SEGMENT DURATION FOR RATE ADAPTATION OF ADAPTIVE HTTP STREAMING

Chenghao Liu[1], Imed Bouazizi[2] and Moncef Gabbouj[1]

[1]Department of Signal Processing, Tampere University of Technology, Finland
[2]Nokia Research Center, Tampere, Finland
chenghao.liu@tut.fi, imed.bouazizi@nokia.com and moncef.gabbouj@tut.fi

ABSTRACT

Recently, 3GPP packet-switched streaming (PSS) specified adaptive HTTP streaming (AHS). The client requests series of media segments and adapt the bitrates of segment to varying network resources. The current state-of-the-art rate adaptation of AHS estimates the end-to-end network capacity using the previous segment reception such as segment fetching time. However, the accuracy and the speed of such rate adaptation highly depend on the segments duration. This paper presents a novel segment duration determination method to provide accurate and fast rate adaptation of AHS. The segment duration is estimated as the minimum duration to produce the smoothed HTTP/TCP rate which represents the current network capacity. Rate adaptation method using the determined segment duration is presented. To prevent buffer draining-up, the segment duration is further restrained as fine-grained duration. Simulation results show that the proposed segment duration enables the rate adaptation algorithm to increase achievable media bitrates and reduce the play-back interruption compared with the state-of-the-art rate adaptation method of AHS.

Index Terms— Segment duration, rate adaptation, adaptive HTTP streaming, DASH, 3GPP

1. INTRODUCTION

Recently, HTTP has been widely used for the delivery of real-time multimedia content over the Internet, such as in video streaming applications. Unlike RTP/UDP, HTTP is easy to configure and is typically granted traversal of firewalls and network address translators (NAT), which makes it attractive for multimedia streaming applications. Real-time media delivery over TCP is firstly reported in the paper [1] which reveals that the instantaneous transmission rate variation of TCP can be smoothed out by receiver-side buffering. Paper [2] presented sender-driven rate adaptation method for media streaming over TCP, however, which results in the scalability issues.

3GPP packet-switched streaming (PSS) [3] specified adaptive HTTP streaming (AHS) for delivery of real-time multimedia over the HTTP. In the AHS, the HTTP-streaming client requests series of media segments from the HTTP-streaming server to progressively download media data. The proxy can be utilized for caching segments. For rate adaptation, the HTTP-streaming client can request different bitrates-encoded segments.

In the HTTP streaming, the instantaneous receiving rate is varying due to TCP congestion control. Recently, a segment fetch time based rate adaptation (SFT-RA) method of AHS is presented in the paper [4]. The segment fetch time denotes spent time to receive a segment. In HTTP-streaming, the segment duration denotes media duration contained in a segment. This paper uses the ratio of segment duration to the segment fetch time as rate adaptation metrics. However, the accuracy and the speed of the SFT-RA method highly depend on the segments duration. Long segment duration results in the slow reaction in the rate adaptation to the bandwidth change. While short segment duration causes the inaccurate estimation of the end-to-end network capacity and incurs non-optimum rate adaptation. But, paper [4] does not present a method to determine the segment duration.

The motivation of this paper is to determine the optimum segment duration in order to provide accurate and fast rate adaptation for the AHS applications. In the HTTP streaming, the achievable streaming media bitrates can be estimated as the average reception HTTP/TCP rate, called smoothed HTTP bitrates, in certain duration. It is the most challenging task to determine the duration in which the smoothed HTTP bitrates represents the end-to-end networks capacity for the rate adaptation usage. Moreover, the duration to produce the smoothed HTTP bitrates varies according to different packet loss rate, round trip time (RTT) and bandwidth. In the literature, the paper [5] reports famous TCP throughput modeling using packets loss rate and RTT. As best of our knowledge, however, none of research works presents a method of theoretically estimating the duration for producing the smoothed HTTP/TCP bitrates for the rate adaptation usage.

This work was supported by the Academy of Finland, (application number 129657, Finnish Programme for Centres of Excellence in Research 2006-2011).

