ADAPTIVE VIDEO TRANSMISSION WITH SUBJECTIVE QUALITY CONSTRAINTS

Chao Chen[†], *Xiaoqing Zhu*[*], *Gustavo de Veciana*[†], *Alan C. Bovik*[†], *Robert W. Heath Jr.*[†],

[†]Department of Electrical and Computer Engineering,
The University of Texas at Austin, Austin, TX 78712-0240, USA
[*] The Advanced Architecture and Research Group,
Cisco Systems, San Jose, CA 95134, USA

ABSTRACT

We conducted a subjective study wherein we found that viewers' Quality of Experience (QoE) was strongly correlated with the empirical cumulative distribution function (eCDF) of the predicted video quality. Based on this observation, we propose a rate-adaptation algorithm that can incorporate QoE constraints on the empirical cumulative quality distribution per user. Simulation results show that the proposed technique can reduce network resource consumption by 29% over conventional average-quality maximized rate-adaptation algorithms.

***Index Terms*—** Quality of Experience, Video transport, Rate adaptation

1. INTRODUCTION

Video traffic is a growing fraction of the data traffic on wireless networks. As reported in [1], video traffic accounted for more than 50% of the mobile data traffic in 2012. Efficiently utilizing network resources to satisfy video users' expectations regarding their Quality of Experience (QoE) is one way to transmit video more efficiently. In this paper, we study approaches to share wireless down-link resources among video users via QoE-based rate adaptation.

Specifically, we focus on a setting in which stored video content is streamed over a wireless network. Because the throughput of a wireless channel varies over time, video rate-adaptation techniques based on scalable video coding or adaptive bitrate switching have been proposed to match the video data rate to the varying channel capacity [2, 3, 4, 5, 6]. Although these rate adaptation techniques can effectively avoid playback interruptions due to throughput variations, the variable bitrate causes quality fluctuations, which, in turn, affect viewers' QoE. In most existing rate-adaptation algorithms such as [7, 8, 9, 10, 11, 12, 13], average video quality is employed as a proxy for QoE. Although the average quality tends to capture the overall video quality to some extent, it does not reflect the impact of quality fluctuations on the QoE, i.e., two videos with the same average quality can have significantly different levels of quality fluctuation.

In this paper, we propose to characterize and predict the users' QoE with a new metric (see definition in Section 2.3). The efficacy of the proposed metric was validated through a subjective experiment reported in [14] and [15]. For each of the 15 quality-varying videos involved in the subjective study, we asked 25 subjects to score its overall quality. The proposed metric achieves a strong linear correlation of 0.84. By comparison, the average video quality only achieves a correlation of 0.57. This lends strong support for the proposed metric as a good proxy for video QoE.

Using the metric, we design adaptive video streaming algorithms that incorporate constraints on the QoE of video users. In particular, we consider a wireless network in which a base station transmits videos to multiple users (see Fig. 1). The user population is dynamic, i.e., users arrive and depart from the network at random times. When a new user joins the network, the base station starts streaming a video to the user. Paralleling prior work such as [12][16][17], we assume a proxy is colocated with the base station. A rate adaptation algorithm is employed by the proxy to control the video data rate of all active video streams according to varying wireless channel conditions.

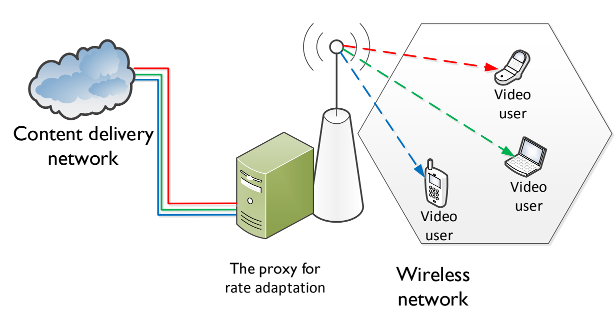


Fig. 1. The considered wireless network.

This research was supported in part by Intel Inc. and Cisco Corp. under the VAWN program.

2. SYSTEM MODEL

In this section, we explain the system models used in the paper. In the followings, the time slots are indexed with $t = 1, 2, \ldots$. Calligraphic symbols such as $\mathcal{A}$ denote sets, while $|\mathcal{A}|$ is the cardinality of $\mathcal{A}$. The function $[x]^+ = \max\{x, 0\}$.

