Computer Networks xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Q1 # A quality of experience-aware cross-layer architecture for optimizing video streaming services

Q2 Qahhar Muhammad Qadir[a,*], Alexander A. Kist[a], Zhongwei Zhang[b]

Q3 [a] School of Mechanical and Electrical Engineering, Australia
[b] School of Agricultural, Computational and Environmental Sciences, Faculty of Health, Engineering and Science, University of Southern Queensland, Australia

A R T I C L E I N F O

Article history:
Received 30 July 2015
Revised 13 February 2016
Accepted 21 February 2016
Available online xxx

Keywords:
Quality of experience
Cross-layer architecture
Optimization
Video

A B S T R A C T

The popularity of the video services on the Internet has evolved various mechanisms that target the Quality of Experience (QoE) optimization of video traffic. The video quality has been enhanced through adapting the sending bitrates. However, rate adaptation alone is not sufficient for maintaining a good video QoE when congestion occurs. This paper presents a *cross-layer architecture* for video streaming that is QoE-aware. It combines adaptation capabilities of video applications and QoE-aware admission control to optimize the trade-off relationship between QoE and the number of admitted sessions. Simulation results showed the efficiency of the proposed architecture in terms of QoE and *number of sessions* compared to two other architectures (*adaptive architecture* and *non-adaptive architecture*).

1. Introduction

The increasing popularity of various video services [1] has made studying the Quality of Experience (QoE) important. The ITU-T defines QoE as a measure to evaluate the service quality as perceived by end users [2]. Various technical and non-technical factors affect this new quality measure [3]. Among these factors are those related to service preparation, delivery and presentation. This makes the task of maintaining QoE at an acceptable level a challenge. Many solutions have been introduced to tackle the challenge of video traffic quality. However, more promising architectures are required to meet the satisfaction of users and preserve the interest of service providers. This common goal has been targeted by various designs. Different approaches focusing on optimization metrics, scope and adaptation methods are available. They can be deployed individually or jointly to achieve the goal which is called cross-layer design in the latter case [4].

Optimization has to resolve the conflict between the interests of end users and network providers. From the end user perspective, maximum quality is expected; whereas low-cost and the number of served users are important from the network providers' perspective. These two can be jointly optimized through an intelligent design. This motivation has promoted the development of cross-layer designs for video transmission that are QoE-aware. The main objective is to utilize network resources efficiently and optimize video quality, throughput or QoE through a joint cooperation between layers and optimization of their parameters. This enables the communication and interaction between layers by allowing one layer to access the data of another layer. For example, having knowledge about the available bandwidth (network layer) helps the sender to adapt the sending rate (application layer). As a result of this cooperation, better quality for as many users as possible can be expected.

Q4 * Corresponding author. Tel.: +61 470534488.
Q5 *E-mail addresses:* safeen.qadir@gmail.com (Q.M. Qadir), kist@ieee.org (A.A. Kist), zhongwei.zhang@usq.edu.au (Z. Zhang).

http://dx.doi.org/10.1016/j.comnet.2016.02.030
1389-1286/

Although dynamic rate adaptation enhances video quality, accepting more sessions than a link can accommodate will degrade the quality. We have investigated how rate adaptation of video sources can provide a better QoE in our previous work [5]. However, the friendly behavior of the Internet's transport protocol accommodates every video session and makes room for everyone. This causes degradation of QoE of all video sessions in a bottleneck link. Adaptive sources attempt to reduce the transmission rate of all video sources in order to share the available link capacity. This process does not consider how much the QoE at the receiving end will be affected by the adaptation process. To overcome this problem, a mechanism is required to maintain the quality of on-going video sessions.

In this paper, we combine the rate adaptation capability of video applications and our previously proposed QoE-aware admission control [6] in a QoE-aware architecture for video streaming. The contribution of this paper is a novel QoE-aware cross-layer architecture for optimizing video streaming services. The proposed architecture addresses the issue of QoE degradation of video traffic in a bottleneck network. In particular, it allows video sources at the application layer to adapt their rate dynamically to the network environment; and the edge of the network at the network layer to protect the quality of active video sessions by controlling the acceptance of new session through a QoE-aware admission control.

