QUALITY-FAIR HTTP ADAPTIVE STREAMING OVER LTE NETWORK

Sergio Cicalò[⋆†], *Nesrine Changuel*[†], *Ray Miller*[‡], *Bessem Sayadi*[†], *Velio Tralli*[⋆]

[⋆]Engineering Department of the University of Ferrara (ENDIF), CNIT, Via Saragat 1 - Italy
[†] Alcatel-Lucent Bell Labs, Route de Villejuste, Nozay, France
[‡] Alcatel-Lucent Bell Labs, 600 Mountain Avenue, Murray Hill, USA

ABSTRACT

In HTTP adaptive streaming (HAS) applications multiple video clients sharing the same wireless channel may experience different video qualities as result of both different video content complexity and different channel conditions. This causes unfairness in the end-user video quality. In this paper, we propose a quality-fair adaptive streaming (QFAS) solution to deliver fair video quality to HAS clients competing for the same resources in an LTE cell. In the QFAS framework the share of radio resource is optimized according to video content characteristics and channel condition. The proposed solution is compared with other state-of-the-art strategies and numerical results in terms of SSIM quality metric shows that it significantly improves the quality fairness among heterogeneous HAS users.

Index Terms— HAS, LTE, Mobile networks, Fairness.

1. INTRODUCTION

Today, we are facing an explosion of the video traffic on wireless network due to the proliferation of multimedia-friendly portable devices [1]. In addition, the emergence of high speed networks provides the infrastructure and the possibility for handling a wide set of new applications among which the multimedia contents delivery. In fact, long term evolution (LTE) [2] substantially improves end-user throughput, sector capacity and reduces user plane latency, thus enabling significantly improved user experience. LTE supports different types of services including web browsing, video streaming, VoIP, online gaming, real-time video, etc. [3] with standardized quality class indicators (QCI) [4]. Each QCI defines a set of requirements, *i.e.*, maximum tolerable delay, packet loss rate and/or guaranteed bit-rate (GBR).

Nowadays much of the video traffic is transmitted over HTTP protocol. A new approach referred to as HTTP adaptive streaming (HAS) [5] is becoming popular. HAS is adaptive in the sense that it allows a client to adaptively switch between multiple bitrates, depending on the bandwidth or data rate available between the server and the client. This is a particularly useful feature for a wireless environment since the data rate of the wireless link can vary over time. Based on TCP, one of the objectives of HAS is keeping the fairness among multiple homogeneous/heterogeneous connections in the network. In fact, fair share of network resources among multiple heterogeneous connections is one of key issues especially for the commercial use of the Internet [6]. However, fair rate may result in unfair received video quality especially when the transmitted videos have different content characteristics. Thus, bringing intelligence into the network is needed to improve the resource allocation, so that to avoid quality fluctuations over time and provide comparable received quality among video flows. When multiple videos are delivered through the same base station, this can be done through content-aware resource allocation as in [7].

Multi-user HAS video delivery optimization has attracted increasingly attention in the last few years. In [8], an overview of the recently standardized quality metrics for HAS and an end-to-end evaluation study has been presented. They concluded that network-level and radio-level adaptation is required for enhancing service capacity and user perceived quality. Recently, Authors in [9] propose an efficient method to optimally and adaptively set up the GBR of each video flow in a LTE network with heterogeneous traffic. The approach is intended to achieve a level of fairness among the video flows while preventing starvation of other data flows. A similar framework was presented in [10], which also aims at improving stability and resource utilization, but without considering heterogeneous traffics. In both frameworks the definition of the utilities is not content-aware and may not lead to the best possible quality fairness among the video flows. To the best of our knowledge only [11] investigated a content-aware multi-user HAS video delivery framework in LTE networks. Similarly to here, a media-aware network element (MANE) is in charge of selecting the streaming rate required by each client in order to maximize the aggregate video utilities under resource constraint. However, differently from here, video quality fairness is not considered and the peculiarities of pull-based delivery strategy of HAS technology are not taken into account.

