COMMITTEE T1 - TELECOMMUNICATIONS STANDARDS CONTRIBUTION

DOCUMENT NUMBER: T1A1.1/2001-037

T1BBS FILES: 1a110370.pdf

DATE: April 30, 2001

STANDARDS PROJECT: T1A1-17

TITLE: Extensions to the E Model to incorporate the effects of time varying packet loss and recency

SOURCE: Telchemy Incorporated

AUTHORS: Alan Clark

CONTACT: Alan Clark
Telchemy Incorporated
3360 Martins Farm Road, Suite 200
Suwanee, GA 30024
Tel +1-770-614-6944
Email alan@telchemy.com

DISTRIBUTION: T1A1.7, T1BBS

ABSTRACT:

This contribution describes an approach to passive or non-intrusive monitoring of Voice over IP transmission quality using an extended version of the E Model, and of then mapping this to an estimated speech quality. The effects of burst packet loss are represented by estimating the parameters of a 4 state Markov model from observed packet loss events and then mapping these parameters onto an Equipment Impairment factor. The mapping to an estimated speech quality also incorporates the effects of Delay and Recency. This approach was incorporated into Annex E of TS 101 329-5 "Quality of Service Measurement Methodologies" [7] by the ETSI TIPHON committee.

Intellectual Property Statement

The individual preparing this contribution knows of patents, the use of which may be essential to a standard resulting in whole or in part from this contribution.

1. Introduction

It is well understood that packet loss, delay, jitter and CODEC characteristics affect the conversational quality of a Voice over IP connection. Current approaches do not model the effects of time varying impairments and can therefore give misleading results. For example, if a 60 second VoIP call using a G.729A CODEC with PLC experiences a 1% packet loss rate then conventional models would suggest that there would be little noticeable degradation in audible quality; if however the 1% of lost packets occurred in a two second period in the call, with zero packet loss for the remaining 58 seconds then the listener would be exposed to a 30% packet loss rate for 2 seconds, which would certainly be noticeable.

The location of a noise burst within a call can also make a substantial difference to the subjective score that a listener would give. AT&T [1] conducted some tests in 1998 in which a 15 second burst of noise was moved from the beginning to the middle and end of a 60 second call. The resulting MOS score was 3.18 if the noise burst was at the end of the call and 3.82 if the noise burst was at the start. This is generally known as "recency" as it has some similarities to the short term memory effect of the same name in which people tend to remember the most recent words or events.

Non-intrusive or passive monitoring systems are able to analyze some characteristics of live speech or data and make performance or quality estimates of the system being tested. In the case of Voice over IP the objective of a passive monitoring system is to determine transmission or speech quality estimates from observations of the digital or analog speech path.

The passive monitoring approach described in this contribution incorporates the effects of burst packet loss and recency, has low computational complexity and could be incorporated into VoIP Gateways or other IP end systems.

2. E Model

The E Model is described in ETSI ETR250 and ITU Recommendation G.107. This is a transmission quality model that can be used to estimate speech quality. The transmission quality factor R typically ranges from 50-90 and is calculated using the equation:

$$R = Ro - Is - Id - Ie + A$$

Where Ro represents the basic signal to noise ration, Is the impairments that occur simultaneously with the signal, Id the impairments that are delayed with respect to the signal (e.g. echo), Ie the Equipment Impairment factor and A the Advantage factor. The Equipment Impairment factor is intended to represent the effects of network equipment, in this case the codec-to-codec connection through the VoIP network.

