Stochastic Packet Loss Model to Evaluate QoE Impairments

Oliver Hohlfeld

Abstract—The estimation of quality for real-time services over telecommunication networks requires realistic models for impairments and failures during transmission. We focus on the classical Gilbert-Elliott model whose second-order statistics is derived over arbitrary time-scales. The model is used to fit packet loss processes of backbone and DVB-H traffic traces. The results show that simple Markov models are appropriate to capture the observed loss pattern and to discuss how such models can be used to examine the quality degradations caused by packet losses.

I. INTRODUCTION

WITH provisioning of broadband access for mass market—even in wireless and mobile networks—multimedia content, especially real-time streaming of high-quality audio and video, is extensively viewed and exchanged over the Internet. Quality of Experience (QoE) aspects, describing the service quality perceived by the user, is a vital factor in ensuring customer satisfaction in today's networks. Creating frameworks for accessing quality degradations in streamed video currently is investigated as a complex multi-layered research topic, involving network traffic load, codec functions and measures of user perception of video quality.

However, the transfer of real-time data for multimedia services over the Internet and channels in heterogeneous packet networks is subject to errors of various types, which will affect the QoS and QoE. On wireless and mobile links temporary and long lasting reductions in the available capacity frequently occur and even in fixed and wired network sectors packets may be dropped at routers and switches in phases of overload. Lost information will affect the perceived quality by impairing the multimedia content. The QoE degradation not only depends on the amount of lost packets, but also on the semantic of the lost information at the application layer.

The impact of packet loss on the user's perception of real-time services can be investigated starting from measurement traces of traffic or generated by finite-state stochastic models, which have been adapted to the characteristics observed in the measurement and thus produce statistically similar traces. Using model based generators for loss processes has several advantages:

- the amount of necessary storage capacity is reduced significantly from several gigabyte to a set of model parameters,
- stochastic models usually include a set of parameters with a clear interpretation, which can be adapted to meet the demands of a considered scenario in which the model is used (e.g. a certain packet loss rate) and makes them more flexible than a measurement trace,
- the length of the generated sequence is independent of the measurement trace used for training,

- stochastic models produce random but statistically consistent sequences.

Both, using real data loss traces—e.g. captured in backbone links—and model generated loss traces have their benefits. The main disadvantage of using model generated loss traces is that statistical properties may not fit to those of a measured trace, as they are likely to be biased by model limitations.

Measurement traces show characteristics on multiple time-scales. Thus, we derive the second-order statistics of finite Markov models, to be used as a parameter estimation technique. This is used to adapt the model to the second-order statistics of the amount of packet losses observed in a given traffic trace on multiple time-scales by moment matching.

The present paper gives a brief introduction on finite Markov models and discusses how these models can be used in the study of QoE impacts on video streams. The aim is to provide a generator for packet loss pattern to be used in the estimation of the degradation in the Quality of Experience for Internet services.

II. STOCHASTIC PACKET LOSS MODELS

Finite-state Markov chains are widely used to characterise error processes in telecommunication systems and for performance evaluation of coding or other measures for error resilience [8], [11]. In this paper, we will use finite-state models to describe the packet loss process found in backbone measurements (wired channels) and DVB-H traces (wireless channels). A discrete Markov chain with a set of M states $S = \{S_1, S_2, \cdots, S_M\}$ characterises the course of the process with regard to the current state, which may change over time at predefined events, e.g. packet arrivals, based on transition probabilities. Each state is associated with different error or packet loss behaviour. Let q_t denote the current state at event time $t, t \in \mathbb{N}_0$. Then the probabilities a_{ij} to change from state $q_{t-1} = i$ to $q_t = j$ are give in the transition matrix A with coefficients

$$a_{ij} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \le i, j \le M,$$

where

$$a_{ij} \ge 0; \quad \sum_{j=1}^{M} a_{ij} = 1.$$

We restrict our considerations to irreducible and aperiodic Markov chains, where each state can be reached from each other with positive probability after a number of transitions and steady state probabilities π_k exist for finding the process to sojourn a state in a long term perspective. The steady state probabilities are invariant with regard to a transition

with matrix A and thus can be computed from the system of linear equations: $\pi_k = \sum_{j=1}^{M} \pi_j a_{jk}, k = 1, \cdots, M$ and $\sum_{j=1}^{M} \pi_j = 1$. Finally, we define error or packet loss rates in each state $E = (e_1, ..., e_M)$; $0 \leq e_j \leq 1$ for $j \in [1, M]$ and the output of the process $O(t)$ as a binary sequence $O(t) \in 0, 1$ indicating an error or loss at an event with $O(t) = 1$, whereas $O(t) = 0$ stands for error free events, respectively. Thus we have $P(O(t) = 1 | q_t = S_j) \stackrel{\text{def}}{=} e_j$.

