Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Networks 48 (2005) 829–855
Computer Networks
www.elsevier.com/locate/comnet

A self-adjusting rate adaptation scheme with good fairness and smoothness properties ☆

P. Balaouras *, I. Stavrakakis

Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, Ilissia 15784, Greece

Received 4 August 2003; received in revised form 8 November 2004; accepted 10 November 2004
Available online 22 December 2004
Responsible Editor: N. Fonseca

Abstract

A novel feedback-based rate adaptation scheme is introduced and investigated in this paper. Its main innovative characteristic is the modulation of the rate increment by the distance between a flow's present rate and an assumed targeted maximum rate as dictated by the associated application. The previous along with the shaping of the rate decrement by the reported flow's losses are responsible for a dynamic and self-adjusting behavior that is shown to improve convergence to fairness, the oscillatory behavior of the rate and the induced packet losses when compared with the basic Additive Increase Multiplicative Decrease (AI/MD) scheme. Numerical results illustrate the good properties and intrinsic advantages of the proposed scheme both under the considered modeling assumptions, as well as under more real networking conditions by employing the ns-2 simulator. A brief comparison of the proposed scheme with the TCP-compatible schemes TFRC, IIAD and the non-AI/MD schemes AIPD, LIMPD, is included as well. Because of the aforementioned induced behavior and assumed flow's characteristics (min and max rates), the proposed congestion control scheme seems to be appropriate for regulating the rate of streaming applications.

Keywords: Elastic flows; Rate adaptation schemes; Congestion control; Multimedia; Smoothness; AI/MD

1. Introduction

Congestion control schemes are necessary in order for shared resource networks to avoid or overcome congestion. Such a scheme—that is based on an Additive Increase Multiplicative Decrease (AI/MD) algorithm [1,3], a linear control algorithm—is employed in today's Internet by the

☆ Work supported in part by the General Secretariat for Research and Technology of Greece under a 2002-4 Joint Greek–Italian Scientific and Technological Cooperation Programme and by the Special Account for Research Grants of the National and Kapodistrian University of Athens.

* Corresponding author. Tel.: +30 210 727 5603; fax: +30 210 727 5601.

E-mail addresses: balaoura@noc.uoa.gr, p.balaouras@noc.uoa.gr, balaoura@di.uoa.gr (P. Balaouras).

1389-1286/$ - see front matter
doi:10.1016/j.comnet.2004.11.006

Transmission Control Protocol (TCP) [2] and is largely responsible for its robustness and stability. TCP has been designed—and successfully used—for unicast reliable data transfer over low bandwidth lines. One of its primary goals for fast recovery from congestion is achieved by halving (multiplicative decrease factor b_D equal to 0.5) the transmission rate upon congestion. The rate halving affects the TCP flows' smoothness, making TCP—and generally all AI/MD schemes with a small value of b_D—unsuitable for the continuous media (CM) streaming applications. Furthermore, the TCP's retransmission mechanism introduces typically unacceptable end-to-end delay and delay variation for CM applications. The requirement for the timely delivery prohibits the recovery of the lost packets through a retransmission scheme as the recovery mechanism typically fails to meet the decoding deadlines [21,22]. This is the reason why the User Datagram Protocol (UDP), that lacks a retransmission mechanism is mainly used to support the CM streaming applications in conjunction with the Real-time Transport Protocol (RTP) and Real-time Transport Control Protocol (RTCP) [5]. Applications that use UDP as transport protocol should also implement an end-to-end congestion control scheme, like TCP, to retain the stability of the Internet. Otherwise, all supported applications will suffer and eventually the network will collapse [6]. Alternative approaches focus on providing for a smoother adaptation while achieving a similar throughput to a TCP flow [4,16,20,23].

This work is motivated by the need of CM flows to smoothly adapt their rate toward the fair share in a dynamic environment where new (existing) CM flows are initiated (terminated). We take into consideration the following characteristics/constraints of the encoders in order to improve the adaptation behavior of the CM flows:

(a) the existence of a minimum transmission rate m that corresponds to the lower acceptable perceptual quality for the end user;
(b) the existence of a maximum transmission rate M that may correspond to one of the following: (i) the maximum rate of the encoder, (ii) a high level of perceptual quality, over which the gain of the perceived quality is relatively low, (iii) the maximum rate that the receiver could receive due to link or decoding capabilities constraints;
(c) the fact that commercial CM streaming applications are rate-based and RTP/RTCP-based, therefore the use of RTCP reports as a feedback mechanism eliminates the need for an additional mechanism and enables the use of *continuous* (as opposed to binary) feedback (i.e., packet loss rate) in the decrease policy;
(d) the encoder's inability to adjust its encoding rate at a high frequency (i.e., once every round trip time (RTT)) and instantaneously, resulting in intervals between successive adaptation point of larger time scale than that of RTT.

Specifically, we explore in the present work a sender-driven rate-based adaptation scheme that exploits the knowledge of (a) a pair of minimum and maximum allowable rates (m, M) in the increase policy, which are considered to be common to all flows in this paper, and (b) the packet loss rate in the decrease policy. The increase function depends on the distance between the current rate and a maximum allowable rate (Distance Weighted Additive Increase (DWAI) policy). The use of the minimum and maximum allowable rates allow for a variable rate increment that is maximized when the flow's rate reaches the minimum allowable rate and minimized when the flow's rate reaches the maximum allowable rate. Therefore, since lower rate flows are more aggressive than higher rate flows, the proposed increase policy converges faster to the fair share than the basic AI which has a neutral impact to the fairness. We use the reported packet loss rate in the decrease policy (Loss rate Dependent Multiplicative Decrease (LDMD) policy) since schemes based on loss dependent decrease policies, such as those presented in [10,29,18], are capable of adjusting the flow's rate according to the loss rate and present a smoother adaptation behavior than the basic multiplicative decrease policy. The new rate shaping is based on the current "effective sending rate" ((1 − current loss rate) × current rate) scaled by a factor d (new rate = $d \times$ (effective sending rate)).