Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Dynamic multimedia stream adaptation and rate control for heterogeneous networks*

SZWABE Andrzej[1], SCHORR Andreas[2], HAUCK Franz J.[2], KASSLER Andreas J.[3]
([1]*Poznan University of Technology, Poznan 60-965, Poland*)
([2]*University of Ulm, Ulm 89069, Germany*)
([3]*University of Karlstad, Karlstad SE-65188, Sweden*)
E-mail: Andrzej.Szwabe@put.poznan.pl; andreas.schorr@uni-ulm.de; franz.hauck@uni-ulm.de; kassler@ieee.org
Received Dec. 1, 2005; revision accepted Feb. 15, 2006

Abstract: Dynamic adaptation of multimedia content is seen as an important feature of next generation networks and pervasive systems enabling terminals and applications to adapt to changes in e.g. context, access network, and available Quality-of-Service (QoS) due to mobility of users, devices or sessions. We present the architecture of a multimedia stream adaptation service which enables communication between terminals having heterogeneous hardware and software capabilities and served by heterogeneous networks. The service runs on special content adaptation nodes which can be placed at any location within the network. The flexible structure of our architecture allows using a variety of different adaptation engines. A generic transcoding engine is used to change the codec of streams. An MPEG-21 Digital Item Adaptation (DIA) based transformation engine allows adjusting the data rate of scalable media streams. An intelligent decision-taking engine implements adaptive flow control which takes into account current network QoS parameters and congestion information. Measurements demonstrate the quality gains achieved through adaptive congestion control mechanisms under conditions typical for a heterogeneous network.

Key words: Stream adaptation, Quality-of-Service (QoS), Heterogeneous networks, Rate control, MPEG-21 DIA
doi:10.1631/jzus.2006.AS0063 **Document code:** A **CLC number:** TN919.8

INTRODUCTION

During the last decade a clear trend towards all-over-IP multimedia communication has evolved, but the heterogeneity of devices and networks and the limited Quality-of-Service (QoS) support in the Internet prevents multimedia-enabled devices from either communicating at all or constrains the communication to low quality. Different devices use different media codecs, and even if they support a common set of media formats, in many situations it is impossible to use a common media format due to resource constraints. For instance, decoding high-resolution video streams may require more CPU resources than available on a Personal Digital Assistant (PDA). A possible solution for these problems is to adapt the media stream (i.e. change the format, resolution or data rate) at the media source. Unfortunately, this is not always possible because of resource constraints of the terminal.

Another promising solution to cope with this heterogeneity problem is to adapt the stream on special Content Adaptation Nodes (CANs) inside the network. In addition to simply converting media from one format to another, CANs located at the edge of an access network may also perform adaptive error and rate control. Adapting error control mechanisms to the current network (link) characteristics can result in better performance of the adaptation process (Schorr *et al*., 2004). In comparison to earlier work on media adaptation (Kassler and Schorr, 2003), our architecture allows use of various different adaptation engines

* Project supported by IST FP6 Integrated Project DAIDALOS (No. IST-2002-506997), and the German Research Foundation (DFG) within the AKOM Framework (No. HA2207/2-3)

for manipulating audio and video streams on-the-fly and includes a transformation engine based on MPEG-21 Bitstream Syntax Description (BSD) (Panis *et al*., 2003), which allows us to efficiently adjust the data rate of scalable media streams. Furthermore, the architecture includes an intelligent Decision Taking Engine (DTE) used for controlling adaptation and rate control parameters. The DTE adopts several advanced solutions of state-of-the-art rate-control algorithms, e.g., TCP-Friendly Rate Control (TFRC) (Handley *et al*., 2003), and also controls the media adaptation engines inside the CAN.

Section 2 of this paper provides an overview of current research activities in the area. Section 3 introduces the architecture of the Content Adaptation Node, developed within the project IST-Daidalos (http://www.ist-daidalos.org). Section 4 describes the rate-control mechanism of the DTE. Section 5 demonstrates quality gains achieved through the adaptive rate-control mechanisms. We conclude the paper in Section 6.