2.1. Wireless Channel Model

We label users according to their arrival times: user u is the u^{th} user to arrive to the network. For each user u, we let A_u and D_u be random variables denoting the arrival and departure times, respectively. The time spent by a user in the network is denoted by $\mathsf{T}_u = \mathsf{D}_u - \mathsf{A}_u + 1$.

We denote by $\mathcal{U}(t)$ the set of video users in the network at slot t. For a video user $u \in \mathcal{U}(t)$, we denote by $\mathrm{w}_u(t)$ the amount of correctly received video data in slot t. The video data rate delivered to the user is thus $\mathrm{r}_u(t) = \mathrm{w}_u(t)/\Delta T$. We call $\mathbf{r}(t) = (\mathrm{r}_u(t) : u \in \mathcal{U}(t))$ the *video rate vector* at time t.

We assume that the set of video rate vectors that can be supported by the wireless channel is given by

$$\mathcal{C}(t) = \{\mathbf{r} : \mathrm{C}_t(\mathbf{r}) \leq 0\}, \tag{1}$$

where $\mathrm{C}_t : \mathbb{R}^{|\mathcal{U}(t)|} \rightarrow \mathbb{R}$ is a time-varying multivariate convex function. The specific form of $\mathrm{C}_t(\cdot)$ depends on the multiuser multiplexing techniques used in the wireless network. For example, in a time-division multiple access (TDMA) system [12][16], the channel is occupied by a single user at any moment. Denote by $\mathsf{P}_u(t)$ the peak transmission rate of user u, i.e., the transmission rate at which user u can be served during slot t. Then, in slot t, video user $u \in \mathcal{U}(t)$ spends $\frac{\mathrm{w}_u(t)}{\mathsf{P}_u(t)}$ seconds to download video. Since the total transmission across all users is less than ΔT, we have $\sum_{u\in\mathcal{U}(t)} \frac{\mathrm{w}_u(t)}{\mathsf{P}_u(t)} \leq \Delta T$. Because $\mathrm{r}_u(t) = \mathrm{w}_u(t)/\Delta T$, we have $\sum_{u\in\mathcal{U}(t)} \frac{\mathrm{r}_u(t)}{\mathsf{P}_u(t)} \leq 1$. Therefore, for TDMA systems, we have $\mathrm{C}_t(\mathbf{r}(t)) = \sum_{u\in\mathcal{U}(t)} \frac{\mathrm{r}_u(t)}{\mathsf{P}_u(t)} - 1$.

2.2. Video Rate-Quality Model

We assume that the quality of the video downloaded in each slot is represented by a Difference Mean Opinion Score (DMOS) [18], which ranges from 0 to 100 where lower values indicate better quality. To represent video quality more naturally, so that higher numbers indicate better video quality, we deploy a Reversed DMOS (RDMOS). Denote by $\mathrm{q}_u^{\mathrm{dmos}}(t)$ the DMOS of the video delivered to user u at slot t, the RDMOS is given by $\mathrm{q}_u^{\mathrm{rdmos}}(t) = 100 - \mathrm{q}_u^{\mathrm{dmos}}(t)$. We employ the following rate-quality model to predict $\mathrm{q}_u^{\mathrm{rdmos}}(t)$ using the video data rate $\mathrm{r}_u(t)$:

$$\mathrm{q}_u(t) = \alpha_u(t)\log(\mathrm{r}_u(t)) + \beta_u(t), \tag{2}$$

where $\mathrm{q}_u(t)$ is the predicted RDMOS. The model parameters $\alpha_u(t)$ and $\beta_u(t)$ can be determined by minimizing the prediction error between $\mathrm{q}_u^{\mathrm{rdmos}}(t)$ and $\mathrm{q}_u(t)$. For stored video streaming, the video file is stored at the CDN. Thus, the model parameters in (2) can be obtained before transmission. Here, we assume the parameters $\alpha_u(t)$ and $\beta_u(t)$ are known *a priori*.

We validated this model on a video database that includes twenty-five videos [14]. The mean prediction error of (2) is less than 1.5, which is visually negligible.

