Packet-loss modelling in IP networks with state-duration constraints

N. Becerra Yoma, C. Busso and I. Soto

Abstract: A Gilbert–gamma topology is proposed to model packet-loss processes in UDP connections. The proposed topology introduces state duration modelling with gamma distributions. When compared with the ordinary Gilbert model the proposed topology substantially improves the likelihood of observed packet-loss processes, and gives reductions as high as 70% in the subjective estimation of speech quality transmitted over IP networks. The results presented can be easily applied to other real-time applications such as audio and video streaming.

1 Introduction

The accurate description of the packet-loss process in IP networks is very relevant to estimating perceptive quality in real-time applications (e.g. VoIP, audio and video streaming), to evaluating speech recognition systems in VoIP conditions, and to estimating the response of coder schemes in simulated UDP connections. In [1] the packet-loss process in the Internet was studied and modelled as a two-state Markov topology or Gilbert model (Fig. 1). This two-state Markov topology has been employed in several papers to analyse the response of TCP [2], to model the response of forward error correction (FEC) [1, 3–6], and to evaluate speech recognition algorithms [7, 8]. However, according to [9], where a two-dimensional discrete-time Markov chain is proposed to model the time interval between two consecutive bursts of lost packets, the ordinary Gilbert topology does not model properly the temporal behaviour of a non-blocking period (i.e. a burst of successfully received packets). Prediction of the subjective quality of speech in VoIP employing network parameters (e.g. packet loss) has been addressed by modelling bursts of lost packets with a four-state Markov chain [10] and with neural-network-based approaches [11, 12].

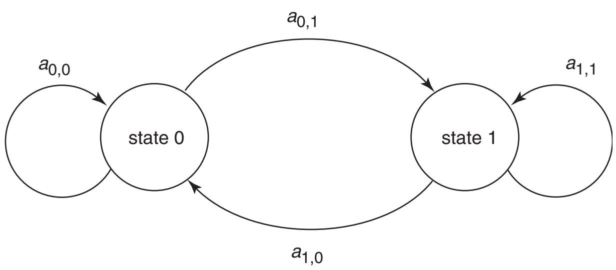

Fig. 1 *Gilbert model*

IEE Proceedings online no. 20040774

doi:10.1049/ip-com:20040774

Paper first received 6th January and in revised form 13th May 2004. Originally published online 13th December 2004

N. Becerra Yoma and C. Busso are with the Electrical Engineering Department, University of Chile, Av. Tupper 2007, P.O. Box 412-3, Santiago, Chile

I. Soto is with the Industrial Engineering Department, University of Santiago, Santiago, Chile

The probability transitions are defined by constants that lead to geometric state-duration distributions in the ordinary Markov chains. However, observed bursts of lost packets in UDP connections show that this model is inaccurate in many cases. The contribution of this paper concerns a Gilbert topology that incorporates state-duration modelling with gamma functions (Gilbert–gamma). When compared with the ordinary Gilbert model the proposed topology substantially improves the likelihood of the observed packet-loss process, and dramatically increases the accuracy of the subjective quality estimation of speech transmitted over IP networks. The results presented here have not been found in the specialised literature and can be employed in other real-time applications such as audio and video streaming.

2 Gilbert model and state-duration distribution

The packet-loss process is modelled here by the packet-loss rate (PL) and the probability distribution of burst length (BL), which in turn corresponds to the number of consecutive lost or received packets. The basic Gilbert model uses state 0 to represent a packet that was lost and state 1 for a packet that successfully reached the destination (Fig. 1). The probability that BL consecutive packets are lost is geometrically distributed and equal to $Pr(\text{burst length} = BL \text{ in state } 0) = (a_{00})^{BL-1} a_{0,1}$, where (see Appendix),