The remaining of the paper is organized as follows: related work is reviewed in Section 2. We introduce our proposed QoE-aware cross-layer architecture in Section 3. The evaluation environment is explained in Section 4. Section 5 presents and discusses the results. Finally, Section 6 concludes the paper.

2. Related work

Extensive research has been done in the area of QoE optimization for video traffic. Some have focused on the optimization of the video's QoE through mechanisms on a single network layer. Classification and survey of these mechanisms can be found in [7–9]. In this section we only focus on cross-layer designs that have been proposed to optimize the QoE of video traffic.

The optimal rate of competing scalable video sources for QoE optimization has been found in [10]. Loss-induced distortion caused by video sources has been minimized and QoE has been maximized by obtaining the optimal rate and capturing the exact effect of packet loss in [11]. The resulting rates from [10,11] are proposed to be used by video encoders for online rate adaptation. In [12], a rate adaptation scheme and the IEEE 802.21 media independent handover are integrated for a single and scalable coding. In [13], the source rate at the application layer and modulation schemes at the radio link layer are optimized for the quality of video streaming using an application-driven objective function. The link adaptation of the high speed downlink packet access and rate adaptation of multimedia applications are integrated in a QoE-based cross-layer framework that is capable of maximizing the QoE [14].

Work in [15] combines link adaptation based on an online QoS to QoE mapping, buffer-aware source adaptation and video-layer dependent packet retransmission techniques to provide delay-constrained scalable video transmissions with optimized perceptual quality. The impact of power allocation on bit error rate and video source coding structure for Scalable Video Coding (SVC) video over Multi-Input Multi-Output (MIMO) with the aim of QoE maximization has been considered in [16].

The work presented in [17] extends the Pre-Congestion Notification (PCN)-based admission control, determines the required redundancy bits for coping with packet loss, and scales video rate to optimize the QoE in multimedia networks. Two different rate adaptation algorithms have been proposed in [18]; an optimal one to adapt the video rate based on the maximization of service provider's revenue or QoE and a heuristic one based on the utility of each connection. Relying on subjective tests, Chen et al. [19] proposes a rate adaptation algorithm and devises a threshold-based admission control strategy to maximize the percentage of video users whose QoE constraints can be satisfied. Per user's QoE constraint was defined by the empirical Cumulative Distribution Function (eCDF) of the predicted video quality.

The cross-layer design presented in [20] has optimized the QoE of the region of interest for mobile physicians by using advanced error concealment techniques. The work in [21] has combined the SVC optimization optimum time slicing for layered coded transmission and adaptive Modulation and Coding Scheme (MCS) to trade between the QoE and energy consumption of wireless broadcast receivers.

In [22], a QoE-aware joint subcarrier algorithm and a power radio algorithm are combined for a QoE-based resource allocation of the heterogeneous Orthogonal Frequency Division Multiple Access (OFDMA) downlink. The model presented in [23], efficiently allocates resources for video applications by mapping between Peak-Signal-to-Noise Ratio (PSNR) and Mean Opinion Score (MOS). Admission control and resource reallocation have been deployed in [24] to increase the session admission rate while maintaining an acceptable QoE of multimedia services in Long-Term Evolution (LTE). The authors of [25] utilized the QoE prediction model of Khan et al. [26] to rate the QoE of multimedia services and allocate resources dynamically.

The QoE-aware cross-layer Dynamic Adaptive Streaming over Hypertext Transfer Protocol (HTTP) (DASH) friendly scheduler introduced in [27], allocates wireless resources for each DASH user. The video quality is optimized based on the collected DASH information through an improved SVC to DASH layers mapping and a DASH proxy. The QoE of multi-user adaptive HTTP video in mobile networks has been optimized by adapting the transmission rate of DASH clients that can be supported by lower layers in [28]. In [29], an efficient video processing, an advanced real-time scheduling and reduced-reference metrics across the application and network layers are combined as components for a QoE-driven cross-layer design of mobile video systems.