2.3. Constraints on the Quality of Experience

We propose to capture video users' QoE using the second order empirical cumulative distribution function [19] (2^{th}-order eCDF) $\mathrm{F}^{(2)}(x; \mathrm{q}_u)$, defined as:

$$\mathrm{F}^{(2)}(x; \mathrm{q}_u) = \frac{1}{\mathsf{T}_u} \sum_{t=1}^{\mathsf{T}_u} \max\{x - \mathrm{q}_u(t), 0\}. \tag{3}$$

Here, $\mathrm{q}_u(t)$ is the video quality given by (2). To justify the use of the $\mathrm{F}^{(2)}(x; \mathrm{q})$ as the QoE metric, we conducted a subjective study [14, 15]. We found that, at $x^* = 37$, $\mathrm{F}^{(2)}(x^*; \mathrm{q}_u)$ achieves a strong correlation of 0.84 with the Mean Opinion Scores obtained from the subjective study. Note that $\max\{x^* - \mathrm{q}_u(t), 0\} > 0$ if and only if $\mathrm{q}_u(t) < x^*$, so $\mathrm{F}^{(2)}(x^*; \mathrm{q}_u)$ captures for how long and by how much the video quality falls below x^*. Thus, we interpret x^* as the users' *video quality expectation*, which is used by the users as a threshold in judging whether the video quality is acceptable or not.

In our subjective study, all subjects viewed the videos in a controlled environment and every subject viewed the videos on the same device. Broadly speaking, the video quality expectation x^* can be environment-dependent. For example, viewers tend to have higher expectation for videos shown on a laptop than videos shown on a smartphone. Therefore, in a practical wireless network, x^* may vary across users. In practice, the video quality expectation x_u^* is difficult to know *a priori*. Thus, in this paper, we assume x_u^* is not known and we impose constraints on all x and for all video users [1]. In particular, we consider the following QoE constraints:

$$\mathrm{F}^{(2)}(x; \mathrm{q}_u) \leq \mathrm{h}(x),\ \forall x \in [0, 100],\ \forall u \in \cup_{t=1}^{\infty}\mathcal{U}(t), \tag{4}$$

where $\mathrm{h}(x)$ is a function of x. Since we cannot apply constraints on all values of $x \in [0, 100]$, we consider a relaxed version of (4) as follows:

$$\mathrm{F}^{(2)}(x_i; \mathrm{q}_u) \leq \mathrm{h}(x_i),\ \forall x_i \in \mathcal{I},\ \forall u \in \cup_{t=1}^{\infty}\mathcal{U}(t). \tag{5}$$

Here, $\mathcal{I}$ is a discrete set of points on $[0, 100]$. It was shown that (5) will approximate (4) if $\mathcal{I}$ is dense. Its proof is given in [20]. Next, we propose a rate adaptation algorithm that maximizes the number of users who satisfy the constraint of (5).

[1]For the case where x_u^* is known *a priori*, please refer to [20] for a extended algorithm design.

3. RATE ADAPTATION ALGORITHM

We first present an off-line problem formulation in which the future channel conditions are assumed to be known. Then, based on the analysis of this offline problem, we propose a new on-line rate adaptation algorithm.

Consider a finite horizon T and assume that the realizations of channel conditions $\mathcal{C}(1), \ldots, \mathcal{C}(T)$ are known. Then the rate adaptation algorithm should solve the following feasibility problem:

$$\begin{aligned}
&\text{find} && \mathbf{r}_{1:T} && \text{(6a)}\\
&\text{subject to:} && \mathbf{r}(t) \in \mathcal{C}(t) && \text{(6b)}\\
& && \mathrm{q}_u(t) = \alpha_u(t)\log(\mathrm{r}_u(t)) + \beta_u(t) && \text{(6c)}\\
& && \mathrm{F}^{(2)}\left(x_i; \mathrm{q}_u\right) \le \mathrm{h}(x_i) && \text{(6d)}\\
& && \forall t \in \{1, \ldots, T\},\ \forall x_i \in \mathcal{I},\ \forall u \in \cup_{t=1}^{T}\mathcal{U}(t).
\end{aligned}$$