RELATED WORK

MPEG-21 Part 7 Digital Item Adaptation (DIA) (Vetro *et al*., 2004) defines tools for the adaptation of Digital Items, e.g. audio and video streams. MPEG-21 Bitstream Syntax Description (BSD) allows retrieving a variety of adapted versions of media streams from a single bit stream by performing efficient editing-style operations. MPEG-21 DIA does not define interactions with existing transport and network technologies. Therefore, an architecture such as the CAN is necessary to perform MPEG-21 DIA on a proxy node in the network.

Different approaches exist to realize scalable media streaming, like Receiver Driven Layered Multicast (McCanne *et al*., 1997) or Stream Switching (Amon and Pandel, 2003). However some approaches are limited to supporting only a small number of different media qualities, or they do not provide the possibility to adapt several dimensions of scalability (e.g. frame rate versus picture quality) at the same time (Kassler, 2001).

Rate smoothness is considered to be an important factor of stream quality (Floyd *et al*., 2000a). In contrast to the Additive-Increase-Multiplicative-Decrease (AIMD) scheme, equation-based rate control respects rate stability and smoothness (Floyd *et al*., 2000a; 2000b). Another factor of streaming media quality is the ability to avoid packet losses. TFRC uses RTT information not only for calculation of TCP-friendly bandwidth but also for oscillation cancellation (Handley *et al*., 2003; Floyd *et al*., 2000a). In result, this mechanism avoids packet losses which would otherwise be caused by rate oscillations.

OVERALL ARCHITECTURE

We propose a generic media CAN which is able to adapt the quality of media streams as well as change their format on the fly. An important design decision was that clients requesting adaptation services should give the CAN some degrees of freedom in choosing adaptation parameters. For instance, a client may ask the CAN to create an MPEG-4 video stream with a data rate of 400 kbps, but accepts also bit streams with lower data rates when resources are scarce and during temporary network congestion. Two decision components inside the CAN are used for adjusting adaptation parameters. A decision taking engine (DTE) is responsible for fine-grained decisions on individual media stream level fine-tuned through QoS feedback information received via Real-time Control Protocol (RTCP) (Schulzrinne *et al*., 2003). Additionally, a central decision component, the so-called Content Adaptation Coordinator (CAC), coordinates all parallel adaptation sessions of the CAN and reacts if resources (like the CPU or total bandwidth of a network interface) become scarce.

In this paper, we focus on the CAN Media Manager (MM) component responsible for receiving and transmitting media samples using the Realtime Transport Protocol (RTP) (Schulzrinne *et al*., 2003). A description of the overall CAN architecture can be found in (Kassler and Schorr, 2003). The MM carries out the actual adaptation operations: format-conversion (transcoding), data-rate scaling, joint error and rate control.

A high-level overview of the MM architecture is shown in Fig.1. Each InChannel instance receives a single flow of RTP packets from its upstream node (another CAN or media source). Depending on the