$$a_{1,0} = \frac{PL}{E_0[BL](1 - PL)} \quad (1)$$

$$a_{0,1} = \frac{1}{E_0[BL]} \quad (2)$$

where BL is the burst length in packets, and $E_0[BL]$ denotes the expected value of BL in state 0. Notice that $a_{0,0} = 1 - a_{0,1}$ and $a_{1,1} = 1 - a_{1,0}$, so the ordinary Gilbert model is completely defined by two variables, $a_{0,1}$ and $a_{1,0}$ which in turn can be estimated with PL and $E_0[BL]$. As a consequence $Var_0[BL]$, the variance of BL in state 0, is univocally determined with PL and $E_0[BL]$. The evaluation of the ordinary Gilbert model was done using SoLPs (sequence of lost packets) that correspond to time sequences that indicate which packets are successfully received and which are lost (Fig. 2). These sequences describe a UDP connection from the packet-loss point of view. The following methodology was implemented: UDP packets were transmitted at a constant bandwidth equal to 100 packets per second over two types of real Internet connection with a client process at the University of Chile in Santiago (UCh), either with a 100 Mbit/s Fast Ethernet or with a 256 kbit/s cable/DSL access. In both cases the server process was implemented in a host at the University of New Mexico in Albuquerque, USA (UNM). For each type of access at UCh, 20 20-minute connections were made and one SoLP per connection was recorded. Altogether, two types of access $\times 20$ connections $= 40$ SoLPs were stored. The parameters PL and $E_0[BL]$were estimated every 30-second frame from the *SoLPs*, and the distribution of BL in packet-loss bursts was evaluated in each 20-minute connection. According to the basic Gilbert model, the probability distribution of BL in packet-loss bursts should be geometric but this model is not accurate in many cases as observed in real Internet connections (Figs. 3 and 4). According to Fig. 3, the distribution of BL seems to be accurately modelled with a geometric function. In contrast the distribution of BL observed in a significant percentage of SoLPs is better fitted by a gamma function, as shown in the example presented in Fig. 4. This result suggests that the original Gilbert model could be improved by introducing state-duration constraints to model more accurately the packet-loss process. Notice that the geometric and gamma distributions are similarly monotonic when the expected value of BL is close to 1. It is worth emphasising that neither the type of access nor PL had an important effect on the probability distribution of BL.

3 Temporal restrictions in the Gilbert model

The problem of introducing state-duration modelling in Markov chains has already been addressed in speech recognition where hidden Markov models (HMMs) have widely been employed. According to [13] state-duration modelling can be included in the Viterbi algorithm by means of the generalisation of transition probabilities $a_{i,j}^{BL} = \mathrm{Prob}(s_{t+1} = j | s_t = s_{t-1} = \ldots = s_{t-BL+1} = i)$. Observe that BL is the number of consecutive packets in state i up to time t; $j = i$, or $j = i + 1$ (if $i = 0$) or $j = i - 1$ (if $i = 1$), given the topology shown in Fig. 1. Then the conditional probabilities $a_{i,i}^{BL}$ and $a_{i,i\pm 1}^{BL}$ can be estimated by

$$a_{i,i}^{BL} = \frac{\mathrm{Prob}(s_{t+1} = i, s_t = s_{t-1} = \ldots = s_{t-BL+1} = i)}{\mathrm{Prob}(s_t = s_{t-1} = \ldots = s_{t-BL+1} = i)} = \frac{D_i(BL) - d_i(BL)}{D_i(BL)} \quad (3)$$

$$a_{i,i\pm 1}^{BL} = \frac{\mathrm{Prob}(s_{t+1} = i \pm 1, s_t = s_{t-1} = \ldots = s_{t-BL+1} = i)}{\mathrm{Prob}(s_t = s_{t-1} = \ldots = s_{t-BL+1} = i)} = \frac{d_i(BL)}{D_i(BL)} \quad (4)$$

where $d_i(BL)$ is the probability of state duration equal to BL packets and $D_i(BL)$ is the probability of state i being active

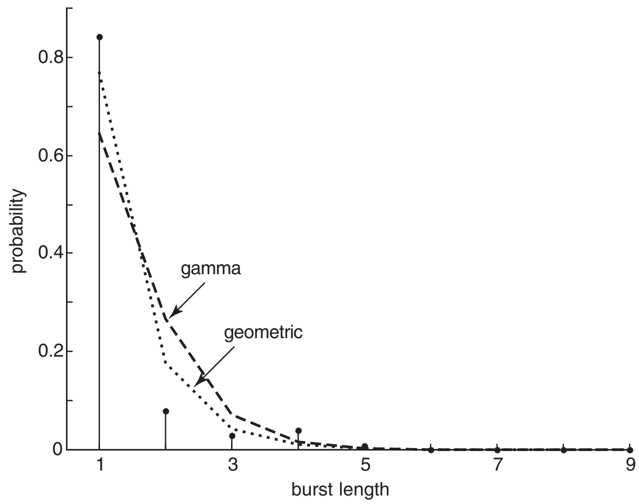

Fig. 3 *Example of burst-length histogram (stems) in packet-loss bursts observed in SoLP*
The corresponding geometric and gamma distributions are compared with the original histogram