In this paper, we propose a quality-fair adaptive streaming (QFAS) solution to deliver fair video quality to HAS clients competing for the same resources in an LTE cell. Using a similar mechanism as [9], the QFAS solution brings intelligence into the network to adaptively select the prescribed GBR of each UE according to the contents characteristics in addition to the channel condition. Such GBR values are derived by solving an optimization problem aimed at maximizing the aggregate video utility under minimum and maximum rate constraints, available resource, and quality-fair constraint across multiple video clients. Numerical evaluations resulting from extensive and detailed *ns2* simulations show that QFAS solution provides significant improvement to the quality received by the end-users demanding more complex video, even when they are experiencing bad channel condition, with a tolerable degradation of the other low-complexity videos. The quality fairness is thus well improved among heterogeneous clients compared to best effort and AGBR approaches.

This paper is organized as follows. Section 2 introduces the system model. In Section 3 we formulate and solve the proposed optimization problem. The performance of the proposed scheme is evaluated in Section 4 and conclusions are drawn in Section 5.

978-1-4799-2893-4/14/$31.00

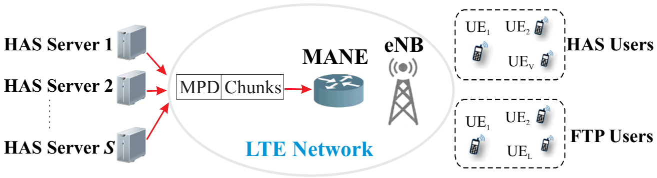

Fig. 1. System Architecture.

2. SYSTEM MODEL AND ASSUMPTIONS

As depicted in Fig. 1, we consider an LTE wireless access network serving a total of K UEs (User Equipments), subdivided in V HAS users indexed by the set $\mathcal{V} = \{1, \ldots, V\}$, and L data users. One or more HAS servers encode the video sequences at multiple bit-rates and, after segmentation, generate a manifest file, also named media presentation descriptor (MPD). We assume that each HAS server extracts synthetic quality information from each segment (also called chunk in the following) and inserts them in the MPD. A MANE, located close to the e-NodeB (eNB), is able to intercept and process the MPD requested by each HAS client in order to get rate and quality information. The eNB allocates the available resources according to a general proportional fair scheduler with minimum bit-rate, *i.e.*, GBR, and maximum bit-rate (MBR) constraints, which are dynamically updated by the MANE.

In the following we omit the index of the client and we details the adaptation process for a single client-server link in an ideal scenario as illustrated in Fig. 2. Let $\mathcal{R}$ be the set of M available rate profiles $r_m, m = 1, \ldots, M$, listed in the MPD and assume that the client is able to follow the GBR provided by the eNB. This means that once a chunk, *i.e.*, chunk $(n-1)$ in Fig. 2, is received, the rate decision algorithm (RDA) at the client completes the measurement of the chunk download rate $\hat{R}[n-1]$ and requests chunk n with a profile rate $r^*[n] = \max_{r_m \leq \hat{R}[n-1]} r_m$, according to a pull-based approach.

When the MANE intercepts the request, it collects the channel state information (CSI) from the eNB and updates the GBR value $R[n]$ that the scheduler at the eNB will use to send chunk n. In case of ideal rate measurement, which requires that the channel state information do not vary significantly between time instant n and time instant $(n+1)$, we would have $R[n] = \hat{R}[n]$, *i.e.*, the GBR value $R[n]$ is then used for the rate request of chunk $(n+1)$. Thus, to avoid mismatch due to video characteristic changing over time, the GBR value $R[n]$ is computed based on the video utility of the chunk $(n+1)$. More details are provided in the next Section.