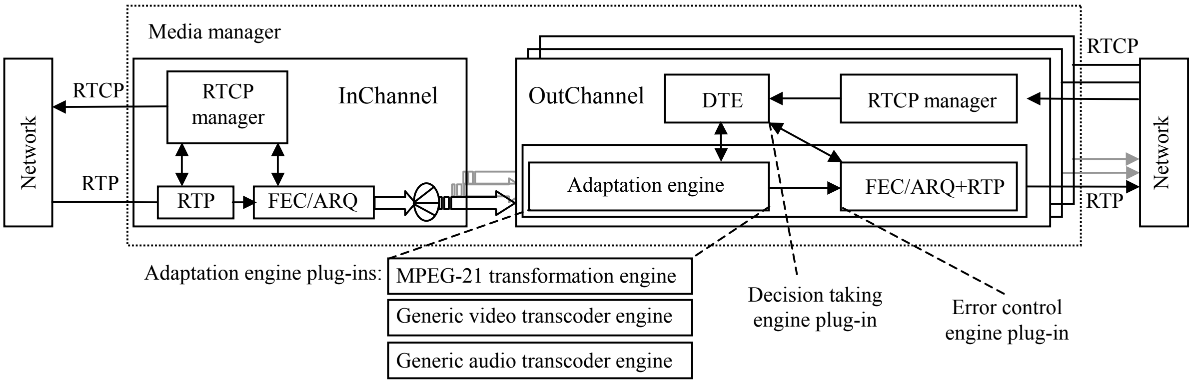

Fig.1 Media manager architecture

number of corresponding downstream nodes, one or more OutChannels are connected with each InChannel, thus enabling point-to-point or point-to-multipoint sessions. The internal components of the OutChannel are implemented as plug-ins. An appropriate Transformation Engine (TE) plug-in performs the media transformation operations, and an Error Control plug-in applies media-independent error control like Forward Error Correction (FEC) (Rosenberg and Schulzrinne, 1999) or ARQ (Rey *et al.*, 2004). The Decision Taking Engine (DTE) adjusts the adaptation parameters according to the current QoS delivery conditions as indicated by the RTCP Manager component based on RTCP Receiver reports.

RATE CONTROL SCHEME

The rate control scheme applied by the DTE follows the equation-based approach of TFRC as specified in IETF RFC 3448 (Handley *et al.*, 2003). Applying solutions of TFRC to the CAN sender-based architecture require several modifications. Four-tap Infinite Impulse Response (IIR) filters (Ifeachor and Jervis, 1993) are used for processing packet loss rate (PLR), round trip time (RTT) and the volume of the stream.

The quality of the decision taking process may be improved by applying a modified version of RTT-based oscillation cancellation mechanism:

$$X_{\mathrm{corr}}(n) = X(n) + \alpha[R_{\mathrm{opt}}(n) - R_{\mathrm{sample}}(n)] \times \frac{X_{\mathrm{rcv}}(n)}{T_{\mathrm{RTCPp}} + R_{\mathrm{sample}}(n)},$$

where if $[PLR(n)=0] \wedge [PLR(n-1)>0]$,

$$R_{\mathrm{opt}}(n) = \beta \frac{R_{\mathrm{sample}}(n-1)}{2} + (1-\beta)R_{\mathrm{sqmean}}(n),$$

and otherwise

$$R_{\mathrm{opt}}(n) = \beta \frac{\max[R_{\mathrm{opt}}(n), R_{\mathrm{sample}}(n)]}{2} + (1-\beta)R_{\mathrm{sqmean}}(n),$$

here X (expressed in bps) is the bit rate before correction, X_{corr} is the corrected bit rate, R_{sample} (expressed in s) is the reported RTT, R_{smean} is the mean RTT value, R_{opt} is estimated optimal RTT value, X_{rcv} (in bps) is an average bit rate received during the last RTCP period, and T_{RTCPp} (in s) is the period between two consecutive RTCP SR packets. The constant denoted as α determines the amount of correction to be applied to the target rate, while constant β determines how strongly the modified mechanism differs from the TFRC-based mechanism.

MPEG-21 DIA-BASED ADAPTATION

In this section, we present the results of the performance evaluation of the implemented rate control mechanisms for MPEG-21 DIA-based adaptation of gBSD BSAC streams.

Testbed architecture

We built a testbed for evaluating the performance of the Media Manager. The testbed included a RTP-based streaming server and a client application equipped with a BSAC decoder. The source streamed gBSD BSAC packets (Feiten *et al.*, 2005), while the payload of the stream adapted by the CAN consisted only of AAC BSAC frames (ISO/IEC, 1998). The bit rate of the non-adapted stream was 128 kbps. The adaptation was performed with a scaling grain of 2 kbps per stereo stream.

