Segment-based Teletraffic Model for MPEG-DASH

Ognen Ognenoski #1, Maria G. Martini #2, Peter Amon *3

WMN Research Group, Kingston University
London, UK
1 ognen.ognenoski@kingston.ac.uk
2 m.martini@kingston.ac.uk

* Imaging and Computer Vision, Siemens Corporate Technology
Munich, Germany
3 p.amon@siemens.com

***Abstract*—This paper proposes a novel segment-based teletraffic model for dynamic adaptive video streaming over HTTP (DASH), based on the MPEG-DASH standard. The functionality of the standard's framework is mapped on the proposed model for which the probabilities of buffer overflow, empty buffer and active buffer are adopted as relevant performance metrics. These metrics describe the MPEG-DASH streaming process from the teletraffic viewpoint and are observed for different segment sizes of the encoded video at the server and different buffer sizes at the client. The results show that small segment sizes of the encoded video increase the probabilities of empty buffer and active buffer, while decreasing the probability of buffer overflow, whereas large segment sizes are suitable for decreasing the probabilities of empty buffer and active buffer, while increasing the probability of buffer overflow. Regardless of the segment sizes, the performance metrics are further improved when the buffer size at the client is increased.**

I. Introduction

Various forms of video are expected to be the dominant service constituting 90% of the global consumer traffic and 66% of the world's mobile traffic by 2015 [1]. Streaming models are becoming more common and they are in particular applicable transmission technologies for video delivery to Internet users. Different models (true streaming, download and play, progressive download and play) [2] conform differently to the variety of video applications. In the context of protocol utilization, streaming models can be classified in the following three groups: *traditional streaming* (RTP/RTCP on top of UDP), *progressive streaming* (HTTP on top of TCP) and *adaptive streaming* (HTTP on top of TCP). Nowadays, adaptive video streaming models support only 17% of the Internet video traffic. However, the study in [3] reports on the growth of the adaptive streaming portion of video traffic, whose expected increase is estimated at an average of 77% a year hence supporting 51% of Internet video by 2015.

Several commercial platforms (Apple, Microsoft, Adobe) provide proprietary solutions for adaptive video streaming. However, these platforms lack the required interoperability features to provide a common integrated ecosystem for video streaming. The emerging MPEG-DASH standard aims to provide a server and client interoperability framework, increasing reliability, performances, and overall Quality of Experience (QoE). The advantages of this technique compared to other streaming techniques (RTP/RTSP streaming, HTTP progressive streaming) are extensively reported in literature [4] indicating towards MPEG-DASH as a promising future framework for adaptive video streaming.

This paper proposes an analysis of MPEG-DASH video streaming from the teletraffic viewpoint. The proposed segment-based teletraffic model is based on the application of probability and queuing theory to the general MPEG-DASH streaming model. The teletraffic model is utilized for probabilistic assessment of the MPEG-DASH system performances through the probabilities of buffer overflow, empty buffer, and active buffer as relevant performance metrics. The metrics are observed for different segment sizes of the encoded video at the server and different buffer sizes at the client. The authors in [5] investigate how different segment sizes affect objective video quality metrics. This paper provides similar insight, however the metrics are observed from the teletraffic viewpoint.

The paper is organized as follows. Section II gives a brief overview on the MPEG-DASH standard. The teletraffic model is elaborated in Section III, which includes the model's boundaries, the adopted segment selection approach, and the performance metrics. Section IV presents the results from the performance analysis. Finally, Section V concludes the paper.

II. MPEG-DASH overview

The MPEG-DASH standard [6] adopts as baseline the specifications introduced in the 3GPP Release 9 PSS standard [7] and the Open IPTV Forum (OIPF) specifications. This standard framework provides definitions that can be classified in two parts. The first part defines the *Media Presentation Description* (MPD), which is a manifest file in XML format used by the client to learn about content availability, location and addititional content characteristics. The second part defines the *content format* in terms of media segments as extensions to 3GP and MP4 media file formats. Each media segment is assigned a unique Uniform Resource Locator (URL), an index, and explicit or implicit start time and duration. The required media stream is composed of consecutive media segments that

MMSP'13, Sept. 30 - Oct. 2, 2013, Pula (Sardinia), Italy.
978-1-4799-0125-8/13/$31.00 *©2013 IEEE.*

are downloaded from a server using the HTTP/TCP protocols, and the video is played out at the client. The proposed teletraffic model in this paper adopts the segment as basic unit in the MPEG-DASH video streaming.