The automatic architecture proposed in [30] monitors quality related parameters such as packet loss, video frame rate and router queue sizes. Proper actions such as lowering bit rate or adding more Forward Error Correction

Please cite this article as: Q.M. Qadir et al., A quality of experience-aware cross-layer architecture for optimizing video streaming services, Computer Networks (2016), http://dx.doi.org/10.1016/j.comnet.2016.02.030

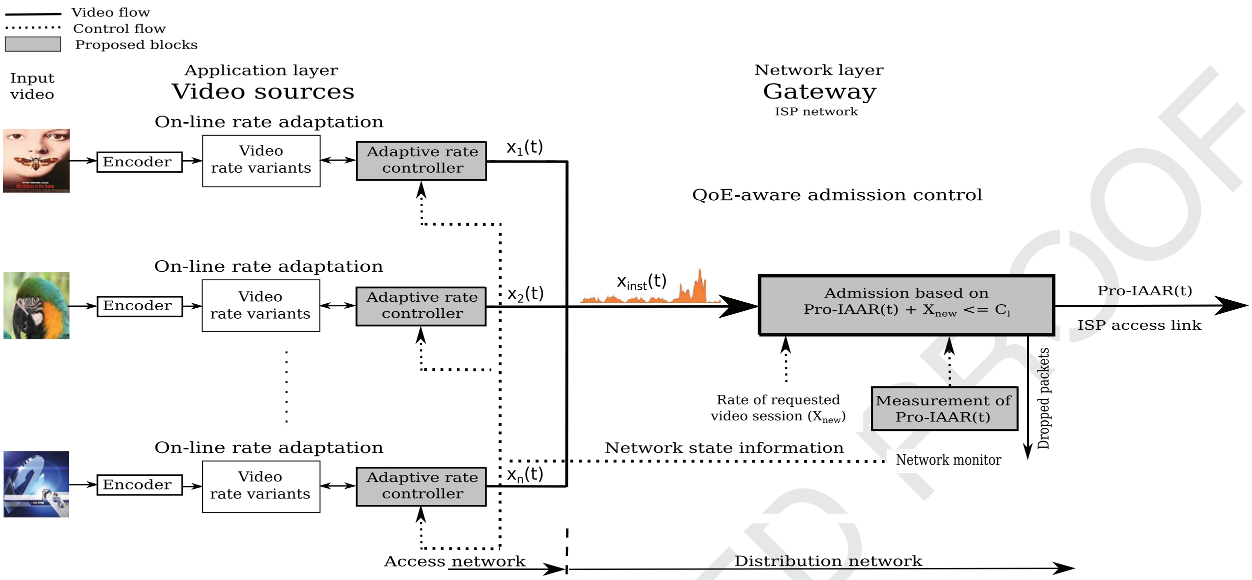


Fig. 1. QoE-aware cross-layer architecture for video traffic.

(FEC) packet are taken to optimize the QoE of multimedia services. In [31], an adaptive cross-layer architecture is presented. The HTTP Adaptive Streaming (HAS)/HTTP-specific media, network QoS and radio QoS are jointly adapted for optimizing the QoE of HAS applications. An end-to-end system for optimizing the QoE in next generation networks is presented in [32]. The QoE/QoS parameters at terminals are reported to the QoE management component for analysis and adjustment. The adjusted QoS/QoE of the end user is then sent to the network and source. A joint framework for video transport optimization over the next generation cellular network that overcomes network congestion, cache failure and user mobility issues is designed in [4]. Path selection, traffic management and frame filtering are mechanisms of the framework for the SVC video streaming over User Datagram Protocol/Real-time Transport Protocol (UDP/RTP). The interface presented in [33] is to enable the ISPs to deliver video contents efficiently and satisfy the user requirement for QoE through dynamic adaptation.

