The Gilbert-Elliott Model for Packet Loss in Real Time Services on the Internet

Gerhard Haßlinger[1] and Oliver Hohlfeld[2]

[1] T-Systems, Deutsche-Telekom-Allee 7, D-64295 Darmstadt, Germany
[2] Darmstadt University of Technology, Merckstr. 25, D-64283 Darmstadt, Germany
`gerhard.hasslinger@t-systems.com,hohlfeld@kom.tu-darmstadt.de`

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 and used to fit packet loss processes of traffic traces measured in the IP backbone of Deutsche Telekom. The results show that simple Markov models are appropriate to capture the observed loss pattern.

1   Introduction

The transfer of real-time data over the Internet and channels in heterogeneous packet networks is subject to errors of various types. A packet can be corrupted—and therefore is unusable for a voice or video decoder—due to unrecoverable bit failures. 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.

Most of the Internet traffic is controlled by the TCP protocol, which provides mechanisms for retransmission of lost or corrupted data and for controlling the load on congested links involving FIFO queues with a Tail-Drop or Random Early Detection (RED) [1] policy. On the other hand, the portion of uncontrolled traffic via the UDP transport protocol has been increased to a level of 5 - 10% in recent time [2], partly since real-time services over IP including voice, video on demand and online gaming are gaining in popularity. The upcoming deployment of IP-TV over VDSL broadband access platforms by Deutsche Telekom and other Internet service providers will strengthen this trend.

In this work we focus on packet loss on Internet links with most traffic controlled by TCP superposed with a considerable contribution of real-time traffic without flow control. Under sufficiently high link load, this causes spontaneous overload peaks causing packet loss. Available traffic traces [2] show, that UDP traffic has a higher variability in the relevant time scales than the total traffic, which at the present stage is dominated by peer-to-peer data exchange.

The impact of transmission errors on the user perception of real-time services can be investigated starting from measurement traces of traffic and loss pattern. In addition, a stochastic model can be set up and used to generate a considered

error process with similar characteristics as observed in the measurement. The Gilbert-Elliott model [3, 4] is one of the most popular examples, which has been preferably applied to bit error processes in transmission channels. Model driven studies usually include a set of parameters with a clear interpretation, which have to be adapted to a considered scenario. Their main advantage lies in an abstraction level, which makes them much more flexible than a fixed measurement trace. Thus the impact of different error rates, burstiness of error pattern etc. can be studied in a common modeling framework.

Both, using real data loss traces—e.g. captured in backbone links—and model generated loss traces has its benefits. The main disadvantage of using model generated loss traces is that statistical properties may not fit and thus traces can be biased by model limitations. The present paper will propose a parameter estimation technique for a 2-state Markov model to adapt the model to the second order statistics observed in a given traffic trace on multiple time scales by moment matching.

In Section 2 we characterise the packet loss pattern observed in traffic traces based on the second order statistics, i.e. the coefficient of variation, in multiple time scales. We consider simple Markov processes to be fitted to the observed second order statistics.

Section 3 summarises classical fitting schemes for the Gilbert-Elliott model [3, 4]. They do not cover the second order statistics, which we found to be non-trivial along the derivation shown in Section 4. In Section 5, a comparison of the model with adapted parameters to the packet loss pattern derived from traffic traces shows that simple Markov processes achieve a fairly close fit to the mean and variances over multiple time scales. Section 6 considers related work.

2 Packet Loss Process in Data Transfer over Multiple Time Scales

We consider a typical scenario found in backbone links of controlled TCP packet flows being superposed with real-time traffic over the UDP protocol, which does not provide error recovery and flow control mechanisms. We refer to measurement traces of traffic taken from a 2.5 Gb/s interface of a broadband access router of Deutsche Telekom's IP platform, which connects residential ADSL access lines to the backbone. Based on the time stamp and the size of each packet, the variability of the traffic rates can be observed in time scale ranging from the accuracy level of the time stamps well below 1 ms up to the 30 minutes length of the traces. As the packet loss process shows characteristic behaviour on multiple time scales, techniques used for describing the variability in traffic rates will be also used for describing the packet loss process in this paper.

Let Δ be a time frame in this range. Then corresponding traffic rates $R_k(\Delta)$ are determined for successive intervals of length Δ by dividing the sum of the size of all packets arriving in a time interval by its length. From the sequence $R_k(\Delta)$ the mean rate μ_R and the variance $\sigma_R^2(\Delta)$ are computed. In this way, the second order statistics is given considering $\sigma_R^2(\Delta)$ over a relevant range of Δ.
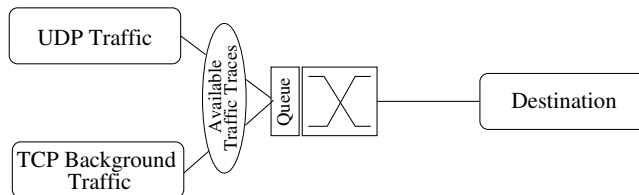
Fig. 1. Measurement topology: TCP backbone traffic is feed from a trace file along with UDP traffic into a router. The traffic is directed over an bottlenecked link to a destination. The loss rate can be arbitrarily chosen by adjusting the capacity of the outgoing, bottlenecked link.