Only in simple cases, e.g. for a 2-state Markov process with $e_1 = 0$ and $e_2 = 1$, the current state can be recovered from the output. The term Hidden Markov Models expresses that $O(t)$ in general leaves uncertainty about the corresponding states $S(t)$ of the Markov chain. A Markov process is completely defined by the transition matrix A, the state specific error rates E and initial state S_0 or, more generally an initial state distribution $\pi_0 = P(S_0 = j)$. We continue with a brief discussion of two state Markov models before the second-order statistics of finite Markov models is derived in general.

A. Gilbert-Elliot: The Classical 2-State Markov Model for Error Processes

In 1960, Gilbert [3] proposed a 2-state Markov chain to characterise a burst-noise channel. The usual notation of the Gilbert model distinguishes at first a good (G) and secondly a bad (B) state with different loss rates $e_G = 1 - k < e_B = 1 - h$. Gilbert [3] started with the special case of an error-free good state $(k = 1)$ and left the extension to include losses generated in both states to Elliott [1]. Dwell times in the states are geometrically distributed with mean $1/p$ for the good and $1/r$ for the bad state, where p and r are the probabilities to change from the good state to the bad and vice versa. The Gilbert-Elliott 2-state Markov approach as depicted in Figure 1 is widely used for describing error patterns in transmission channels [12], [8] and for analysing the efficiency of coding for error detection and correction. For applications in data loss processes, we interpret an event as the arrival of a packet and an error as a packet loss. Thus, the transition matrix and the steady state probabilities are

$$A = \begin{pmatrix} 1 - p & p \\ r & 1 - r \end{pmatrix}; \quad \pi_G = \frac{r}{p+r}; \quad \pi_B = \frac{p}{p+r}$$

with a total error rate $e = \pi_G e_G + \pi_B e_B = \frac{r(1-k)+p(1-h)}{p+r}$.
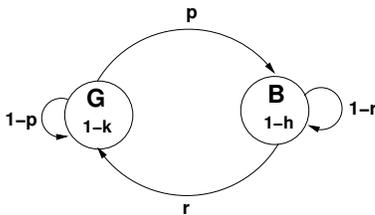
Fig. 1. The 2-state Markov model introduced by Gilbert and Elliott

III. VARIANCE OF THE ERROR PROCESS OVER MULTIPLE TIME-SCALES

Second-order statistics in multiple time-scales are a standard approach to capture and to describe traffic variability

including long-range dependencies and self-similarity [10], [7]. Following this trend, we next derive the second-order statistics of the number of packet losses over a range of relevant time frames. Based on our work in [4], [5], where we discussed the derivation in-depth, we will briefly discuss the concrete result for the Gilbert-Elliott model. We discussed the general eigenvalue solution for N-state Markov models in [6]. Although Markov models do not exhibit self-similar properties, they have been successfully adapted to self-similar traffic [2], [9] and are still popular since they often lead to simple analytical treatment.

In order to capture a packet loss process generated by an M-state Markov model, we can set up recursive equations for the distribution function of losses in a considered sequence of packets. Let $G_N(z)$ $(B_N(z))$ denote the generating function $X(z) \stackrel{\text{def}}{=} \sum_i P\{X = i\}z^i$ for the number of packet drops in a sequence of N packet arrivals, leaving the Markov chain in the last step at state G (B). Iterative relationships can be set up to compute $G_{N+1}(z)$ from $G_N(z)$, taking into account the state transitions and factors $(k + (1-k)z)$ and $(h + (1-h)z)$ for possible drop of the $(N+1)$-th packet with state dependent probabilities $1 - k$ and $1 - h$, respectively:

$$G_{N+1}(z) = (1-p)(k+(1-k)z)G_N(z) + r(h+(1-h)z)B_N(z)$$

$$B_{N+1}(z) = p(k+(1-k)z)G_N(z) + (1-r)(h+(1-h)z)B_N(z)$$

Starting in steady state conditions, we initialise $G_0(z) = r/(p+r)$ and $B_0(z) = p/(p+r)$. The corresponding distributions $G_N(z)$, $B_N(z)$ remain defective $G_N(1) = r/(p+r)$ and $B_N(1) = p/(p+r)$ $\forall N \in \mathbb{N}$. We finally evaluate complete distributions given by $G_N(z) + B_N(z)$, where $G_N(1) + B_N(1) = 1$ independent of the final state.

