

2017-08-31

Semantic-aware adaptation scheme for soccer video over MPEG-DASH

Hu, S

<http://hdl.handle.net/10026.1/14417>

10.1109/icme.2017.8019541

2017 IEEE International Conference on Multimedia and Expo (ICME)

IEEE

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

SEMANTIC-AWARE ADAPTATION SCHEME OVER MPEG-DASH

Shenghong Hu^{1,2,3}, Lingfen Sun³, Chunxia Xiao^{1,2*}, Chao Gui²

1. State Key Lab of Software Engineering, Wuhan University, Wuhan, Hubei 430072, China

2. Computer School, Wuhan University, Wuhan, Hubei 430072, China

3. Information Engineering School, Hubei University of Economics, Wuhan 430205, China

4. School of Computing and Mathematics, Plymouth University, Plymouth PL4 8AA, U.K.

wuhanhush@126.com, lingfen.sun@plymouth.ac.uk, cxxiao@whu.edu.cn, gui_chao@126.com

ABSTRACT

In recent years, quality of experience (QoE) has been investigated and proved to have both influential factors on user's visual quality and perceptual quality, while the perceptual quality means user's requirement on personalized content should be acquired in optimized quality. That's to say, those segments holding user interested content such as highlights need to be allocated more network resource in a resource-limited streaming scenario. However, all the existing HTTP-based adaptive methods only focus the content-agnostic bitrate adaptation according to limited network resources or energy resource, since they ignored user perceived semantics on some important segments, which suffered less quality on the important segments than on those ordinary ones, so as to hurt the overall QoE. In this paper, we have proposed a new semantic-aware adaptation scheme for MPEG-DASH services, which decides how to preserve bandwidth and buffering time depending on content descriptors for the perceived important content to users. Further, a semantic-aware probe and adaptation (SMA-PANDA) algorithm has been implemented in a DASH client to compare with conventional bitrate adaptations. Preliminary results show that SMA-PANDA achieves better QoE and flexibility on streaming user's interested content on MPEG-DASH platform, and it also aggressively helps user interested content compete more resource to deliver high quality presentation.

Index Terms—MPEG-DASH, quality of experience, content-aware, video personalization

1. INTRODUCTION

Nowadays, Internet streaming prefers the delivery of video service over HTTP adaptive streaming (HAS) for its performance of adaptive rate switching and free to firewall. Among the all available HAS standards, MPEG's Dynamic Adaptive Streaming over HTTP (MPEG-DASH) is the best one to handle varying bandwidth conditions during a streaming session^[1]. Most of rate adaptation based methods

are designed for DASH to decide how to probe the available bandwidth accurately, and switch or adapt efficiently to improve the fairness, efficiency and stability of network performance^[2-5], as well as trying to achieve satisfaction on end-users' QoE with subjective utility originated from video bit rate^[5-7]. In our previous work^[6], a content-aware adaptation scheme has been introduced to improve QoE for high-motion scenes, where the segments was allocated a certain proportion of buffered time depend on the motion rank to download the higher quality representation than the actual bandwidth. However, all of those for DASH streaming are irrespective of semantic features of the segmented video, and the adaptation decisions are employed to optimize just for the average quality of all scenes presented to viewers. This may scale down the perceived quality for the interesting scenes according to users' preference, and thus achieve less optimized QoE for the delivered video streaming services.

In this paper, we investigated the temporal structure of soccer video and take the shot as an adaptation unit ranked by semantic importance and motion activity. We further proposed an innovative content-aware QoE optimized adaptation decision, which encourages the highlights of soccer video, namely interesting content, to be allocated more resources and prevented from stalling and buffer starvation. We implemented an optimization algorithm of SMA-PANDA over the popular MPEG-DASH to evaluate our algorithm. Preliminary results show that content-aware adaptation decisions achieve better QoE when compared with conventional PANDA scheme and motion-aware PANDA scheme.