This statistics is a standard description method for traffic and is equivalent to the autocorrelation function over the considered time scales. Long range dependent traffic patterns are classified as exact or second order self-similar depending on the autocorrelation of the process [5, 6].

Table 1 shows the second order statistics for $\Delta = 1$ ms, 10 ms, 100 ms, 1 s and 10 s measured for the UDP and the total traffic. The coefficients of variation $c_v(\Delta) = \sigma(\Delta)/\mu$ are observed to be about twice as high for UDP as for the total traffic.

	Mean Rate	$c_v(1 \text{ ms})$	$c_v(10 \text{ ms})$	$c_v(100 \text{ ms})$	$c_v(1 \text{ s})$	$c_v(10 \text{ s})$
UDP traffic	$\mu = 50.8$ Mb/s	0.3209	0.1220	0.0531	0.0433	0.0394
Total traffic	$\mu = 753.9$ Mb/s	0.1689	0.0635	0.0322	0.0259	0.0216

Table 1. Second order statistics for $\Delta = 1$ ms, 10 ms, 100 ms, 1 s and 10 s for the UDP and the total traffic

In this paper, we adhere to the second order statistics for describing the packet loss process. The traffic traces are at a load level of about 30% and originally do not exhibit packet losses in the considered time scales. However, at higher load, i.e. for reduced capacity < 2.5 Gb/s, overload phases occur above some medium load level and we can easily compute the resulting packet loss process corresponding to the trace at any sufficiently high load level. In general, the loss pattern is evaluated for a predefined capacity C (versus load) including a buffer of limited size B, assuming that an arriving packet is lost by tail drop each time when it does not fit into the remaining buffer. The loss pattern obtained in this way are adequate for uncontrolled UDP traffic, but do not regard the TCP retransmission and source rate adaptation. However, the TCP control does not respond on the 1 ms, but on essentially larger time scales. We assume that TCP will establish a stabilised non-excessive load level without much data loss and will focus on the UDP traffic portion with regard to TCP background traffic.

We obtain the packet loss process from the traces at a predefined load level and calculate its second order statistics. Since the loss rate is monotonously increasing with the load, we can adjust the load in order to approach a considered packet loss rate.

Next, we study simple Markov models again with focus on their second order statistics. 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 (QoE) for Internet services.

3 Gilbert-Elliott: The Classical 2-State Markov Model for Error Processes

We consider the 2-state Markov approach as introduced by Gilbert [3] and Elliott [4], which is widely used for describing error patterns in transmission channels [7–15] and for analysing the efficiency of coding for error detection and correction [16]. We follow the usual notation of a good (G) and bad (B) state. Each of them may generate errors as independent events at a state dependent error rate $1 - k$ in the good and $1 - h$ in the bad state, respectively. The model is shown in Figure 2. For applications in data loss processes, we interpret an event as the arrival of a packet and an error as a packet loss. The transition matrix A is given by the two transitions

$$p = P(q_t = B | q_{t-1} = G); \quad r = P(q_t = G | q_{t-1} = B); \quad A = \begin{pmatrix} 1 - p & p \\ r & 1 - r \end{pmatrix},$$
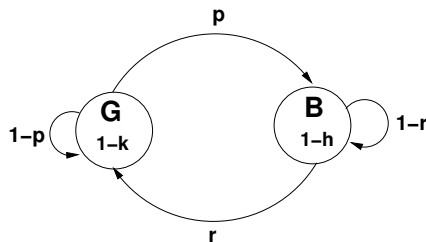
(1)

where q_t denotes the state at time t.

Fig. 2. The Gilbert-Elliott model generating a 2-state Markov modulated failure process

The stationary state probabilities π_G and π_B exist for $0 < p, r < 1$ [16] from which the error rate p_E is obtained in steady state:

$$p_E = (1 - k)\pi_G + (1 - h)\pi_B; \quad \pi_G = \frac{r}{p + r}, \quad \pi_B = \frac{p}{p + r}.$$

(2)

In 1960, Gilbert [3] proposed a model to characterise a burst-noise channel. It adds memory to the Binary Symmetric Channel coded into two states of the Markov chain. Gilbert considered the special case of an error-free good state ($k = 1$) and suggested to estimate the model parameters from three measurable instances of a binary error process $\{E_t\}_{t \in \mathbb{N}}$, where $E_t = 1$ indicates an error:

$$a = P(1), \quad b = P(1|1), \quad c = \frac{P(111)}{P(101) + P(111)}.$$ (3)