3. OPTIMIZATION PROBLEM AND SOLUTIONS

The objective of our quality-based approach is to derive the rate which allows to maximize the overall video quality under quality fairness constraint and according to UEs channel condition. Let n be the chunk index requested by UE k, we define U_k as the *utility* of requesting chunk $(n+1)$ in terms of video quality metric. The following parametric rate-utility model is used to describe the evolution of the utility U_k as a function of the rate R_k:

$$U_k = f(\mathbf{a}_k, R_k), \tag{1}$$

where $\mathbf{a}_k \in \mathcal{A} \subset \mathbb{R}^{N_a}$ is a time-varying and content-dependent parameter vector. For all values of $\mathbf{a}$ belonging to the set of admissible parameter values $\mathcal{A}$, $f(\mathbf{a}_k, R)$ is assumed to be a continuous, invertible and strictly increasing function of R. The model (1) may represent the variation of the PSNR, the SSIM, or any other strictly increasing quality metric as function of the encoding rate [12].

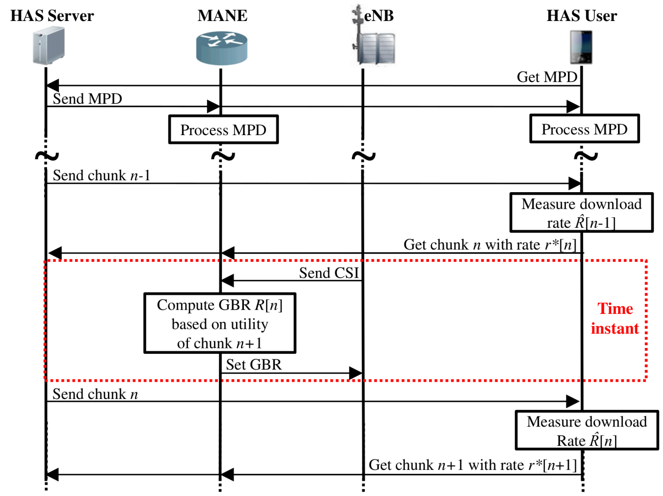

Fig. 2. Proposed approach for a single HAS video delivery optimization.

Following the approach in [9, 11], we consider a simplified air interface model where the maximum achievable rate for each UE is estimated according to its average channel condition. Let γ_k be the average signal-to-noise plus interference ratio (SNIR) experienced by UE k. As in [9], we define $w_k = [\log_2(1+\gamma_k)]^{-1}$ as the inverse of the estimated average rate per unit bandwidth. The optimization problem is then stated as follows:

$$\max \sum_{k \in \mathcal{V}} f(\mathbf{a}_k, R_k) \tag{2a}$$

$$s.t.\ A_k \leq R_k \leq B_k, \forall k \in \mathcal{V} \tag{2b}$$

$$\sum_{k \in \mathcal{V}} w_k R_k \leq \Pi \tag{2c}$$

$$\Delta(U_i, U_j) = 0 \quad \forall i, j \in \mathcal{V}, i \neq j \tag{2d}$$

where A_k, B_k are the minimum and maximum rates from the MPD of the video requested by UE k. The value of Π defines the amount of resources dedicated to the HAS UEs, which can be statically configured or dynamically computed at each time transmission interval (TTI) based on number of UEs and scaling factors [9].

The utility-fairness metric in the constraint (2d) is defined as:

$$\Delta(U_i, U_j) = \begin{cases} 0 & \text{if } U_i = f(\mathbf{a}_i, A_i) \wedge U_j < U_i \\ 0 & \text{if } U_i = f(\mathbf{a}_i, B_i) \wedge U_j > U_i \\ |U_i - U_j| & \text{otherwise.} \end{cases} \tag{3}$$

This metric was introduced in [13] in terms of video distortion and extends the simple fairness metric $|U_i - U_j|$ to the case where R_i, R_j are constrained to their minimum and maximum values. In fact, in presence of rate constraints, if a video achieves its maximum utility, it is reasonable to use the available resources to increase the utilities of other videos. On the other hand, in a case of scarce resources, if decreasing the rate of the i-th video is not possible since its minimum utility value has been already reached, it is necessary to decrease the rate of the other videos, at the price of decreasing the related utility.

