A self-extracting accurate modeling for bounded-delay video services

Marcelo Dias de Amorim[a,b,*], Otto Carlos M.B. Duarte[b]

[a]*LIP6 Laboratory Université Pierre et Marie Curie 8, rue du Capitaine Scott, 75015 Paris, France*
[b]*GTA COPPE/Poli, Universidade Federal do Rio de Janeiro, P.O. Box 68504 21945-970 Rio de Janeiro, RJ, Brazil*

Abstract

This article proposes the XGOP-B deterministic traffic model for MPEG video services requiring strict bounds on the quality of service provided by the network. The XGOP-B model takes advantage of two characteristics of MPEG video traffic by performing a two-level analysis of such sources. In our scheme, the traffic parameters are directly obtained by a single-input single-output parameter extracting algorithm. The proposed algorithm is autonomous and leads to a single set of parameters for each video stream avoiding user interaction determine the most appropriate set of parameters. We show through analysis of several video traces that these original features of the XGOP-B model yield accurate parameterizations at a low complexity cost. Our results show that the accuracy of the XGOP-B model leads to significant improvements when compared with the accuracy of the other traffic models. Furthermore, the complexity is evaluated through the number of traffic parameters and we show that it can be substantially decreased.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Bounded-delay services; MPEG video; Deterministic characterization; Modeling accuracy

1. Introduction

The development of efficient communication systems to transport variable bit rate (VBR) video has been a major challenge in recent years. An efficient use of the bandwidth and the provision of the required quality of service (QoS) are two important keys for the success of oncoming applications. Mechanisms to access and manage network resources must be provided to avoid bandwidth misuse and to improve the quality of the application at the receivers [1–4].

Most video applications require strict QoS guarantees from the network. Delay bounds, throughput, jitter, and loss rate are important parameters that must be controlled by the communication infrastructure to provide such applications with the required QoS degree [1,5–11]. Three components are used by the network service in order to provide QoS guarantees: traffic model, packet scheduler, and connection admission control (CAC) [12–15]. Nevertheless, neither the packet scheduler nor the CAC can provide by themselves an efficient service if an accurate traffic model is not supplied [7,16–19].

A number of traffic models for statistic services have been proposed in the literature [20–24]. These models are mostly based on Markov chains, self-similarity, and other stochastic techniques. Such models, however, either are extremely complex for practical use or they do not capture traffic characteristics accurately. Moreover, they cannot be used for services that require deterministic guarantees from the network (e.g. strict video applications) since there exists a nonzero probability that the QoS parameters are violated.

In a deterministic service, all the packets are guaranteed to be delivered within the specified QoS bounds [7,13–19, 25,26]. The traffic characteristics are parameterized by a deterministic traffic model. A traffic model is deterministic when the corresponding traffic behavior function $M(t)$, which denotes the number of bits generated by the modeled source, is always equal to or greater than the actual number of bits generated by the source. A deterministic traffic model often makes use of the concept of traffic constraint function [7,18,27]. Let $\mathscr{G}[t, t + \tau]$ be the number of bits transmitted by the source in the interval $[t, t + \tau]$. A function $\mathscr{F}(\cdot)$ is a traffic constraint function if

$$\mathscr{G}[t, t + \tau] \le \mathscr{F}(\tau), \quad t \ge 0, \tag{1}$$

* Corresponding author. Tel.: +33-1-44278772; fax: +33-1-44277495.
E-mail addresses: amorim@rp.lip6.fr (M.D. de Amorim); otto@gta.ufrj.br (O.C.M.B. Duarte).

i.e. if during any interval of length τ the source does not transmit more than $\mathscr{F}(\tau)$ bits.

The only requirement that must be respected is that $(\cdot)$ must be an upper bound on the traffic behavior. This implies that an infinite number of traffic constraint functions exists. The most accurate traffic constraint function that can be obtained is called empirical envelope [7,18], defined as

$$\zeta(\tau) = \sup \mathscr{G}[t, t + \tau], \quad t \geq 0. \tag{2}$$

In practice, representing the behavior of VBR sources by the empirical envelope may be extremely complex, due to the large number of parameters that would be necessary. The solution is to approximate $\zeta(\cdot)$ by an upper bound $\mathscr{F}(\tau) \geq \zeta(\tau), \forall \tau$, which is mathematically easier to represent than $\zeta(\cdot)$.

An efficient deterministic characterization aims at finding a function $\mathscr{F}(\cdot)$ as close as possible to $\zeta(\cdot)$. An accurate traffic model keeps as low as possible the inter-model distance Δ_M, given by

$$\Delta_M(\tau) = \mathscr{F}(\tau) - \zeta(\tau), \tag{3}$$

for all τ.

In the (σ, ρ) model [27,28] for an interval of length τ, the source is constrained to transmitting no more than $\sigma + \rho\tau$ bits. The $(\vec{\sigma}, \vec{\rho})$ model [7], which is an extension of the (σ, ρ) model, defines a two-dimensional array of (σ_k, ρ_k) pairs resulting in a piecewise concave function. The discrete model [29] uses four parameters $(x_{min}, x_{ave}, I, s_{max})$ to model a traffic, where x_{min} is the minimum interval between two consecutive packet arrivals, x_{ave} is the maximum average interval between two packet arrivals within an interval of length I, and s_{max} is the maximum packet size. The D-BIND (Deterministic Bounding INterval Dependent) model [30] characterizes a source by a $R(I)$ function, which gives the maximum transmission rate within any interval of length I. The $R(I)$ function is parameterized by the pairs (R_k, I_k), $k = 1, 2, ..., N$, and results in a piecewise linear traffic constraint curve. Table 1 summarizes the traffic constraint functions for each one of the models.

The abovementioned models aim at characterizing general VBR video sources. When parameterizing MPEG video [31], which is a de facto standard for networked multimedia applications, they capture neither the intrinsic property of burstiness caused by the intra-picture frames nor the long-term deviations in the average traffic rate due to

changes in scenes. Moreover, for some models, the procedures to extract parameters from the sources cannot be efficiently performed if the user does not interfere to decide which parameters seem to be the most adequate.

This paper proposes the XGOP-B (eXtended GOP-Based) model to deterministically characterize MPEG video sources. The XGOP-B model is based on Group of Pictures (GOP), which is the second-level hierarchy of MPEG video streams. The model definition relies on a two-level analysis of MPEG traffics, which results in a detailed description of such sources in the frame and GOP levels. The proposed model uses the results of this two-level analysis to obtain a simple and accurate characterization. Furthermore, in order to obtain the XGOP-B parameters, we develop the single-input single-output (SISO) parameter extracting algorithm that computes one and only one set of traffic parameters for each video stream. This implies that the final parameters are not result of erroneous choices incurred by interactions between the algorithm and the user. With the proposed parameter extracting algorithm, the number of parameters, and, consequently, the complexity of the model, can be calibrated by tuning a special variable. This variable allows a fine control on the tradeoffs between the accuracy and the complexity of the model, so that badly behaved sources (which require a large number of parameters to achieve high accurate modeling) can be modeled at an acceptable complexity level.

We use several traces of MPEG video streams to demonstrate the improvements achieved by the XGOP-B model. First, by taking into account specific characteristics of MPEG video sources, the accuracy of the XGOP-B model is shown to be up to 70% greater than the ones of the other models. Second, these values are achieved at no extra complexity in terms of the number of parameters. This is mainly attributed to the autonomous characteristic of the parameter extracting algorithm.