978-1-61284-350-6/11/$26.00 ©2011 IEEE

This paper presents a novel method to determine the segment duration to produce the smoothed HTTP bitrates. First, the variance of the congestion window mean in certain duration is estimated firstly, as TCP transmission rate is controlled by the congestion window. Second, the segment duration is determined as the minimum duration in which the variance of the mean of the congestion window is lower than a predefined threshold. A method to deploy the segment duration in the rate adaptation of AHS is presented in this paper. The optimum rate adaptation speed and rate adaptation accuracy can be obtained since the deployed segment duration is the minimum duration to produce the smoothed HTTP bitrates which represents the end-to-end network capacity.

2. TCP CONGESTION WINDOW STATISTICS

This section describes the statistics of the congestion window presented in the famous paper [5]. The TCP congestion algorithm uses the congestion window to control the amount of data that TCP can send. The congestion window expectation is

$$E[W] \approx \sqrt{\frac{8}{3bp}} \tag{1}$$

where b denotes the number of packets that are acknowledged by a received ACK and p denotes the packets loss rate.

The TCP congestion avoidance is modeled in the term of rounds. A round starts with the back-to-back transmission of W packets, wherein W denotes current congestion window size. Once all packets in the current congestion window have been sent, the TCP sender stops sending any further packets until it receives the first ACK. The round ends with the ACK's reception. The duration of a round is equal to the round trip time and assumed to be independent of the window size. X_i denotes the account of round where the first packet loss is detected by triple duplicate ACKs. For the algorithm of congestion avoidance, the following relationship between X_i and the congestion window exists

$$W_i = \frac{W_{i-1}}{2} + \frac{X_i}{b} \tag{2}$$

where W_i and W_{i-1} denote the congestion window size at round X_i and round X_{i-1}.

The expectation of X_i is modeled with b and p

$$E[X] \approx \sqrt{\frac{2b}{3p}} \tag{3}$$

3. PROPOSED SEGMENT DURATION DETERMINATION

Our paper focuses on derivation of the minimum duration to produce the smoothed HTTP bitrates for the rate adaptation usage. The variance of the mean of the congestion window in certain duration is used to scale the smoothed factor of HTTP/TCP rate since the HTTP/TCP bitrates is controlled by the congestion window. The evolution of W_i is evaluated

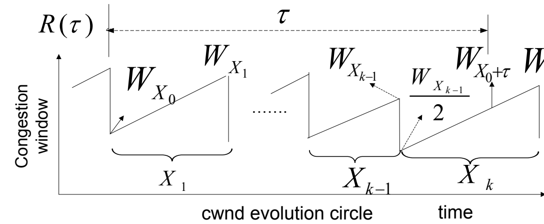

Fig. 1. Evolution of congestion window

in terms of rounds as shown in Fig. 1.

According to the Eq. (2), $E[W^2]$ can be derived as

$$E[W^2] = \frac{4}{3}\left(\frac{1}{b}E[W].E[X] + \frac{1}{b^2}\left(E[X]^2 + Var[X]\right)\right) \tag{4}$$

where $E[W]$ and $E[X]$ can be calculated as (1) and (3), X_i denotes the account of round before occurring of first packet loss in a congestion window (cwnd) evolution circle which is assumed to meet the poisson distributed. In poisson distribution, the $Var[X]$ is equal to$E[X]$, therefore we have

$$E[W^2] = \frac{4}{3b^2}\left(\frac{2b}{p} + \sqrt{\frac{2b}{3p}}\right) \tag{5}$$

From (1) and (5), the variance of the congestion window is

$$Var[W] = \frac{4}{3b^2}\sqrt{\frac{2b}{3p}} \tag{6}$$

In the congestion avoidance algorithm, the congestion window is correlated thus the variance of the average congestion window $Var[\bar{W}]$ is

$$Var[\bar{W}] = \frac{Var\,[W]}{n} + \frac{(n-1)RVar\,[W]}{n} \tag{7}$$

where $Var[W]$ denotes the variance of congestion window, n denotes the number of the congestion window samples and R denotes the mean of the autocorrelation of W.