While the discussion has covered similar aspects, [17–19] have specifically combined rate adaptation and admission control in a cross-layer design for QoE optimization. In [17], the rate of layered video flows is re-scaled and protected through a number of changes to the original PCN. In contrast, our architecture accounts for the QoE of video sessions through a QoE-aware admission control. Latré and De Turck [18] integrates an existing standardized Measurement-Based Admission Control (MBAC) system with a novel video rate adaptation, while our work integrates the existing rate adaptation capability of multimedia applications with a QoE-aware admission control. Furthermore, our architecture optimizes the link utilization considering the QoE of video sessions whereas [18] accounts for QoE as the output of the system. Finally, [19] incorporates the QoE constraints into the rate adaptation algorithm, but our proposal incorporates QoE in the rate measurement algorithm and admission control.

3. QoE-aware cross-layer architecture

Much of the research discussed in the literature proposed rate adaptation for layered videos such as SVC. The video content (base and enhancements layers) generated by a layered encoder is injected to the network, then the network decides whether they are forwarded or dropped. In contrast, this paper proposes online rate adaptation for single layer videos. Instead of sending the whole video content to the network, video sources based on the condition of the network, decide at what rate to transmit the content. By using this strategy, the rate is adjusted on the fly and additional redundant data is not sent to the network during times of congestion. This is in contrast to offline coding which completely relies on coarse network state assumptions [34].

Rate adaptation attempts to adapt the sending rate of all video sources to share the available link capacity without considering how much the received QoE will be affected by the adaptation. Therefore, there is a need for a mechanism to control the number of video sessions which can be accommodated with an acceptable QoE.

Fig. 1 shows the proposed architecture which focuses on the optimization of QoE in relation to the *number of sessions* on the ISP access links (ISP links which are directly connected to and controlled by the gateway in Fig. 1). The video sources share the ISP access links of the distribution network which is controlled by the gateway. The rate adaptation is performed at the application layer and QoE-aware admission control at the network layer. More specifically, The QoE-aware admission control is implemented at the ISP gateway while the sources perform rate adaptation based on the available bandwidth of the ISP access links.

Unlike current MBACs, the QoE-aware admission control considers the bursty characteristic of video flows as the burstiness of individual video flow can be compensated by the silence of others [35]. Bursty traffic refers to inconsistency of the traffic level. It is at high level sometimes

Please cite this article as: Q.M. Qadir et al., A quality of experience-aware cross-layer architecture for optimizing video streaming services, Computer Networks (2016), http://dx.doi.org/10.1016/j.comnet.2016.02.030

Table 1
Notations used for modeling the QoE-aware cross-layer architecture. Q6

Notation	Description
$x_i(t)$	Instantaneous video rate of session i at time t
x_{new}	Rate of requested video session
n	Number of video sessions
C_l	Link capacity
β	A parameter defines the upper limit of the aggregate rate that can exceed C_l while maintaining the QoE of enrolled video sessions
Pro-IAAR(t)	Measured QoE-aware aggregate rate at time t
$\mu_s(t)$	Expected value of the total aggregate rate at time t
$p_i(t)$	Active or inactive probability of the session i at time t
ϵ	The positive number defined by the Hoeffding inequality theorem [36]
α and δ	Coefficients of β model defined by Eq. (5)
k	A local variable/counter, where $k \in \{2–31\}$
QP	Quantization parameter of video encoder, where $QP \in \{2–31\}$

while is at low level at some other times. The model and implementation of the QoE-aware measurement algorithm and admission control were presented in [6]. The proposed architecture employs parameters from relevant layers; application and network layers in this paper. The key parameters to be considered for the cross-layer optimization are the instantaneous video rate of session i at time t; $x_i(t)$ and rate of requested video session; x_{new} from the application layer, while at the network layer, the link capacity; C_l, *number of sessions*; n, parameter β (explained later in this section), and the measured QoE-aware aggregate rate; Proposed-Instantaneous Aggregate Arrival Rate (*Pro-IAAR(t)*) are taken into account. Table 1 summarizes the notations used for structuring the QoE-aware cross-layer architecture. The architecture assumes that there are efficient and reliable routing protocols to route the video traffic through the ISP intra-domain links once they have been placed on the *ISP access link* by the gateway. It also assumes that there is sufficient bandwidth on the access (between video sources and ISP gateway) and core (Internet) networks.