This paper is organized as follows. In Section 2, we perform a detailed analysis of MPEG video streams and establish the specific properties intrinsic to such sources. We describe then the fundamental requirements to achieve a simple and accurate deterministic traffic modeling and propose the XGOP-B model, showing how it fulfils the requirements. Section 3 presents the single-input single-output parameter extracting algorithm for the XGOP-B model. In Section 4, we analyze several MPEG video traces in terms of the level of accuracy and present numerical results concerning the tradeoffs between the adjustment of the model and the respective accuracy. Finally, Section 5 concludes the paper.

2. The XGOP-B model

We present in this section the XGOP-B model that captures specific characteristics of MPEG video sources and uses these characteristics to achieve high modeling

Table 1
Traffic constraint function for the models

Model	Traffic constraint function
(σ, ρ)	$\mathscr{F}(\tau) = \sigma + \rho\tau$
$(\vec{\sigma}, \vec{\rho})$	$\mathscr{F}(\tau) = \min_i\{\sigma_i + \rho_i\tau\}$
Discrete	$\mathscr{F}(\tau) = \left\lfloor \dfrac{\tau}{I} \right\rfloor \dfrac{I s_{max}}{x_{ave}} + \min\left\{ \left\lceil \left(\dfrac{\tau}{I} - \left\lfloor \dfrac{\tau}{I} \right\rfloor \right) \dfrac{I}{x_{min}} \right\rceil, \dfrac{I}{x_{ave}} \right\} s_{max}$
D-BIND	$\mathscr{F}(\tau) = \dfrac{R_k I_k - R_{k-1} I_{k-1}}{I_k - I_{k-1}}(\tau - I_k) + R_k I_k, \quad I_{k-1} \leq \tau \leq I_k$

accuracy. It is well suited to MPEG video distribution, which may handle hundreds of highly variable streams, because it is a simple specific traffic model that can be used to maximize the bandwidth utilization and to optimize the admission control.

The XGOP-B model is based on a two-level analysis of MPEG compressed video streams. In the *frame-level* analysis, we investigate the rate differences between intra-picture frames and inter-picture frames. We show that due to the larger size of intra-picture frames, the traffic constraint function assumes a staircase behavior, which is characterized by a simple (η, ν) pair of parameters. In the *GOP-level* analysis, we find that the long-term behavior of the traffic constraint function is perturbed by deflection points that affect the average increase rate of the curve. The (η, ν) parameters are then modified to adapt to the different homogeneous regions outlined by these deflection points.

2.1. Frame-level observation

An MPEG coder [31] produces three frame types: intra-picture I frames and inter-picture P and B frames. These frames are generated in a cyclic sequence called GOP, which contains an initial I frame and some P and B frames, in conformity with a specified pattern, for instance IBBPBBPBB. Most MPEG video streams have I frames much larger than P or B frames, which results in traffic bursts at the beginning of each GOP. In the corresponding traffic constraint function, this is equivalent to periodical stairs in the curve, as shows the example of a real MPEG traffic constraint function of Fig. 1.

2.2. GOP-level observation

The staircase behavior is valid for the short-term but not for the long-term analysis of the traffics. Our next
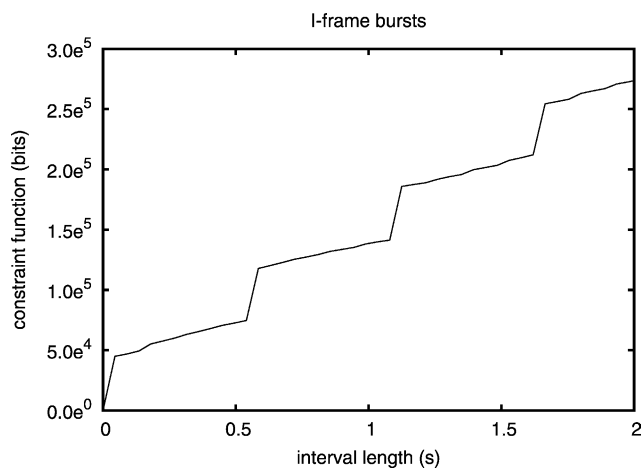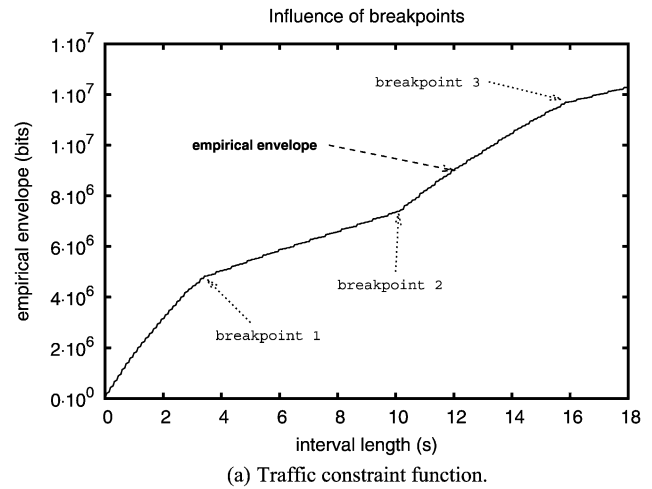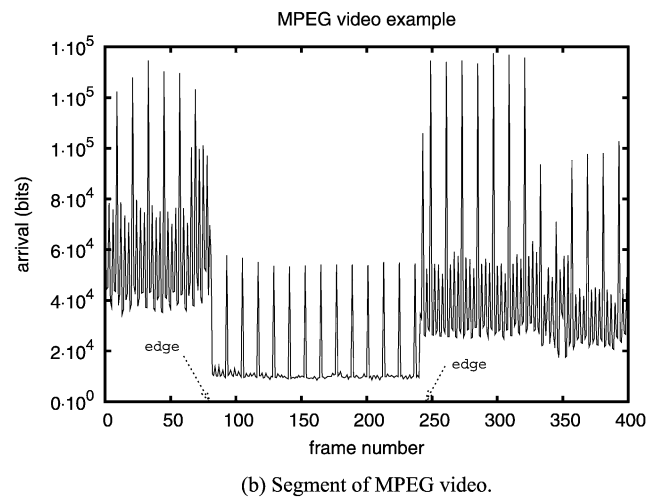

Fig. 1. Bursts due to intra-picture frames.

(a) Traffic constraint function.

(b) Segment of MPEG video.

Fig. 2. Influence of breakpoints on the average increasing rate of the traffic constraint function.

investigation relies on the GOP-level observation of the video characteristics. As it will be seen, this analysis is the foundation for the modeling of the source's long-term behavior. To illustrate the methodology, Fig. 2(a) shows a 18-second sequence of the traffic constraint function corresponding to the MPEG sequence shown in Fig. 2(b). Note that the empirical envelope presents deflection points in the average increasing rate of the curve. We refer to these points as *break points* and let k of them be denoted by $\xi_1, \xi_2, ..., \xi_k$.

Breakpoints are caused by changes in scenes, when the average rate of the video traffic varies beyond a given threshold. Fig. 2(b) shows a segment of an MPEG video stream that exhibits this characteristic. Edges in the arrival process correspond to breakpoints in the traffic constraint function.