The optimization problem in (2) admits a feasible solution under the condition $\sum_{k\in\mathcal{V}} w_k A_k \leq \Pi$. By considering the trivial condition $\sum_{k\in\mathcal{V}} w_k B_k \geq \Pi$, it has been proved in [13] that the problem (2) collapses in a constraint-satisfaction problem where the objective is achieved by fulfilling constraint (2c) with an equality constraint. Optimal solution can be derived by relaxing constraint (2b) with two boolean variables and applying a procedure with quadratic complexity in the worst case. More specifically, let $x_k, y_k \in \{0,1\}$, $k \in \mathcal{V}$, with $(x_k, y_k) \neq (0,0)$, be binary variables that indicate whether (1) or not (0) the two constraints $R_k \geq A_k$ and $R_k \leq B_k$, respectively, are satisfied. We define the function

$$\Gamma(\boldsymbol{x},\boldsymbol{y},U) = \sum_{k\in\mathcal{V}} x_k y_k w_k f^{-1}(\mathbf{a}_k, U) - \Pi(\boldsymbol{x},\boldsymbol{y}) \qquad (4)$$

where

$$\Pi(\boldsymbol{x},\boldsymbol{y}) = \Pi - \sum_{k\in\mathcal{V}} w_k \big[(1-x_k)A_k + (1-y_k)B_k\big], \qquad (5)$$

and f^{-1} is the inverse function of f. Since $f(\boldsymbol{a}, R)$ is a continuous and strictly increasing function of R, $f^{-1}(\boldsymbol{a}, U)$ is continuous and strictly increasing function of U. The algorithm can be obtained as described below and is derived by following the methods applied to the problem presented in [13], which has a similar structure.

Algorithm 1 Pseudo code to solve problem (2)

if $\sum_{k\in\mathcal{V}} w_k A_k \geq \Pi$ **then**
    report infeasibility
else if $\sum_{k\in\mathcal{V}} w_k B_k \leq \Pi$ **then**
    Set $\tilde{R}_k = B_k, \forall k \in \mathcal{V}$
else
    $y_k = 1, \forall\, k \in \mathcal{V}$;
    repeat
        $cond_{\mathrm{B}} =$ false; $x_k = 1, \forall\, k \in \mathcal{V}$;
        repeat
            $cond_{\mathrm{A}} =$ false;
            Compute $\tilde{U} : \Gamma(\boldsymbol{x},\boldsymbol{y},\tilde{U}) = 0$;
            for all $k \in \mathcal{V} : x_k y_k = 1$ **do**
                $\tilde{R}_k = f^{-1}(\boldsymbol{a}_k, \tilde{U})$;
                if $\tilde{R}_k < A_k$ **then**
                    $\tilde{R}_k = A_k$; $x_k = 0$; $cond_{\mathrm{A}} =$ true;
                end if
            end for
        until $cond_{\mathrm{A}}$ is false
        for all $k \in \mathcal{V} : x_k y_k = 1$ **do**
            if $\tilde{R}_k > B_k$ **then**
                $\tilde{R}_k = B_k$; $y_k = 0$; $cond_{\mathrm{B}} =$ true;
            end if
        end for
    until $cond_{\mathrm{B}}$ is false
end if

4. NUMERICAL RESULTS

We consider 4 video sequences extracted from real time programs, *i.e.*, *Interview*, *Sport*, *Big Buck Bunny* and *Home*, in 4CIF format at 30 fps, each one comprising 2800 frames. The sequences are looped 10 times and encoded through DASH Encoder [14] with 10 profiles with rate ranging from 150 *kbps* to 5 *Mbps*. Chunk duration is set to 2 seconds. We have considered both peak-SNR (PSNR) and Structural SIMilarity (SSIM) metric [15] to assess the video quality. Due to the lack of space, we here provide results only in terms of SSIM. Specifically, to model the dependency between the utility (here SSIM) and the rate, we consider a logarithmic SSIM to rate continuous utility in the interval of interest $[A_k, B_k]$, *i.e.*,

$$U_k = f(\mathbf{a}_k, R_k) = a_1 \log(a_2 R_k + a_3) \qquad (6)$$

where the parameters a_1, a_2, a_3 depends on the spatial and temporal complexity of each chunk and are derived through curve-fitting over the actual discrete empirical points. The validation results of the model (6) have shown almost perfect correlation with a Pearson coefficient always higher than 0.99 for each chunk of the considered video sequences. The parameters values of the SSIM-Rate model are derived off-line and inserted in each MPD as optional information.