3. Network Impairments

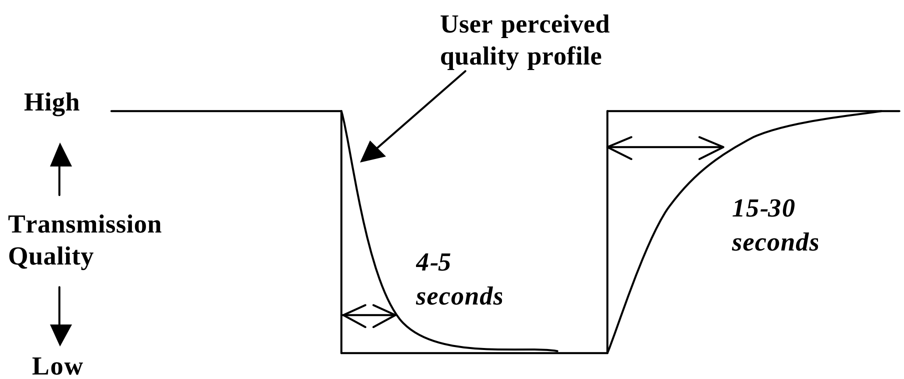
The primary network impairments that affect transmission quality are:-

(i) CODEC effects
(ii) Packet loss, which can be somewhat mitigated by PLC. In general packet loss rates of 2% or better are not noticeable and packet loss rates of over 20% are
(iii) Jitter, which in general is translated to additional packet loss and delay by the jitter buffer
(iv) Delay, which causes some difficulty in interactive conversation
(v) Temporal clipping effects due to voice activity detectors
(vi) Echo

The passive monitoring approach described in this contribution currently incorporates CODEC, packet loss, jitter and delay, the inclusion of clipping and echo effects are under study.

4. Effects of Burst Packet Loss

Within this contribution burst packet loss is defined as a period of time during which a high percentage of packets are lost. Bursts of high packet loss rate are separated by gaps, during which the packet loss rate is very low.
Intuitively, it seems reasonable to expect that a sudden change in packet loss rate would not necessarily result in a sudden change in perceived quality. For example, if packet loss rate abruptly changed from 1% to 25% then the perceived quality would initially be unchanged however would gradually worsen until some the listener either learnt to communicate through the noise and distortion or disconnected. Similarly it would seem reasonable to expect that a sudden improvement in packet loss would result in a gradual improvement in perceived quality.



In listening tests conducted by France Telecom [5] the rate of packet loss on a VoIP connection was varied from 0 to 30% for periods of time ranging from 15 to 60 seconds during a 3 minute test call. Listeners were asked to provide feedback on instantaneous quality during the call. The test results showed that the time constant for the "good to bad" change was of the order of 4-5 seconds whereas the time constant for the "bad to good" change was of the order of 10-15 seconds. The reported call quality correlated with the average of the corrected profile.

5. An Approach to Measuring Burst Characteristics

Burst packet loss distribution can be modeled using a Markov model, for example the Gilbert Model [6]. The approach described in [7] provides a low complexity method of gathering packet loss information in the form of a 4 state Markov model.

The Markov model states used are Gap-no loss, Gap-loss, Burst-no loss and Burst-loss. A simple packet loss event driven process is used to count some of the key transition events during the VoIP call. At the end of the call these counts are used to derive remaining transition event counts and then normalized to give the transition probabilities.

6. Extending the E Model to incorporate Burst Packet Loss Effects

The Markov model parameters are used to determine an estimate of the burst and gap duration, and of the packet loss density in each state. Provisional mappings between packet loss and the E Model's Equipment Impairment factor are given in G.113 [3]. The transitions between burst and gap states are corrected, to model the effect described in 3, using exponential decays with time constants of 5 seconds for the gap to

burst transition and 15 seconds for the burst to gap transition. The average of the corrected profile is used as the Equipment Impairment factor for the call.
If the Equipment Impairment values for burst and gap are Ieb and Ieg respectively, and the burst and gap durations b and g seconds then:

Gap to burst transition

$$I_1 = I_{eb} - (I_{eg} - I_2)\, e^{-b/t1} \qquad \text{where } t_1 \text{ is typically 5}$$

Burst to gap transition

$$I_2 = I_{eg} + (I_1 - I_{eg})\, e^{-g/t2} \qquad \text{where t2 is typically 15}$$

Integrating the expressions for I_1 and I_2 to give a time average gives

$$I_e(av) = (\, b\, I_{eb} + g\, I_{eg} - t_1\, (I_{eb} - I_2)\, (1 - e^{-b/t1}) + t_2\, (I_1 - I_{eg})\, (1 - e^{-g/t2})\,)\, /\, (b + g)$$

Default values are assumed for most of the E Model parameters, giving an initial R factor of 94 from which the impairment values are subtracted.