A breakpoint can be of two types. We define a positive breakpoint the one that increases the average increase rate of the traffic constraint function. Similarly, a negative breakpoint corresponds to a point where the average increase rate

of the traffic constraint function decreases. In the example of Fig. 2(a), breakpoints 1 and 3 are negative, whereas breakpoint 2 is positive.

Breakpoints have a negative impact on the accuracy of general traffic models. The (σ, ρ) and discrete models have traffic constraint functions with constant average increasing rate. At negative breakpoints, the traffic constraint functions for these models diverge from the empirical envelope. The implication is that the inter-model distance increases, which consequently leads to a degradation of the model's accuracy. For the $(\vec{\sigma}, \vec{\rho})$ model, if a negative breakpoint is followed by a positive breakpoint (which results in a convex segment), the model cannot track the empirical envelope since the $(\vec{\sigma}, \vec{\rho})$ curve is strictly concave. These issues are addressed by the XGOP-B model through its average rate adaptive characteristic. The concave and convex aspects are adjusted according to the type and intensity of the breakpoint.

2.3. Formalization of the XGOP-B model

Based on our two-level analysis of MPEG video sources, we can now conclude that an accurate traffic model for such traffics should be composed of two components:

- *Short-term component*: this component characterizes the traffic at the frame level. It must take into account the differences between I-frames and P- and B-frames.
- *Long-term component*: this component characterizes the traffic at the GOP level. The resulting model must behave such as to reflect the presence of breakpoints.

These two components are the basis of the XGOP-B model. We detail them in the following.

2.3.1. Modeling the short-term component

To model the short-term component of the video traffics, it suffices to find a function that presents the staircase behavior observed in Section 2.1. To obtain this characteristic, the XGOP-B model defines the pair (η, ν) as traffic parameters. The η factor, called *impulse factor*, models the bursts caused by intra-frames. This factor is added to the traffic constraint function for interval lengths proportional to the GOP size[1]. The *average factor* ν parameterizes the average increasing of the traffic constraint curve. These parameters are combined in order to generate a traffic constraint function with the staircase characteristic observed in the empirical envelope.

Fig. 3 shows the approximation obtained through the (η, ν) parameters. Fig. 3(a) shows the actual traffic

---

[1] The GOP size is the number of frames in a GOP.
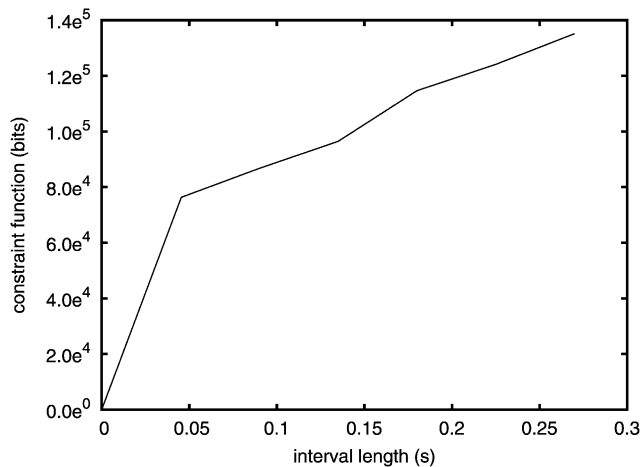
(a) Actual GOP.
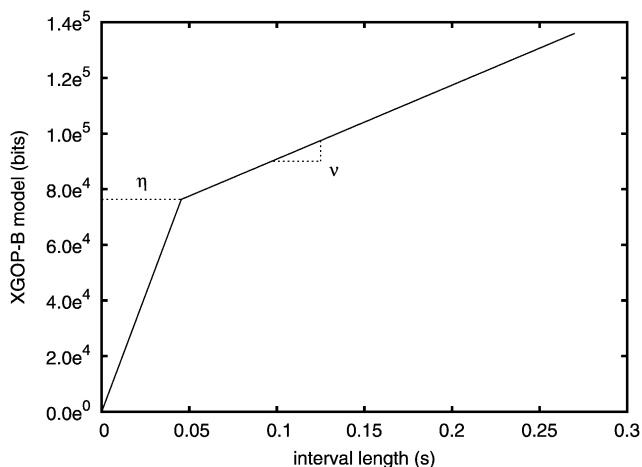
(b) Modeled GOP.

Fig. 3. Similarities between the actual behavior of an MPEG group of pictures and the corresponding XGOP-B model.

constraint function (one GOP is represented). The resulting modeled curve is depicted in Fig. 3(b).

We can now define the exact curve using the (η, ν) pair, as

$$\mathscr{X}_{\text{short}}(\tau) = \left\lceil \frac{\tau + T}{TD} \right\rceil \eta + \nu\tau, \tag{4}$$

where T is the inverse of the video's sample rate and D is the GOP size.

2.3.2. Modeling the long-term component

The use of a single pair (η, ν) results in a good approximation of the traffic constraint function if and only if the average increasing rate of this latter is constant. Nevertheless, we have observed that in practice this is not true, because of the presence of breakpoints. Thus, in order to capture the breakpoint property, we must change average inclination of the modeling curve at each breakpoint. A pair (η, ν) must then be defined for each area delimited by the breakpoints. The empirical envelope is then subdivided in

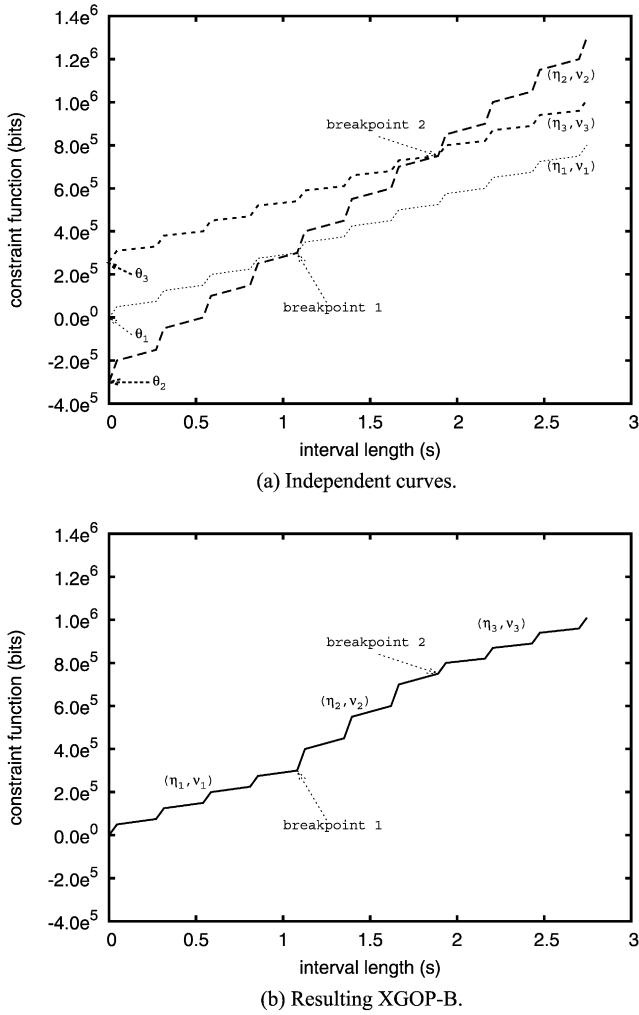(a) Independent curves.