Scen.	Average Rate [Mbps]				Average SNIR [dB]			
	Traffic	BE	AGBR	QFAS	*Bunny*	*Home*	*Interv.*	*Sport*
A	FTP	1.19	0.93	0.94	8.1	16.8	10.2	19.6
	HAS	0.76	1.11	1.26				
B	FTP	1.13	0.94	0.97	16.8	8.1	19.6	10.2
	HAS	0.73	1.11	0.95				

Table 1. Average SNIRs in dB of HAS clients and resulting average rates for FTP and HAS users.

Simulations are carried out on the *ns2*-platform which includes HAS servers and clients, LTE radio interface and radio resource management as well as the different protocol layers (TCP/IP, PDCP and RLC). Specifically, we consider a single cell where a total of 20 UEs (L=16 FTP UEs and V = 4 HAS UEs) are uniformly distributed in one cell with a radius of 1 km. The system bandwidth is set to 10 MHz. The radio channel is modeled according to the ITU extended pedestrian A model [16] and UEs are also affected by log-normal shadowing (std. deviation: 8 dB) with an exponential auto-correlation (correlation distance: 100 m). Each HAS client requests one of the video sequences mentioned above. The amount of available resources Π dedicated to HAS UEs is derived according to the on-line implementation proposed in [9]. A maximum receiver buffer of 40 s is considered allowing to absorb the possible mismatch between the rate at which the chunk is encoded and the rate actually resulting after transmission. The RDA at the client requests a chunk immediately after the previous chunk is downloaded, and the LTE scheduler sets for each client the MBR equal to the GBR, thus allowing each HAS client to fairly estimate the bandwidth [17].

In order to better assess the goodness of the proposed framework we investigate two different scenarios; in the first scenario (A) UEs requesting high complex videos, *i.e.*, *Home* and *Sport*, are in good channel conditions while in the second (B) such UEs experience bad channel conditions. We compare QFAS with the two following approaches: (i) best effort (BE), where all UEs are non-GBR (QCI equal to 9) [4]; (ii) AGBR approach [9] where the GBR values are updated every 2 seconds for each HAS UE (QCI equal to 4).

Table 1 reports the average SNIRs experienced by each client as well as the average MAC rate provided by the three approaches in the two scenarios. Fig. 3 shows the received chunk-by-chunk SSIM at the client in the first scenario (A) for each strategy, while Table 2 reports the overall average and the standard deviation of the SSIM. We can note how AGBR approach allows to increase the quality of the high complex video with respect to BE approach by increasing the average rate provided to HAS clients. However, both approaches (BE and AGBR) experience less than ideal quality fairness with a standard deviation of the SSIM at the client higher than

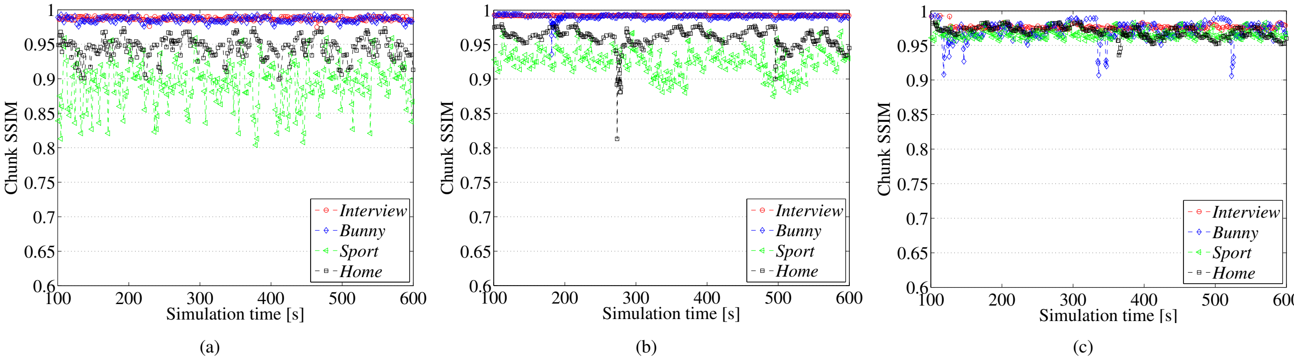

Fig. 3. Chunk-by-chunk SSIM at the client for scenario (A) resulting from BE (a), AGBR (b) and the proposed QFAS (c).