For calculating the autocorrelation of cwnd (W) one sample is selected as start of a cwnd circle denoted as $W_{X_0}$and second one, denoted as $W_{X_0+\tau}$, is moving to cover the whole evolution wherein the distance between two congestion window samples is denotes as τ. In the case of $\sum_{i=1}^{k-1} X_i < \tau \leq \sum_{i=1}^{k} X_i$, we have the general equation to calculate the $W_{X_0+\tau}$

$$W_{X_0+\tau} = \frac{W_{X_{k-1}}}{2} + \left(\tau - \sum_{i=1}^{k-1} X_i\right) \tag{8}$$

where $W_{X_0+\tau}$ denotes the congestion window in $X_0 + \tau$ round, $W_{X_{k-1}}$ denotes the peak of congestion window at X_{k-1} which is the round count in the evolutionary circle $k-1$, τ denotes the distance between $W_{X_0+\tau}$ and W_{X_0} in terms of round count. In addition, W_{X_k} has relationship with $W_{X_{k-1}}$ as

$$W_{X_k} = \frac{W_{X_{k-1}}}{2} + X_k \tag{9}$$

Using the autocorrelation formula, Eq. (8) and (9), the autocorrelation $R[\tau]$ is

$$R[\tau] = \begin{cases} \frac{1}{2^k}, & \sum_{i=1}^{k-1} X_i < \tau \leq \sum_{i=1}^{k} X_i\,,\ k \geq 2 \\ 1, & \tau \leq X_1 \end{cases} \tag{10}$$

where τ denotes distance between samples in terms of round count, the k denotes the cwnd evolution circle index and X_k denotes the round count in the cwnd evolution circle k.

Then the mean of the autocorrelation of W is

$$R = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{2^{k-1}}X_k \quad (11)$$

where n denotes the number of round count and X_k denotes the round count in the cwnd evolution circle k. Typically a smoothed TCP rate contains a large amount of rounds, so n is a large value.

It can be further estimated as

$$R \approx \frac{2E[X]}{n} \quad (12)$$

where $E[X]$ denotes the expectation of X_k.

From (7) and (12), the variance of the average congestion window $Var[\bar{W}]$ is

$$Var[\bar{W}] = \frac{(n+2(n-1)E(X))}{n^2}Var[W] \quad (13)$$

The smoothing factor of the average congestion window is quantified with the variance. If variance is less than the upper limit then we call the congestion window as mid-term smoothed congestion window.

$$Var[\bar{W}] \leq \varepsilon \quad (14)$$

where ε denotes the upper limit of the variance of the avera ge congestion window. It follows that

$$n \geq \frac{(2E[X]+1)Var\,[W]+\sqrt{(2E[X]+1)^2Var\,[W]^2-8\varepsilon E\,[X]Var\,[W]}}{2\varepsilon} \quad (15)$$

From (6) and (15) we have that

$$n \geq \frac{8+\frac{6p}{b}\sqrt{\frac{2b}{3p}}+6\sqrt{\frac{16}{9}-4bp\,\varepsilon+\frac{2p}{3b}(4\sqrt{\frac{2b}{3p}}+1)}}{9bp\,\varepsilon} \quad (16)$$

where n denotes the account of the round of the congestion window evolution to produce the mid-term smoothed congestion window and also denotes the account of the round to produce the mid-term smoothed TCP rate, p denotes the packet loss rates, b denotes the number of packets that are acknowledged by a received ACK and ε denotes the upper limit of the variance of the average congestion window with round account n. So left term of the (16) is the minimum account of the round of the congestion window evolution to produce the mid-term smoothed congestion window wherein the smoothing factor is measured with ε and lower ε denotes the higher smoothing factor.

The time unit of congestion window evolution process is the round trip time denoted as RTT. So the segment duration is determined as the minimum duration to produce the smoothed congestion window, which can be obtained by multiplying n of (16) by the RTT. And the $d_{smoot\ h}^{min}$ is

$$SD = RTT \;\; \frac{8+\frac{6p}{b}\sqrt{\frac{2b}{3p}}+6\sqrt{\frac{16}{9}-4bp\,\varepsilon+\frac{2p}{3b}(4\sqrt{\frac{2b}{3p}}+1)}}{9bp\,\varepsilon} \quad (17)$$

where SD denotes the segment duration.