(b) Resulting XGOP-B.

Fig. 4. Utility of the offsets for the configuration of the final XGOP-B curve.

'zones' bounded by the breakpoints, where each zone k is considered separately and is assigned to a specific pair (η_k, ν_k).

Each pair must be associated with an offset θ_k (see Fig. 4), where θ_1, i.e. the offset of the first zone, is zero.

The equation that represents the long-term component is given by

$$\mathcal{X}_{\text{long}}^k(\tau) = \theta_k + \mathcal{X}_{\text{short}}^k(\tau). \tag{5}$$

2.4. The XGOP-B model

We can now define the global equation of the XGOP-B model by combining Eqs. (4) and (5). We must determine, for each zone bounded by the breakpoints, an offset θ and a pair (η, ν).

Thus, the set of parameters that completely define the XGOP-B model is the number of zones, Z, and the 4-tuple $(\eta_i, \nu_i, \theta_i, \xi_i)$ for zone i bounded by breakpoints ξ_{i-1} and ξ_i. We also have $\xi_0 = 0$.

The final equation for the XGOP-B model is given by equation:

$$\mathcal{X}(\tau) = \theta_i + \left\lceil \frac{\tau + T}{TD} \right\rceil \eta_i + \nu_i \tau, \quad \xi_{i-1} < \tau < \xi_i, \quad 1 \leq i \leq Z. \tag{6}$$

One of the most significant advantage of the XGOP-B model relies on its injective parameter extracting algorithm. For each video traffic, the algorithm obtains one and only one set of parameters. To achieve an efficient and accurate characterization, the computation of the traffic parameters must not be based on intuitive considerations. Consider for example the (σ, ρ) model and the empirical envelope shown in Fig. 5. There are no conceptual constraints on what values should be assigned to σ and ρ. In the example, we show three possibilities for the (σ, ρ) curve, where each one is more suitable to a particular range of interval lengths. If case 1 is chosen, the traffic constraint function is close to the empirical envelope for interval lengths below 3 s but does not capture the long-term behavior of the source. Case 3, in turn, is suited for larger interval lengths but captures poorly the traffic behavior for short intervals. An intermediate choice is represented by case 2. None of them, however, can serve as an accurate approximation for all interval lengths.

The above example shows that different parameters can be obtained for the same traffic source. A fine traffic model should be associated with an autonomous extracting algorithm that computes the traffic parameters without incurring a decision stage as in the example of the (σ, ρ) model. To be independent of a subjective choice of such parameters, the traffic model must then be defined to lead to a single set of parameters for each input traffic.

For the XGOP-B model, users need not to specify their traffic parameters since the parameter extracting algorithm determines itself the most adequate values. Furthermore, the structure of the XGOP-B model results in a single characterization for each MPEG traffic.
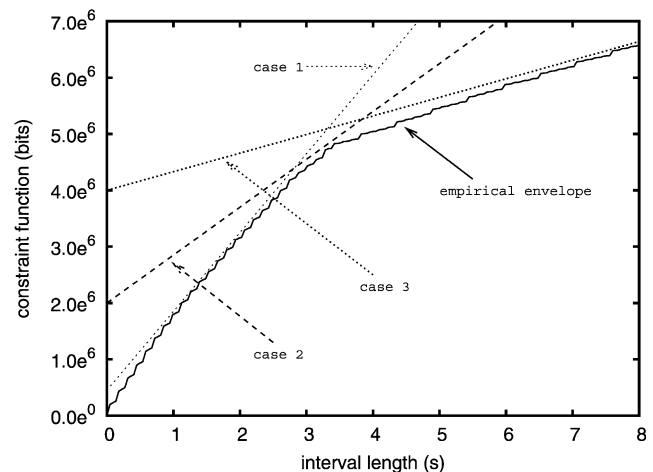

Fig. 5. Possible solutions for the σ and ρ parameters.

3. Single-input single-output parameter extracting algorithm

The single-input single output (SISO) parameter extracting algorithm comprises two main steps: the reckoning of breakpoints and the computation of the effective parameters. The first step is to develop a mathematical procedure to identify the breakpoints. Let $\zeta(\cdot)$, called the *first derivative function*, (Fig. 6) be the function that represents the number of bits generated by the source during one GOP. An example is shown in Fig. 6(a). Based on this function, the algorithm computes a discrete instance of the non-zero variations of the first derivative function, which results in the *second derivative function* $\zeta^{**}(\cdot)$. Fig. 6(b) shows the second derivative functions relative to Fig. 6(a). The impulses given by the second derivative function will help us to identify the breakpoints.

The amplitudes of the impulses plotted in Fig. 6(b) point out the rate variation of two consecutive GOPs in the empirical envelope. It must be distinguished among these

impulses which of them should be identified as breakpoints. Note that there is no exact quantitative definition of what is a breakpoint. The challenge is to define an impulse amplitude (threshold) to correspond to a breakpoint. This identification is performed according to the amount of accuracy specified to the parameter extracting algorithm. Reducing the threshold implies more breakpoint identifications. The model is then more accurate since a greater number of parameters is specified. On the other hand, if fewer points are selected, the model is simpler but less accurate. For this purpose the parameter extracting algorithm defines the β variable, called quality of the model. The algorithm is performed in conformity with the following propositions:

Proposition 1. *Given an inclination, a breakpoint ξ is identified if this inclination varies, negatively or positively, of a factor β, with $\beta > 0$:*

$$\zeta^{**}(\tau) \geq \beta.$$

In this case, there is a breakpoint at $t = \tau$.

Proposition 2. *A breakpoint ξ is also identified if the sum of an uninterrupted sequence of second derivatives varies, negatively or positively, with regard to the first inclination of this sequence, of a factor β, with $\beta > 0$:*

$$\left| \int_{\tau}^{\tau+\Delta\tau} \zeta^{**}(t)\delta t \right| \geq \beta, \quad \tau, \Delta\tau > 0.$$

In this case, a breakpoint is identified at $t = \tau + \Delta\tau$.

Note that a breakpoint is also identified when sum of a sequence of small impulses is equal to or greater than the threshold. Thus, the parameter extracting algorithm also considers that the traffic constraint function has changed its average increasing rate beyond the defined threshold.

The SISO parameter extracting algorithm computes η and ν from the input video traffic as follows. For the sake of simplicity, we present the algorithm for the first zone, but it can be easily extended to the other zones by simple translation of the $x-y$ axis.

The basic input to the algorithm is naturally the empirical envelop $\zeta(\cdot)$. First of all, we obtain its first derivative through a simple function given by

$$\zeta^*(t) = \frac{\zeta(nDT)}{DT}, \tag{7}$$

where $n = \lceil t/DT \rceil$.

The computation of the second derivative function is straightforward and given by

$$\zeta^{**}(t) = \zeta^*(t) - \zeta^*(t - T), \tag{8}$$

where, by definition, $\zeta^*(t) = 0$ for $t \leq 0$.