2. SEMANTIC ANALYSIS AND CONTENT DESCRIPTOR

When parsing a video program semantically, a scene is recognized as a complete story unit. A scene consists of a sequence of shots, and the shot is a set of frames with the similar features. Normally, a scene is ranked with different importance according to the story's ongoing tempo, and a shot is also ranked with different importance to the scene

according to the cameraman's purpose. In an encoded MPEG4/H.264 video stream, high motion content with the same QoE often results into a big amount of data size in the same duration^[6-7]. According to that, we think the shot is the most appropriate semantic unit for adaptation purpose. In this section, we introduce some work on detection of video highlights and rank the excitement of each streamed segment.

2.1. Semantic analysis for soccer video

We only need hard cut, and then divide the soccer video into a sequence of shots according to the color histogram based methods^[9]. The view types of the shots were recognized as long view, medium view, close-up view and out-field view by grassland ratio based methods^[10]. The sequence of shots was clustered into three types of normal-plays, highlights and replays guided by temporal producing structures as Fig. 1.

- Normal-play (NP): where a long view focuses on the field, or changes to a medium view or close-up only for demonstrating special information about the match.
- Highlight (HL): highlights encloses a user's excitement moment when goals, sudden shots or such semantic events occur. We automatically located highlights by analyzing excitement time curve aroused by affective features^[11].
- Replay (RP): We detected replays from producer-specific logos^[10], mostly they are inserted in the end of a highlight by the producer.

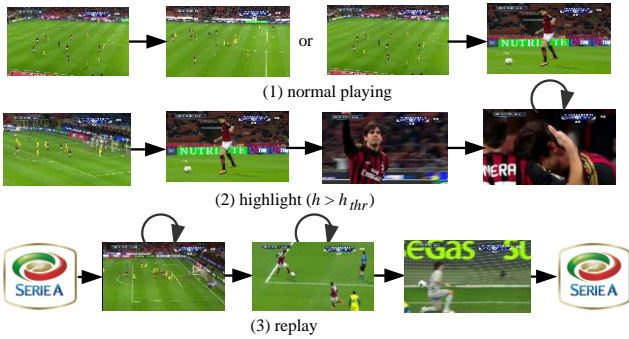


Fig. 1. Scene classification of soccer video.

2.1.1. Semantic importance of shots

As the production actions, the scenes and shots are hired to play both semantic utilities: one is for excitement, and the other is for information^[12]. As users viewing video programs, their perceptual demands also express infotainment, sometimes they request exciting content, and sometimes they only request information. Then, we can rank the importance of the content for different semantic purposes. Without loss of generality, we take the users' excitement as our optimized purpose.

- The long view demonstrates little excitement in a normal-play scene, but the medium and close-up views often hold special information about the game tempo.

- All the shots included in a highlight are the most exciting content except the out-field shots.
- A replay is not more exciting as a highlight, but it is more important than normal-play scene for it holds a lot of information about the highlight.
- All out-field shots are less important than the other shots in the same scene because they are often sidelights.

Assuming the semantic importance is 3 levels, the strategy to rank semantic importance (si_i) of each shot (sh_i) is denoted as:

$$si_i = \begin{cases} 1, & sh_i \in Normal-play \vee st_i = 3 \\ 2, & sh_i \in replay \vee sh_i \in Normal-play \wedge st_i \in \{2, 3\} \\ 3, & sh_i \in highlight \wedge st_i \leq 3 \end{cases} \quad (1)$$

where the view type of each shot is denoted as st_i , and its value is assigned as 1: close-up view, 2: medium view, 3: long view and 4: else.

2.1.2. Motion type of shots

Motion activity describes whether the intensity of activity, action, or motion is low, normal or high in the video content. The motion activity is evaluated by motion features derived from motion vector of macro-blocks. Suppose frame k is divided into $M \times N$ macro-blocks, where M and N are the numbers of macro-blocks in row and column. Let $mv_{mn}(k)$ denote the motion vectors of block (m, n) . Motion activity (ma_k) of the frame k is denoted as

$$ma_k = \frac{1}{M \cdot N} \cdot \sum_{m=1}^M \sum_{n=1}^N (|mv_{mn}(k)|) \quad (2)$$

If a shot is composed of K frames, its motion type is ranked by the mean and standard deviation of motion activity of all the frames it includes. We use k -mean cluster to classify the motion type (mt_i) of each shot into 3 types: slow moving ($mt_i = 1$), gentle walking ($mt_i = 2$), and rapid moving ($mt_i = 3$).