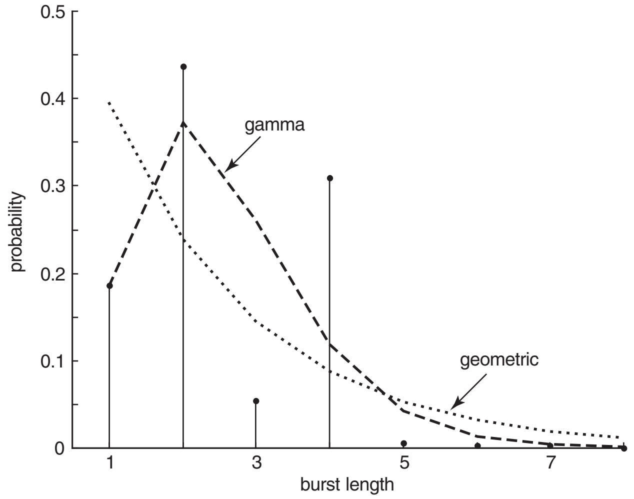

Fig. 4 *Example of burst-length histogram (stems) in packet-loss bursts observed in SoLP*
The corresponding geometric and gamma distributions are compared with the original histogram

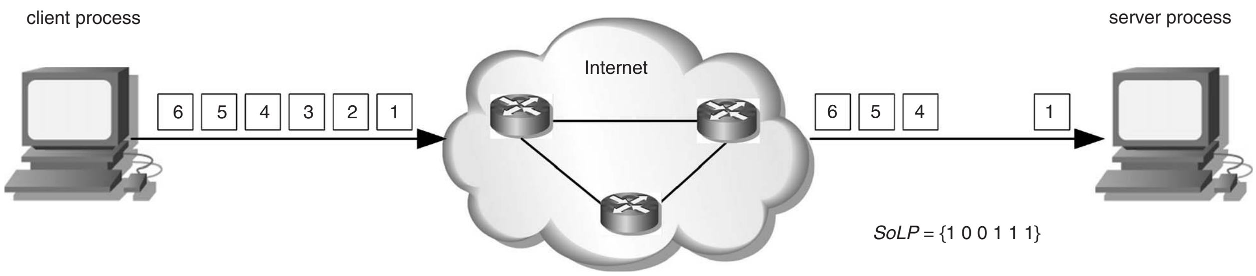

Fig. 2 *Procedure to generate sequences of lost packets (SoLP)*

for $t \geq BL$:

$$D_i(BL) = \sum_{t=BL}^{\infty} d_i(t) \quad (5)$$

To include the possible min and max durations, $\min_i(BL)$ and $\max_i(BL)$, respectively, the transition probabilities were modified to

$$a_{i,i}^{BL} = \begin{cases} 1 & \text{if} \quad BL < \min_i(BL) \\ 0 & \text{if} \quad BL > \max_i(BL) \\ \dfrac{D_i(BL) - d_i(BL)}{D_i(BL)} & \text{otherwise} \end{cases} \quad (6)$$

$$a_{i,i\pm l}^{BL} = \begin{cases} 1 & \text{if} \quad BL < \min_i(BL) \\ 0 & \text{if} \quad BL > \max_i(BL) \\ \dfrac{d_i(BL)}{D_i(BL)} & \text{otherwise} \end{cases} \quad (7)$$

where $d_i(BL)$ is modelled with the discrete gamma distribution given by:

$$d_i(BL) = K_i e^{-\alpha_i BL} BL^{p_i - 1} \quad (8)$$

where $BL = 1, 2, 3 \ldots$ is the duration of a given state i in number of packets; $\alpha_i = E_i[BL]/\mathrm{Var}_i[BL]$ and $p_i = E_i^2[BL]/\mathrm{Var}_i[BL]$; $E_i[BL]$ and $\mathrm{Var}_i[BL]$ are, respectively, the mean and variance of BL in state i; and K_i is a normalising term. The mean $E_i[BL]$ and variance $\mathrm{Var}_i[BL]$, and the min and max durations were computed for state $i = 0$ and $i = 1$ by means of directly observing the bursts of lost (state 0) and successfully received (state 1) packets. Notice that the Gilbert–gamma model uses statistics of bursts for both lost and successfully received packets. In contrast, the basic Gilbert model employs only PL and the expected value of BL in packet-loss bursts. It can easily be shown that, according to the Gilbert–gamma model, the packet-loss rate is given by