A user's QoE (in terms of MOS) for video streaming services can be defined by a utility function [14]. MOS as a function of the aggregate bitrate is given by a simplified utility function in Eq. (1)

$$U = f(Pro-IAAR(t)), \quad f: Pro-IAAR(t) \rightarrow MOS \tag{1}$$

where *Pro-IAAR(t)* [6] is the upper limit of the total aggregate rate that can exceed a specific link capacity considering the QoE of ongoing video sessions and is given by Eq. (2)

$$Pro\text{-}IAAR(t) = \mu_s(t) + n\epsilon \tag{2}$$

$\mu_s(t)$ is the expected value of the total aggregate rate given by Eq. (3) and ϵ as a positive number of the Hoeffding inequality theorem [36] is quantified by Eq. (4) [6]. The probability of the session i to be active or inactive is represented by p_i in Eq. (3)

$$\mu_s(t) = \sum_{i=1}^{n} x_i(t)\ p_i(t) \tag{3}$$

$$\epsilon = \beta\mu_s(t)\frac{n-1}{n} \quad 0 < \beta \le 1. \tag{4}$$

Parameter β, modeled by Eq. (5), defines the upper limit of the total aggregate rate that can exceed C_l while maintaining the QoE of enrolled video sessions. The value of β determines the level of video quality. The values of coefficients α and δ are determined by video contents [6].

$$\beta = \alpha + (\frac{C_l}{\delta * n}). \tag{5}$$

Encoders that provide quality variability such as MPEG-4 can be used to produce different video quality from the video scenes. The rate controller (Fig. 1) adapts the transmission rate based on *Pro-IAAR(t)*. The load is monitored by the network monitor and *Pro-IAAR(t)* is estimated, the information is then sent back to the rate controller via the acknowledgment packet of Transport Control Protocol (TCP) Friendly Rate Control (TFRC) as an extension of TCP. TFRC can be utilized for this purpose. TFRC is a congestion control mechanism for unicast transmission over the Internet. In addition to fairness when competing with other flows, it has a much lower variation of throughput over time compared with TCP. This makes TFRC more suitable for applications which require smooth sending rate such as video streaming [37]. The significance of TFRC for media applications has been growing remarkably [34]. The rate controller selects a suitable video quality among available bit rates (video rate variants in Fig. 1) for each Group of Picture (GoP) based on the information on the network state received from the network monitor. An open loop Variable Bit Rate (VBR) controller requires access to the video content and network state information. The Explicit Congestion Notification (ECN) bit in the acknowledgment packet of the TFRC header is utilized for the purpose of network monitoring and thus no additional overhead is introduced. The rate controller at the sender side reduces its transmission rate by selecting a lower video rate variant if ECN 1 is detected in the acknowledgment packet.

The rate controller switches to the next rate by selecting the next quantizer scale at the start of the next GOP. This may delay the new rate up to the duration of one GOP. A leaky bucket can be used to control the target bit rate and allowed bit rate variability. It acts as a virtual buffer, therefore it does not introduce additional delay to video packets. Leaky bucket algorithms are widely used by rate controllers to control traffic to packet-switched and ATM-based networks [38].

The QoE-aware admission control component measures the network load and based on that makes the admission decision. The new requested session will be admitted only if the sum of *Pro-IAAR(t)* on the link plus x_{new} is less than or equal to C_l. The details of a possible scenario is explained in the next paragraph.