2.1.3. Decision to rank segments

Since a segment is not divided by the boundary of the shot, the segment should normally be extended to the shot which it belongs to. If the majority of the frames (f_0, \dots, f_{k-1}, f_k) in a segment (seg_n) belongs to a shot (sh_m), then that segment is regarded as belonging to that shot.

$$SID_n = m, \text{ for } \max_{m=1 \dots M} \left\{ \sum (f_k \in seg_n \wedge f_k \in sh_m) \right\} \quad (3)$$

where SID_n is the shot identification of seg_n , and then the semantic importance (ssi_n) and motion rank (sma_n) of seg_n will inherit from that shot it belongs to.

$$ssi_n = si_{SID_n}, \quad sma_n = ma_{SID_n} \quad (4)$$

2.2. Semantic markers for segments

In our approach of SMA-PANDA we want to show an easy and straightforward way of using DASH for high quality video streaming services. Hence, we require to allow DASH-compliant communication mechanisms between the

conventional clients and the servers. Therefore, the MPD file should be modified to include the content descriptors in the definition of each segment property. Since the content in each segment now are ranked by the vector $\langle SID, ssi, sma \rangle$, we add a property ("Marker='SID: ssi: sma'") of the segment element in a MPD file and update the function of MPD parser. The client can directly extract the content descriptor of each segment when downloading MPD file from the streaming server, then the adaptation engine makes an optimization decision on which quality level to download on condition of network status.

3. THE PROPOSED ADAPTATION SCHEME

In a MPEG DASH system, MPD file identifies the various content components and the location of all alternative streams. The conventional DASH player dynamically chooses appropriate segments to adapt current bandwidth and buffer status, and safe factor is used for determining how the segments consuming link capacity. Since we have extended the MPD file to hold a new marker on semantic ranks, most of the users' desired content can be downloaded intelligently, we should focus the optimization designing on the semantic-aware manners.

3.1. Designing Goals

The goal of content-aware adaptation was to asymmetrically improve the content quality according to the semantic importance while network resource is insufficient for all segments to be downloaded with the best quality. We identify the following optimization goals critically as:

- Segments with higher semantic importance should be optimized with higher utility to assure that users' expected scenes are performing at higher QoE under the limited resources.
- Each segment should be allocated resources associated with semantic and motion features. So to preserve available buffer time for highlights, or to assign the lowest safe factor to permit more bandwidth to them.
- To monitor buffer fullness and adjust the safe factors dynamically for different type of content in order to prevent from buffer under-runs, as high motion segments appear bigger data size on the average quality in a representation.
- Minimize the number of video quality oscillations, and try to keep constant quality in the same scene.

3.2. Essentials for online algorithm

In order to implement our designing goals for semantic content, the principal bit rate adaption named as PANDA^[3] are modified to optimize streaming quality according to content descriptors.

3.2.1. Bandwidth estimation

Current bandwidth (bw_n) is probed by the latest downloading, and smoothed to the historic bandwidth (bw_{n-1}) by

$$bw_n = \mu \cdot \frac{size(Seg_{n-1})}{download_time(Seg_{n-1})} + (1 - \mu) \cdot bw_{n-1} \quad (5)$$

3.2.2. Semantic and motion based resources management

Current buffer (B_n) fullness is measured by the number of remaining segments. In addition, we use lowest buffer (B_{low}) status to prevent under-run, and use upper buffer (B_{up}) to provide fairness or overrun limited memory. The safe factor (sf_n) and detected buffer status will be adjusted by semantic importance and motion activity of each segment as

$$\begin{aligned} & \text{if } (B_n \leq B_{low}) \{ repNo = 0 \} \\ & \text{else if } (B_n < \varepsilon \cdot B_{up}) \{ sf_n = \phi(sma_n), B_{avl} = \phi(ssi_n) \cdot B_n \} \\ & \text{else } \{ sf_n = \phi'(sma_n), B_{avl} = \phi'(ssi_n) \cdot B_n \} \end{aligned} \quad (6)$$