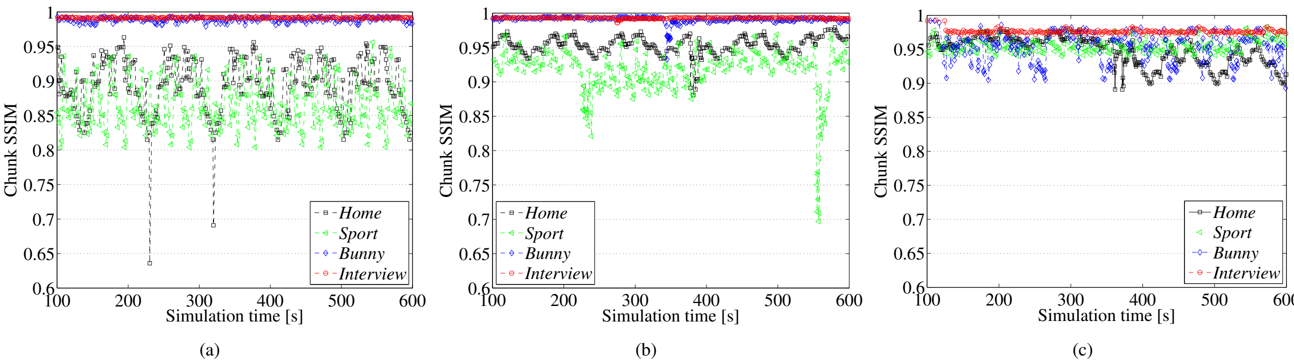

Fig. 4. Chunk-by-chunk SSIM at the client for scenario (B) resulting from BE (a), AGBR (b) and the proposed QFAS (c).

Scen.	SSIM	BE	AGBR	QFAS
A	Average	0.946	0.966	0.971
	Std. Dev.	0.042	0.031	0.005
B	Average	0.934	0.960	0.959
	Std. Dev.	0.063	0.045	0.012

Table 2. Overall average and standard deviation of the SSIM at the clients for two different scenarios resulting from BE, AGBR and the proposed QFAS.

0.031. Moreover, high quality fluctuations are experienced by users requesting high-complexity video, although they are in good channel condition. Our proposed QFAS strategy allows to significantly increase the quality of the high-complexity video up to 0.058 in average SSIM and 0.011 in overall SSIM, while keeping reasonable high quality to the low-complexity ones. However some quality drops in QFAS are still experienced by the *Bunny* client due to the gap between the limited number of available rate profiles and the continuous utility. The benefits of our QFAS approach are also significant in the scenario B, as showed in Fig. 4. Due to the unfavorable channel condition, BE causes intolerable average and instantaneous quality degradation. AGBR improves video quality over BE. However, compared to QFAS, AGBR still results in lesser quality to *Sport*, exhibits higher fluctuations, and provides unfair quality in contrast to *Interview*. QFAS better distributes the available resources, as confirmed by the resulting average rates in Table 1, by providing a chunk SSIM equal or higher than 0.9 to all video programs which ensure a good quality level for all users. As reported in Table 2, similar overall SSIM are provided by both approaches but the fairness in terms of standard deviation is highly improved. We also verified that both AGBR and QFAS approaches maintains similar buffer stability at the client. However, some drop in quality experienced by *Home* client at time index 370 s in QFAS and *Sport* client in AGBR are due to the RDA at the client, which is selecting the minimum rate profile to prevent buffer underflow.