7. Incorporating a Recency Model

The Recency Effect seems to involve a gradual improvement in perceived quality over a period of 30 seconds. This can be modeled using an exponential decay in perceived quality which starts at the end of the last significant burst of packet loss and asymptotically approaches the average quality level for the call.

The "exit" value from the last burst is assumed to be I_1

$$I_e\,(\text{end of call}) = I_e(av) + (k(\, I_1 - I_e(av)))\, e^{-y/t3}$$

Where k is a constant, assumed to be 0.7, y is the time delay since the last burst of packet loss and t3 is a time constant, assumed to be 30 seconds.

8. Reporting Metrics

In order to avoid the proliferation of metrics it is recommended that two values are reported:

R1 – R factor incorporating the effects of CODEC, packet loss but not delay or recency
This value is more indicative of the impact of network performance factors that can be affected by network configuration

R2 – R factor incorporating the effects of delay and recency
This value is more representative of the quality that the end user would report.

To minimize network traffic the metrics can be embedded into the user data field in the H.323 DRQ message or equivalent end of call message, periodically reported through RTCP or retrieved using SNMP.

9. Summary

This contribution described an approach to passive monitoring of Voice over IP calls that is suitable for integration into VoIP Gateways. The approach has been incorporated into an ETSI standard on QoS Measurement and is being introduced to ITU SG12. It is proposed that TIA TR41.4 consider the application of this technique to VoIP Gateways.

10. References

[1] ANSI T1A1.7/98-031: “Testing the quality of connections having time varying impairments”, AT&T

[2] ITU-T Recommendation G.107 (05/00): “The E Model, A Computational Model for use in Transmission Planning”

[3] ITU-T Recommendation G.113, Appendix 1: “Transmission Impairments - Appendix I: Provisional Planning Values for the Equipment Impairment factor Ie”

[4] ITU-T Recommendation G.114, Appendix 1

[5] ITU-T SG12 D.139: “Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequences: influence of a recency effect”, France Telecom

[6] M. Yajnik, J. Kurose, D. Towsley, “Packet loss correlation in the MBone multicast network: Experimental measurements and Markov chain models”, Technical Report UM-CS-95-115, University of Massachusetts, August 1995.

[7] ETSI TIPHON TS 101329-5 QoS Measurement Methodologies

TS 101 329-5 Annex E

Method for determining an Equipment Impairment Factor using Passive Monitoring

E.1 Introduction

This annex describes a method of passively monitoring a Voice over IP stream that produces an Equipment Impairment (I_e) factor that may be used with the E-Model to calculate an R factor. The Ie factor determined using this approach incorporates the effects of the Voice over IP CODEC, packet loss, packet loss distribution, jitter and recency. This process can be applied on a per-call or continuous session basis.

Passive monitoring comprises the extraction of performance metrics from an in-process call or session. This process is non-intrusive, i.e. does not interfere with the data stream, and adds little or no overhead to network traffic. The real time elements of the algorithms described below have been designed to be computationally efficient.

The preferred location of the monitoring point is at the Voice over IP CODEC, at which the monitoring function has access to information such as CODEC type, jitter buffer size and post-jitter buffer packet loss events. The monitoring point may be placed at other locations however may not have access to all the necessary information.

Note:
The methodology described in this Annex requires further validation by subjective testing. Based on these tests numerical constants used in this Annex may be modified. After validation of the methodology and some more practical experience Annex E should become normative.