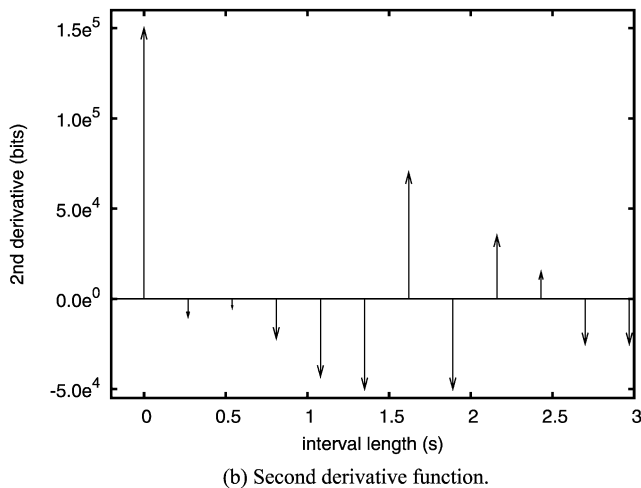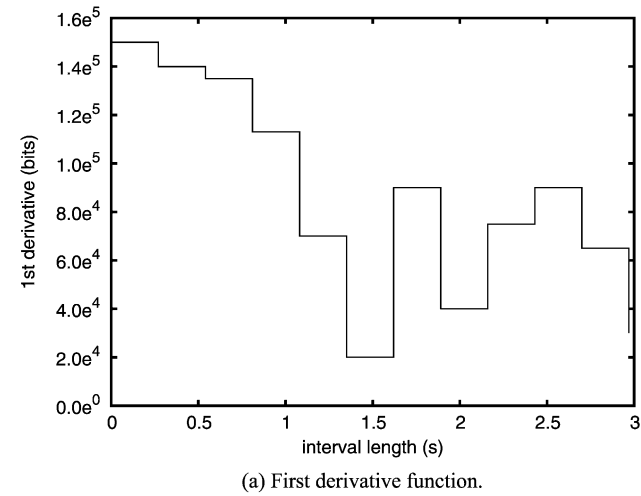(a) First derivative function.

(b) Second derivative function.

Fig. 6. Examples of derivative functions.

The η and ν values must be computed for each zone. Let z be the zone bounded by breakpoints ξ_{z-1} and ξ_z, $1 \leq z \leq Z$ (Z is the number of zones).

We first derive the impulse factor η_z. Since it corresponds to the largest impulse within the zone, we have that

$$\eta_z = \max \zeta^{**}(t), \quad \xi_{z-1} \leq \xi_z. \qquad (9)$$

For the computation of ν_z, we define the first point of $\mathscr{X}$ each zone based on the empirical envelop as

$$\mathscr{X}(\xi_{z-1}) = \zeta(\xi_{z-1}) + \eta_z. \qquad (10)$$

These points are fix and they will be used for determining ν_z and θ_z as detailed in the following.

For the computation of ν_z, the algorithm initially sets ν_z as the average slope of the first GOP in the zone, i.e.

$$\nu_z = \frac{\zeta(\xi_{z-1} + DT) - \mathscr{X}(\xi_{z-1})}{T(D-1)}. \qquad (11)$$

Note that this does not guarantee that this curve is a deterministic characterization of the source. It must be assured that the difference $\mathscr{X}(\tau) - \zeta(\tau)$ is positive for all τ. Fig. 7(b) shows an example where this requirement is not
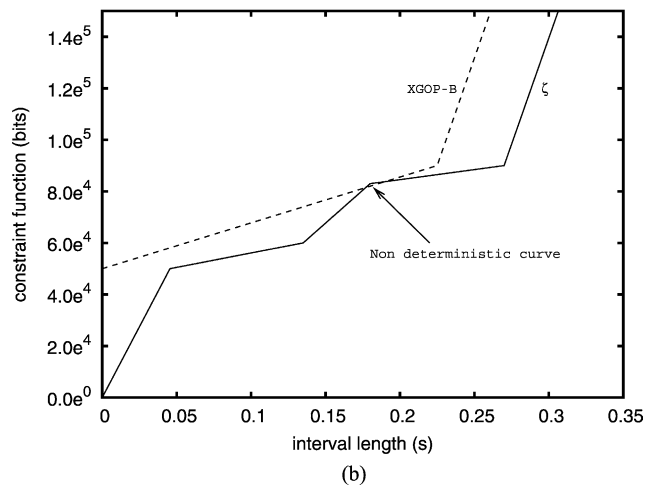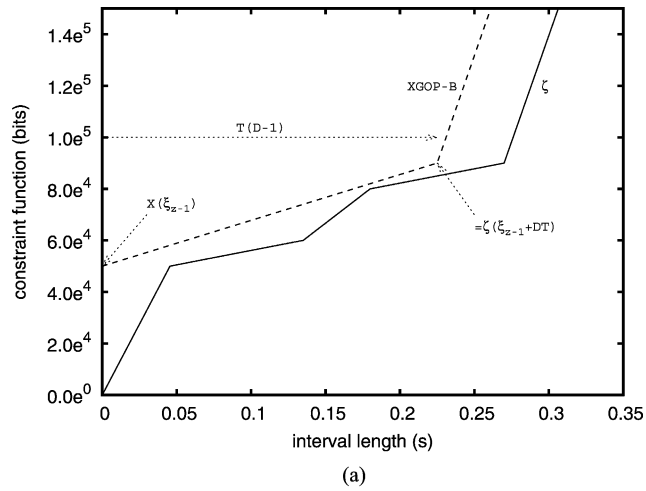
satisfied. To eliminate this possibility, the algorithm performs a loop that checks the traffic constraint function for each $\xi_{z-1} \leq t < \xi_z$. If at some time we have $\mathscr{X}(\tau) < \zeta(\tau)$ then the algorithm assures a deterministic characterization by modifying the average factor ν. For each $\xi_{z-1} \leq t \leq \xi_z$, if $[\mathscr{X}(t) - \zeta(t)] < 0$, then a new ν_z is computed through

Pseudo-code 1.

for $t = \xi_z$ down to ξ_{z-1}

    if $(\mathscr{X}(t) - \zeta(t)) < 0$ then

$$\tau = \lfloor \frac{t}{TD} \rfloor TD;$$
$$\nu_z = \frac{\zeta(t) - \mathscr{X}(\tau)}{t - \tau};$$

    end if

end for

This verification is repeated until the entire zone has been verified and no violation of the deterministic rule has been detected.

The next step is to determine the offset θ_z of zone z. The first point of the XGOP-B model for zone z is

$$\mathscr{X}(\xi_{z-1}) = \zeta(\xi_{z-1}) + \eta_z. \qquad (12)$$

For time $t = \xi_{z-1}$, we can rewrite Eq. (6) as

$$\mathscr{X}(\xi_{z-1}) = \theta_z + \left\lceil \frac{\xi_{z-1} + T}{TD} \right\rceil \eta_z + \nu_z \xi_{z-1}. \qquad (13)$$

Since we know $\mathscr{X}(\xi_{z-1})$, η_z, and ν_z, we manipulate Eq. (13) and obtain

$$\theta_z = \mathscr{X}(\xi_{z-1}) - \left\lceil \frac{\xi_{z-1} + T}{TD} \right\rceil \eta_z - \nu_z \xi_{z-1}, \qquad (14)$$

subject to $\theta_0 = 0$.