The NIST Net tool (Carson and Santay, 2003) was used to emulate a bottleneck bandwidth and a factor of delay (emulated end-to-end link-level delay and queuing delay introduced by network nodes). The effects of a variable queuing delay and queuing losses were obtained by means of a Derivative Random Drop (DRD) mechanism. The maximum queue size was set to 64 packets.

TCP-friendly rate control algorithms developed for the Internet (i.e. not supported by QoS management system) tend to behave oscillatorily in uncongested networks (Floyd *et al*., 2000a; 2000b). It is crucial to evaluate the behavior of a rate control mechanism in the single bottleneck scenario (Floyd *et al*., 2000a; 2000b). Therefore a comparison of various DTE's rate control schemes was made based on quantitative evaluation of their behavior against network conditions dominated by a single drop-tail queue.

Quality measures

Performance evaluation included measurements of PLR-based parameters (average PLR, the highest recorded PLR), average sending bit rate and coefficient of variation (CoV)—a rate variability measure proposed for evaluation of rate control schemes (Floyd *et al*., 2000a).

Evaluated algorithms

Evaluation of the DTE's rate-control algorithm was done by comparing four enhanced versions of the algorithm with four algorithms not featuring the enhancements proposed by the authors. In addition, the comparison presents results obtained by using non-adaptive streaming (Null_test).

The most representative TFRC-based options—TFRC1a, TFRC2a—were obtained by straightforward adaptation of TFRC's features (Handley *et al*., 2003) to a sender-based architecture of the MM. While the algorithms TFRC1a and TFRC1b used simple one-tap IIR filter to process PLR data, TFRC2a and TFRC2b used a 4-tap filter to process the PLR signal. In the case of TFRC2a, the algorithm included TFRC's oscillation cancellation.

The difference between the two most advanced versions of the enhanced algorithms (EDTE1a, EDTE1b) was a different value of a constant determining an amplitude of oscillation cancellation. The comparison also includes two algorithms which feature PLR correction driven by stable bandwidth estimate but were not supported by the modified version of oscillation cancellation mechanism (EDTE2a, EDTE2b). One of these versions featured oscillation cancellation mechanism as specified in TFRC (EDTE2b) while the other one did not feature any RTT-driven oscillation cancellation (EDTE2a). Table 1 presents an overview of the evaluated algorithms' capabilities.

Table 1 Capabilities of the evaluated algorithms

Option	1-tap filter	4-tap filter	PLR correction	Oscillation cancellation
TFRC1a	×	–	–	–
TFRC1b	×	–	–	–
TFRC2a	–	×	–	–
TFRC2b	–	×	–	–
EDTE1a	–	×	×	Modified
EDTE1b	–	×	×	Modified
EDTE2a	–	×	×	–
EDTE2b	–	×	×	TFRC-like

Results of the algorithms' comparison

As the CAN is used for high-quality video streaming services, the DTE's rate control algorithm should be able to achieve low PLR for both average ('Average PLR') and maximum ('Highest reported PLR'). From this perspective, the best of the evaluated options are those which feature PLR correction driven by stable bandwidth estimate used together with any RTT-driven oscillation cancellation mechanism, i.e., options EDTE1a, EDTE1b and EDTE2b. This advantage can be seen in Fig.2. Especially low PLR values were achieved with the algorithms which featured the modified RTT-driven oscillation cancellation (EDTE1a, EDTE1b), although oscillation cancellation adopted from TFRC (EDTE2b) allowed for trading some amount of quality regarding PLR for even higher average send rate.

