Adaptive per-GOP Bandwidth allocation for H.264 Video Transmission over Differentiated Services Networks

Fabio De Vito *Student member IEEE*, Erhan Yilmaz *Student member IEEE*,
Murat Tekalp *Fellow IEEE* and Juan Carlos De Martin *Member IEEE*

***Abstract*— While transmitting over differentiated services networks, in case of severe congestion also the most privileged classes may experience losses. In those cases, and especially in case of video transmission, protecting a higher fraction of traffic can have the effect of decreasing the quality, due to the overload of high-priority classes. We propose a method to compute, at source side, the allocation of video traffic over the available classes to ensure the lowest decoder-side distortion and provide traffic friendliness.**

To show this algorithm performance, the simple case of Poisson traffic with a bottleneck shared-buffer router is shown. The same approach can be extended to other traffic characteristics and router architectures. [1]

***Index Terms*— Differentiated services networks, video traffic, bandwidth allocation, minimum distortion.**

I. Introduction

Continuous media transmission over packet networks is growing in importance in modern communications. For multimedia signals transmission, the network should take into account the strict bounds on loss probability and delay jitter imposed by speech, audio and video (but also synchronized text) playout. Providing quality of service (QoS) to multimedia flows will result in a higher overall quality of the received stream, as perceived by the user. In this direction, differentiated services networks (DiffServ, [1]) are one of the most promising solutions; they offer different quality levels to packets according to the class of service to which they are assigned. The packet classification is performed according to the amount of distortion that each packet would introduce at the decoder in case of loss. The concealment technique used must be taken into account as well.

This distortion defines the *perceptual importance* of the packet. Such a solution is particularly useful in the case of multimedia transmission, where the perceptual importance noticeably varies among packets, allowing a profitable usage of *unequal error protection* (UEP) techniques.

Distortion computation in these cases is a key issue. Several approaches have been proposed; in the *analysis-by-synthesis* (AbS [2]) it is exactly computed by means of complete sequence decoding at encoder side; in *recursive optimal per-pixel estimate* (ROPE [3]), it is estimated by recursively following the motion vectors. In this work, we will use a lightweight model-based estimate, described in [4].

All of these methods consider errors in isolation, so that the effects of multiple errors are considered summable in a first approximation. This is true as long as the loss probability is low. Approaches to multiple losses problem can be found in [5] and [6].

After distortion estimation, packets are divided into classes to be transmitted onto the network.

In this paper, we present a method to determine what fraction of the video traffic should be sent on each class, in order to minimize the distortion of the received video. In this work we will not optimize the allocation given price constraints, as in [7], [8]; we will focus mainly on distortion minimization, and we will how this approach results also in advantages under the price and traffic friendliness points of view.

Whenever congestion occurs, routers react discarding packets from their queues; DiffServ routers behave differently with respect to best-effort counterparts in the discarding policy, since they drop packets starting from low-priority queues. In case of network overload, sources could choose to protect more information from losses, sending more packets in the high-priority classes; this can result in an increased loss rate for those privileged classes, since the number of losses depends on the queue load, therefore producing the effect of reducing the quality.

If a feedback channel from the network is present, this can be used to signal what is the condition of the bottleneck router in terms of per-class traffic load and queue length; the source can then estimate the allocation of traffic in classes (how much of the video bitrate should be assigned to each class) that minimizes the receiver-side distortion.

In this paper, we will suppose the presence of this feedback information. We will describe the results under the hypothesis of Poisson traffic, but the same formulation could be used for any other case.

This paper is organized as follows: in section II we will develop the theory used to compute the expression of total distortion, given an arbitrary bandwidth subdivision in

Fabio De Vito is with Dipartimento di Automatica e Informatica, Politecnico di Torino, Torino 10129 Italy; currently he is visiting PhD student at Department of Electrical Engineering, Koc University, Sariyer, 34450, Istanbul, Turkey (email: fabio.devito@polito.it, fdevito@ku.edu.tr).