5. CONCLUSIONS

In this paper, a quality-oriented optimized video delivery framework called QFAS is proposed for HAS clients competing for radio resources in the same LTE cell. By adding intelligence in the network, *i.e.*, through the use of a MANE, the proposed approach is able to control the rate provided to each HAS user in order to obtain fair video quality among multiple HAS clients. This is achieved even when HAS users are requesting programs with significant differences in video complexity and are experiencing different channel conditions. Numerical results have shown that, compared to other state-of-the-art approaches, our proposed QFAS solution provides significant improvement in the overall quality delivered to users demanding complex video with a tolerable degradation of the other low-complex videos. However, some quality fluctuations dependent on the RDA at the client, are still present. Future works will consider the possibility of optimizing the client rate request, in order to improve video quality fairness and buffer stability.

6. REFERENCES

[1] "Cisco visual networking index: Forecast and methodology, 2008-2013," Tech. Rep., Cisco, 2009.

[2] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From theory to practice*, Wiley, 2009.

[3] "Top 10 considerations for a successful evolved packet core (EPC) deployment," Tech. Rep., Cisco Syst., 2010.

[4] 3GPP TS 23.203, "Policy and charging control architecture," Tech. Rep., version 10.7.0 Release 10, 2012.

[5] T. Stockhammer, "Dynamic adaptive streaming over HTTP design principles and standards," *in Proc. ACM conference on Multimedia systems*, pp. 133–144, 2011.

[6] G. Hasegawa and M. Murata, "Survey on fairness issues in TCP congestion control mechanisms," *IEICE Trans. on Communications*, vol. E84-B, no. 8, pp. 1461–1472, 2001.

[7] Hyunggon Park and M. van der Schaar, "Fairness strategies for wireless resource allocation among autonomous multimedia users," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 20, no. 2, pp. 297–309, Feb 2010.

[8] O. Oyman and S. Singh, "Quality of experience for HTTP adaptive streaming services," *IEEE Communications Magazine*, vol. 50, no. 4, pp. 20–27, 2012.

[9] D. De Vleeschauwer, H. Viswanathan, A. Beck, S. Benno, G. Li, and R. Miller, "Optimization of HTTP adaptive streaming over mobile cellular networks," in *INFOCOM, 2013 Proceedings IEEE*, 2013, pp. 898–997.

[10] Jiasi Chen, Rajesh Mahindra, Mohammad Amir Khojastepour, Sampath Rangarajan, and Mung Chiang, "A scheduling framework for adaptive video delivery over cellular networks," in *Proc. of the 19th Annual Int. Conf. on Mobile Computing; Networking*, 2013, MobiCom '13.

[11] Ali El Essaili, Damien Schroeder, Dirk Staehle, Mohammed Shehada, Wolfgang Kellerer, and Eckehard Steinbach, "Quality-of-experience driven adaptive HTTP media delivery," in *IEEE Int. Conf. on Commun. (ICC 2013)*, Budapest, Hungary, Jun 2013.

[12] K. Seshadrinathan, R. Soundararajan, A.C. Bovik, and L.K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Trans. on Image Processing*, vol. 19, no. 6, pp. 1427 – 1441, June 2010.

[13] S. Cicalò and V. Tralli, "Distortion-fair cross-layer resource allocation for scalable video transmission in OFDMA wireless networks," *IEEE Transactions on Multimedia*, vol. 16, no. 3, April 2014.

[14] Stefan Lederer, Christopher Müller, and Christian Timmerer, "Dynamic adaptive streaming over HTTP dataset," in *Proceedings of the 3rd Multimedia Systems Conference*, New York, NY, USA, 2012, MMSys '12, pp. 89–94, ACM.

[15] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Trans. on*, vol. 13, no. 4, pp. 600–612, 2004.

[16] 3GPP TS 36.521-1, "User equipment (UE) conformance specification, radio transmission and reception. part 1: Conformance testing," Tech. Rep., version 11.0.1 Release 11, 2013.

[17] Junchen Jiang, Vyas Sekar, and Hui Zhang, "Improving fairness, efficiency, and stability in HTTP-based adaptive video streaming with FESTIVE," in *Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies*. 2012, CoNEXT '12, pp. 97–108, ACM.