By knowing a, b and c, the three model parameters can be computed in the following manner

$$1 - r = \frac{ac - b^2}{2ac - b(a + c)}; \quad h = 1 - \frac{b}{1 - r}; \quad p = \frac{ar}{1 - h - a}.$$ (4)

Gilbert argues that the c measurement may be avoided by choosing $h = 0.5$ and using $1 - r = 2b$. Furthermore, he showed that the method introduced above can lead to ridiculous parameters ($p, r, h < 0$, or $p, r, h > 1$), if the observation (the trace) is too small. Morgera et al. [17] also conclude that the method proposed by Gilbert is more appropriate for longer traces. In case of shorter observations, better results can be obtained when considering the Gilbert model as Hidden Markov Model trained by the Baum-Welch algorithm [18–20].

Parameters of an even simplified Gilbert model with $h = 0$ can also be estimated with the method presented by Yajnik et al. [8]

$$p = P(1|0); \quad r = P(0|1).$$ (5)

A more intuitive parameter estimation technique can be found by considering the Average Burst Error Length (ABEL) to determine $r = 1/\text{ABEL}$ and the average number of packet drops to determine p_E. Equation (2) leads to $p = p_E \cdot r/(h - p_E)$. Gilbert's model was extended by Elliott [4] in 1963 including errors in both states as in Figure 2.

Model	Parameter	Training Complexity	Simplification
Simple Gilbert	p, r	simple	$k = 1, h \in \{0, 0.5\}$
Gilbert	p, r, h	medium	$k = 1$
Gilbert-Elliott	p, r, h, k	high	/

Table 2. Comparison of simplified two-state Markov channel models

4 Variance of the Error Process over Multiple Time Scales

The second order statistics of the 2-state Markov process can be derived via generating functions. While it is straightforward to compute the distribution

function of errors in time frames of length $N + 1$ iteratively from the result for length N, a non-iterative direct solution is less obvious already for the 2-state Markov model [21]. To the authors knowledge, explicit expressions for the variance of the number of errors during a time frame of fixed length, are not given in the literature, although there is a large volume of work involving the Gilbert-Elliott model, as partly discussed in Section 6 on related work. However, most of this work is devoted to error detecting and correcting codes and the residual error probabilities of coding schemes, rather than on traffic or packet loss characterisation. Second order statistics in multiple time scales is a standard approach in teletraffic modelling [5, 6].

Although Markov models do not exhibit self-similar properties, they have been successfully adapted to self-similar traffic [22] and are still popular since they often lead to simple analytical results. Following this trend, we next derive the variance of the number of packet drops as errors in the 2-state Gilbert-Elliott model over a range of relevant time frames.

4.1 Generating Functions

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:

$$G_{N+1}(z) = (1 - p)(k + (1 - k)z)G_N(z) + r(h + (1 - h)z)B_N(z) \qquad (6)$$

$$B_{N+1}(z) = p(k + (1 - k)z)G_N(z) + (1 - r)(h + (1 - h)z)B_N(z) \qquad (7)$$

Starting in steady state conditions we initialise

$$G_0(z) = \frac{r}{p + r}; \quad B_0(z) = \frac{p}{p + r}. \qquad (8)$$

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 [23, 24]: $E[X^k] = \frac{\partial}{\partial z^k} X(z)|_{z=1}$. The mean $\mu_X = E(X)$ 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.

$$G'_{N+1}(z) = (1 - p) \cdot ((1 - k) \cdot G_N(z) + (k + (1 - k)z) \cdot G'_N(z))$$
$$+ r \cdot ((1 - h) \cdot B_N(z) + (h + (1 - h)z) \cdot B'_N(z)) \qquad (9)$$

$$G''_{N+1}(z) = (1 - p) \cdot (2(1 - k) \cdot G'_N(z) + (k + (1 - k)z) \cdot G''_N(z))$$
$$+ r \cdot (2(1 - h) \cdot B'_N(z) + (h + (1 - h)z) \cdot B''_N(z)) \qquad (10)$$

$$B'_{N+1}(z) = p \cdot ((1-k) \cdot G_N(z) + (k + (1-k)z) \cdot G'_N(z))$$
$$+ (1-r) \cdot ((1-h) \cdot B_N(z) + (h + (1-h)z) \cdot B'_N(z)) \tag{11}$$

$$B''_{N+1}(z) = p \cdot (2(1-k) \cdot G'_N(z) + (k + (1-k)z) \cdot G''_N(z))$$
$$+ (1-r) \cdot (2(1-h) \cdot B'_N(z) + (h + (1-h)z) \cdot B''_N(z)) \tag{12}$$