E.2 Passive QoS Monitor Framework

The Equipment Impairment factor determined using the methodology described in this annex contains the following elements:-

(i) I_e(packet loss)
The distribution of packets lost and received is measured from observation of the received packet stream and modelled using a Markov process. The parameters of the Markov process are mapped onto an I_e factor using the CODEC specific curves described in G.113.

(ii) I_e(PDV)
The packet delay variation level is determined during the call and is assumed to be constant throughout the call and bounded by the jitter buffer level and discard thresholds. In many implementations the jitter buffer is of sufficient depth that received packets are either properly aligned in time or discarded, which would render this step un-necessary.

(iii) I_e(CODEC)
The CODEC type is assumed to be constant throughout the call and is mapped to an Ie value using the parameters specificed in G.113, Appendix 1

(iv) Delay
The estimated one-way delay, including transmission, jitter buffer and CODEC related delays, is estimated.

E.3 Determining Equipment Impairment Factor for Packet Loss

IP network packet loss distribution can be modeled using a Markov process. The resulting model can be used in both analytical and numerical performance estimation and has well known and understood properties.

In typical Voice over IP implementations packet loss can occur if packets are excessively delayed. It is therefore preferable to measure packet loss after the receive jitter buffer or with prior knowledge of the packet delay which would cause packets to be discarded. If packet loss is measured before the jitter buffer then it is preferable to measure the per-packet jitter and to assume that any packets that are delayed by more than the jitter buffer level are discarded.

The channel is assumed to have high packet loss (burst) and low packet loss (gap) conditions. During the Voice over IP call packet loss events and inter-loss gaps are counted. At the end of the call, or on request from a service management system, the transition probabilities of the Markov model are determined and used to compute an R factor for the call.

If the number of packets received between two successive lost packets is less than a minimum value g_{min} then the sequence of the two lost packets and the intervening received packets is regarded as part of a burst. If a sequence of g_{min} or more packets are correctly received the sequence is regarded as being part of a gap.

The Markov model is defined as having the following states and associated transitions:

State 1 - gap - no loss
- p_{11} - *packet received*
- p_{13} - *packet loss (start of burst)*
- p_{14} - *isolated packet loss*

State 2 - burst - no loss
- p_{22} - *packet received within burst*
- p_{23} - *packet lost within burst*

State 3 - burst - packet loss
- p_{31} - *packet received (end of burst)*
- p_{32} - *packet received within burst*
- p_{33} - *packet lost*

State 4 - gap - packet loss
- p_{41} - *packet received*

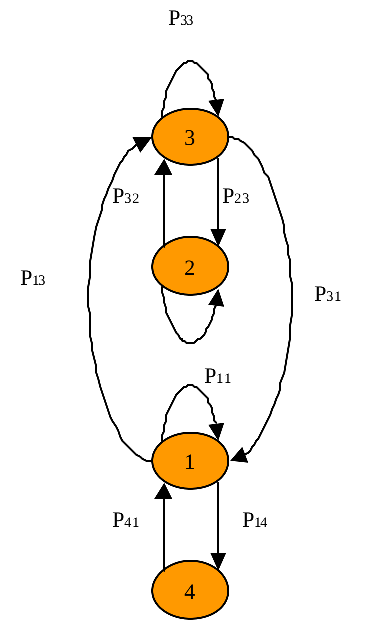


Fig. E 1: 4-stage Markov model

This model can be constructed either by accumulating packet loss information during fixed sampling intervals or at packet loss events. An example of a computationally efficient method for determining the parameters of the Markov model is given below.