Fig. 8 shows an example of the XGOP-B traffic constraint curve. Observe the adaptive characteristic of the model at the breakpoints.
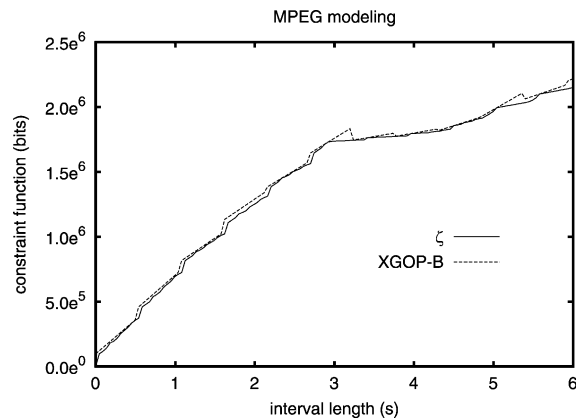

Fig. 7. (a) Deterministic and (b) non deterministic characterization.

Fig. 8. Example of the XGOP-B traffic constraint curve.

4. Numerical results

In this section, we evaluate the XGOP-B model by performing simulation analysis over several MPEG video traces. First, we compare the accuracy of the XGOP-B model with the results obtained for the other traffic models. We use the empirical envelope as a benchmark to estimate the accuracy of the models. Second, we assess the complexity of the models through the number of parameters. Finally, we investigate the tradeoffs between the quality adjustment given by the β factor and the number of breakpoints.

The analyzed MPEG video traces represent a wide range of video applications coded from films, sport events, and TV shows Ref. [32]. The samples have the following characteristics:

- GOP pattern: IBBPBBPBBPBB (GOP size = 12)
- Definition: 384 × 288 pels
- 12 bits of color information
- Number of frames: 40000
- Sample rate: 22 frames/s → $T = 45.5$ ms.

In order to compare the accuracy of the models, we define a variable called Modeling Approach Loss (MAL) that is computed as the difference between the area under the traffic constraint curve, $S_{\mathcal{F}} = \int_0^\varepsilon \mathcal{F}(\tau)d\tau$, and the area under the empirical envelope, $S_\zeta = \int_0^\varepsilon \zeta(\tau)d\tau$, where ε is equivalent to the period of 10 GOPs in our simulations. The equation for the modeling approach model is then

$$\text{MAL} = \left(\frac{S_{\mathcal{F}}}{S_\zeta} - 1 \right)100\%. \tag{15}$$

As a consequence, the smaller is MAL, more accurate is the function. In the particular case where MAL = 0%, the two curves meet each other and the maximum accuracy is obtained, since the empirical envelope is the most accurate function. The goal, however, is to have a low MAL and still maintain a reasonable complexity, since representing MPEG streams with the empirical envelope is not practical due to the large number of required parameters.

We formalized some considerations about the procedures to extract the parameters for the models proposed in the literature. This is necessary to make a coherent and

normalized basis to compute the respective traffic parameters and to make the comparisons fair. For the (σ, ρ) model, we select the *maximum* distance $(\tau_2 - \tau_1)$ and the parameters (σ, ρ) such that

$$\zeta(\tau_1) = \sigma + \rho\tau_1, \tag{16}$$

$$\zeta(\tau_2) = \sigma + \rho\tau_2, \tag{17}$$

$$\zeta(u) \le \sigma + \rho u, \ \forall u. \tag{18}$$

For the D-BIND model, we limit the interval I_k to $I_k \ge I_{\max}, \forall k$, where $I_{\max} = 32T$. We make $I_0 = 0$ and $R_0 = 0$ and compute I_k and R_k as

$$I_k = \left\{ t \Big|_{I_{k-1} \le t \le I_{k-1}+I_{\max}} \max \left[\frac{\zeta(t) - \zeta(I_{k-1})}{t - I_{k-1}} \right] \right\}, \tag{19}$$

$$R_k = \left\{ \frac{\zeta(I_k)}{I_k} \right\}. \tag{20}$$

In Ref. [7], the authors present an algorithm that uses the empirical envelope to construct the $(\vec{\sigma}, \vec{\rho})$ traffic constraint function. For a given time τ, a parameter pair (σ_i, ρ_i) and a new time value $0 \le \tau' < \tau$ are selected such that $\zeta(\tau) = \sigma_i + \rho_i\tau$, $\zeta(t) = \sigma_i + \rho_i t$ for all $0 \le t \le \tau$, and $\zeta(\tau') = \sigma_i + \rho_i\tau'$. This procedure is repeated with a newly calculated time value τ' as long as τ' is positive. For the $(x_{\min}, x_{\text{ave}}, I, s_{\max})$ model, we assume the interval I equivalent to one GOP size and $s_{\max} = 48$ bytes. The parameters x_{ave} and $x_{\min}$ are given by

$$x_{\text{ave}} = \min\left\{ \frac{Is_{\max}}{\zeta(\tau+I) - \zeta(\tau)} \right\}, \tag{21}$$

$$x_{\min} = \min\left\{ \frac{Ts_{\max}}{\zeta(\tau+T) - \zeta(\tau)} \right\}. \tag{22}$$

Table 2 shows the results for the analyzed video traces when the quality of the model for the XGOP-B model is $\beta = 0.01$. The results show the high accurate characterization obtained by the XGOP-B model and prove that the XGOP-B model actually captures the intrinsic properties of MPEG video sources, leading to modeling approach losses up to 70% lower than the ones of the other models.

Fig. 9 shows an example of the traffic constraint curves for a cartoon video trace. Note that the static property of the $(x_{\min}, x_{\text{ave}}, I, s_{\max})$ and (σ, ρ) models inhibits an accurate characterization. The $(\vec{\sigma}, \vec{\rho})$ model is a good approximation
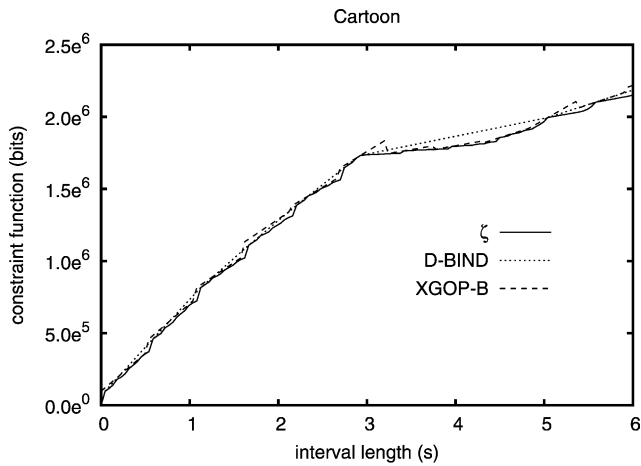
Table 2
Modeling approach loss for the models

Video	Modeling approach loss (%)				
	(σ, ρ)	$(\vec{\sigma}, \vec{\rho})$	Discrete	D-BIND	XGOP-B
James Bond	44.02	2.16	31.75	1.48	0.41
Lambs	30.56	3.42	20.31	2.23	1.63
News	7.44	2.56	33.67	2.47	1.58
Soccer	7.21	2.20	39.58	1.53	1.22
Conference	6.79	5.42	13.49	5.18	1.93
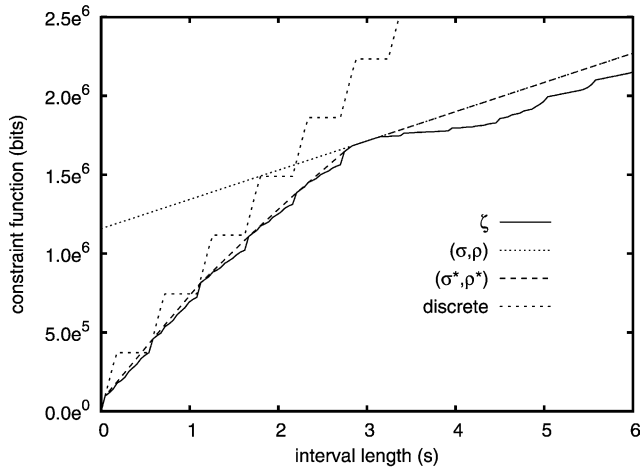
Cartoon