4.2 Mean Values

The mean values are given by $\mu_N^G = G'_N(1)$ and $\mu_N^B = B'_N(1)$, which leads to the following expressions

$$\mu_{N+1}^G = (1-p) \cdot \left(\frac{(1-k)r}{p+r} + \mu_N^G \right) + r \cdot \left(\frac{(1-h)p}{p+r} + \mu_N^B \right), \tag{13}$$

$$\mu_{N+1}^B = p \cdot \left(\frac{(1-k)r}{p+r} + \mu_N^G \right) + (1-r) \cdot \left(\frac{(1-h)p}{p+r} + \mu_N^B \right). \tag{14}$$

Considering the sum of μ_{N+1}^G and μ_{N+1}^B leads to the expected result of $N+1$ times the failure rate in the steady state:

$$\mu_{N+1} = \mu_{N+1}^G + \mu_{N+1}^B = (N+1) \left(\frac{(1-k)r}{p+r} + \frac{(1-h)p}{p+r} \right) = (N+1)p_E. \tag{15}$$

To eliminate the reference to the opposite term, μ_{N+1}^B can be rewritten as

$$\mu_{N+1}^B = \frac{pr(1-k)}{p+r} + p \cdot \left(\underbrace{\mu_N^G + \mu_N^B}_{(15)} - \mu_N^B \right) + (1-r) \cdot \left(\frac{p(1-h)}{p+r} + \mu_N^B \right)$$
$$= (N+1) \cdot \left(\frac{pr(1-k)}{p+r} + \frac{p^2(1-h)}{p+r} \right) + [1 - (p+r)] \cdot \left(\frac{p(1-h)}{p+r} + \mu_N^B \right). \tag{16}$$

Next, we structure the above equation according to their dependence on N and μ_N^B with abbreviations for the main terms α, β_B and γ_B:

$$\mu_{N+1}^B = (N+1) \cdot \beta_B + \alpha \cdot (\mu_N^B + \gamma_B); \quad \alpha := 1 - (p+r); \tag{17}$$

$$\beta_B := \frac{pr(1-k)}{p+r} + \frac{p^2(1-h)}{p+r}; \quad \gamma_B := \frac{(1-h)p}{p+r}. \tag{18}$$

Computing a series of the first mean values

$$\mu_1^B = \beta_B + \alpha\gamma_B, \quad \mu_2^B = 2\beta_B + \alpha(\beta_B + \gamma_B) + \alpha^2\gamma_B, \tag{19}$$

suggests the general result, which is proven by induction over N:

$$\mu_N^B = -\gamma_B + \sum_{j=0}^{N} \alpha^j (\gamma_B + \beta_B \gamma_B (N - j))$$

$$= \beta_B \frac{N}{1 - \alpha} + \left(\gamma_B - \frac{\beta_B}{1 - \alpha}\right) \frac{\alpha}{1 - \alpha} \left(1 - \alpha^N\right), \quad \text{for } \alpha \neq 1 \qquad (20)$$

The case $\alpha = 1$, which means $p = r = 0$ implies a reducible and thus non-ergodic Markov chain, which is not relevant for modelling purposes.

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 $\mu_N^G(p, r, h, k) = \mu_B^N(r, p, k, h)$. Considering the sum $\mu_N^B + \mu_N^G$ again leads to Equation (15).

4.3 Explicit Solution for the Variance

Using the mean values, the variance of the number of packet losses in a time frame of size N can be derived as follows

$$G_{N+1}''(1) + B_{N+1}''(1) = 2(1 - k)\mu_N^G + 2(1 - h)\mu_N^B + G_N''(1) + B_N''(1)$$

$$= 2(1 - k) \sum_{i=1}^{N} \mu_i^G + 2(1 - h) \sum_{i=1}^{N} \mu_i^B. \qquad (21)$$

The sum of the mean values yields

$$\sum_{i=1}^{N} \mu_i^s = \sum_{i=1}^{N} \beta_s \frac{i}{1 - \alpha} + \left(\gamma_s - \frac{\beta_s}{1 - \alpha}\right) \frac{\alpha}{1 - \alpha} \left(1 - \alpha^i\right), \quad s \in \{G, B\}$$

$$= -\beta_s \frac{N(N + 1)}{2(1 - \alpha)} + \left(\gamma_s - \frac{\beta_s}{1 - \alpha}\right) \frac{\alpha}{1 - \alpha} \left(N - \frac{\alpha(1 - \alpha^N)}{1 - \alpha}\right).$$

Based on the relationship $G_N''(1) + B_N''(1) = \mu_N^2 + \sigma_N^2 - \mu_N$, the previous solution for $G_{N+1}''(1) + B_{N+1}''(1)$ yields the standard deviation $\sigma_N = \sigma_N^B + \sigma_N^G$ of the number of lost packets as well as the coefficient of variation $c_v(N) = \sigma_N / \mu_N$. We finally arrive at the following result for $c_v(N)$ expressing the second order statistics of the number of errors or packet losses in a sequence of length N generated by the Gilbert-Elliott model:

$$c_v(N) = \frac{\sqrt{G_N''(1) + B_N''(1) - \mu_N^2 + \mu_N}}{\mu_N}; \qquad \omega := (1 - h)p + (1 - k)r$$

$$= \frac{1}{\sqrt{N}} \sqrt{\frac{hp + kr}{\omega} + \frac{2pr(1 - p - r)(h - k)^2}{\omega^2(p + r)} \left(1 - \frac{1 - (1 - p - r)^N}{N(p + r)}\right)}.$$

$$(22)$$

The solution is comprehensible enough to interpret the influence of the model parameters. 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 $1 - (1 - p - r)^N = 1 - e^{\ln(1 - p - r) \cdot N}$.

