model (all E-model input parameters of default values, which are shown in [3]) and calculated using following expression:

$$I_d = 0.65 + (0.1T_a - 15.93) \cdot \delta(T_a - 165)\,, \qquad (2)$$

where: $\delta(x) = 0$ for $x < 0$, $\delta(x) = 1$ for $x \geq 0$.

The absolute end-to-end delay in packet connections is given by:

$$T_a = T_{enc} + T_p + T_{dec} + T_n\,, \qquad (3)$$

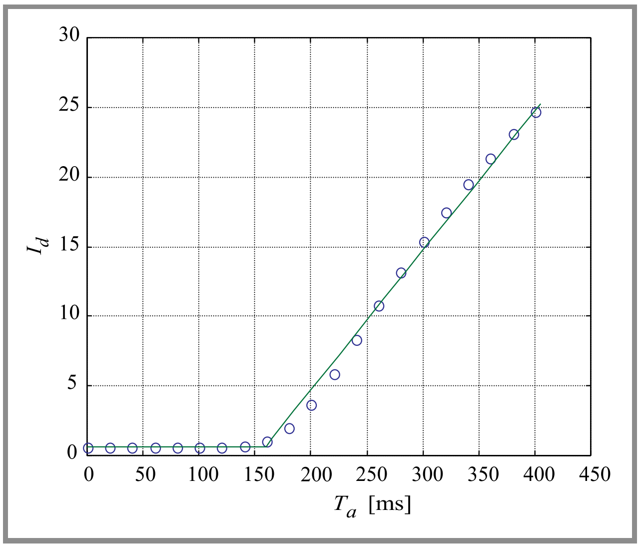

Fig. 1. The relationship between the I_d factor and the T_a delay; ∘ – based on E-model, curve fit – Eq. (3).

where: T_{enc} is the encoding delay, T_p is the packetization delay, T_{dec} is the decoding delay and T_n is the so-called network delay, which is equal to the sum of the propagation, queuing, service and dejittering delay [4].

3. Packet loss – I_e factor

No analytical expressions for the equipment impairment factor (I_e) are directly available today. Instead, the I_e factor must be obtained from subjective tests of voice quality for particular codec and various operating conditions (e.g. packet loss, number of voice frames per packet) [5]. ITU-T Rec. [6] gives measured value of the I_e factor only for G.711, G.723 and G.729 codecs.

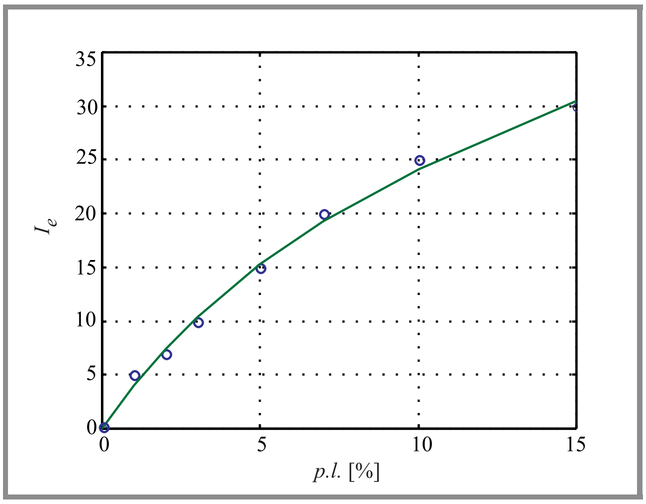

Fig. 2. The equipment impairment factor (I_e) versus the packet loss for G.711 codec (random packet loss); ∘ – based on [6], curve fit – Eq. (4).

Figure 2 shows results for the I_e factor for G.711. Also shown in Fig. 2 is a curve fit, which is derived from an expression of the form:

$$I_e = 22 \ln \left(1 + 0.2 \cdot p.l.\right), \tag{4}$$

where $p.l.$ is the packet loss in percent.

The results for the I_e factor and G.723 codec are shown in Fig. 3. The curve fit in this figure is derived from the following expression:

$$I_e = 15 + 33 \ln(1 + 0.15 \cdot p.l.). \tag{5}$$

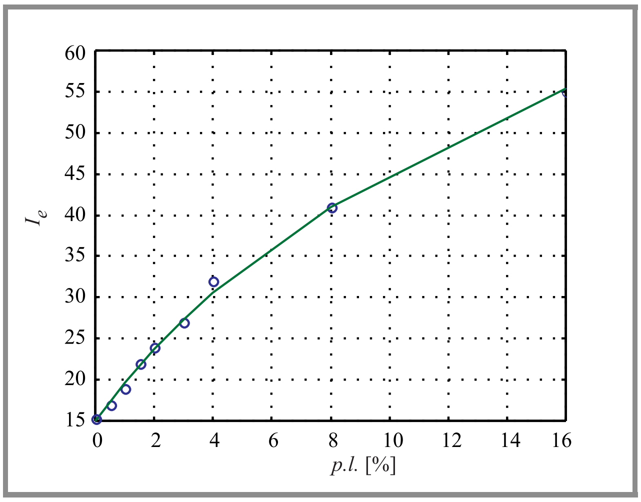

Fig. 3. The equipment impairment factor (I_e) versus the packet loss for G.723 codec (number of voice frames per packet: 1); ∘ – based on [6], curve fit – Eq. (5).