$$PL \approx \frac{E_0[BL]}{E_0[BL] + E_1[BL]}$$

where $E_0[BL]$ and $E_1[BL]$ are the expected value of BL in lost and successfully received packet bursts, respectively.

4 Gilbert model as generator of SoLP

Each SoLP can be decomposed as a sequence of frames. The wth frame or window in the nth SoLP is defined as $SoLP_{n,w} = (SoLP_{n,w,1}; SoLP_{n,w,2}; SoLP_{n,w,3}; SoLP_{n,w,l}; \ldots; SoLP_{n,w,L})$ where L is the frame length in packets and $SoLP_{n,w,l} = 0$ or $SoLP_{n,w,l} = 1$ if packet l was lost or successfully received, respectively. Every frame $SoLP_{n,w}$ could be described by the ordinary Gilbert or the Gilbert–gamma model, which in turn can be seen as process generators. Consequently the log-likelihood of frame $SoLP_{n,w}$ given the corresponding Gilbert model log-likelihood($SoLP_{n,w}$|Gilbert model in w) can be expressed as

$$\begin{aligned} &\text{log-likelihood}(SoLP_{n,w}|\text{Gilbert model in } w) \\ &\quad = \log[\mathrm{Prob}(SoPL_{n,w,1})] \\ &\quad + \sum_{l=2}^{L} \log[A_{s(n,w,l-1),s(n,w,l)}] \end{aligned} \quad (9)$$

where

$$A_{s(n,w,l-1),s(n,w,l)} = \begin{cases} a_{s(n,w,l-1),s(n,w,l)}, & \text{transition prob. in ordinary Gilbert model} \\ a_{s(n,w,l-1),s(n,w,l)}^{BL}, & \text{transition prob. in Gilbert–gamma model given by (6), (7)} \end{cases}$$

$s(n, w, l)$ is the state at packet l in frame w in the n^{th} SoLP; and $\mathrm{Prob}(SoPL_{n,w,1})$, the probability of $SoPL_{n,w,1}$, is equal to PL or to $1-PL$ if the first packet was lost or successfully received, respectively. As a consequence the log-likelihood of the nth SoLP given the Gilbert model, log-likelihood ($SoLP_n$ |Gilbert model), is

$$\begin{aligned} &\text{log-likelihood}(SoLP_n|\text{Gilbert model}) \\ &\quad = \sum_{w=1}^{W} \text{log-likelihood}(SoLP_{n,w}|\text{Gilbert model in } w) \end{aligned} \quad (10)$$

where W is the number of frames in $SoLP_n$; and log-likelihood ($SoLP_{n,w}$|Gilbert model in w) is defined in (9).

5 Results

The introduction of temporal restrictions in the ordinary Gilbert model was evaluated by computing the probability to generate SoLPs and by estimating the perceptual evaluation of speech quality (PESQ) as shown in Fig. 5. Notice that PESQ refers to the score given by the ITU-T P.862 standard, which has already shown acceptable accuracy with CELP codecs in transmission with packet loss [14]. Moreover, PESQ has also been employed by several authors to validate speech quality evaluations in packet-loss environments [12, 15, 16]. The 40 20-minute SoLPs described in section 2 were employed. In the first case the following procedure was adopted: PL, $E_i[BL]$, $Var_i[BL]$, and $\min_i(\tau)$ and $\max_i(\tau)$ in (6) and (7) were estimated every 30-s window in each SoLP. Secondly, the probability transitions $a_{0,1}$ and $a_{1,0}$ in the ordinary Gilbert model were computed according to (1) and (2). Thirdly, the parameters α_i and p_i of the gamma function defined in (8) that model the burst of received and lost packets were computed. Finally the log-likelihoods of each SoLP given the ordinary Gilbert and Gilbert–gamma models were evaluated with (10). The results which correspond to the average log-likelihoods within the set of 40 SoLPs are