Fig.3 presents the evaluation regarding measures related to send rate after adaptation. Good quality of rate control provided by the most advanced options

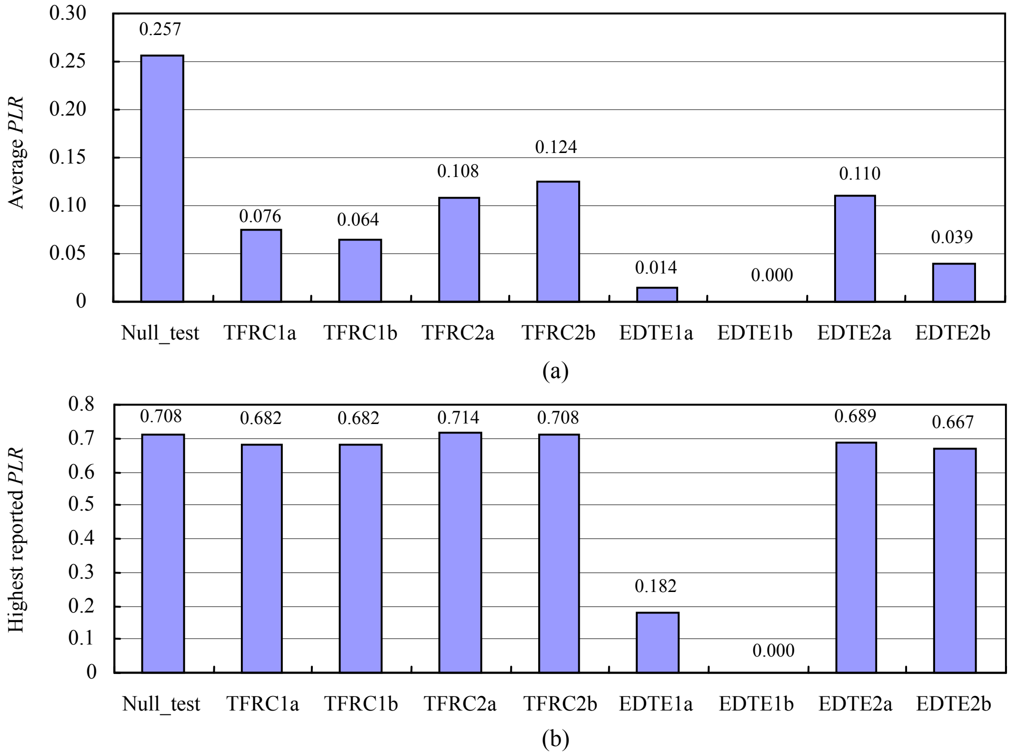

Fig.2 Measures related to packet losses. (a) Average *PLR*; (b) Highest reported *PLR*