Assume a counter *pkt* tracks the number of received packets, *lost* tracks the number of lost packets in a burst, g_{min} is the minimum gap size, and that an event can be generated if a packet loss is detected:

```
Packet loss event->
    c5=c5 + pkt
    if pkt >= gmin then
        if lost = 1 then
            c14 = c14 + 1
        else
            c13 = c13 + 1
        lost = 1
        c11 = c11 + pkt
    else
        lost = lost + 1
        if lost > 8 then c5 = 0
        if pkt = 0 then
            c33 = c33 + 1
        else
            c23 = c23 + 1
```

$$c_{22} = c_{22} + pkt$$

$$pkt = 0$$

The series of counters c_{11} to c_{14} are used to determine the corresponding Markov model transition probabilities (i.e c_{11} is used to calculate p_{11}). Counter c_5 is used to measure the delay since the last "significant" burst of lost packets. Parameter g_{min}, the minimum gap size, is typically 16.

The key metrics needed for determining application performance are:-

$c_{31} = c_{13}$
$c_{32} = c_{23}$
$c_{11} = c_{11} + c_{14}$ (for simplicity - combine states 4 and 1)
$p_{11} = c_{11} / (c_{11} + c_{13})$
$p_{13} = 1 - p_{11}$
$p_{31} = c_{31} / (c_{31} + c_{32} + c_{33})$
$p_{32} = c_{32} / (c_{31} + c_{32} + c_{33})$
$p_{33} = 1 - p_{31} - p_{32}$
$p_{22} = c_{22} / (c_{22} + c_{23})$
$p_{23} = 1 - p_{22}$

$d = (p_{23}\ \ p_{31} + p_{13}\ \ p_{32} + p_{13}\ \ p_{23})$

$p_1 = p_{31}\ \ p_{23} / d$
$p_2 = p_{13}\ \ p_{32} / d$
$p_3 = p_{13}\ \ p_{23} / d$

frame size	F = frame size (in seconds)	
average packet loss rate	$L = 100\ p3$	percent
gap length	$g = F / (1 - p_{11})$	seconds
gap loss density	$D_g = 100\ c_{14} / c_{11}$	percent
burst length	$b = F\ (1 - p1) / (p1\ \ p13)$	seconds
burst loss density	$D_b = 100\ p23 / (p23+p32)$	percent
delay since last burst	$y = F\ \ c_5$	seconds

An estimate of the published "Provisional Planning values for the Equipment Impairment Factor" is given by the equation below:

$I_e(\text{Loss}) = 0$ $\quad D < 0.5$ percent
$I_e(\text{Loss}) = d_1\ D$ $\quad 0.5 < D < d_2$ percent
$I_e(\text{Loss}) = d_3 + d_4\ D$ $\quad d_2 < D$ percent

This can be separately applied to the packet loss rates for the gap and burst state (D_g and D_b), giving I_{eg} and I_{eb}.

CODEC	d_1	d_2	d_3	d_4
G.723.1+VAD 6.3k	4.25	4.8	12	1.75
G.729A				

E.4 Determining Equipment Impairment Factor for Packet Delay Variation

The packet delay variation is bounded by the jitter buffer, which removes small amounts of variation by increasing delay, and by the discard threshold. In many implementations the discard threshold is effectively equal to the jitter buffer delay, which would result in packets either be properly retimed or discarded - in this case the I_e(PDV) value would be 0..

The effect of low levels of packet delay variation on voice quality is substantially less than that of packet loss. Packet loss and packet delay variation are often correlated as high levels of packet delay variation will lead to an increased level of packet discard.

Let the packet inter-arrival time be t, the jitter buffer delay be denoted t_{jb}, the discard delay be denoted $t_{discard}$ and the adjusted packet inter-arrival time be t' - all given in milliseconds.

$t' = 0 \qquad t < t_{jb}$
$t' = t - t_{jb} \qquad t_{jb} < t < t_{discard}$
omit measurement $\qquad t > t_{discard}$

adjusted PDV = average(t')

I_e(PDV) = 0.1 adjusted PDV

E.5 Measuring Delay

The monitoring function estimates the round trip transmission delay using an echo mechanism, for example RTCP. This value is divided by two to give the estimated one way transmission delay t_{owtd} (Note: this assumes that the delay is symmetric).