Erhan Yilmaz is with Department of Electrical Engineering, Koc University, Sariyer, 34450, Istanbul, Turkey (email: eyilmaz@ku.edu.tr).

Murat Tekalp is with Department of Electrical and Computer Engineering, University of Rochester, Rochester, 14627, NY, USA and with Department of Electrical Engineering, Koc University, Sariyer, 34450, Istanbul, Turkey (email: mtekalp@ku.edu.tr).

Juan Carlos De Martin is with Dipartimento di Automatica e Informatica, Politecnico di Torino, Torino, 10129, Italy (email: demartin@polito.it).

[1]This work is supported by Centro Supercalcolo Piemonte (CSP), Torino, Italy.

classes. In section III we will describe the network parameters used for this study; section IV will show simulation results to confirm the minimization of distortion, and traffic friendliness analysis will also be presented. Conclusions will be outlined in section V.

II. Background

The number of classes available is fixed and depends on the network architecture; the choice of the amount of traffic sent on each class is usually free and limited only by the bandwidth pricing, therefore each source can decide to use only a subset of the available service classes.

On a DiffServ network, if all of the sources assign their traffic to the most privileged class, the effect is the same as sending together all the traffic in one of the other the non-privileged classes. DiffServ networks take advantage from exploitation of packet diversification, so that there should be a way of splitting the traffic which minimizes the distortion.

The aim of this work is to build a tool to discover where this point is placed. We will now present the basic ideas of this work.

For every GOP, we will call *allocation* a vector A of N elements, as many as the number of available classes, showing the fraction of video traffic for each class:

$$A = \left(a_1, a_2, \cdots, a_{N-1}, 1 - \sum_{i}^{N-1} a_i \right).$$

Supposing in a first approximation losses to be uncorrelated, the total distortion for a given allocation can be written as:

$$E(D(A)) = \sum_{i=1}^{N} p_i(A)\overline{D_i(A)} \qquad (1)$$

where $E(D)$ is the expected distortion, p_i and $\overline{D_i}$ represent respectively the loss probability and the average per-packet distortion for class i. In the following we will not indicate the dependence on A. The terms of the sum can be explicitly written for the particular setting being studied, and the sum minimized.

For each allocation, the quantities $\overline{D_i}$ can be computed after storing each GOP's packets. They are reordered according to their importance, and then assigned starting from the highest-importance class; finally, the average for each class is easily computed. The reordering has to be performed only once.

The computation of loss probabilities p_i is instead more complex, since it is not only correlated to the video characteristics but also to the interfering traffic per-class load.

The dependency of the per-class loss probabilities on the allocation A involves communication of some parameters by the network to the sources. In principle, the network should feedback the relationship of each per class loss probability with all the possible allocations, but this would require the implementation of an excessive intelligence in the router nodes, slowing down their performance.

A simpler solution can be communicating only the instantaneous per-class load ρ_i for each class, together with some parameters like the queue and the output link capacities. Those parameters will then be used by the source nodes to compute the expressions of $p_i(A)$.

We will show this computation in the case of a DiffServ enabled router at the bottleneck, with shared buffer memory for packet storage. Poisson traffic is considered to carry out this computation. Shared buffers with several classes can be solved using a multidimensional Markov chain, with one dimension per class. Unfortunately, the number of equations to be written can be excessively high due to the high number of states; for example, the number of states, so the number of equations, in the case of 256 packets of queue dimension and only two classes is in the order of 33 thousands.

We will not solve the multidimensional chain; we will adopt an iterative approach to the problem, moving from some considerations on the discarding policy behavior.