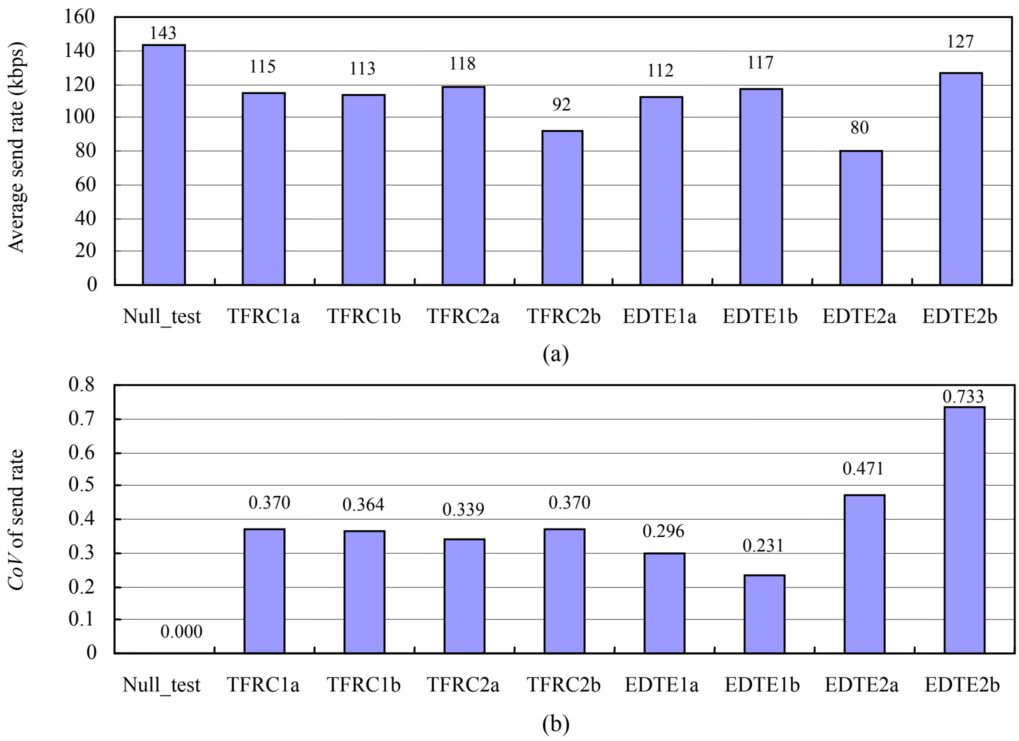

Fig.3 Measures related to send rate. (a) Average send rate (kpbs); (b) *CoV* of send rate

(EDTE1a and EDTE1b) has been additionally confirmed by high average sending rates, low values of CoV of the sending rate as well as good results regarding highest reported PLR. In the case of option EDTE2b, the modified oscillation cancellation mechanism was able to reduce packet losses to zero through strong queue overflow avoidance.

Options not featuring the modified oscillation cancellation mechanism produced much less stable rate (i.e. of high CoV) either due to too slow rate increase after strong rate reduction (EDTE2a), or as a result of too prompt reaction to RTT changes (EDTE2b). In the case of option EDTE2a, slow rate increase (not accelerated by any RTT-driven correction) resulted in very low average send rate.

Behavior of the TFRC-based algorithm

Fig.4a presents the behavior of algorithm TFRC1a, which uses the same 1-tap filter for PLR and RTT processing (Handley *et al*., 2003). This algorithm responds quickly to congestion. However, the effectiveness of oscillation cancellation is reduced because the mean RTT value constantly follows the actual RTT sample and thus it does not really repre sent the optimal RTT.

As shown in Figs.2 and 3, a configuration of a multi-tap filter's coefficients may result in less oscillatory behavior of the algorithm. However, such a filter configuration significantly reduces rate-probing performed in periods without packet losses (TFRC2b). The TFRC-based algorithm applies the oscillation cancellation to a theoretical bandwidth rather than to the real bandwidth (Fig.4a). This results in limited effectiveness of oscillation cancellation.

Analysis of the optimized algorithm

Fig.4b shows that when rapid and strong oscillation canceling correction was applied to the target rate, the algorithm (i.e. the option EDTE1b) achieved fully successful congestion avoidance, i.e. no packets were lost (despite high queue utilization observed during large parts of the test). A cost of the avoiding packet losses was too prompt rate decrease experienced during the period of queue length decrease visible around 27th second of the test (Fig.4b).