The constraint (6b) is associated with the achievable rate region (1). The constraint (6c) is because of the rate-quality model (2). The constraints (6d) are the QoE constraints (5) that were discussed in Section 2.3. For each admitted user, a series of QoE constraints are applied to the 2^{nd}-order eCDF at discrete points in $\mathcal{I} = \left\{x_1, \cdots, x_{|\mathcal{I}|}\right\}$. Since the rate-quality function (6c) is concave, according to [21], the problem (6) is equivalent to the following convex optimization problem:

$$\begin{aligned}
&\underset{\mathbf{r}_{1:T},\ (\hat{\mathrm{q}}_u)_{1:T}}{\text{maximize}} && 0 && \text{(7a)}\\
&\text{subject to:} && \mathbf{r}(t) \in \mathcal{C}(t) && \text{(7b)}\\
& && \hat{\mathrm{q}}_u(t) \le \alpha_u(t)\log(\mathrm{r}_u(t)) + \beta_u(t) && \text{(7c)}\\
& && \mathrm{F}^{(2)}\left(x_i; \hat{\mathrm{q}}_u\right) \le \mathrm{h}(x_i) && \text{(7d)}\\
& && \forall t \in \{1, \ldots, T\},\ \forall x_i \in \mathcal{I},\ \forall u \in \cup_{t=1}^{T}\mathcal{U}(t).
\end{aligned}$$

where $\hat{\mathrm{q}}_u(t)$ are virtual variables introduced here to make the constraint (7c) convex. Since (7) is a convex problem, if it is feasible, there exists a set of Lagrange multipliers $\lambda^*_{u,i} \ge 0$ for the constraints in (7d) such that a solution of (7) can be obtained by solving the following problem (see [22]):

$$\begin{aligned}
&\underset{\mathbf{r}_{1:T},\ (\hat{\mathrm{q}}_u)_{1:T}}{\text{minimize}} && \sum_{\forall u}\sum_{\forall x_i}\lambda^*_{u,i}\left[\mathrm{F}^{(2)}\left(x_i; \hat{\mathrm{q}}_u\right) - \mathrm{h}(x_i)\right] && \text{(8a)}\\
&\text{subject to:} && \mathbf{r}(t) \in \mathcal{C}(t) && \text{(8b)}\\
& && \hat{\mathrm{q}}_u(t) \le \alpha_u(t)\log(\mathrm{r}_u(t)) + \beta_u(t) && \text{(8c)}\\
& && \forall t \in \{1, \ldots, T\},\ \forall u \in \cup_{t=1}^{T}\mathcal{U}(t).
\end{aligned}$$

Define a function

$$\mathrm{s}_{u,i}(t) = \frac{1}{\mathsf{T}_u}\left(\left[x_i - \hat{\mathrm{q}}_u(t)\right]^+ - \mathrm{h}(x_i)\right). \tag{9}$$

Now expanding $\mathrm{F}^{(2)}\left(x_i; \hat{\mathrm{q}}_u\right) - \mathrm{h}(x_i)$, we have

$$\begin{aligned}
\mathrm{F}^{(2)}\left(x_i; \hat{\mathrm{q}}_u\right) - \mathrm{h}(x_i) &= \textstyle\sum_{t=\mathsf{A}_u}^{\mathsf{D}_u} \frac{1}{\mathsf{T}_u}\left(\left[x_i - \hat{\mathrm{q}}_u(t)\right]^+ - \mathrm{h}(x_i)\right)\\
&= \textstyle\sum_{t=1}^{T} \mathrm{s}_{u,i}(t).
\end{aligned} \tag{10}$$

Substituting (10) in (8a) and changing the order of summation, the optimization in (8) becomes

$$\begin{aligned}
&\underset{\mathbf{r}_{1:T},\ (\hat{\mathrm{q}}_u)_{1:T}}{\text{minimize}} && \sum_{t=1}^{T}\left(\sum_{u\in\mathcal{U}(t)}\sum_{x_i\in\mathcal{I}}\lambda^*_{u,i}\mathrm{s}_{u,i}(t)\right)\\
&\text{subject to:} && \mathbf{r}(t) \in \mathcal{C}(t)\\
& && \hat{\mathrm{q}}_u(t) \le \alpha_u(t)\log(\mathrm{r}_u(t)) + \beta_u(t)\\
& && \forall t \in \{1, \ldots, T\},\ \forall u \in \mathcal{U}(t).
\end{aligned} \tag{11}$$

Note that, except for the Lagrange multipliers, the optimization in (11) does not involve variables that depend on the entire process $\{\hat{\mathrm{q}}_u(t) : 1 \le t \le T\}$. Thus, if it is possible to estimate the Lagrange multiplier $\lambda^*_{u,i}$, then (7) can be solved by greedily choosing the rate vector $\mathbf{r}(t)$ at each time slot as the minimizer of $\sum_{u\in\mathcal{U}(t)}\sum_{x_i\in\mathcal{I}}\lambda^*_{u,i}\mathrm{s}_{u,i}(t)$.