4. PROPOSED SEGMENT DURATION BASED RATE ADAPATION METHOD

A rate adaptation is presented using the segment duration presented in the section 3 to demonstrate the usage of segment duration in rate adaptation of AHS. Fig. 2 shows the flowchart of the rate adaptation deploying the proposed segment duration determination method at the client. After receiving current segment, the client estimates the smoothed HTTP bitrates as the bits of the segments divided by the spent time to receive the current segment. The smoothed HTTP bitrates can represent end-to-end networks capacity as the current segment duration covers feasible duration to smooth varying instantaneous HTTP rate. The next segment bitrates is selected as the first bitrates of the server provided (in descending order) which is lower than the smoothed HTTP bitrates. The media bitrates for the client to select is obtained from the media presentation description (MPD).

To prevent the buffer-underflow, the segment duration should selected as minimum of the segment duration estimated using (17) and the minimum buffer draining time due to the possible bandwidth decrease. In (17), average packet loss rates (p) and round trip time (RTT) are estimated at the transport layer. The estimated values are passed to the application layer of the client. The available buffered media time to drain is equal to the current buffered media time minus the minimum buffered media time. And the condition to calculate the minimum buffer draining time is that the achievable HTTP bitrates is decreased from the current estimated media bitrates to the lowest media bitrates.

The client requests the next segment with the determined bitrates and the segment duration. The accurate and fast rate adaption can be obtained because the proposed segment duration is the minimum duration to produce the smoothed HTTP bitrates.

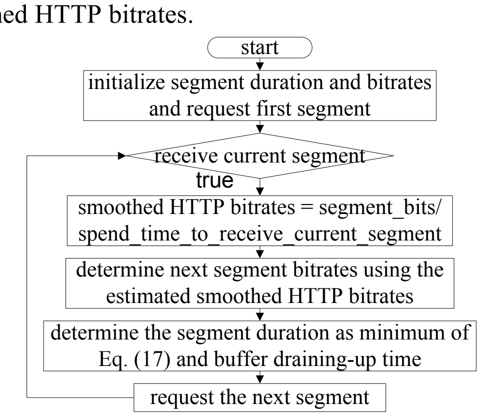

Fig. 2. Proposed segment duration based rate adaptation flowchart

5. SIMULATION RESULTS

To demonstrate the usage of the proposed segment duration in the rate adaptation of AHS, we implemented the proposed segment duration based rate adaptation method (SD-RA) and the recently presented segment fetch time based rate adaptation method (SFT-RA) [4] in ns2 [6]. For media data, 10 sets of different bitrates encoded bit-streams are provided by the HTTP-streaming server to the client for rate adaptation. The bitrates covers from 300Kbits/s to 1200Kbits/s with distance of 100Kbits/s. The initial and minimum buffered media time is selected as the 20s and 10s. In (17), the ε is set as 0.3.

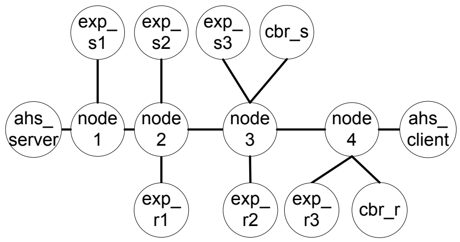

Fig. 3. Network topology

Fig. 3 shows the network topology used in the simulations. The bandwidths of the links between node 1, 2, 3 and 4 are set as 2.5 Mbits/s and all others are set as 2.0 Mbits/s. And all link delays are set as 2ms. Three exponential (EXP/UDP) background traffic senders are attached to node 1, 2 and 3 and receivers are attached to node 2, 3, and 4, respectively, to simulate the dynamic bandwidth, delay and loss rate. One constant bitrates (CBR/UDP) traffic sender and receiver is attached to node 3 and node 4 with different bitrates to evaluate the rate adaptation. All above traffics start from 0s and end at 900s except the CBR/UDP sender which starts from 300s and ends at 600s. For each method, the 5 set of simulation was run with 5 different CBR/UDP bitrates from 0.8 Mbits/s to 1.6 Mbits/s. The achievable HTTP-streaming bitrates is not equal to the bottleneck bandwidth, but related to bandwidth, RTT and packet loss rate etc.