A. Media Presentation Description

The *MPD file format* consists of one or more periods. A period is a program interval along the temporal axis which has start time, duration, and contains one or multiple adaptation sets. The adaptation set provides information about one or more media components and its encoded alternatives (e.g. same media component but different bitrates of the video/audio component). Adaptation sets include multiple representations. A representation contains alternatives of the same multimedia content, but encoded with different bitrate and resolution, that are organized in multiple segments. The segment is an entity body of the response to the DASH client that makes an HTTP-GET request to retrieve media from the server. Segments contain the actual media information and are located on specific URLs in the network.

B. Segment formats and types

The *segment format* specifies the syntax and semantics of the resources associated with the URL identified by the MPD file. The HTTP-GET request to a resource in MPD is responded with an HTTP response that contains information conforming to a segment format. The standard focuses on segment formats based on MPEG container formats and it describes the usage of two segment formats: Media Segments based on the ISO Base Media File Format (ISO/IEC 14496-12) and Media Segments based on the MPEG-2 Transport Stream (ISO/IEC 13818-1). Both segment formats are defined such that the Media Segment format complies with the respective container formats. The standard defines the following segment types:

- *Initialization*. Segments containing initialization information for access to different Representation;
- *Media*. Segments containing encoded media content components;
- *Index*. Segments primarily containing indexing information for Media Segments;
- *Bitstream Switching*. Segments containing essential data to switch to the Representation to which it is assigned.

III. System model

The mapping of the MPEG-DASH streaming model on the proposed teletraffic model is depicted in Fig. 1, where network and teletraffic views are presented to capture the video streaming in both domains.

From the network side, the DASH encoded video segments are transfered from the server through the Internet, then the segments are received through the access network and played at the DASH client. It should be noted that the WLAN AP shown in Fig. 1 is selected as an example, since the model is more general and can refer to different types of access networks.

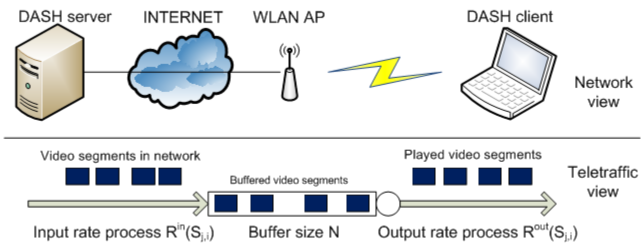

Fig. 1. System model: network and teletraffic view.

From the teletraffic viewpoint, the arrival of the video segments from the network and their playout at the client is modeled by the input and output processes $R^{\text{in}}(S_{\text{j,i}})$ and $R^{\text{out}}(S_{\text{j,i}})$, expressed in bits/second. These processes correspond to the average arrival/serving rates, which are used to calculate the throughput $\gamma(S_{\text{j,i}})$ for the observed segment $S_{\text{j,i}}$. The model used for the teletraffic analysis is based on the Schwartz diagram [8] and it is depicted in Fig. 2.

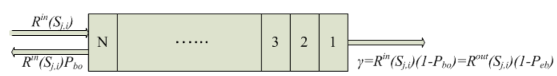

Fig. 2. Arrivals and serving rates in the queue.

We adopt a finite buffer (queue) size for which the arrival and serving rate vary for every segment. Furthermore, we assume a simple M/M/1 queuing system as a reasonable approximation for the arrival and serving of the segment's packets. The buffer size N shows the number of packets that can be buffered at the client. Regarding the client's buffer we observe the following three cases:

- buffer is full, hence the arriving packet is dropped;
- buffer is active, hence the arriving packet is stored in the buffer and played out at its respective display time;
- buffer is empty, there is no arriving packet and playout is stopped.

The case of full buffer indicates that packets in the system are dropped, hence retransmissions occur. In the case of acitve buffer, i.e., when the buffer contains buffered segments, the arriving packets are played in their respective display time. The state of empty buffer causes stopping in the playout process. The effects of the three states are captured through the following proposed performance metrics for the teletraffic model:

- *Probability of buffer overflow*;
- *Probability of empty buffer*;
- *Probability of active buffer*.