where $\phi(sma_n)$, $\phi(ssi_n)$, $\phi'(sma_n)$ and $\phi'(ssi_n)$ are the factors to regulate safe factor and buffer consumption according to different content features. Current buffer is lower than the lowest, we only need to switch to the representation with lowest quality ($repNo = 0$). The safe factor is adjusted by motion activity of the segment to be downloaded, while high motion segment will be assigned lower sf_n to consume more bandwidth. The available buffer time (B_{avl}) is also adjusted by semantic importance so that the highlights consumes majority of remaining buffers time to download those representations with better QoE but higher bitrate. If the buffer status is not so good ($B_n < \varepsilon \cdot B_{up}$), higher sf_n and lower B_{avl} should be served for downloading a segment.

3.2.3. Reconcile rate selection in a scene

To prevent bit rates from switching too often which may affect user's experience in a shot, the selected rate of this segment should be reconciled if it is in the same shot as the last segment:

$$repNo(Seg_{n+1} | SID_{n+1} = SID_n) = repNo(Seg_n) \quad (7)$$

3.2.4. Time to schedule next downloading

Finally, if the download process reaches fullness, the download time for next segment should be delayed by a distribution:

$$\text{if } (B_n > B_{up}) \text{ delay}_n \in [\tau - \gamma(ssi_n) \cdot \tau, \tau + \gamma(ssi_n) \cdot \tau] \quad (8)$$

where τ is the duration of each segment, it is often fixed to 1s, 2s, 4s, ..., 10s, 20s in a list of segments.

3.3. Real-time and semantic-aware PANDA for DASH

The PANDA was generally used for quality-oriented HTTP adaptive streaming. The original PANDA was designed as content-agnostic bit-rate adaptation but it worked well in a real-time streaming scenario. In this section, we extend it to incorporate QoE optimized adaptation with content features since the MPD file has content markers. With the probing part of essential resources, our online algorithm determines how to download the next segment depending on its maker value under current network status. Assuming Seg_n was the current segment being fetched, the semantic-aware probe and adapt (SMA-PANDA) was described in Algorithm 1.

Algorithm 1 Semantic-aware PANDA for DASH

At the beginning of each downloading of segment n :

- 1) whether Seg_n and Seg_{n-1} belong to the same shot
 if $SID_n = SID_{n-1} \wedge bw_n \geq bw_{n-1}$, goto step5
 else $\{count(SID_{n+1} = SID_n) \text{ as } K \text{ until } (SID_{n+K} \neq SID_n), \text{ goto step2}\}$ (9)
- 2) probe the bandwidth as equation (5)
- 3) probe the buffer status, then adjust available buffer and safety factor to prevent stalling as equation (6)
- 4) If there are M representations with bitrate ordered by increasing from $br[0], \dots, br[M-1]$, assume all the left segments of shot SID_n to be served the same bitrate, maximize the representation index by
 $repNo_n = m, \text{ for } \max(m) \wedge K \cdot br[m] \cdot sf_n \leq (B_{avl} + K - B_{low}) \cdot bw_n$ (10)
- 5) Schedule the next downloading with the $delay_n$ computed as equation (8) on condition of buffer fullness:

$$TimeToDownload = \begin{cases} 0, & B_n < B_{up} \\ delay_n, & \text{otherwise} \end{cases} \quad (11)$$

4. EXPERIMENTS & EVALUATIONS

To evaluate the performance of the proposed algorithms in real-world scenarios, we have implemented the C-PANDA in player prototypes included in libdash3.0 library [14]. The win2003 server with IIS 6.0 and win7 clients are running on virtual PC installed in a VMware 9.0 workstation. The bandwidth of adapter of a server can be changed when the server is running. This enables us to evaluate the performance of the online adaptation algorithms. We selected several clips from the *Serie A 2014* and *World Cup 2010*, and both of them were the most famous soccer matches. One game was *AC Milan vs Chievo*, and the other was *Brazil vs North Korea*. They are all 3 minutes long with 4500 frames, and were encoded into six representations of 256kbps, 384kbps, 512kbps, 768kbps, 1Mbps and 1.34Mbps with the resolution of 640x360 and 25fps. The encoded videos were then transferred into 2-second segment MPD manifest file with “marker” properties. Both the two MPD files include 100 segments. The marker values of all segments of the two videos were illustrated in Fig. 2 & 3.