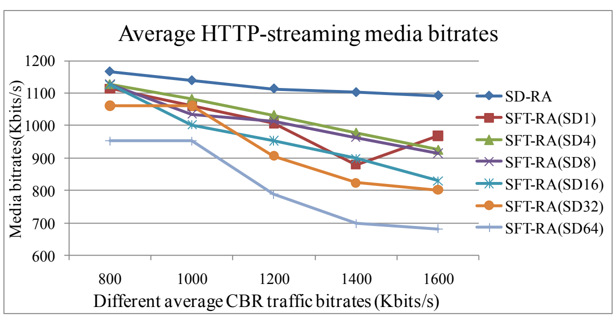

Fig. 4. Average HTTP-streaming media bitrates

Fig. 4 shows the average bitrates of the proposed SD-RA and SFT-RA method with the different CBR/UDP average bitrates. For the SD-RA the segment duration is determined using the proposed method. For the SFT-RA, the 6 different lines are plotted in the Fig. 4 corresponding to the results with 6 different segment durations including 1s (SFT-SD1), 4s (SFT-SD4), 8s (SFT-SD8), 16s (SFT-SD16), 32s (SFT-SD32) and 64s (SFT-SD64). SFT-RA does not present a method to determine the segment duration.

Simulation results show that the average HTTP streaming media bitrates of the proposed SD-RA are higher than all of the 6 results of the SFT-RA method. The proposed method estimates the smoothed HTTP streaming bitrates and matches the media bitrates to the estimated smoothed HTTP streaming rate. So the proposed SD-RA method can enhance bitrates quickly instead of the step-wise switching up used in the SFT-RA. In the proposed method, the segment duration is determined to produce the smoothed HTTP bitrates which represents the end-to-end network capacity, so the matching media bitrates to the bandwidth can be used. In contrast the SFT-RA method uses the step wise rate adaptation method as SFT-RA method cannot figure out if the segment duration is set long enough to smooth the sawtooth-shaped TCP rate. In addition, the SFT-RA method shows the different average HTTP-streaming bitrates for the different segment durations as we expected.

The client buffer draining-up did not occur in the proposed method, but occurred one time in the SFT-RA with segment duration 1s (SFT-RA SD1) when the bottleneck bandwidth decreased at time 300s.

6. CONCLUSION

In this paper, we present a novel method to determine the segment duration for the rate adaptation in adaptive HTTP streaming (AHS). In order to use the average bitrates of previously received segment to estimate the bandwidth, the segment duration should be set wisely to smooth the saw-tooth shaped HTTP/TCP rate. The proposed method determines the segment duration as the minimum duration in which the average reception rate can represent the end-to-end networks capacity for the usage of the rate adaptation. Simulation results show that the proposed segment duration based rate adaptation outperforms the rate adaptation method without using the segment duration determination method.

7. REFERENCES

[1] C. Krasic, K. Li, and J. Walpole, "The Case for Streaming Multimedia with TCP," *In Proceedings of IDMS*, Lancaster. UK, September 2001.

[2] L.S. Lam, J.Y.B. Lee, S.C. Liew, W. Wang, "A transparent rate adaptation algorithm for streaming video over the internet," *In 18th International conference on advanced information networking and applications*, Fukuoka, Japan, March 2004.

[3] 3GPP TS 26.234, "Transparent End-To-End Packet-Switched Streaming Service (PSS): protocols and codecs," (Release9), Dec. 2010.

[4] C. Liu, I. Bouazizi, M. Gabbouj, "Rate adaptation for adaptive HTTP streaming," *ACM MMSys 2011, Special Session: Modern Media Transport, Dynamic Adaptive Streaming over HTTP (DASH)*, San Jose, California, Feb. 23-25, 2011.

[5] Padhye,J., Firoiu, V., Towsley, D. and Kurose, J, "Modeling TCP Reno performance: a simple model and its empirical validation," *IEEE/ACM Transactions on Networking*, vol. 8, no. 2, pp. 133-145, April 2000.

[6] The University of Southern California, "The Network Simulator - ns-2," July 2006.