Fig. 9. Traffic constraint functions for the *Cartoon* video trace.
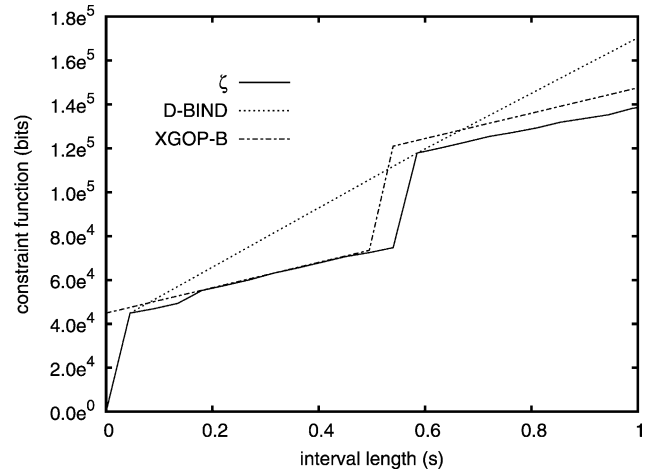

Fig. 10. Zoom in the traffic constraint functions.

of traffic parameters. In one side, the admission control algorithms and the policing mechanism are more complex since they have to manipulate more traffic parameters. On the other hand, if the model is more accurate, the bandwidth utilization can be improved. Fig. 11 shows the effect of β on the number of breakpoints for some video streams. We can also have from the same figure a qualitative view of the variable characteristics of the traffics. For example, for the *cartoon* video trace, even for high β (between 0.6 and 0.9) the SISO algorithm detects two breakpoints, which means that the equivalent source presents high fluctuations over long-term intervals.

The D-BIND and the XGOP-B are the models that better capture the variable characteristic of the sources, as it can be observed in Table 2. The complexity of these two models can be compared in terms of the required number of parameters. Table 3 presents the required number of parameters for the XGOP-B model for $\beta = 0.01$. The number of parameters for the D-BIND model is obtained by adjusting the same level of accuracy. The results show that the staircase characteristic of the XGOP-B model leads to an

between 0 and 3 s. After that, the empirical envelope is convex and consequently the model does not follow the curve since its traffic constraint curve is strictly concave. Even though the XGOP-B and the D-BIND models can result in equivalent levels of accuracy, the characterization of the former is straightforward. Furthermore, the β variable can be modified to adjust the number of traffic parameters and consequently the level of accuracy. The accuracy adjustment is simple because the users need not to look into the traffic stream in order to change the traffic parameters, but only adjust the value of β. The smaller is β, more breakpoints are identified, which results in greater accuracy.

Analyzing in details, the behaviors of the D-BIND and XGOP-B models, we observe that the XGOP-B better approximates the empirical envelope. To illustrate this statement, we show in Fig. 10 a zoom in a one-second segment of video. As we can see, the XGOP-B leads to a tighter approximation of the empirical envelope.

In the second part of the analysis, we evaluate the tradeoffs between the adjustment of the β factor and the number of breakpoints. This tradeoff is important because as the number of breakpoints increases, greater is the number
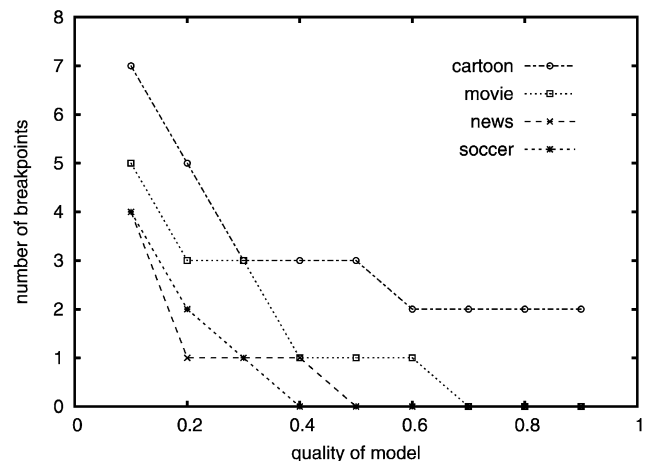

Fig. 11. β versus the number of breakpoints.

Table 3
Complexity in terms of the number of parameters

	XGOP-B	D-BIND
Movie	20	28
News	16	16
Soccer	16	26

improvement accuracy with a lower number of traffic parameters.

The (σ, ρ) model is the simplest model, but may not be accurate depending on the traffic characteristics. The $(\vec{\sigma}, \vec{\rho})$ model defines a piecewise curve which improves the (σ, ρ) model, but does not characterize well concave sources. The D-BIND model is also a piecewise curve model which overcomes the effect caused by concave sources on $(\vec{\sigma}, \vec{\rho})$, resulting in a better characterization. For the D-BIND model, the greater is the number of parameters, the better is the characterization and, as consequence, more complex is the processing.

XGOP-B is also a piecewise curve model. As well D-BIND, the XGOP-B characterizes well concave sources being more accurate than (σ, ρ) and $(\vec{\sigma}, \vec{\rho})$ models. It is not easy to compare XGOP-B with D-BIND, because they have different approaches. The D-BIND model improves the characterization decreasing the I interval whereas XGOP-B increases the considered number of breakpoints. We can say that the XGOP-B model gives a better tuned improvement, because it increases the number the parameters when and where is required. Table 3 shows that the XGOP-B model requires less parameters than D-BIND for the same quality of characterization and, as a consequence, being less complex. Furthermore, the parameter extraction of the XGOP-B is direct obtained.

5. Conclusion

In this article, we proposed the XGOP-B traffic model to provide deterministic characterization of MPEG-compressed video sources. The advantages of the XGOP-B traffic model are twofold. First, it leads to an accurate modeling of the video source because the model takes into account some intrinsic characteristics of MPEG videos. Second, the traffic parameters are obtained in a straightforward fashion, i.e. for a given video source, the parameter extracting algorithm returns a unique set of traffic parameters. The simulation results show that these two features together lead to good results.

The XGOP-B model is based on a two level analysis of MPEG video traffics. The frame-level analysis identifies that MPEG video sources generate bursts due to intra-picture frames. This happens because intra-picture frames are normally much greater than inter-picture frames. This corresponds to a staircase characteristic in the traffic constraint curve. In the GOP-level analysis, the model captures the variable long-term behavior of the sources that reflects the variation of the average video rate due to changes in scenes. These variations are identified as breakpoints. Both the short-term and the long-term analysis were used to obtain an accurate characterization with a small number of traffic parameters. We have succeeded to define a simple and accurate model consisting of four parameters.