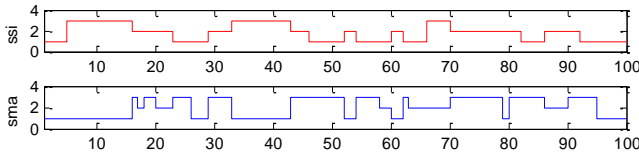


Fig. 2. Marker values of *AC Milan vs Chievo*.

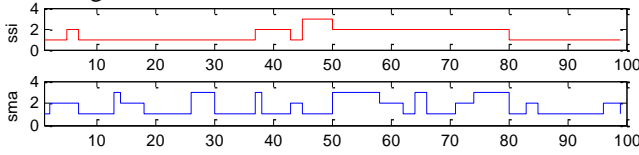


Fig. 3. Marker values of *Brazil vs North Korea*.

The parameters used in above algorithms are set as the default values as listed in Table 1. In the previous work, we have introduced a content-aware adaptation scheme with only motion activity [6], which proved that efficient resource planning on different motion content could improve the overall QoE. Here we name the motion-aware adaption algorithm as MA-PANDA. We compare the performance of our solution both with the content-agnostic PANDA algorithm and MA-PANDA algorithm.

Table 1. Default client parameters

Parameters	Default values
$\mu, \varepsilon, B_{low}, B_{up}, \tau$	0.85, 0.5, 3, 20, 2s
$\phi(sma_n=1), \phi(sma_n=2), \phi(sma_n=3)$	1.6, 1.4, 1.1
$\phi(ssi_n=1), \phi(ssi_n=2), \phi(ssi_n=3)$	0, 0.5, 0.8
$\phi'(sma_n=1), \phi'(sma_n=2), \phi'(sma_n=3)$	1.3, 1.2, 1.1
$\phi'(ssi_n=1), \phi'(ssi_n=2), \phi'(ssi_n=3)$	0.4, 0.6, 0.9
$\gamma(ssi_n=1), \gamma(ssi_n=2), \gamma(ssi_n=3)$	1/3, 1/6, 0

We have considered two scenarios shown in Fig. 4: (1) The video of *AC Milan vs Chievo* is streamed into a bottleneck of 0.6Mbps/1Mbps separately with our proposed C-PANDA and conventional PANDA, and this scenario is aimed at showing the dynamic behavior on content-aware QoE optimization and network performance of the proposed algorithms. (2) Two connections sharing a broad-wave like link whose available bandwidth vary from 1Mbps to 2Mbps, and this scenario is aimed at showing the competition ability between different semantic content in SMA-PANDA.

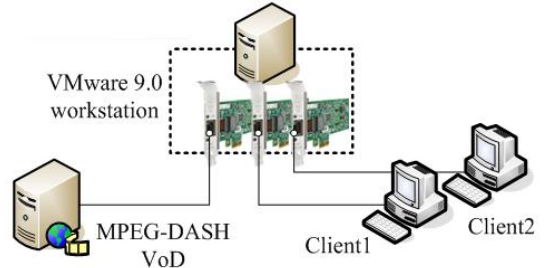


Fig. 4. The evaluation scenario.

4.1. Video streaming into a bottleneck

Fig. 5 and Fig.6 illustrate how PANDA and C-PANDA select different representations under bandwidth of 0.6 and 1.0 Mbps. The conventional PANDA selects the proper quality representation compared to the historical throughput, MA-PANDA selects better quality representation for the high motion content (e.g. segments 15-20 and segments 50-58), and SMA-PANDA selects better quality representation for the highlight content (e.g. segments 30-35 and segments 70-90). But PANDA just loses good quality on some highlight segments, as it won't reserve resources for them, and PANDA even wins higher representation for the normal scenes than the other algorithms. PANDA also varies very often on representation selection especially in a high

bandwidth environment, and it is regarded to hurt users' experience. Because the MA-PANDA plays adaptations on a shot which has been known as a smaller semantic unit than the highlight, then MA-PANDA results more fluctuations on rate-switching than that in SMA-PANDA. On the other hand, SMA-PANDA performs better rate-switching in a high bandwidth environment as Fig.6 for there was sufficient resource that can be planned for the highlight.