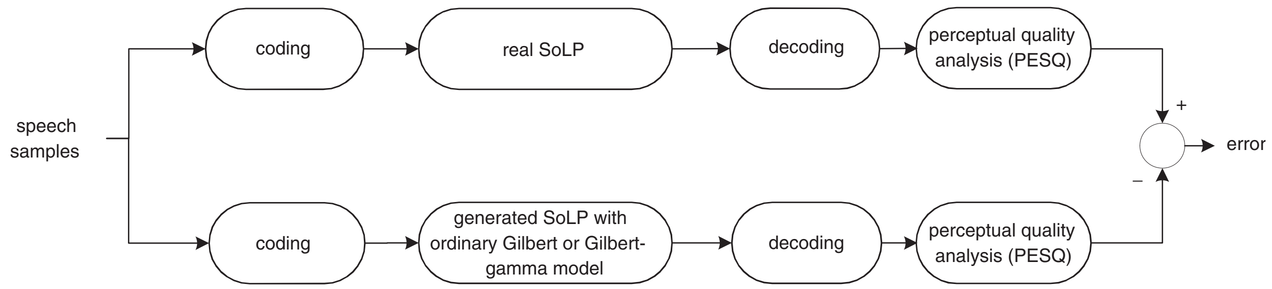

Fig. 5 *Scheme employed to compare original Gilbert and Gilbert–gamma models with respect to accuracy to map packet-loss process to PESQ score in UDP connection*

presented in Table 1. According to Table 1 the Gilbert–gamma model provides a significant increase in log-likelihood when compared to the ordinary Gilbert topology. This result strongly suggests that the introduction of state duration modelling with gamma distribution significantly improves the accuracy of the ordinary Gilbert model. Observe that the distribution of *BL* in successfully received packet bursts is also modelled with a gamma function. Moreover, the truncation of the conditional transition probabilities according to (6) and (7) also contributes to improve the description of the distributions of lost and received bursts.

Table 1: Average log-likelihoods within the set of 40 SoLPs

log[Gilbert]	log[Gilbert–gamma]
−120.5466806	−118.0527024

The Gilbert–gamma model was also evaluated according to its ability to map the packet-loss process to the PESQ score. Speech samples from 400 utterances (40 speakers and 10 utterances per speaker) from LATINO-40 database [17] were employed. Three coders were considered: 64 kbit/s PCM G.711 [18]; 32 kbit/s ADPCM G.726 [19]; and, 8 kbit/s CS-ACELP G.729 [20]. The procedure presented in Fig. 5 is described as follows: first, each observed SoLP was employed to simulate a real UDP connection by discarding lost packets after coding, and then the PESQ score was evaluated after decoding. Secondly, the same procedure was repeated by replacing the original *SoLP* with a sequence generated with the corresponding ordinary Gilbert or Gilbert–gamma models. Finally, the error in PESQ score was computed. Consequently the original Gilbert and Gilbert–gamma models were used as process generators to replace the observed *SoLP* as mentioned: *PL*, $E_i[BL]$, $\mathrm{Var}_i[BL]$, $\min_i(\tau)$ and $\max_i(\tau)$, and α_i and p_i of the gamma function defined in (8) were all estimated every 30-s window in each SoLP. As a result, each real SoLP was represented by sequences of ordinary Gilbert and Gilbert–gamma models as described in Section 4, and these sequences were employed to generate one simulated SoLP each. The average errors in PESQ score are shown in Table 2.

According to Table 2, the average errors in PESQ score are 3.83, 3.37 and 3.10 with the 64 kbit/s PCM G.711, 32 kbit/s ADPCM G.726 and 8 kbit/s CS-ACELP G.729 coders, respectively, when the ordinary Gilbert model was employed to generate the SoLPs. These errors are reduced to 1.08, 0.96 and 1.12 with the 64 kbit/s PCM G.711, 32 kbit/s ADPCM G.726 and 8 kbit/s CS-ACELP G.729 coders, respectively, when the ordinary Gilbert model is replaced with the Gilbert–gamma topology to generate the simulated SoLPs. Consequently the Gilbert–gamma model gave an overall reduction of 70% in PESQ error. This result strongly suggests that the Gilbert–gamma model describes a packet-loss process more precisely than the ordinary Gilbert model to reproduce the PESQ score given by the real SoLP. In other words, the proposed Gilbert–gamma topology dramatically improves the accuracy to predict the PESQ score from a packet-loss process.