The processing delay through the "transmitter" and "receiver" is:

CODEC encoder delay t_{enc}
Framing delay t_{frame}

Jitter buffer delay t_{jb}
Decoding delay t_{dec}

The values of these parameters for typical CODECs can be found in ITU Recommendation G.114 Appendix 1.

The overall one-way delay is therefore

$$t_d = t_{owtd} + t_{enc} + t_{frame} + t_{jb} + t_{dec}$$

E.6 Determining Equipment Impairment Factor for CODEC

ITU Recommendation G.113, Appendix 1 gives the following Equipment Impairment factors for certain CODECs

CODEC	G.711	G.729A + VAD	G.723.1 + VAD 6.3kbit/s
Ie (CODEC)	0	11	15

E.7 Determining Overall Equipment Impairment Factor

E.7.1 Determining Average Equipment Impairment Factor

It is generally acccepted that perceived quality does not change abruptly but exponentially "decays" from one level to another. This is intuitively obvious, as a 100mS burst of noise would be less annoying than a 10S burst.

Determine the Equipment Impairment value for the burst condition and the gap condition as:

$$I_{eg} = I_{eg}(LOSS) + I_e(PDV) + I_e(CODEC)$$

$$I_{eb} = I_{eg}(LOSS) + I_e(PDV) + I_e(CODEC)$$

Let I_1 be the quality level at the change from burst condition I_{eb} to gap condition I_{eg} and let I_2 be the quality level at the change from I_{eg} to I_{eb}

$$I_1 = I_{eb} - (I_{eg} - I_2)\, e^{-b/t1} \text{ where } t_1 \text{ is typically 5}$$

$$I_2 = I_{eg} + (I_1 - I_{eg})\, e^{-g/t2} \text{ where t2 is typically 15}$$

Combining these gives

$$I_2 = (\, I_{eg}\, (\, 1 - e^{-g/t2}) + I_{eb}\, (\, 1 - e^{-b/t1})\, e^{-g/t2}\,) / (1 - e^{-b/t1-g/t2})$$

Integrating the expressions for I_1 and I_2 to give a time average gives

$$I_e(av) = (\, b\, I_{eb} + g\, I_{eg} - t_1\, (I_{eb} - I_2)\, (1 - e^{-b/t1}) + t_2\, (I_1 - I_{eg})\, (1 - e^{-g/t2})\,) / (b + g)$$

E.7.2 Recency Effect

It has been noted by a number of researchers that the perceived quality of a call varies with the location of impairments. Impairments occurring late in a call have more effect than those occurring early in the call.

ANSI T1A1.7/98-031 described an experiment in which both mutes and noise bursts were introduced at the beginning, middle and end of a 60 second call. For the "high burst" result given:-

burst at start of call	MOS = 3.82
burst at middle of call	MOS = 3.28
burst at end of call	MOS = 3.18

ITU-T SG12 D.139 conducted an experiment in which a burst of high packet loss of duration 15, 30 or 60 seconds was introduced at the start, middle and end of a 180 second call and noted similar effects.

It is proposed that a simplified "adjustment" for recency be used, to minimise complexity. The delay since the last burst of packet loss is given above as y. It is assumed that the value of I_e at the end of the previous burst is given by I_1 and that the adjusted average quality approaches $I_e(av)$ exponentially.

$$I_e \text{ (end of call)} = Ie(av) + (k(I_1 - I_e(av)))\, e^{-y/t3}$$

where y is the delay to the previous burst, t_3 is a time constant (assumed to be 30) and k is a constant (assumed to be 0.7).

E.8 Use of the E Model

The Equipment Impairment value and the estimated one way delay determined above may be used as inputs to the E-Model (G.107) in order to calculate an R factor for the call. If other parameters required for the E-Model are unavailable then they should be set to their default values.