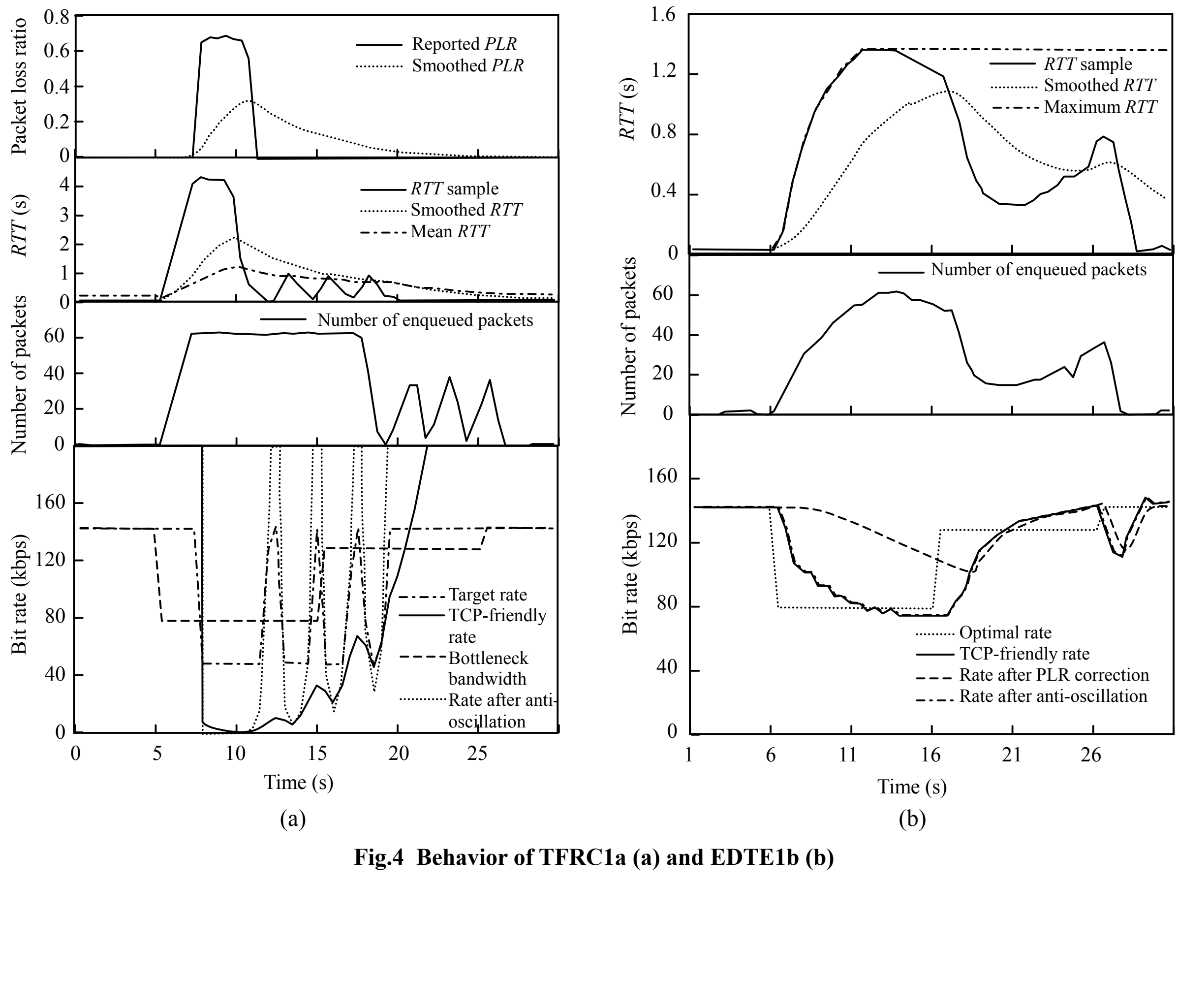

Fig.4 Behavior of TFRC1a (a) and EDTE1b (b)

CONCLUSION

In this work, we studied how rate control and adaptation can work together to support heterogeneous media delivery. The contributions of this paper are manifold. In particular, we (1) presented the architecture of an MPEG-21-DIA based CAN which supports functionalities typically offered by streaming-oriented overlay networks: stream format conversion, stream rate adaptation and application-layer multicast; (2) demonstrated how the CAN enables an innovative integration of rate control with functionalities related to multimedia session setup and control; (3) proposed several enhancements to rate control mechanisms suitable for adaptive multimedia streaming, and integrated those schemes with MPEG-21 DIA; (4) evaluated our proposed schemes over a variety of scenarios and showed significant advantages gained by combining oscillation cancellation with stable bandwidth estimation.