We introduce a method to approximate the Lagrange multiplier $\lambda^*_{u,i}$. We know that the Lagrange multiplier $\lambda^*_{u,i}$ indicates the difficulty of satisfying the constraint $\mathrm{F}^{(2)}_u(x_i; \hat{\mathrm{q}}_u) \le \mathrm{h}(x_i)$ [21]. Inspired by prior work in [23] and [24], we employ a virtual queue to capture this difficulty. For each admitted user u and each $x_i \in \mathcal{I}$, define the virtual queue as

$$\mathrm{v}_{u,i}(t) = \left[\mathrm{v}_{u,i}(t-1) + \mathrm{s}_{u,i}(t)\right]^+. \tag{12}$$

From (10) it follows that, if the summation of $\mathrm{s}_{u,i}(t)$ is large, then it is difficult to satisfy the constraint $\mathrm{F}^{(2)}_u(x_i; \hat{\mathrm{q}}_u) \le \mathrm{h}(x_i)$. The virtual queue captures the cumulative summation of $\mathrm{s}_{u,i}(t)$ up to slot t. Hence, the virtual queue reflects the level of difficulty in satisfying $\mathrm{F}^{(2)}_u(x_i; \hat{\mathrm{q}}_u) \le \mathrm{h}(x_i)$. We replace the Lagrange multipliers in (11) with virtual queue $\mathrm{v}_{u,i}(t)$ and our online rate adaptation algorithm is summarized in Algorithm 1. In every slot, we maximize the weighted sum of $\mathrm{s}_{u,i}(t)$, where the weight is given by $\mathrm{v}_{u,i}(t-1)$. Thus users with larger virtual queues tend to be allocated more network resources. This helps users satisfy their QoE constraints.

Algorithm 1 Online algorithm for video data rate adaptation

for $t = 1 \to \infty$ **do**

Choose rate vector $\mathbf{r}(t)$ that solves the problem

$$\begin{aligned}
&\underset{\mathbf{r}(t),\ \hat{\mathrm{q}}_u(t)}{\text{minimize}} && \sum_{u\in\mathcal{U}(t)}\sum_{x_i\in\mathcal{I}}\mathrm{v}_{u,i}(t-1)\mathrm{s}_{u,i}(t)\\
&\text{subject to:} && \mathbf{r}(t) \in \mathcal{C}(t)\\
& && \hat{\mathrm{q}}_u(t) \le \alpha_u(t)\log(\mathrm{r}_u(t)) + \beta_u(t)\\
& && \forall u \in \mathcal{U}(t),
\end{aligned} \tag{13}$$

For $\forall u \in \mathcal{U}(t), \forall x_i \in \mathcal{I}$, update virtual queues with

$$\mathrm{v}_{u,i}(t) = \left[\mathrm{v}_{u,i}(t-1) + \mathrm{s}_{u,i}(t)\right]^+.$$

end for

4. SIMULATION RESULTS

Below, we evaluate our rate-adaptation algorithm via numerical simulations. We assume the duration of a time slot is $\Delta T = 1$ second. Video users arrival as a Poisson process with an average arrival rate of $\frac{1}{20}$ users/second. The time spent by a video user in the network is exponentially distributed with a mean value of 200 seconds. To simulate variations of the rate-quality characteristics in each video stream, we assume the rate-quality parameters $(\alpha_u(t), \beta_u(t))$ of each slot are independently sampled from the rate-quality parameters in the video database [14].