The k-th moment can be derived from the generating function by considering the k-th derivative: $E[X^k] = \frac{\partial}{\partial z^k} X(z)|_{z=1}$. The mean, given by $\mu_N^G = G'_N(1)$ and $\mu_N^B = B'_N(1)$, and the second moment $E(X^2)$ are sufficient to derive the second-order statistics involving the first and second derivative of the generating functions. Due to the symmetry of both states G and B, $G_N(z)$ can be obtained from $B_N(z)$ by swapping the parameters $p \leftrightarrow r$ and $h \leftrightarrow k$ and vice versa. Thus, $G_N(p, r, h, k, z) = B_N(r, p, k, h, z)$ and $\mu_N^G(p, r, h, k) = \mu_N^B(r, p, k, h)$.

The second-order statistics—expressed by the coefficient of variation $c_v = \sigma/\mu$—of the number of packet losses in a sequence of length N generated by the Gilbert-Elliott model is

$$c_v(N) = \frac{1}{\sqrt{N}} \sqrt{\frac{hp + kr}{\omega} + \alpha \frac{2pr(1-p-r)(h-k)^2}{\omega^2(p+r)}} \quad (1)$$

with

$$\alpha \stackrel{\text{def}}{=} \left(1 - \frac{1 - (1-p-r)^N}{N(p+r)}\right); \qquad \omega \stackrel{\text{def}}{=} (1-h)p + (1-k)r$$

The solution is comprehensible enough to interpret the influence of the model parameters and leads to a new parameter adaptation technique, where model parameters can be obtained by numerically fitting (1) to the empirical c_v-curve of a particular trace. Note that the evaluation of the term

$1 - (1 - p - r)^N$ may cause numerical instability for small p, r, which can be improved by implementing the equivalent form $1 - (1 - p - r)^N = 1 - e^{\ln(1-p-r) \cdot N}$.

IV. EVALUATION

The following section will discuss the evaluation of the proposed parameter adaptation technique based on second-order statistics in multiple time-scales for 2-state Markov models and compares the results to classical fitting methods.

The evaluation of the trained 2-state Markov models using the coefficient of variation $c_v(N)$ is shown in Figure 2 for a backbone trace and in Figure 3 for a DVB-H trace. The Poisson process provides a linear lower bound for $c_v(N) = c_v(1)/\sqrt{N}$ without any autocorrelation.

The parameters of the Gilbert model and its simplified version for $k = 1, h = 0$ have been estimated from the given loss traces using the traditional methods described in [3] and [12]. Moreover, the simplified Gilbert model with only two parameters and the Gilbert-Elliott model have been trained based on the second-order statistics over multiple time-scales $N \in [1, 10^5]$ (cf. Section III), as shown in Figure 2. The model parameters were estimated by fitting the coefficient of variation curve to the one obtained from the corresponding trace using the Levenberg-Marquardt algorithm for numeric optimisation of non-linear functions.
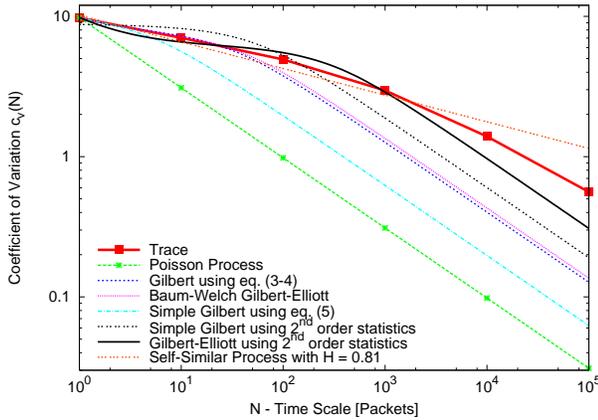
Fig. 2. Evaluation of the trained 2-state Markov models using the coefficient of variation $c_v(N)$ for backbone traffic with a mean packet loss rate of 1%

Considering the coefficient of variation curve as variance-time plot [7], the Hurst parameter can be obtained by fitting $c_v(1) \cdot N^{H-1}$ to the $c_v(N)$ curve for $N \in [1, 10^5]$. For both traces, a Hurst parameter $H \approx 0.8$ has been found, as can be seen in Figure 2. The loss process of the considered traces shows a faster decay than self-similar processes for a time-scale of $N > 10^3$ packets, suggesting that Markov models are appropriate to capture this behaviour.