The simulation results for the XGOP-B model showed a significant improvement of the characterization. The proposed model results in levels of accuracy that are up to 70% greater than the ones obtained by other traffic models. For the majority of the traces, the XGOP-B model has presented negligible difference from the optimum characterization given by the empirical envelope. It means that the XGOP-B model results in a traffic constraint function that almost meets the optimal characterization. Furthermore, the XGOP-B model introduces a quality parameter that allows to balance complexity and accuracy. We also observed that the XGOP-B model achieves fine modeling with less parameters than other accurate models.

The single-input single-output (SISO) parameter extracting algorithm is a main contribution of this article. A difficult task of general deterministic traffic modeling is to automatically define a unique set of parameters for each video traffic. This is due to loose nature of the traffic constraint functions. In fact, the only restriction is that the traffic constraint function must be an upper bound on the empirical envelop. Thus, there is no inherent requirements for defining an accurate constraint function. The XGOP-B model is associated with the SISO parameter extracting algorithm to compute the traffic parameters. With the SISO algorithm, each parameter is assigned a specific equation that leads to unambiguous results. We showed that it is important to obtain an accurate characterization without the user's intervention to help the choice of the most adequate parameters.

References

[1] W.C. Poon, K.T. Lo, The transportation of VBR-encoded video using deterministic characterization and bandwidth renegotiation, Computer Communications 25 (8) (2002) 731–740.

[2] M. Mandjes, D. Mitra, W. Scheinhardt, Simple models of network access, with applications to the design of joint rate and admission control, in IEEE Infocom, New York, USA, 2002.

[3] Y. Yang, S.S. Hemami, Rate control for VBR video over ATM: simplification and implementation, IEEE Transactions on Circuits and Systems for Video Technology 11 (9) (2001) 1045–1058.

[4] S. Shakkottai, How good are deterministic fluid mod els of internet congestion control?, in IEEE Infocom, New York, USA, June, 2002.

[5] V. Parthasarathy, J.W. Modestino, K.S. Vastola, Reliable transmission of high-quality video over ATM networks, IEEE Transactions on Imaging Processing 8 (3) (1999) 361–374.

[6] M.D. deAmorim, O.C.M.B. Duarte, G. Pujolle, An improved MPEG behavioral analysis with autonomous parameter-extracting algorithm

for strict video applications, in IEEE ICC', New Orleans, USA, June (2000) 2000.

[7] D.E. Wrege, E.W. Knightly, H. Zhang, J. Liebeherr, Deterministic delay bounds for VBR video in packet-switching networks: fundamental limits and practical tradeoffs, IEEE/ACM Transactions on Networking 4 (3) (1996) 352–362.

[8] J.-T. Wang, P.-C. Chang, Error-propagation prevention technique for real-time video transmission over ATM networks, IEEE Transactions on Circuits and Systems for Video Technology 9 (3) (1999) 513–523.

[9] E. Steinbach, N. Farber, B. Girod, Standard compatible extension of H.263 for robust video transmission in mobile environments, IEEE Transactions on Circuits and Systems for Video Technology 7 (6) (1997) 872–881.

[10] M.D. deAmorim, O.C.M.B. Duarte, A novel deterministic traffic model based on a two-level analysis of MPEG video sources, in IEEE Globecom, Sydney, Australia, 1998.

[11] A. Ziviani, B.E. Wolfinger, J.F. deRezende, O.C.M.B. Duarte, S. Fdida, Joint adoption of QoS schemes for MPEG streams, Multimedia Tools and Applications (2003) in press.

[12] V. Elek, G. Karlsson, R. Rönngren, Admission control based on end-to-end measurements, in IEEE Infocom, Tel-Aviv, Israel, 2000.

[13] S. Sahu, V. Firoiu, D. Towsley, J. Kurose, Traffic models and admission control for variable bit rate continuous media transmission with deterministic services, in SPIE Symposium Voice, Video and Data Communications, 1998.

[14] I. Chlamtac, A. Faragó, H. Zhang, A. Fumagalli, A. deterministic, approach to the end-to-end analysis of packet flows in connection-oriented networks, IEEE/ACM Transactions on Networking 6 (4) (1998) 422–431.

[15] J. Liebeherr, D.E. Wrege, D. Ferrari, Exact admission control for networks with a bounded delay service, IEEE/ACM Transactions on Networking 4 (6) (1996) 885–901.

[16] C.-S. Chang, On deterministic traffic regulation and service guarantees: a systematic approach by filtering, IEEE Transactions on Information Theory 44 (3) (1998) 1097–1110.

[17] S.-K. Kweon, K.G. Shin, Providing deterministic delay guarantees in ATM networks, IEEE/ACM Transactions on Networking 6 (6) (1998) 838–850.

[18] C. Chang, Stability, queue length, and delay of deterministic and stochastic queueing networks, IEEE/ACM Transactions on Automatic Control 39 (5) (1994) 913–931.

[19] H. Zhang, D. Ferrari, Improving utilization for deterministic service in multimedia communication, in 1994 International Conference on Multimedia Computing and Systems, Boston, MA, May, 1994, pp. 295–304.

[20] M. Krunz, A. Makowski, Modeling video traffic using M/G/infinity input processes: a compromise between Markovian and LRD models, IEEE Journal on Selected Areas in Communications 16 (5) (1998) 733–748.

[21] E. Knightly, H-BIND: A new approach to providing statistical performance guarantees to VBR traffic, in IEEE Infocom, San Francisco, CA, March, 1996.

[22] I. Nikolaidis, R. Onvural, A bibliography on performance issues in ATM networks, Computer Communications Review 22 (5) (1996) 8–23.

[23] D. Heyman, T. Lakshman, What are the implications of long-range dependence for VBR video traffic engineering, IEEE/ACM Transactions on Networking 4 (3) (1996) 301–317.

[24] V. Sivaraman, F. Chiussi, Providing end-to-end statistical delay guarantees with earliest deadline first scheduling and per-hop traffic shaping, in IEEE Infocom, Tel-Aviv, Israel, March, 2000.

[25] Z.-R. Chang, I. Lee, C.-S. Chang, C.-H. Li, B.-L. Sui, A. novel, A novel scheme using the information of departure processes for delay guarantees of distributed VBR traffic, IEEE/ACM Transactions on Networking 9 (4) (2001) 452–463.

[26] J. Qiu, E. Knightly, Measurement-based admission control with aggregate traffic envelopes, IEEE/ACM Transactions on Networking 9 (2) (2001) 199–210.

[27] R.L. Cruz, A. calculus, A calculus for network delay, part I: network elements in isolation, IEEE Transactions on Information Theory 37 (1) (1991) 114–121.

[28] R.L. Cruz, A. calculus, A calculus for network delay, part II: network analysis, IEEE Transactions on Information Theory 37 (1) (1991) 121–141.

[29] D. Ferrari, D. Verma, A scheme for real-time channel establishment in wide-area networks, IEEE Journal on Selected Areas in Communications 8 (3) (1990) 368–379.

[30] E.W. Knightly, H. Zhang, D-BIND: An accurate traffic model for providing QoS guarantees to VBR traffic, IEEE/ACM Transactions on Networking 5 (2) (1997) 219–231.

[31] D. Le Gall, A video compression standard for multimedia applications, Communications of the ACM 34 (4) (1991) 305–313.

[32] O. Rose, Statistical properties of MPEG video traffic and their impact on traffic modelling in ATM systems, in 20th Annual Conference on Local Computer Networks, Minneapolis, MN, Oct, 1995, pp. 397–406.