4.4 Simple Cases

In case of $h = k$, both states are indistinguishable and the Markov chain collapses to a single state leading to the simplified result

$$c_v(N) = \sqrt{\frac{h}{(1-h)N}}. \tag{23}$$

This corresponds to a binomial distribution $G_N(z) + B_N(z) = [h + (1-h)z]^N$ of independent random packet losses generated by a memoryless process.

If $p + r = 1$, the Markov chain again generates a memoryless process, since the transition probabilities, e.g. to state B, are the same starting from B or G:

$$P(q_t = B|q_{t-1} = G) = p; \quad P(q_t = B|q_{t-1} = B) = 1 - r = p. \tag{24}$$

Again, the coefficient of variation is simplified:

$$c_v(N) = \sqrt{\frac{1}{N}\frac{hp + kr}{(1-h)p + (1-k)r}}. \tag{25}$$

The precondition $p + r \ll 1/N$ also leads to a simpler representation of the form, since the last fraction in Equation (22) approaches 0 in that case:

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

4.5 Parameter Impact on the Second Order Statistics of the Gilbert-Elliott Model

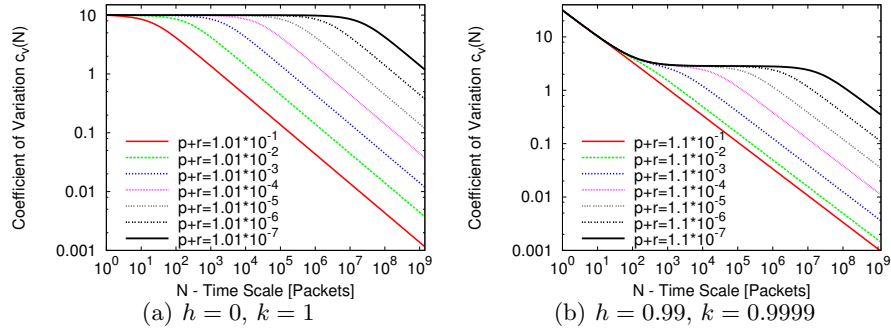(a) $h = 0$, $k = 1$

(b) $h = 0.99$, $k = 0.9999$

Fig. 3. Parameter impact on the second order statistics of the Gilbert-Elliott model

Based on the analytical result in Equation (22) for $c_v(N)$, the main properties of the second order statistics of the Gilbert-Elliott model become visible.

1. The starting point of the curves for $c_v(N)$ is given by $c_v(1) = \sqrt{(hp + kr)/\omega} = \sqrt{1/p_E - 1}$ and thus only depends on the packet loss rate.
2. Figure 3(a) shows curves of $c_v(N)$ for $h = 0$ and $k = 1$ such that the bad state generates bursts of subsequent packet losses and $p_E = \pi_B = r/(p + r)$. In all examples of Figure 3(a) we keep the ratio $r/p = 1/100$ constant such that $p_E = 1/101 \Rightarrow c_v(1) = 10$. The curves are characterised by a horizontal part, which holds the variance on the initial $c_v(1)$ value followed by a declining part. The length of the part at constant level depends on $p + r$, i.e. on the intensity of transitions between the states, which is different but fixed for each curve in Figure 3(a). The sojourn times of the good and bad state are geometrically distributed with mean $(1 - p)/p$ and $(1 - r)/r$, respectively. For $\lim_{p+r \to 0}$ the mean holding times of the states are extended on longer times scales.
3. Then the correlation in the modelling process persists over about the same time scale and the transition point from the constant to the declining part of the $c_v(N)$ curve is shifted in the range between $1/p$ and $1/r$.
   The decreasing part soon approaches the same slope as is valid for a memoryless process with independent random losses at a given rate, such that $c_v(kN)/c_v(N) \to \sqrt{k}$.

Figure 3(b) shows results, where $p + r$ is again stepwise reduced by a factor 10 as in 3(a), but this time with $h = 0.99$, which means a low loss probability of 1% in the bad state and by setting $r/p = 1/10$ the total loss probability is kept at $p_E = 1/1100 \Rightarrow c_v(1) = \sqrt{1099}$. Again the coefficient of variation stays at a constant level over multiple time scales for small $p + r$, but essentially below the initial $c_v(1)$ value.