The performance of the fitted models can be expressed by the degree to which they replicate the empirical c_v-curve of the considered trace. The proposed fitting technique shows a closer fit than classical methods. However, when we look at the distribution of the length of packet losses in a series, then the classical fitting procedures seem to be in favour, as experienced from first evaluations. This is not unexpected,
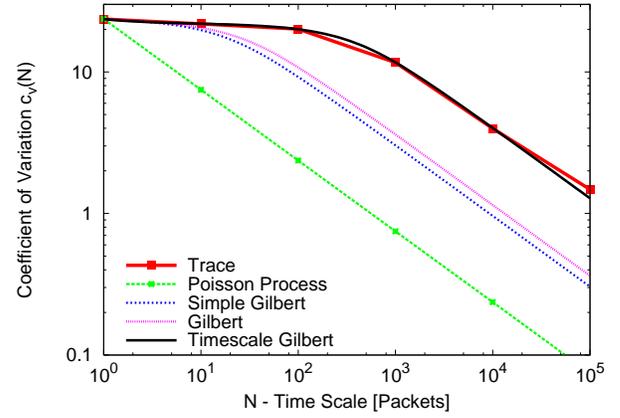
Fig. 3. Evaluation of the trained 2-state Markov models using the coefficient of variation $c_v(N)$ for DVB-H traffic with a mean packet loss rate of 1.8%

since they are closer related to error burst lengths whereas the second order statistics can include long-range correlation.

V. VIDEO QUALITY OF EXPERIENCE

In the following section, we discuss the impacts of packet loss on video quality metrics. We discussed related work in [5] in detail. The integration of the previous packet layer loss models into a common framework with the video transmission and impairments visible to a viewer are for further study in an ongoing project.

A. Impairments at the Video Layer caused by Packed Loss

We focus mainly on the H.264/AVC ((MPEG-4 Part 10) video coding standard. In video compression, sequences are decomposed into still pictures (*frames*) which are grouped into a *Group of Pictures* (GoP). Frames are segmented into *macro blocks*—usually of size 16x16 pels—and organized into groups of blocks called *slice*. Macro blocks are typically positioned in scan-line order, but due to Flexible Macroblock Ordering in H.264 arbitrary order is also possible. A frame is coded in one of three modes: *I frames* contain only intra coded macro blocks, *P frames* can contain intra or predicted macro blocks and *B frames* can also contain bidirectional predicted macro blocks. When using prediction, the displacement of a macro block is expressed as motion vector relative to the reference frame.

During the decoding process, the decoder may have to cope with errors due to packet loss, where the concrete error concealment strategy is decoder dependent. When only a small fraction of the slice is corrupted, the decoder often is not aware of the concrete error position and thus may discard the slice. The perceived visual quality of an erroneous received but decoded and displayed slice may be worse than in case of discarding and copying the slice from a previous frame. Information introduced by copying from a previous slice may not always be noticeable by a viewer. An example may be comic strips which usually have a very static background and thus losing a slice that holds only background information is often unnoticed if the lost data is copied from the last slice.

As most currently used encoders are not aware of the underlying transport protocol, slices are typically larger than the payload of a single IP packet. When MPEG2 transport streams (TS) are used to packetize the video data into IP packets, slice boundaries are usually inside an IP packet, as MPEG-2 TS considers video data as endless bitstream. Thus, losing such a packet can discard two slices at once.

Losing a parameter set due to packet loss will have a severe effect on the decoding process. Thus, they may be transmitted out-of-band or in-band but periodically, e.g. with each GoP, or prioritized using differentiated services.

B. Markov Models as Basis for Subjective Evaluations

Finite Markov models can be used as a basis for subjective video quality evaluations where subjects are asked to rate given video sequences using the Mean Opinion Score (MOS), as currently conducted at T-Labs Berlin for high-definition video sequences. During the test preparation, H.264 videos are packetised using an MPEG2 transport stream (TS) and impaired with a Markov model. The impaired video sequence is rated by the subject during test. The advantage of Markov models is their flexibility to be adapted to arbitrary loss rates and loss burst lengths, such that different impairments can be studied.