The following notations are adopted for the teletraffic model:

- $S_{\text{j,i}}$ refers to the segment from representation j in the MPD file indexed with i;
- $B_{\text{j,i}}$ refers to the number of bytes used for the encoding of the segment $S_{\text{j,i}}$;
- $T_{\text{j,i}}$ refers to the time duration (i.e., size) of the segment $S_{\text{j,i}}$;

- $t(S_{j,i})$ refers to the variable time needed for downloading segment $S_{j,i}$;
- r_j is the encoded rate of the jth representation, where $j = 1, 2, 3, ..J$;
- M is the total number of segments that compose the video $i = 1, 2, 3, ..M$;
- $R^{in}(S_{j,i})$ refers to the rate of the input process for segment $S_{j,i}$;
- $R^{out}(S_{j,i})$ refers to the rate of the output process for segment $S_{j,i}$;
- $\gamma(S_{j,i})$ refers to the throughput in the model.

The following assumption for the model is adopted: the number of bytes in segments with fixed size from a particular representation is equal ($B_{j,1} = B_{j,2} = ... = B_{j,M}$, for all j) and depends only on the representation level (j) and the segment size T. Henceforth, B_j is used to denote the number of bytes in a segment depending on the representation j.

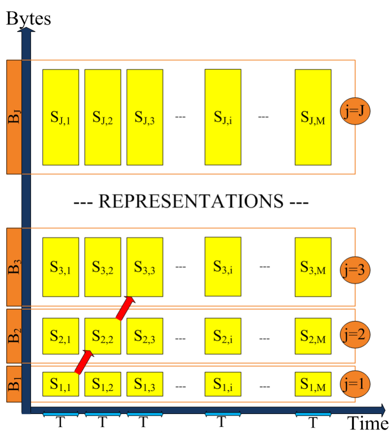

Fig. 3. DASH representations.

In Fig. 3 we show a simple diagram of the DASH representations over time, with fixed segment size T. The segments from each representation j, contain B_j bytes, which are streamed to the client. The switching among different representations is performed at segment level, i.e., by every T seconds.

A. Model boundaries

The boundaries of the input and output process in the teletraffic model are directly dependent on the MPD encoding of the video (representation levels, segment sizes) and on the network conditions/protocol efficiency (available throughput, round trip time, re-transmissions, etc.). The average arrival rate of the input process for the segment $S_{j,i}$ is calculated according to (1), whereas the average serving rate of the output process for the obtained segment $S_{j,i}$ is calculated according to (2):

$$R^{in}(S_{j,i}) = \frac{B_j}{t(S_{j,i})}; \tag{1}$$

$$R^{out}(S_{j,i}) = \frac{B_j}{T}. \tag{2}$$

Each segment from the MPD is delivered utilizing the HTTP/TCP protocols, where the packet size L of the Protocol Data Unit (PDU) is 1430 bytes. The segment size within one MPD is fixed regardless of the representation and the segment index ($T_{j,i} = T$). The following set of integers $\{K_1, K_2, K_3, .., K_J\}$ is introduced to represent the number of packets that are needed for the delivery of a particular segment that corresponds to representation r_j, where $B_j = K_j L$, regardless of the segment's index (i).

The maximum arrival rate for the input process is obtained when a segment i from the highest representation level r_J is downloaded for the minimal time supported by the network $t_{min}(S_{J,i})$, whereas the minimum arrival rate for the input process is obtained when a segment i from the lowest representation level r_1 is downloaded for the maximum time possible in the network $t_{max}(S_{1,i})$. The t_{max} and t_{min} parameters are the maximum and minimum values needed for delivery of a packet in the network. Hence, the arrival rate in the input process is bound according to (3) and (4):

$$R^{in}_{max}(S_{J,i}) = \frac{B_J}{t_{min}(S_{J,i})} = \frac{K_J L}{\sum_{k=1}^{K_J} t_{min}}; \tag{3}$$

$$R^{in}_{min}(S_{1,i}) = \frac{B_1}{t_{max}(S_{1,i})} = \frac{K_1 L}{\sum_{k=1}^{K_1} t_{max}}. \tag{4}$$