5 Evaluation

The evaluation of the trained 2-state Markov models using the coefficient of variation $c_v(N) = \sigma_N/\mu_N$ is shown for two backbone traces with different packet loss rates in Figure 4 and 5. The Poisson process provides a linear lower bound for $c_v(N)$ without any autocorrelation. The parameters of the simple Gilbert ($h = 0$, $k = 1$) and the Gilbert model have been estimated from the given traces using the traditional methods proposed by Yajnik et al. [8] and Gilbert [3] as introduced in Section 3.

Moreover, the simplified Gilbert, the Gilbert and the Gilbert-Elliott model have been trained based on the second order statistics over multiple timescales $N \in [1, 10^5]$, as shown in Figure 4 and 5. 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. Initial trial values for the parameters were estimated from the study of the impact of different model parameters discussed in Section 4.5.
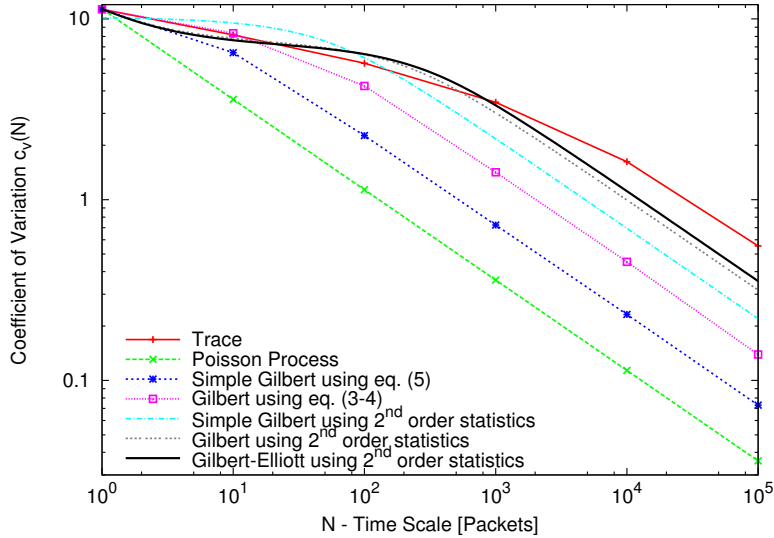
Fig. 4. Evaluation of the trained 2-state Markov models using the coefficient of varia-tion $c_v = \sigma/\mu$ for a backbone trace with a mean packet loss rate of 0.7%.

Trace	Trained Model	p	r	h	k	p_E
0.7 %	Simple Gilbert	0.000401648	0.0414789	0	1	0.95%
	Gilbert	0.000196854	0.0109547	0.563513	1	0.77%
	Elliott	0.000132253	0.00811837	0.559691	0.999372	0.71%
0.1 %	Simple Gilbert	$2.9365\cdot10^{-5}$	0.0229754	0	1	0.13%
	Gilbert	$1.3343\cdot10^{-5}$	0.00601795	0.555044	1	0.098%
	Elliott	$1.33308\cdot10^{-5}$	0.00601795	0.554949	0.999999	0.098%

Table 3. Estimated model parameters for both traces using second order statistics. The mean packet loss rate of the first trace is 0.7% and 0.1% in case of the second.

The distance between different model curves as shown in Figure 4 and 5 and the trace curve is measured by the the Mean Square Error (MSE)

$$\text{MSE(model)} = 10^{-5} \sum_{N=1}^{10^5} (c_v^{\text{model}}(N) - c_v^{\text{model}}(N))^2, \tag{26}$$

where a smaller MSE indicates a better fit. Table 4 compares the considered models.

The model parameters resulting from the adaption to the coefficient of varia-tion found in the trace are given in Table 3. This trend confirms the assignment of $h = 0.5$ by Gilbert [3]. The packet loss rate p_E of the simple Gilbert model with $h = 0$ essentially deviates from the trace.
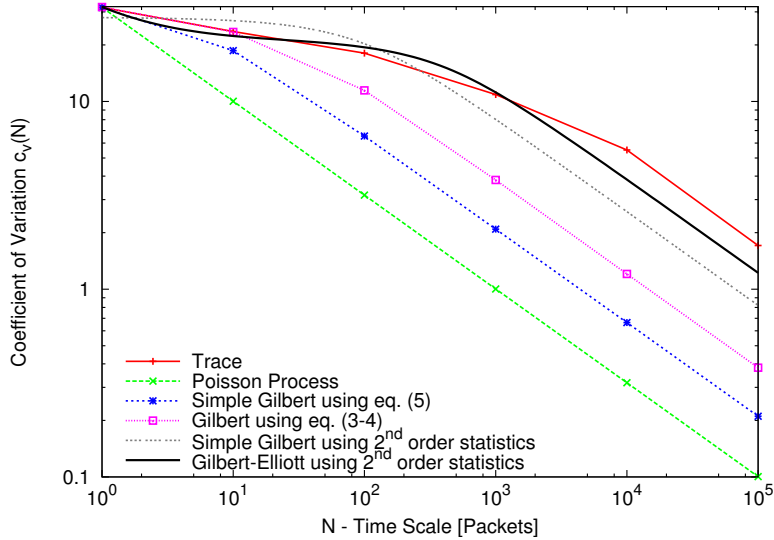
Fig. 5. Evaluation of the trained 2-state Markov models using the coefficient of varia-
tion $c_v = \sigma/\mu$ for a backbone trace with a mean packet loss rate of 0.1%.