C. Objective Evaluation of Impaired Video Sequences

The analytical study of loss pattern is supplemented with a qualitative investigation of their effects on QoE metrics for video transmission in [5]. In general, impairments of video sequences are sensitive to different assumptions on loss processes yielding different reaction in coding of slices and frames, which stresses the importance of referring to the most relevant pattern. A comparison of 2400 video sequences, impaired using the proposed packet loss generator, led to strong evidence that uniform packet losses produce a larger distortion than busty packet losses at the same and sufficiently high loss-rate, since the former affect more frames on the video coding layers. Moreover, the results suggest that impairments at the video layer are affected by the applied parameter estimation technique.

VI. CONCLUSION

Quality of Experience aspects are a vital factor in ensuring customer satisfaction in today's network services. In order to study quality degradations in video streams caused by packet loss, a Markovian error pattern generator can be used to simulate physical channels of various type. We derived the second-order statistics for the distribution of the number of lost packets over multiple time-scales, which in general can be recursively determined for increasing time frames and via explicit terms for the Gilbert-Elliott model. The fitting procedure leads to a closer match in multiple time-scales than classical methods. The proposed approach gives more flexibility to include information from different time-scales enabling a simple and useful fit for long traces of traffic and packet loss processes. We discussed the impact of packet loss on streamed video and showed how Markovian error pattern generators can be used to study quality degradations.

REFERENCES

[1] E. O. Elliott. Estimates of Error Rates for Codes on Burst-Noise Channels. *Bell System Technical Journal*, 42:1977–1997, September 1963.

[2] Anja Feldmann and Ward Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance Evaluation*, 31(3):245–279, 1998.

[3] E. N. Gilbert. Capacity of a Burst-Noise Channel. *Bell System Technical Journal*, 39:1253–1265, September 1960.

[4] Gerhard Haßlinger and Oliver Hohlfeld. The Gilbert-Elliott Model for Packet Loss in Real Time Services on the Internet. In *Proceedings of the 14th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems (MMB)*, Dortmund, Germany, March 2008.

[5] Oliver Hohlfeld. Statistical error model to impair an H.264 decoder. Master's thesis, TU Darmstadt, March 2008. KOM-D-0310.

[6] Oliver Hohlfeld, Rüdiger Geib, and Gerhard Haßlinger. Packet Loss in Real-Time Services: Markovian Models Generating QoE Impairments. In *Proc. of the 16th International Workshop on Quality of Service (IWQoS)*, pages 261–270, June 2008.

[7] Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson. On the self-similar nature of ethernet traffic. *ACM SIGCOMM Computer Communication Review*, 23(4):183–193, 1993.

[8] Jussi Poikonen. Geometric Run Length Packet Channel Models Applied in DVB-H Simulations. In *Proceedings of the IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2006.

[9] Paulo Salvador, Rui Valadas, and Antonio Pacheco. Multiscale fitting procedure using markov modulated poisson processes. *Telecommunication Systems*, 23:123–148, 2003.

[10] Boris Tsybakov and Nicoals D. Georganas. On self-similar traffic in atm queues: definitions, overflow probability bound, and cell delay distribution. *IEEE/ACM Trans. Netw.*, 5(3):397–409, 1997.

[11] William Turin. *Performance Analysis and Modeling of Digital Transmission Systems (Information Technology: Transmission, Processing and Storage)*. Springer New York, Inc., Secaucus, NJ, USA, 2004.

[12] Maya Yajnik, Sue B. Moon, James F. Kurose, and Donald F. Towsley. Measurement and modeling of the temporal dependence in packet loss. In *IEEE INFOCOM*, pages 345–352, 1999.

Oliver Hohlfeld is a research assistant in the intelligent networks research group at Deutsche Telekom Laboratories / TU Berlin. He studied computer science at Darmstadt University of Applied Sciences, Institut Eurecom (France) and Darmstadt University of Technology with a focus on computer networks. From 2004 to 2006 he was with the Fraunhofer IGD where he worked on network schemes for telemedical applications. From 2007 to 2008 he was with T-Systems where he worked on his master's thesis entitled *Statistical Error Model to Impair an H.264 Decoder*, which has been awarded by the KuVS group.