Similarly, for the maximum/minimum serving rate of the output process the boundaries are dependent on the segment size and are calculated according to (5) and (6):

$$R^{out}_{max}(S_{J,i}) = \frac{B_J}{T} = \frac{K_J L}{T}; \tag{5}$$

$$R^{out}_{min}(S_{1,i}) = \frac{B_1}{T} = \frac{K_1 L}{T}. \tag{6}$$

B. Segment selection

The rate of the input process depends on the representation level from which a segment is selected and the throughput variations in the network during the download of the segment's bytes. The adopted approach in this paper selects a segment from a representation based on the averaged network throughput experienced in a time interval equal to the segment size (T). The averaging is done across the measured throughput values in a time window equal to the segment size, in which the packets from the segment are received. Hence, if the average throughput is G, then the i-th segment is selected from the representation j for which $G - r_j$ is the minimal positive number. The selection of the segment determines the rate of the output process.

C. Model metrics

The metrics for the teletraffic model are derived based on the approach in [9], which provides a description of a general streaming model. This model proposes a consumption function $A(t)$ at the player, a transmission schedule function $S(t)$, and a buffer function $b(t)$, which are used to observe the accumulated data as function of time. These functions correspond to the segment-based functions in the teletraffic model, i.e., the output process $R^{\text{out}}(S_{\text{j,i}})$, the input process $R^{\text{in}}(S_{\text{j,i}})$, and the buffer for the segments, which is modeled with its size N. Furthermore, the general streaming model proposes an overflow limit function $B(t)$ as a sum of the consumption and the buffer function. This function determines the upper bound of the model for which the streaming process is preserved.

The performance metrics for the teletraffic model are derived based on [10]. The probability of buffer overflow $P_{\text{bo}}(S_{\text{j,i}})$ corresponds to the overflow limit function in the general streaming model and it is calculated as a blocking probability for finite queue, according to (7):

$$P_{\text{bo}}(S_{\text{j,i}}) = \Big(\frac{R^{\text{in}}(S_{\text{j,i}})}{R^{\text{out}}(S_{\text{j,i}})}\Big)^N \frac{1 - \dfrac{R^{\text{in}}(S_{\text{j,i}})}{R^{\text{out}}(S_{\text{j,i}})}}{1 - (\dfrac{R^{\text{in}}(S_{\text{j,i}})}{R^{\text{out}}(S_{\text{j,i}})})^{N+1}}. \qquad (7)$$

The throughput $\gamma(S_{\text{j,i}})$ in the teletraffic model can be calculated considering the average arrival rate and the probability of buffer overflow according to (8), or considering the average serving rate and the probability of empty buffer according to (9):

$$\gamma(S_{\text{j,i}}) = R^{\text{in}}(S_{\text{j,i}})(1 - P_{\text{bo}}(S_{\text{j,i}})); \qquad (8)$$

$$\gamma(S_{\text{j,i}}) = R^{\text{out}}(S_{\text{j,i}})(1 - P_{\text{eb}}(S_{\text{j,i}})). \qquad (9)$$

The probability of empty buffer is given in (10), and is derived from (8) and (9).

$$P_{\text{eb}}(S_{\text{j,i}}) = 1 - (\frac{R^{\text{in}}(S_{\text{j,i}})(1 - P_{\text{bo}}(S_{\text{j,i}}))}{R^{\text{out}}(S_{\text{j,i}})}). \qquad (10)$$

In the end, the probability of active buffer is calculated according to (11):

$$P_{\text{ab}}(S_{\text{j,i}}) = 1 - P_{\text{eb}}(S_{\text{j,i}}) - P_{\text{bo}}(S_{\text{j,i}}). \qquad (11)$$

In order to investigate the effect of different segment sizes of the encoded video, an averaging approach for these probabilities is introduced. The averaged probabilities of buffer overflow, empty buffer, and active buffer are calculated according to equations (12)-(14):

$$P_{\text{bo}}^{\text{av}} = \frac{1}{M}\sum_{i=0}^{M} P_{\text{bo}}(S_{\text{j,i}}); \qquad (12)$$

$$P_{\text{eb}}^{\text{av}} = \frac{1}{M}\sum_{i=0}^{M} P_{\text{eb}}(S_{\text{j,i}}); \qquad (13)$$

$$P_{\text{ab}}^{\text{av}} = 1 - P_{\text{bo}}^{\text{av}} - P_{\text{eb}}^{\text{av}}. \qquad (14)$$

IV. Performance analysis

The performance analysis is derived utilizing files obtained from tracing real MPEG-DASH streaming over a 54 Mbps WLAN network with the open-source platforms provided in [11]. From this trace we obtain the available data rate in the network during the download of the video segments with the considered segment sizes. These values are used to determine the average arrival/serving rate for the input and output processes, which are used as an input to the teletraffic model implemented in Matlab. The serving rate of the output process is determined by the segment selection approach described in Section III-B.