Trace / Model	Simple Gilbert (eq. 5)	Gilbert (eq. 3-4)	Simple Gilbert	Gilbert	Elliott
0.7%	4.02	1.3	0.98	0.2	0.18
0.1%	43.42	19.04	7.96	1.09	1.09

Table 4. Mean Square Error (MSE) distance between different trained models and the
two traces. The Markov models were trained using classical techniques (eq. 5 and 3-4)
and the second order statistics as described in Section 4.

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, since they are closer related to
error burst lengths whereas the second order statistics can include long range
correlation. The extraction of the most relevant information in measurement
traces to be used for the fitting of model parameters with regard to the Quality
of Experience aspects (QoE) is still for further study. The relevance of bursts
surely increases with the observed mean failure burst length in a considered
traffic flow.

6 Related Work

Cuperman [21] derives the generating function $H_m(z)$ for m errors in a binary
series of arbitrary length, where $P(m, n)$ denotes the probability of m errors in

a series of length n.

$$H_m(z) = \sum_{n=m}^{\infty} P(m,n)z^n = \frac{p}{p+r}z\left(\frac{1-g(z)}{1-z}\right)^2 [g(z)]^{m-1} \qquad (27)$$

$$g(z) = \sum_{k=1}^{\infty} P(0^{k-1}1/1)z^k; \quad 1 \le m \le n$$

However, the result is not directly transferable to obtain the second order statistics, since it is summing up over an infinite range of the length n rather than focusing on a fixed length n.

Girod et al. [7] found a simple Gilbert model ($k = 1$, $h = 0$) useful to describe the characteristics of packet losses in Internet connections and to derive an error model for Internet video transmissions on top, as lost packets will affect the perceived quality of the video transmission. Huitika et al. [25] extended the simple Gilbert model by adapting it to the datagram loss process in the scope of real-time video transmissions, by adding a third state to describe out-of-order packets. Zhang et al. [9] use a simple Gilbert model to describe a cell discard model for MPEG video transmissions in ATM networks, where the cell losses are caused by excessive load at ATM multiplexers.

McDougall et al. [26] proposed a 4-state Markov model with a hypergeometrical distribution of the sojourn time in the good and bad state as approximation of an IEEE 802.11 channel. Poikonen et al. [14] [15] compared finite state Markov models, such as the McDougall model, in order to simulate the packet error behaviour of a DVB-H system. The McDougall model and the Markov-based Trace Analysis (MTA) [27] outperformed the Gilbert model, as the latter was unable to reproduce the variance in burst error lengths. Yajnik et al. [8] point out that the simple Gilbert model is suitable if the error gap length of the traces is geometrically distributed, but can be outperformed by considering high-order Markov chains.

Tang et al. [12] used a simple Gilbert model to create a multicast loss model in IEEE 802.11 channels. Hartwell et al. [13] compared five finite-state Markov models to create a frame loss model for IEEE 802.11 indoor networks and found out that high order models trained by the Baum-Welch algorithm outperformed the Gilbert model. McDougall et al. [11] were able to reproduce the packet error rate and the average burst error length of an IEEE 802.11 channel using the simple Gilbert model, but failed to replicate the variance in error burst lengths and therefore suggested to use Gamma based state durations, as in [28]. McDougall et al. [11] also suggest that the restriction of geometrically distributed state lengths due to the Gilbert-Elliott model can be overcome and, for example, the Gamma distribution can be used.

7 Conclusion

This work provides a method to adapt the parameter set of a 2-state Markovian error pattern generator to match the second order statistics over multiple time

scales. The generating functions approach provides recursive relationships for the distribution of the number of lost packets, which finally leads to an explicit and clearly structured solution for the second order statistics. Special cases of the model as well as the impact of its parameters are discussed. Naturally, fitting procedures based on second order statistics yield a closer match in multiple time scales than classical adaptation schemes, which on the other hand are better in modelling error bursts.

Therefore it depends on the purpose of the model and it partly remains for further study, which statistical indicators should be involved in the fitting procedure. However, 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. Several Markov approaches have been proposed providing more states and parameters, which improve the accuracy of the fit to the observed process characteristics on account of more complex adaptation schemes.