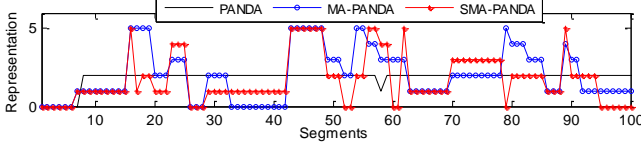


Fig. 5. Representation selection of *ACMilan vs Chievo* under 600kbps.

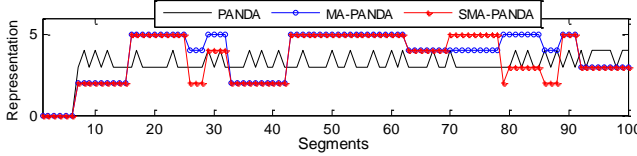


Fig. 6. Representation selection of *ACMilan vs Chievo* under 1Mbps.

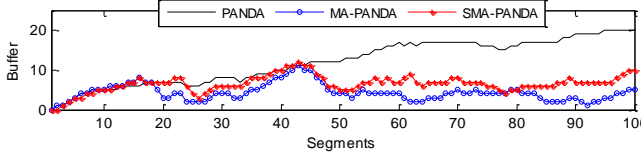


Fig. 7. Buffer status of *ACMilan vs Chievo* under 600kbps.

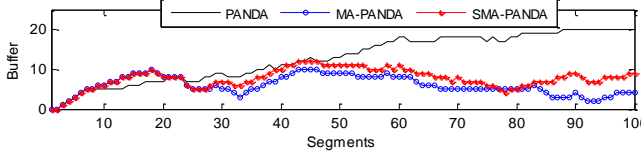


Fig. 8. Buffer status of *ACMilan vs Chievo* under 1Mbps.

Fig. 7 and Fig.8 shows results for buffer consumption status for the three adaptation algorithms. PANDA maintains a certain portion of buffer occupancy to prevent rebuffering after the streaming enters the stable situation. MA-PANDA and SMA-PANDA attempts to exploit the buffer occupancy on the segments according to the content features, so that the reserved resource ensures the streaming segments to cover the radical resource requirements of highlights. However, the conventional PANDA just maintains a certain amount of buffer time best effort without any consideration to consume for the important scenes. This will normally reduce the efficiency on resource management and pull down the overall QoE of the streamed video. We also notice that in all the runs no buffer under-runs occurred in streaming by all the three algorithms, while the MA-PANDA was more aggressive on buffer consumption than SMA-PANDA, for the high motion content is also included in normal scenes and the actual requirement of highlight will be suppressed.

We used average Signal-to-Noise Ratio (PSNR) and MOS (Mean Opinion Score) converted as [15] to compare average visual quality on a segment-to-segment manner.

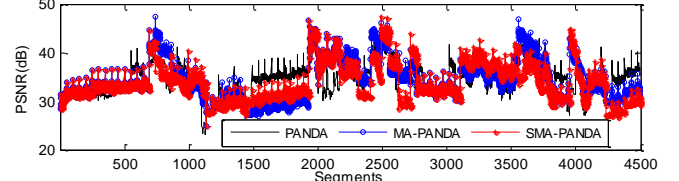


Fig. 9. PSNR values of each segment in *ACMilan vs Chievo* under 600kbps.

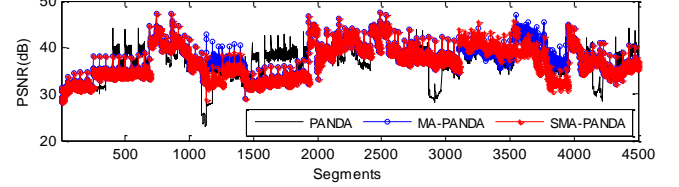


Fig. 10. PSNR values of each segment in *ACMilan vs Chievo* under 1Mbps.