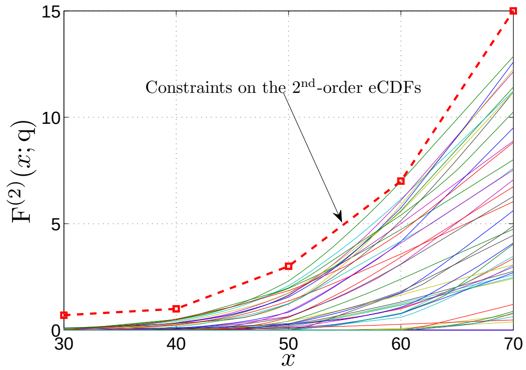
We assume the wireless system is a TDMA system. The rate region $\mathcal{C}(t)$ is that introduced in Section 2.1. We model the peak transmission rate $\mathsf{P}_u(t)$ as the product of two independent random variables, i.e., $\mathsf{P}_u(t) = \mathsf{P}_u^{\mathrm{avg}} \times \mathsf{P}_u^*(t)$. The random variable $\mathsf{P}_u^{\mathrm{avg}}$ is employed to simulate the heterogeneity of the channel condition across users and remains constant during a user's sojourn. We assume that $\mathsf{P}_u^{\mathrm{avg}}$ is uniformly distributed on $[1250\gamma, 3750\gamma]$ kbps, where the parameter γ is used to scale the channel capacity in our simulations. The random variable $\mathsf{P}_u^*(t)$ is employed to simulate channel variation across time slots. We assume that $\{\mathsf{P}_u^*(t) : t \in \mathbb{N}^+\}$ is an i.i.d. process with $\mathsf{P}_u^*(t)$ being uniformly distributed on $[0.5, 1.5]$. In our simulations, we set $\mathcal{I} = \{30, 40, 50, 60, 70\}$. Correspondingly, for $x_i = 30$, 40, 50, 60, and 70, we let the constraints $\mathrm{h}(x_i) = 0.7$, 1.0, 3.0, 7.0 and 15.0, respectively.

We set the scaling parameter $\gamma = 12$ and simulate Algorithm 1 until 100 users have arrived and departed the network. We plot the 2^{nd}-order eCDFs of the video users in Fig. 2(a). It may be seen that, using Algorithm 1, the 2^{nd}-order eCDFs of the video users all satisfy the constraints. By comparison, if we adapt the rate vector to maximize the sum of the average-quality of all users[2], the QoE constraint is violated by many users (see Fig. 2(b)).

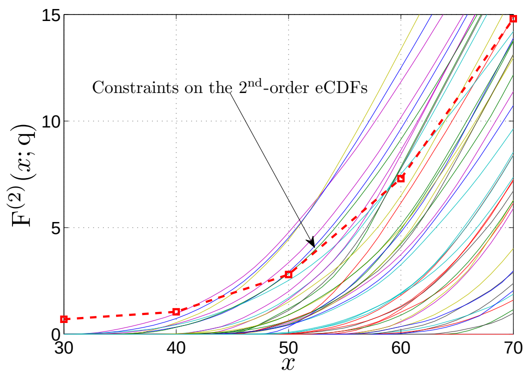
We scale the channel scaling parameter from $\gamma = 6$ to $\gamma = 16$. The percentage of video users whose video qualities satisfy the QoE constraints is shown in Fig. 3. When compared with the average-quality-maximized rate-adaptation algorithm, the percentage of video users who satisfy the QoE constraints is improved significantly. For example, at a moderate channel condition of $\gamma = 14$, about 90% of the video users satisfy the QoE constraints when the average-quality maximizing algorithm is applied. The proposed algorithms achieve the same performance at $\gamma = 10$, reducing the consumption of resources by $(14 - 10)/14 \approx 29\%$.

5. CONCLUSIONS

We created a new QoE metric based on the second-order empirical cumulative distribution function (eCDF) of time-varying video quality. We then proposed an online rate adaptation algorithm to improve the number of video users who satisfy the QoE constraints on the second-order cumulative distribution function. Simulation results showed that combining the proposed approaches leads to a 29 % reduction in wireless network resource consumption.



(a) Proposed rate-adaptation algorithm



(b) Averge-quality maximized rate adaptation

Fig. 2. The 2^{nd}-order eCDFs of video users with (a) proposed rate adaptation algorithm and (b) average-quality maximized rate adaptation algorithm.

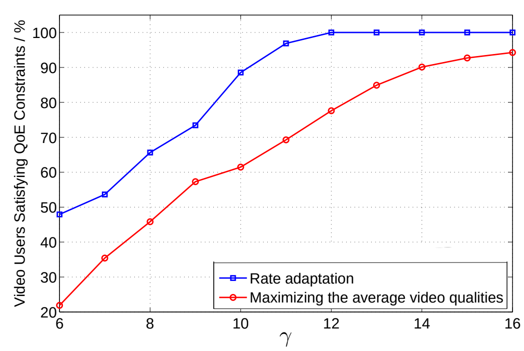


Fig. 3. Simulation results of the proposed algorithms under different channel scaling parameters.