Acknowledgements

References

1. Floyd, S., Jacobson, V.: Random early detection gateways for congestion avoidance. IEEE/ACM Trans. Netw. **1** (1993) 397–413
2. Haßlinger, G., Mende, J., Geib, R., Beckhaus, T., Hartleb, F.: Measurement and characteristics of aggregated traffic in broadband access networks. In Mason, L., Drwiega, T., J.Yan, eds.: 20th International Teletraffic Congress (ITC20), Ottawa, Canada, Springer LNCS 4516 (2007) 998–1010
3. Gilbert, E.N.: Capacity of a Burst-Noise Channel. Bell System Technical Journal **39** (1960) 1253–1265
4. Elliott, E.O.: Estimates of Error Rates for Codes on Burst-Noise Channels. Bell System Technical Journal **42** (1963) 1977–1997
5. Tsybakov, B., Georganas, N.D.: On self-similar traffic in atm queues: definitions, overflow probability bound, and cell delay distribution. IEEE/ACM Trans. Netw. **5** (1997) 397–409
6. Leland, W.E., Taqqu, M.S., Willinger, W., Wilson, D.V.: On the self-similar nature of ethernet traffic. ACM SIGCOMM Computer Communication Review **23** (1993) 183–193
7. Girod, B., Stuhlmüller, K., Link, M., Horn, U.: Packet Loss Resilient Internet Video Streaming. In: Proceedings of SPIE Visual Communications and Image Processing. (1999) 833–844
8. Yajnik, M., Moon, S.B., Kurose, J.F., Towsley, D.F.: Measurement and modeling of the temporal dependence in packet loss. In: IEEE INFOCOM. (1999) 345–352

9. Zhang, L., Chow, D., Ng, C.H.: Cell loss effect on QoS for MPEG video transmission in ATM networks. In: IEEE International Communication Conference (ICC). (1999) 147–151

10. Lemmon, J.J.: Wireless link statistical bit error model. US National Telecommunications and Information Administration (NTIA) Report 02-394 (2002)

11. McDougall, J., Miller, S.: Sensitivity of wireless network simulations to a two-state Markov model channel approximation. In: IEEE Globecom. (2003)

12. Tang, C., McKinley, P.K.: Modeling multicast packet losses in wireless LANs. In: 6th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems (MSWIM), New York, NY, USA, ACM Press (2003) 130–133

13. Hartwell, J.A., Fapojuwo, A.O.: Modeling and characterization of frame loss process in IEEE 802.11 wireless local area networks. In: IEEE Vehicular Technology Conference. (2004) 4481–4485

14. Poikonen, J., Paavola, J.: Comparison of Finite-State Models for Simulating the DVB-H Link Layer Performance. In: 2nd International Symposium on Wireless Communication Systems (IEEE ISWCS). (2005)

15. Poikonen, J., Paavola, J.: Error Models for the Transport Stream Packet Channel in the DVB-H Link Layer. In: IEEE International Conference on Communications (ICC). (2006) 1861–1866

16. Osmann, C.: Bewertung von Codierverfahren für einen störungssicheren Datentransfer. PhD thesis, Universität Duisburg-Essen (1999)

17. Morgera, S.D., Simard, F.: Parameter estimation for a burst-noise channel. In: International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Washington, DC, USA (1991) 1701–1704

18. Turin, W.: Performance Analysis and Modeling of Digital Transmission Systems (Information Technology: Transmission, Processing and Storage). Springer New York, Inc., Secaucus, NJ, USA (2004)

19. Sivaprakasam, S., Shanmugan, K.: An Equivalent Markov Model for Burst Errors in Digital Channels. IEEE Transactions on Communications **43** (1995) 1347–1355

20. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. (1990) 267–296

21. Cuperman, V.: An Upper Bound for the Error Probability on the Gilbert Channel. IEEE Transactions on Communication Technology **17** (1969) 532–535

22. Feldmann, A., Whitt, W.: Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. Performance Evaluation **31** (1998) 245–279

23. Wilf, H.S.: Generatingfunctionology. Academic Press (1992)

24. Grinstead, C.M., Snell, L.J.: Introduction to Probability. American Mathematical Society (1997)

25. Huitika, T., Driessen, P.: Datagram loss model for noninteractive real time streaming video. In: IEEE Pacific Rim Conference on Communications. Computers and Signal Processing (PACRIM). (2003) 756–759

26. McDougall, J., John, J.J., Yu, Y., Miller, S.L.: An improved channel model for mobile and ad-hoc network simulations. In Hamza, M.H., ed.: Communications, Internet, and Information Technology, IASTED/ACTA Press (2004) 352–357

27. Konrad, A., Zhao, B.Y., Joseph, A.D., Ludwig, R.: A Markov-based Channel Model Algorithm for Wireless Networks. Wireless Networks **9** (2003) 189–199

28. Yoma, N., Busso, C., Soto, I.: Packet-loss modelling in IP networks with state-duration constraints. In: IEE Proceedings Communications. Volume 152. (2005)