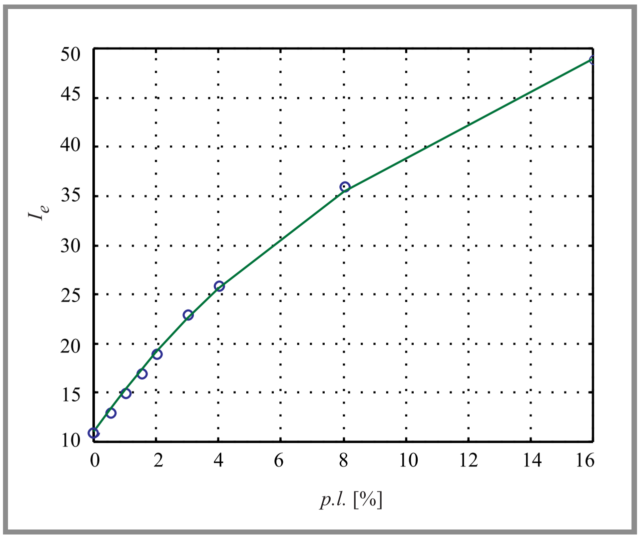

Fig. 4. The equipment impairment factor (I_e) versus the packet loss for G.729 codec (number of voice frames per packet: 2); ∘ – based on [6], curve fit – Eq. (6).

The results for the I_e factor and G.729 codec are shown in Fig. 4. The curve fit in this figure is derived from the following expression:

$$I_e = 11 + 31 \ln(1 + 0.15 \cdot p.l.). \qquad (6)$$

4. Impact of packet loss and network delay on the factor R

Choosing E-model default values [3] we can reduce the expression (1) to:

$$R = 93.33 - I_d - I_e\,. \qquad (7)$$

Substituting Eqs. (2)–(6) into (7), the factor R can be rewritten as:

- for G.711 codec:

$$R = 92.68 - (0.1\,T_{n\max} - 15.90) \times \times \delta(T_{n\max} - 164.75) + -22\ln(1 + 0.2 \cdot p.l.), \qquad (8)$$

- for G.723 codec:

$$R = 77.68 - (0.1\,T_{n\max} - 9.18) \times \times \delta(T_{n\max} - 97.50) + -33\ln(1 + 0.15 \cdot p.l.)\,, \qquad (9)$$

- for G.729 codec:

$$R = 81.68 - (0.1\,T_{n\max} - 12.43) \times \times \delta(T_{n\max} - 130) + -31\ln(1 + 0.15 \cdot p.l.)\,, \qquad (10)$$

where $T_{n\max}$ is the maximum of the network delay.

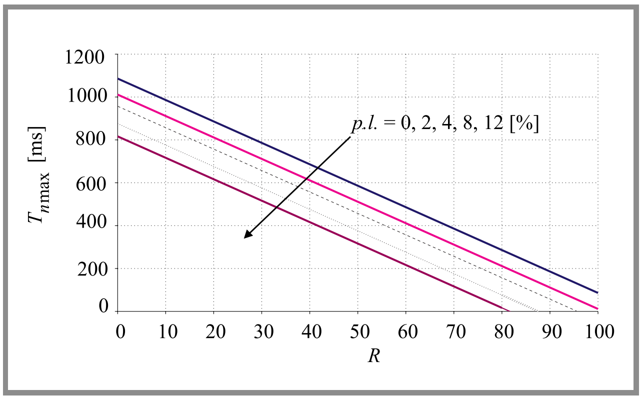

Fig. 5. Maximum of the network delay versus the factor R (G.711 codec).

It is assumed in Eqs. (8)–(10) that the minimum values of the sum of the encoding delay, the packetization delay and the decoding delay are [7]:

- for G.711 codec: 0.25 ms,
- for G.723 codec: 67.5 ms (number of voice frames per packet: 1),
- for G.729: 35 ms (number of voice frames per packet: 2).

Figures 5, 6, 7 show the relationship between the maximum of the network delay ($T_{n\max}$) and the R factor and G.711, G.723 and G.729 codecs, respectively.

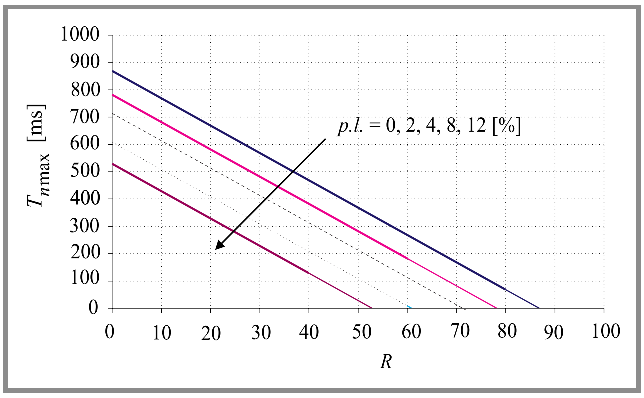

Fig. 6. Maximum of the network delay versus the factor R (G.723 codec).