Table 2: Average error (%) between PESQ score achieved with real SoLP and those obtained with simulated SoLPs generated with ordinary Gilbert and Gilbert–gamma models

Coder	Gilbert model error [%]	Gilbert–gamma model error [%]
G.729 (8 kbit/s)	3.0968	1.1193
G.726 (32 kbit/s)	3.3654	0.9583
G.711 (64 kbit/s)	3.8236	1.0782
Average	3.4286	1.0519

Number of real SoLPs was equal to 40

6 Conclusions

The Gilbert–gamma model has attempted to overcome the fact that the transition probability is represented by a constant in the ordinary two-state Markov model, which in turn leads to a geometric probability distribution for state duration. This model is not accurate enough to describe bursts of lost packets observed in UDP connections. The method proposed here models the state duration distributions with gamma functions and introduces conditional transition probabilities that are truncated according to the minimum and maximum observed burst lengths in lost or successfully received packet bursts. The results presented show that the proposed Gilbert–gamma model substantially increases the likelihood of observed SoLP when compared with the ordinary Gilbert topology.

As far as the ability to map packet-loss processes to PESQ scores is concerned, the Gilbert–gamma topology dramatically improves the accuracy of the mapping from a packet-loss process to a PESQ score. According to the results presented in this paper, the Gilbert–gamma model can give reductions as high as 70% in PESQ error when compared to the original gamma topology. As a consequence, the model proposed here could be applied to estimate the perceptive quality of speech in VoIP or to optimise the coding scheme according to the observed packet-loss process in real-time applications over IP networks. The evaluation of speech recognition systems with an artificially introduced packet-loss process to simulate VoIP conditions should also benefit from the results discussed in this paper. The Gilbert–gamma model could also be applied to estimate the perceptive quality of audio and video streaming in IP networks.

7 Acknowledgment

The authors would like to thank Dr. Philip Jackson, University of Surrey, UK, for having proofread this manuscript and for his suggestions. The authors also wish to thank Prof. Ramiro Jordan and John Salas, from the University of New Mexico, USA, for having provided the remote host employed in the experiments reported. The research described here was supported by FONDECYT no. 1030149.

8 References

1 Bolot, J.-C., Fosse-Parisis, S., and Towsley, D.: 'Adaptive FEC-based error control for Internet telephony'. Proc. 18th Annual Joint Conf. of IEEE Computer and Communications Societies, INFOCOM'99, New York, NY, USA, March 1999, Vol. 3, pp. 1453–1460

2 Abouzeid, A.A., Roy, S., and Azizoglu, M.: 'Stochastic modelling of TCP over lossy links'. Proc. 19th Annual Joint Conf. of IEEE Computer and Communications Societies, INFOCOM 2000, 2000, Vol. 3, pp. 1724–1733

3 Frossard, P.: 'FEC performance in multimedia streaming', *IEEE Commun. Lett.*, 2001, **5**, (3), pp. 122–124

4 Wah, L.K., and Heng, T.C.: 'Adaptive bandwidth-constrained media-specific forward error correction for voice over Internet Protocol'. Proc. 10th IEEE Int. Conf. on Networks, ICON 2002, 27–30 Aug. 2002, pp. 455–460