[2]This is achieved by maximizing $\sum_{u \in \mathcal{U}(t)} \mathrm{q}_u(t)/\mathsf{T}_u$ in each slot.

6. REFERENCES

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

[2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sept. 2007.

[3] Microsoft Corporation, "IIS smooth streaming technical overview," `http://www.microsoft.com/en-us/download/default.aspx`, Sept. 2009.

[4] R. Pantos and E. W. May, "HTTP live streaming," *IETF Internet Draft, work in progress*, Mar. 2011.

[5] Adobe Systems, "HTTP dynamics streaming," `http://www.adobe.com/products/hds-dynamic-streaming.html`, Mar. 2012.

[6] MPEG Requirements Group, "*ISO/IEC FCD* 23001-6 Part 6: Dynamics adaptive streaming over HTTP (DASH)," `http://mpeg.chiariglione.org/working_documents/mpeg-b/dash/dash-dis.zip`, Jan. 2011.

[7] C.E. Luna, Y. Eisenberg, R. Berry, T.N. Pappas, and A.K. Katsaggelos, "Joint source coding and data rate adaptation for energy efficient wireless video streaming," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1710–1720, 2003.

[8] J. Chakareski and P. Frossard, "Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 207–218, 2006.

[9] J. Huang, Z. Li, M. Chiang, and A. K. Katsaggelos, "Joint source adaptation and resource allocation for multi-user wireless video streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 5, pp. 582–595, 2008.

[10] X. Zhu and B. Girod, "Distributed media-aware rate allocation for wireless video streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1462–1474, 2010.

[11] K.C. Lin, Wei-Liang Shen, Chih-Cheng Hsu, and Cheng-Fu Chou, "Quality-differentiated video multicast in multirate wireless networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 1, pp. 21–34, 2013.

[12] H. Hu, X. Zhu, Y. Wang, R. Pan, J. Zhu, and F. Bonomi, "Proxy-based multi-stream scalable video adaptation over wireless networks using subjective quality and rate models," *to appear on IEEE Trans. Multimedia*, 2013.

[13] C. Chen, R. W. Heath Jr., A. C. Bovik, and G. de Veciana, "A markov decision model for adaptive scheduling of stored scalable videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 6, pp. 1081–1095, 2013.

[14] C. Chen, L. K. Choi, G. de Veciana, C. Caramanis, R. W. Heath Jr., and A. C. Bovik, "A dynamic system model of time-varying subjective quality of video streams over HTTP," *IEEE Int'l Conf. Acoust. Speech Signal Process*, May 2013.

[15] C. Chen, L. K. Choi, G. de Veciana, C. Caramanis, R. W. Heath Jr., and A. C. Bovik, "A model of the time-varying subjective quality of http video streams," June 2013.

[16] V. Joseph and G. de Veciana, "Jointly optimizing multi-user rate adaptation for video transport over wireless systems: Mean-fairness-variability tradeoffs," in *2012 Proceedings IEEE INFOCOM,*, 2012, pp. 567–575.

[17] Wei Pu, Zixuan Zou, and Chang Wen Chen, "Video adaptation proxy for wireless dynamic adaptive streaming over http," in *19th Int'l Packet Video Wkshp. (PV)*, 2012, pp. 65–70.

[18] ITU-R Recommendation BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," `http://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf`, Jan. 2012.

[19] D. Dentcheva and A. Ruszczynski, "Optimization with stochastic dominance constraints," *SIAM Journal on Optimization*, vol. 14, no. 2, pp. 548–566, 2003.

[20] C. Chen, X. Zhu, G. de Veciana, A. C. Bovik, and R. W. Heath Jr., "Rate adaptation and admission control for video transmission with subjective quality constraints," *IEEE J. Sel. Topics Signal Process.*, submitted, Available at: http://arxiv.org/abs/1311.6453.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, NY, USA, 2004.

[22] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1996.

[23] A. L. Stolyar, "Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm," *Queueing Systems*, vol. 50, no. 4, pp. 401–457, 2005.

[24] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*, Morgan & Claypool, 2010.