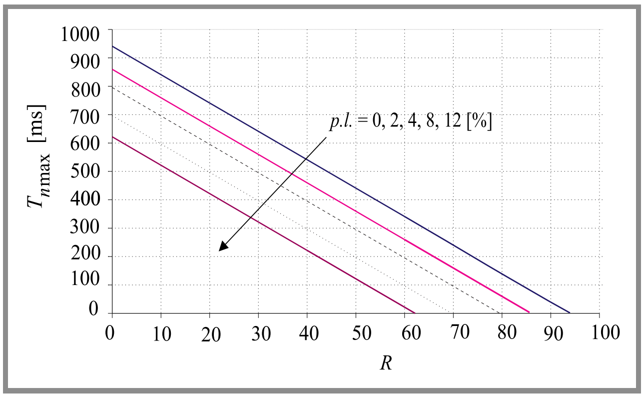

Fig. 7. Maximum of the network delay versus the factor R (G.729 codec).

Table 2
The value of $T_{n\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 (G.711 codec)

Maximum of the network delay $T_{n\max}$ [ms]					
R	Packet loss [%]				
	0	2	4	8	12
50	585.8	511.8	456.5	375.6	316.6
60	485.8	411.8	356.5	275.6	216.6
70	385.8	311.8	256.5	175.6	116.6
80	285.8	211.8	156.5	75.6	16.6
90	185.8	111.8	56.5	—	—

The value of $T_{n\,\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 under conditions of packet loss for G.711, G.723 and G.729 codecs are given in Tables 2, 3, and 4, respectively.

Table 3
The value of $T_{n\,\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 (G.723 codec)

Maximum of the network delay $T_{n\,\max}$ [ms]					
R	Packet loss [%]				
	0	2	4	8	12
50	368.6	282.0	213.5	108.4	28.8
60	268.6	182.0	113.5	8.4	—
70	168.6	82.0	13.5	—	—
80	68.6	—	—	—	—
90	—	—	—	—	—

Table 4
The value of $T_{n\,\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 (G.729 codec)

Maximum of the network delay $T_{n\,\max}$ [ms]					
R	Packet loss [%]				
	0	2	4	8	12
50	441.1	359.8	295.4	196.7	121.9
60	341.1	259.8	195.4	96.7	21.9
70	241.1	159.8	95.4	—	—
80	141.1	59.8	—	—	—
90	41.1	—	—	—	—

5. Conclusions

The E-model has been used to study the quality of packetized voice calls. The overall transmission quality rating (user satisfaction) decreases as the packet loss and the delay increases. The quality "Very satisfied" ($R > 90$) is reached only for G.711 and G.729 codecs. In the case of G.711 codec $R > 90$ is reached at: $p.l. = 0$ and $T_n \leq 185.5$ ms, $p.l. = 2$ and $T_n \leq 111.8$ ms, $p.l. = 4$ and $T_n \leq 56.5$ ms. In the case of G.729 codec $R > 90$ is only reached at $p.l. = 0$ and $T_n \leq 41.1$ ms. It can be seen that for G.723 codec the greatest quality is "Satisfied" ($R > 80$). It is reached at $p.l. = 0$ and $T_n \leq 68.6$ ms. For the packet loss 12% the G.711 codec has a rating R equal to 80 ($T_{n\,\max} = 16.6$ ms). For this packet loss the G.723 codec has the R rating equal to 50 ($T_{n\,\max} = 28.8$ ms), and G.729 codec has $R = 60$ ($T_{n\,\max} = 21.9$ ms).

References

[1] ETSI TS 101 329-2 V1.1.1, "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON)"; "End to End Quality of Service in TIPHON Systems"; Part 2: "Definition of Quality of Service (QoS) Classes", Sophia Antipolis Cedex – France, ETSI, 7/2000.

[2] ETSI EG 201 050 V1.2.2, "Speech Processing, Transmission and Quality Aspects (STQ)"; "Overall Transmission Plan Aspects for Telephony in a Private Network", Sophia Antipolis Cedex – France, ETSI, 2/1999.

[3] ITU-T Rec. G.107, "The E-model, a computational model for use in transmission planning", Geneva, ITU, 5/2000.

[4] D. De Vleeschauwer and J. Janssen, "Quality bounds for packetized voice transport", *Alcatel Telecommun. Rev.*, no. 1, pp. 19–23, 2000.

[5] ITU-T Rec. P.833, "Methodology for derivation of equipment impairment factors from subjective listening – only tests", Geneva, ITU, 2/2001.

[6] ITU-T Rec. G.113, "Transmission impairments", Appendix I: "Provisional planning values for the equipment impairment factor I_e", Geneva, ITU, 9/1999.

[7] ITU-T Rec. G.114, "One-way transmission time", Geneva, ITU, 5/2000.

Krzysztof Perlicki
e-mail: perlicki@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

The author's fee was financed by the Association for Authors Rights Collective Administration of Scientific and Technical Works KOPIPOL with a seat in Kielce from the remuneration collected on the basic of Art. 20 of the Law on Author Right and Related Rights.