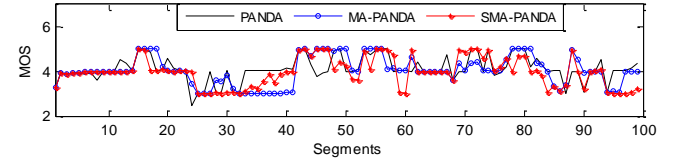


Fig. 11. MOS values of each segment in *ACMilan vs Chievo* under 600kbps.

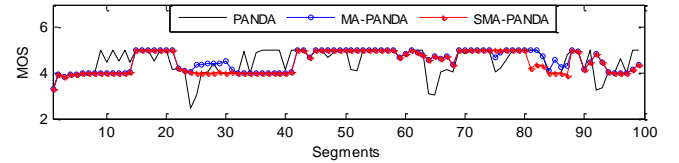


Fig. 12. MOS values of each segment in *ACMilan vs Chievo* under 1Mbps.

In Fig. 9 and Fig.10, we can see that the conventional PANDA only obtain better PSNR quality on lower motion content than MA-PANDA and SMA-PANDA (e.g. segments 6-15 and segments 33-40 in Fig. 9 and Fig.10), and the fluctuation of objective quality of PANDA streaming is higher than MA-PANDA and SMA-PANDA. While it achieves less MOS quality on higher motion content than SMA-PANDA, for it consumes too much resource on normal scenes that achieves less improvement on highlight. SMA-PANDA reserves more resources for highlight segments so that the minimum quality is maximized and there is no any segment fell to unaccepted QoE (MOS < 3.5) in Fig. 11 and Fig.12. For highlight scenes, C-PANDA achieves higher MOS values when compared to the PANDA (e.g. segments 35-45 and segments 40-50 in Fig. 11 and Fig.12), and the amplitude of quality switching with PANDA is higher than both MA-PANDA and SMA-PANDA. Even MA-PANDA achieves higher MOS values on high motion content than the other two algorithms, SMA-PANDA achieves stable MOS values

compared to MA-PANDA in lower bandwidth as Fig.11, for MA-PANDA is more sensitive to bandwidth fluctuations.

4.2. Two videos share broad-wave like link

In the shared broad-wave link, streaming of *Brazil vs North Korea* start 3 seconds earlier than the video of *ACMilan vs Chievo*. As shown in Fig. 13, the two streams compete with each other for the limited bandwidth, and they share available link according to the semantic importance after segment 10. The highlight segments will win the competition and achieve more bandwidth than those who are not (e.g. segments 30-33 and segments 80-92 of *ACMilan vs Chievo*; segments 50-58 of *Brazil vs North Korea*).

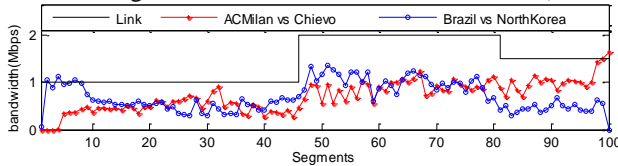


Fig. 13. Bandwidth sharing of the two streams.

5. CONCLUSIONS

In this work, we have proposed content features working on both semantic level and signal level for QoE optimized MPEG-DASH streaming. The performance of the SMA-PANDA was compared with the conventional PANDA algorithm and MA-PANDA, in terms of representation selection, buffer consumption, and video quality based on PSNR and QoE. Preliminary results show that semantic-aware adaptation schemes are able to adapt video bit rate based on both network bandwidth and video semantics, the highlight proves better quality than those normal content, and achieves stable representation switching and MOS variations in the dynamic streaming. The schemes are able to reserve buffer resources from low importance segments when streaming into a bottleneck and enhance the opportunities of high importance segments winning more bandwidth in the sharing link. Considering that users' personalized needs, the content-aware adaptation schemes achieve higher QoE representations for more important content. Besides the soccer video, this work also encourages a roadmap to design customized DASH applications for delivery of surveillance video^[16] and remote sensing video^[17] in the future.