5 Miyoshi, M., Sugano, M., and Murata, M.: 'Performance improvement of TCP on wireless cellular networks by adaptive FEC combined with explicit loss notification'. Proc. IEEE 55th Vehicular Technology Conf., 6–9 May 2002, Vol. 2, pp. 982–986
6 Yuk, S.W., Kang, M.G., Kim, D.K., Shin, B.C., and Cho, D.H.: 'Adaptive redundancy control for systematic erasure code based real-time data transmission in Internet'. Proc. ICC 2000, June 2000, Vol. 3, pp. 1769–1773
7 Peláez-Moreno, C., Gallardo-Antolín, A., and Díaz-de-María, F.: 'Recognizing voice over IP: a robust front-end for speech recognition on the World Wide Web.', *IEEE Trans. Multimed.*, 2001, 2, (3), pp. 209–218
8 Quercia, D., Docio-Fernandez, L., Garcia-Mateo, C., Farinetti, L., and De, C.: 'Performance analysis of distributed speech recognition over IP networks on the AURORA database'. Proc. ICASSP'02, 2002, Vol. 4, pp. IV-3820–IV-3823
9 Zhu, Q., Kneuer, J.G., and Huang, W.-M.: 'A study of temporal behavior between the packet-loss bursts in packet-switched systems'. Proc. ICC'95, 1995, Vol. 3, pp. 1930–1936
10 Clark, A.D.: 'Modeling the effects of burst packet loss and recency on subjective voice quality'. Proc. IPTEL' 2001, New York, NY, USA, April 2001, pp. 123–127
11 Mohamed, S., Cervantes-Perez, F., and Afifi, H.: 'Integrating networks measurements and speech quality subjective scores for control purposes'. Proc. 20th Annual Joint Conf. IEEE Computer and Communications Societies, INFOCOM 2001, Vol. 2, pp. 641–649
12 Sun, L., and Ifeachor, E.C.: 'Perceived speech quality prediction for voice-over-IP-based networks'. Proc. IEEE Int. Conf. on Communications, ICC 2002, Vol. 4, pp. 2573–2577
13 Yoma, N.B., McInnes, F., Jack, M., Stump, S., and Ling, L.: 'On including temporal constraints in Viterbi alignment for speech recognition in noise', *IEEE Trans. Speech Audio Process.*, 2001, **9**, (2), pp. 179–182
14 ITU-T 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' Feb. 2001
15 Ding, L., and Goubran, R.A.: 'Assessment of effects of packet loss on speech quality in VoIP'. Proc. 2nd IEEE Int. Workshop on Haptic, Audio and Visual Environments and Their Applications, 20–21 Sept. 2003, pp. 49–54
16 Lee, M.K., Jung, S.K., Kang, H.G., Park, Y.C., and Youn, D.H.: 'A packet-loss concealment algorithm based on time-scale modification for CELP-type speech coders'. Proc. ICASSP'03, 2003, Vol. 1, pp. I-116–I-119
17 LDC: 'Latino-40 database provided by linguistic data consortium (LDC)', University of Pennsylvania: http://www.ldc.upenn.edu/Catalog/LDC95S28.html
18 ITU-T G.711: 'Recommendation G.711, Pulse code modulation (PCM) of voice frequencies' Nov. 1988.
19 ITU-T G.726: Recommendation G.726, '40-,32-,24-, and 16-Kb/s adaptive differential pulse code modulation' Dec. 1990
20 ITU-T G.729: 'Recommendation G.729-Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)' Mar. 1996

9 Appendix

Given the Gilbert model shown in Fig. 1,

$$PL = \text{Prob}[s(n, w, l) = 0] \tag{11}$$

where $s(n, w, l)$ is the state at packet l in frame w in the nth *SoLP* as mentioned in Section 4. Moreover,

$$\begin{aligned}\text{Prob}[s(n, w, l) = 0] &= a_{1,0}\text{Prob}[s(n, w, l-1) = 1]\\ &+ a_{0,0}\text{Prob}[s(n, w, l-1) = 0]\end{aligned} \tag{12}$$

If the packet-loss process is considered stationary in frame w

$$\text{Prob}[s(n, w, l) = 0] = \text{Prob}[s(n, w, l-1) = 0] \tag{13}$$

$$\text{Prob}[s(n, w, l) = 1] = \text{Prob}[s(n, w, l-1) = 1] \tag{14}$$

From (12), (13) and (14), PL can be expressed as

$$PL = \frac{a_{1,0}}{a_{1,0} + a_{0,1}} \tag{15}$$

The probability that BL consecutive packets are lost is geometrically distributed and equal to

$$\text{Prob}(BL) = (a_{0,0})^{BL-1} a_{0,1} \tag{16}$$

By making use of the expression

$$\sum_{j=0}^{\infty} A^{j-1} j = 1/(1-A)^2$$

with $|A| < 1$, it can easily be shown that the expected value of the burst length of consecutive lost packets $E_0[BL]$ is equal to

$$E_0[BL] = \sum_{BL=0}^{\infty} a_{0,1}\left(1 - a_{0,1}\right)^{BL-1} BL = \frac{1}{a_{0,1}} \tag{17}$$

Expressions (1) and (2) are derived from (15) and (17).