Packets belonging to the first class are lost only if the queue is filled up with first-class packets. If packets of lower priority are present, losses will happen in the other classes in case of congestion. The loss probability of the first class, under Poisson traffic, can then be obtained by means of a unidimensional Markov chain. In Kendall notation this queue will be represented as an $M/M/1/K$ queue, where K is the number of packets that can be stored in queue:

$$P_{loss,1} = \frac{(1-\rho_1)\rho_1^K}{1-\rho_1^{K+1}} \qquad (2)$$

Similarly, losses occur in the first two classes if the queue is full of packets belonging to class one and two; no packet losses will be introduced in these classes if lower priority packets are present. Again, the joint loss probability of class one and two can be computed via a unidimensional chain, whose load is the sum of loads for class one and two:

$$P^c_{loss,2} = \frac{(1-\rho^c_2)(\rho^c_2)^K}{1-(\rho^c_2)^{K+1}} \qquad (3)$$

where $P^c_{loss,2}$ is the loss rate for the cumulation of classes up to the second, and in the same way ρ^c_2 is the cumulative load offered to the router by the traffic in the same classes. Given the Poisson hypothesis, it is:

$$\rho^c_i = \sum_{j=1}^{i} \rho_j \qquad (4)$$

This approach can be iterated recursively to compute the cumulative loss probabilities $P^c_{loss,i}$ for each class i. The last step is to extract the per-class loss probabilities p_i from $P^c_{loss,i}$, to be used in Formula 1.

This process of extraction can be performed as follows. Cumulative loss probability is referred to the aggregate of packets up to a given class, so that for class 1 computation is trivial, since it is $p_1 = P^c_{loss,1}$.

In the following, we will indicate $P^c_i = P^c_{loss,i}$, to allow lighter notation.

For the other priority classes, the following consideration holds: the number of lost packets in the aggregate of the first i classes is equal to the number of lost packets in the aggregate of the first $i-1$ classes plus the number of packets discarded in the i-th class in isolation. In formulas:

$$n^c_{lost,i} = n^c_{lost,i-1} + n_{lost,i} \tag{5}$$

where again $n^c_{lost,i}$ represents the lost packets in the cumulation of the first i classes, and $n_{lost,i}$ is the number of lost packets within the class i in isolation. We will indicate $n^c_i = n^c_{lost,i}$.

The number of lost packets is given by the number of packets present in queue on average, multiplied by the loss probability; in this way Equation (5) becomes:

$$P^c_i n^c_i = P^c_{i-1} n^c_{i-1} + p_i n_i \tag{6}$$

where n^c_i is the number of packets transmitted on average in the first i classes. p_i can then be computed by means of Equation (6):

$$p_i = \frac{P^c_i n^c_i - P^c_{i-1} n^c_{i-1}}{n_i} \tag{7}$$

and since $n_i = n^c_i - n^c_{i-1}$ and considering n^c_i proportional to ρ^c_i, we have:

$$p_i \simeq \frac{P^c_i \rho^c_i - P^c_{i-1} \rho^c_{i-1}}{\rho^c_i - \rho^c_{i-1}} \tag{8}$$

which represents the solution we searched for.

As it is clear form Equation (8), in general the per-class loss probabilities are not constant, but they strongly depend on the load of all the classes with index lower or equal to their one (so higher or equal priority). The above formulas allow to simply compute the loss probabilities under the assumption of Poisson traffic.

Following the same method, computation can be extended to other buffer architectures, queueing and discard policies, and traffic characteristics.

The above computation can be performed for every GOP, to better follow the fast network dynamics.

A. Approximations

Formula (1) is an approximation of the distortion we expect from network simulation. Before any further experiment, it is necessary to consider the different factors that may affect the results.

In developing the theory, we made some assumptions which may introduce errors in our estimate. They are, in descending order of importance:

- the effect of correlation among losses, due to motion prediction;
- the effect of using a simplified interfering traffic model (Poisson); our traffic is "poissonized" in the sense that our encoder is able to output packets whose length in bytes follows a memoryless distribution, even if the effect is limited by the macroblock-based granularity of packets;
- the packet importance is already an estimate of the actual value.

Given these approximations, we expect the curve computed by means of Formula (1) to show a gap with respect to the simulated results. We will show that this still allows us to reach our goal of determining the most convenient allocation.