Acknowledgement

This work was partly supported by the NSFC (No. 61472288, 61672390, 61572012), NCET (NCET-13-0441), the Fundamental Research Funds for the Central Universities (2042015kf0181) and the State Key Lab of Software Engineering (SKLSE-2015-A-05). Foundation of Key Research Institute of Humanities and Social Science at Universities, Chinese Ministry of Education (16JJD870002). The Key Natural Science Foundation of Hubei Province of China (No. 2014CFA055). Chunxia Xiao is the corresponding author.

6. REFERENCES

- [1] T. Christian and M. Christopher, "HTTP Streaming of MPEG Media", In Proceedings of the Streaming Day, Udine, Italy, September 16-17, 2010.
- [2] C. Liu, B. Imed and G. Moncef. "Rate adaptation for adaptive HTTP streaming," MMSys '11, San Jose, California, USA, pp.169-174, Feb. 2011
- [3] Z. Li, X. Zhu, J. Gahm, et al. "Probe and Adapt: Rate Adaptation for HTTP Video Streaming At Scale," IEEE Journal on Selected Areas in Communications, vol. 32, no. 4, pp.719-733, Apr. 2014.
- [4] J. M. Batalla, P. Krawiec, A. Beben, P. Wisniewski, "Adaptive Video Streaming: Rate and Buffer on the Track of Minimum Rebuffering," IEEE Journal on Selected Areas in Communications, vol.34, no.8, pp.1-1, June 2016.
- [5] J. Junchen, S. Vyas. and Z. Hui. "Improving Fairness, Efficiency, and Stability in HTTP-Based Adaptive Video Streaming with Festive," CoNEXT '12, pp.97-108, Dec. 2012.
- [6] S. Hu, L. Sun, C. Gui, E. Jammeh and I-H. Mkwawa, "Content-aware Adaptation Scheme for QoE Optimized DASH Applications," IEEE GLOBECOM 2014, Austin, TX, USA, Dec. 2014.
- [7] K. Asiya and L. Sun, Emmanuel I. "QoE Prediction Model and its Application in Video Quality Adaptation over UMTS Networks," IEEE Transaction on Multimedia, vol.14, no.2, pp.431-442, Apr. 2012.
- [8] S. Tavakoli, S. Egger, M. Seufert, R. Schatz, K. Brunnström and N. García, "Perceptual Quality of HTTP Adaptive Streaming Strategies: Cross-Experimental Analysis of Multi-Laboratory and Crowdsourced Subjective Studies," IEEE Journal on Selected Areas in Communications, vol. 34, no.8, pp. 2141-2153, June 2016.
- [9] J. Mas, G. Fernandez, "Video Shot Boundary Detection based on Color Histogram," Notebook Papers TRECVID, 2003.
- [10] Z. Wang, J. Yu, Y. He and T. Guan. "Affection arousal based highlight extraction for soccer video," Multimedia Tools and Applications, vol.73, no.1, pp.519-546, Nov. 2014.
- [11] A. Hanjalic, "Adaptive extraction of highlights from a sport video based on excitement modeling," IEEE Transaction on Multimedia, vol.7, no.6, pp.1114-1122, 2005.
- [12] G. Gheorghita and P. Johnson. "Quality of Perception: User Quality of Service in Multimedia Presentations," IEEE Transaction on Multimedia, vol.7, no.4, pp.786-789, 2005.
- [13] S. F. Chang, A. Vetro. "Video adaptation: Concepts, technologies and open issues," Proceedings of the IEEE, vol.93, no.1, pp.148-158, 2005.
- [14] libdash, online: <http://www.bitmovin.net/libdash.html>
- [15] J. Klaue, B. Tathke and A. Wolisz. "Evalvid-A framework for video transmission and quality evaluation," In Proc. Of the 13th International Conference on Modeling Techniques and Tools for Computer Performance Evaluation, Urbana, Illinois, USA, pp.255-272, 2003.
- [16] Y. Nie, C. Xiao, H. Sun et al. "Compact video synopsis via global spatiotemporal optimization", IEEE Transactions on Visualization & Computer Graphics, 2013, 19(10):1664-76.
- [17] Q. Zhang, C. Xiao. "Cloud Detection of RGB Color Aerial Photographs by Progressive Refinement Scheme", Geoscience & Remote Sensing IEEE Transactions on, 2014, 52(11): 7264-7275.