Fig. 4 shows the arrival and serving rates for the input and output process as a function of the number of video segments. The representations of the *Big Buck Bunny* sequence available in [11] are considered for the rate of the output process. It should be noted that different rates for the input and output process would change the observed values, however the trends would be preserved.

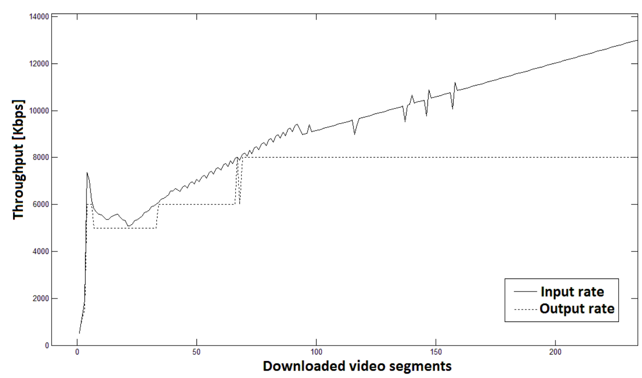

Fig. 4. Snapshot of the model's input trace.

The average probabilities of buffer overflow, empty buffer and active buffer for different segment sizes and buffer sizes are plotted in Figs. 5-7.

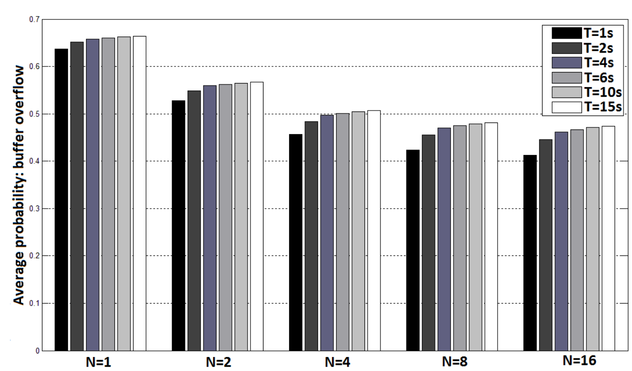

Fig. 5. Average probability of buffer overflow.

The configurations with higher buffer size N result as expected in decreased probabilities of buffer overflow, and empty buffer. As a result, the probability of active buffer is

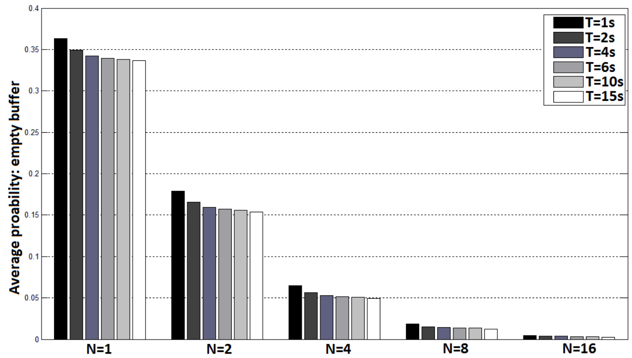

Fig. 6. Average probability of empty buffer.

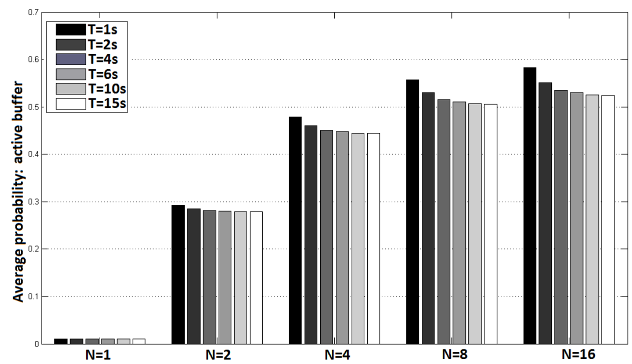

Fig. 7. Average probability of active buffer.