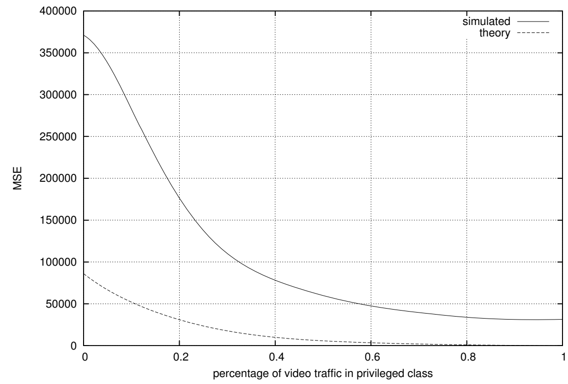

Fig. 1. Comparison between theory and simulation; parameters are described in the text, sequence is *foreman*.

III. Simulation Setup

The video compression algorithm chosen for this work is ITU-T H.264 standard [9]. Simulations have been carried out to check the correctness of the model. The experiment has been set up for the case of two classes, with link capacity ranging among the values of 1 and 1.5 Mbps. The (Poisson) interfering traffic sums up to 1 Mbps and it is allocated in the first class for an amount ranging between 250 and 1000 kbps. The queue capacity is set to values between 16 and 48 packets, in order to introduce a significant amount of losses.

The video sequences are encoded using a codec modified to obtain nearly-memoryless packet lengths. Video bitrate has been set to 250 kbps. With these settings, the total load of the bottleneck will range between 0.83 and 1.25.

IV. Results

In this section we will present some comparisons between the performance obtained using a DiffServ simulator for a bottleneck network as described in Section III, and the performance estimated with the method presented in Section II. The not shown results behave in the same manner.

In Figure 1 the comparison is shown for a core router with queue of 32 packets, two classes and an output link capacity equal to 1 Mbps. The interfering traffic is 1 Mbps, whose traffic is split for 25% in class 1 and 75% in class 2. Being the video traffic 250 kbps, the total load of this network will be $\rho = 1.25$, so generating a high loss percentage, as discussed below.

The two curves show a wide gap between them, as an effect of the approximations outlined in Section II-A. On the other hand, except for a rescaling factor, the two curves show a very similar slope. We can derive from the theory curve the point where the slope stops decreasing significantly. This point indicates in both curves an estimate of the best splitting of video traffic in classes. It is possible to identify this point as $a_1 = 0.5$.

Injecting more video traffic in the most privileged class, will not give significant advantages under the distortion point of view, but it will only increase the price payed to protect video data.

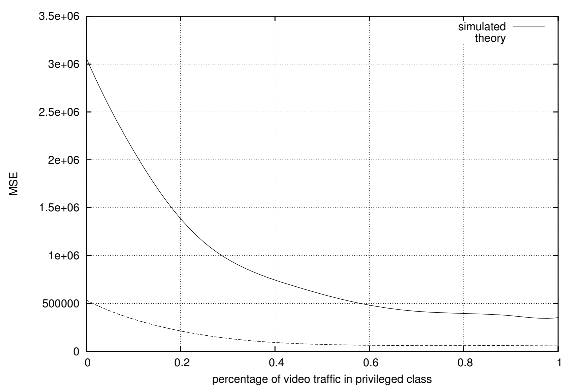

Fig. 2. Comparison between theory and simulation; parameters are described in the text, sequence is *foreman*.

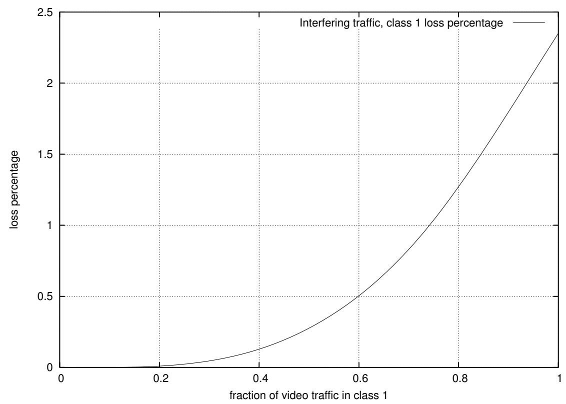

Fig. 3. Class 1 interfering traffic loss rate with respect to the portion of our video sent in class 1; settings are the same as in Figure 1.