The rate control scheme presented in this paper successfully estimates TCP-friendly rate and effectively minimizes packet losses of a stream transmitted in a QoS-enabled, heterogeneous network. The implementation presented in the paper can be regarded as an innovative application of advanced equation-based rate control to a streaming system compliant to the MPEG-21 DIA (http://clabprj.ee.nctu.edu.tw/~mpeg21tb/).

References

Amon, P., Pandel, J., 2003. Evaluation of Adaptive and Reliable Video Transmission Technologies. Proc. of the 13th Packet Video Workshop. Nantes, France.

Carson, M., Santay, D., 2003. NIST net—a linux-based network emulation tool. *ACM SIGCOMM Computer Communications Review*, **33**(3):111-126. [doi:10.1145/956993.957007]

Feiten, B., Wolf, I., Guenkova-Luy, T., Schorr, A., 2005. New Mode for rfc3640: AAC-BSAC with MPEG-21 gBSD. draft-feiten-avt-bsacmode-for-rfc3640-00.txt, IETF AVT WG.

Floyd, S., Handley, M., Padhye, J., 2000a. A Comparison of Equation-Based and AIMD Congestion Control. ACIRI Technical Report. Http://www.aciri.org/tfrc/.

Floyd, S., Handley, M., Padhye, J., Widmer, J., 2000b. Equation-Based Congestion Control for Unicast Applications: the Extended Version. ICSI TR-00-003. Http://www.aciri.org/tfrc/.

Handley, M., Floyd, S., Padhye, J., Widmer, J., 2003. TCP Friendly Rate Control (TFRC): Protocol Specification. IETF RFC3448.

Ifeachor, E.C., Jervis, B.W., 1993. Digital Signal Processing. A Practical Approach, Addison Wesley Publishing Company.

ISO/IEC, 1998. Final Draft International Standard 14496-3: MPEG-4 Audio, ISO/IEC JTC1/SC29/WG11 N2503.

Kassler, A., 2001. Video Adaptation within a Quality of Service Architecture. Dissertation, University of Ulm, Germany.

Kassler, A., Schorr, A., 2003. Generic QoS Aware Media Stream Transcoding and Adaptation. Proc. of the 13th Packet Video Workshop. Nantes, France.

McCanne, S., Vetterli, M., Jacobson, V., 1997. Low complexity video coding for receiver-driven layered multicast. *IEEE Journal on Selected Areas in Computing*, **15**(6): 983-1001. [doi:10.1109/49.611154]

Panis, G., Hunter, A., Heter, J., Hellwagner, H., Kosch, H., Timmerer, C., Devillers, S., Amielh, M., 2003. Bitstream syntax description: A tool for multimedia resource adaptation within MPEG-21. *EURASIP Signal Processing: Image Communication Journal*, **18**(8):721-747. [doi:10.1016/S0923-5965(03)00061-4]

Rey, J., Leon, D., Miyazaki, A., Varsa, V., Hakenberg, R., 2004. RTP Retransmission Payload Format. Draft-ietfavt-rtp-retransmission-10.txt. Http://www.ietf.org/internet-drafts/.

Rosenberg, J., Schulzrinne, H., 1999. An RTP Payload Format for Generic Forward Error Correction. IETF RFC2733.

Schorr, A., Kassler, A., Petrovic, G., 2004. Adaptive Media Streaming in Heterogeneous Wireless Networks. Proc. of IEEE International Workshop on Multimedia Signal Processing (MMSP2004). Siena, Italy.

Schulzrinne, H., Casner, S., Frederick, R., Jacobson, V., 2003. RTP: A Transport Protocol for Real-Time Applications. IETF, RFC 3550.

Vetro, A., Timmerer, C., Devillers, S., 2004. Information Technology—Multimedia Framework (MPEG-21)—Part 7: Digital Item Adaptation. ISO/IEC, Tech. Rep. 21000-7:2004.