increased, hence a larger client buffer produces better overall performances since it improves the proposed performance metrics. When the buffer size is fixed, the effect of the video segment size is observed. The results show that small segment sizes of the encoded video are preferable for decreasing the probability of buffer overflow at the cost of increased probability of empty buffer. In addition, in this case the probability of active buffer is increased which would result in smooth playout of the buffered segments. However, this provides limitation in the adaptation since the already buffered segments can be with lower quality that the network can support at a particular moment in time. Furthermore, the performance analysis shows that large segment sizes are preferable for decreasing the probabilities of empty buffer and active buffer at the cost of increased probability of buffer overflow.

Besides segment size, the mode of segmentation of the video has a significant effect on the streaming performances [12]. In our analysis, the I-frame position is fixed by restricting the GOP-size of the encoder to the segment size. Therefore, smaller segment sizes result in a higher number of I-frames, thus, in order to keep the bitrate of the stream corresponding to the particular representation, less bits can be used to encode the following P frames and the average errorless PSNR for the small segments decreases as reported in [5]. As future work we plan to investigate different segmentation techniques.

V. Concluding remarks

This paper proposes an approach for mapping the functionality of the standard MPEG-DASH model into a segment-based teletraffic model, for which appropriate probability metrics (i.e., probabilities of buffer overflow, empty buffer, and active buffer) are proposed and investigated for different video segment sizes at the DASH server and different buffer sizes at the DASH client. Small segment sizes of the encoded video increase the probabilities of empty buffer and active buffer, while decreasing the probability of buffer overflow, whereas large segment sizes are suitable for decreasing the probabilities of empty buffer and active buffer, while increasing the probability of buffer overflow. Furthermore, as expected, the performance metrics are improved when the buffer size at the client is increased.

The MPEG-DASH streaming in the context of protocol utilization is based on HTTP/TCP, where the impairments in the video streaming occur during large round trip times of the TCP packets and due to retransmissions when packets are lost, which influence the network throughput and hence the selected video representation from the DASH server. The analysis in this paper shows that by increasing the size of the advertised TCP window (i.e., larger buffer size) at the client and selecting appropriate segment sizes the overall performances of the streaming can be improved. In the end, we outline that different content preparation/segmentation techniques should be adopted in parallel with streaming adaptation in order to further enhance the quality of the video stream.

Acknowledgements

The research leading to these results has received funding from the European Unions Seventh Framework Programme ([FP7/2007-2013]) under grant agreement N^{o} 288502 (CONCERTO).

References

[1] "Cisco visual networking index: Global mobile data traffic forecast update 2011-2016," White Paper, Cisco, February 2012.

[2] W. Simpson, *Video over IP: Internet video, H.264, P2P, web TV & streaming*. Focal Press, 2008.

[3] "The TDG website, available at: http://tdgresearch.com/."

[4] I. Sodagar, "The MPEG-DASH standard for multimedia streaming over the internet," *Multimedia, IEEE*, vol. 18, no. 4, pp. 62–67, April 2011.

[5] S. Lederer, C. Müller, and C. Timmerer, "Dynamic adaptive streaming over HTTP dataset," in *Proceedings of the 3rd Multimedia Systems Conference*. New York, NY, USA: ACM, February 2012, pp. 89–94.

[6] "ISO/IEC 23009-1, Information technology - Dynamic adaptive streaming over HTTP (DASH) - Part 1: Media presentation description and segment formats," 2012.

[7] *3GPP TS 26.234: Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs*, 3GPP Std. TS 26.234, 2010.

[8] M. Schwartz, *Telecommunication networks: protocols, modeling and analysis*. Addison-Wesley Reading, 1987, vol. 7.

[9] J. Y. B. Lee, *Scalable Continuous Media Streaming Systems: Architecture, Design, Analysis and Implementation*. John Wiley "&" Sons, Ltd, 2005.

[10] V. B. Iversen, *Teletraffic Engineering and Network Planning*. Lecture Notes, Technical University of Denmark, 2010.

[11] "The ITEC-DASH website, available at: http://www-itec.uni-klu.ac.at/dash/."

[12] V. Adzic, H. Kalva, and B. Furht, "Optimizing video encoding for adaptive streaming over HTTP," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 397–403, 2012.