The same behavior can be seen with different settings, like in Figure 2, where the queue capacity is set to 48 packets, and the interfering traffic is loading only class 1.

Also in this case, the point where the gain in distortion becomes negligible can be identified around $r_1 = 0.6$.

Different settings will have this point at different values, but it results to be always in the range $r_1 \in [0.4, 0.6]$. For each sequence and setting, the observation of the estimated behavior indicates the fraction of video traffic beyond which there is no significant improvement in distortion; on the other hand, it turns to be cheaper than sending all of the traffic in the privileged class and also traffic friendly, as we will outline in the following.

A. *Traffic friendliness*

We showed that to obtain the minimum video distortion, in case of congestion, sending all the traffic in the most privileged class does not give advantages. This will push the sources to split the traffic over all the classes allowed by the network setting. This behavior will turn to be also traffic friendlier, since it will not overload the high-importance class without degrading excessively the video quality.

To show this traffic friendliness, the loss percentage for interfering traffic in class 1 as a function of a_1 is shown in Figure 3. Settings are the same as in Figure 1.

It can be seen there that this loss probability starts growing significantly around the point we indicated for best compromise between video quality and price for the video source (0.5).

For all the other experiments, the point in which distortion as defined in Formula 1 stops decreasing significantly represents also the point in which the loss probability of the first class starts growing noticeably.

The same considerations hold for packet loss rates of lower-priority classes.

V. Conclusions

In this paper, we proposed a simple modeling of distortion for video transmission over DiffServ networks, based on Poisson traffic hypothesis. The aim of this work is to build a tool to compute how to split the video over the available classes, given the per-packet distortion for the sequence and the network parameters.

We developed formulas to get this approximation as function of the above parameters, and performed simulations to test the correctness of the results. We showed that the proposed formula is able to indicate an interval of values for the video traffic allocation, whose points guarantee nearly the lowest distortion and also low price for the service.

Furthermore, we studied the traffic friendliness of the allocation, and showed that splitting the video using the allocation indicated, the interfering traffic can benefit of lower packet loss rates in all classes.

References

[1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," *RFC 2475*, December 1998.

[2] E. Masala and J.C. De Martin, "Analysis-by-synthesis distortion computation for rate-distortion optimized multimedia streaming," in *Proc. IEEE Int. Conf. on Multimedia & Expo*, Baltimora, MD, July 2003, vol. 3, pp. 345–348.

[3] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 966–976, June 2000.

[4] F. De Vito, D. Quaglia, and J.C. De Martin, "Model-based distortion estimation for perceptual classification of video packets," in *Proc. IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, Siena, Italy, September 2004, vol. 1, pp. 79–82.

[5] Y. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Does burst loss matter?," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Hongkong, China, April 2003, vol. 1, pp. 79–82.

[6] J. Chakareski, J. Apostolopoulos, Wai tian Tan, S. Wee, and B. Girod, "Distortion chains for predicting for video distortion for general loss patterns," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Montreal, Canada, May 2004, vol. 1, pp. 79–82.

[7] A. Sehgal, P. A. Chou, "Cost-distortion optimized streaming media over diffserv networks," in *Proc. IEEE Int. Conf. on Multimedia & Expo*, Lausanne, Switzerland, August 2002, vol. 1, pp. 857–860.

[8] F. Zhai, C.E. Luna, Y. Eisenberg, T.N. Pappas, R. Berry, and A.K. Katsaggelos, "A novel cost-distortion optimization framework for video streaming over differentiated services networks," in *Proc. IEEE Int. Conf. on Image Processing*, Barcelona, Spain, October 2003, vol. 2, pp. 293–296.

[9] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "DRAFT H.264 Standard ISO/IEC 14496-10 (MPEG-4 Part 10)," June 2003.