A video source prior to transmitting, sends a request to the ISP gateway indicating its intended sending rate (highest bit rate) as well as other possible bit rates (30 bit rates in total). Existing session signaling protocols such as Session Initiation Protocol (SIP) is currently used by Internet telephone calls and it also can be utilized for video distribution [39]. The gateway upon receiving the request, calculates $\mu_s(t)$ using Eq. (3), β using Eq. (5), *Pro-IAAR*(t) using Eq. (2) and checks *Pro-IAAR(t)* + $x_{new} \leq C_l$. The new session is accepted with its intended bit rate x_{new} only if the condition meets. If it does not however, the gateway checks the next bit rate (from higher to lower) that satisfies the condition. The gateway acknowledges the potential source should any other bit rate meets the condition which is then adopted by the source. If none of the bit rates satisfies the condition however, the request is rejected. The video sources are able to switch to a higher bit rate after they have been successfully accepted when bandwidth becomes available. Since only the acceptance/rejection admission policy was the target of this paper, post-acceptance bit rate switching was not addressed by our algorithm. The pseudocode for the implementation of the proposed QoE-aware cross-layer architecture for video admission is summarized in Algorithm 1.

Q7

Algorithm 1 Implementation of the QoE-aware cross-layer architecture for video admission.

Given C_l, n, α, *and* δ
1: **for** *Every video session request* **do**
2: *Compute* $\mu_s(t)$ *from* Eq. (3)
3: *Compute* β *from* Eq. (5)
4: *Compute* ϵ *from* Eq. (4)
5: *Compute Pro-IAAR*(t) *from* Eq. (2)
6: $k = 2$
7: $x_{new} = Highest\ bit\ rate\ (QP = k)$
8: **if** *Pro-IAAR*$(t) + x_{new} \leq C_l = True$ **then**
9: *Session accepted with rate* x_{new}
10: *Send the QP/k that satisfies accepted* x_{new}, *to the source*
11: **else**
12: **if** $k \leq 31$ **then**
13: *Increment* k
14: $x_{new} = Next\ bit\ rate\ (QP = k)$
15: *goto* **line 8**
16: **else**
17: *Session rejected*
18: **end if**
19: **end if**
20: **end for**

QoE is included into Algorithm 1 through parameter β which controls the total bitrate on a specific link based on the QoE of current sessions. On the other hand, the rate controller makes the architecture flexible by offering 30 different bit rates-with preference from high to low-assuming that they do not cause noticeable artifacts.

Table 2
Description of the video sequences used in this paper.

Description	Video sequence 1	Video sequence 2
Name	Mother and daughter (MAD)	Grandma
Description	A mother and daughter speaking at low motion	A woman speaking at low motion
Frame size	CIF (352 × 288)	QCIF (176 × 144)
Duration (s)	30	28
Number of frames	900	870

Table 3
Simulation parameters.

	Parameter	Value
Encoder	Frame rate (fps)	30
	GoP	30
	Video quantizer scale	2 (*non-adaptive architecture* traffic) 2–31 (*adaptive architecture* and *cross-layer architecture* traffic)
Network	C_l(Mbps)	32 (MAD) 7 (Grandma)
	β	0.9 (MAD, *cross-layer architecture*) 0.78 (Grandma, *cross-layer architecture*)
	VBR sources	24
	FTP sources	48
	Packet size (byte)	1052
	UDP header size (byte)	8
	IP header size (byte)	20
	Queue size (packet)	300 (MAD) 100 (Grandma)
	Link delay (ms)	1
	Queue management	Droptail
	Queue discipline	FIFO (First In First Out)
	Simulation time (s)	500

Algorithm 1 is jointly implemented by the video sources and ISP gateway relying on the available communication messages of the TCP/IP protocol suite for showing the interest to send, notification of the sender and network monitoring as explained earlier in this section. It therefore, does not demand additional requirements. We assume that each media content is encoded with 30 video rate variants. This allows for a wide range of playback rates (30) exploiting the capability of the ffmpeg encoder. The assumption is justifiable for video streaming services and the dropping cost of storage on media servers. Other studies have chosen videos files dynamically in response to channel conditions and screen forms under a limited storage budget through intelligent algorithms [40].

Using Big O notation metric, the complexity of Algorithm 1 is determined by counter k of the iteration loop in line 12 as well as fundamental operations in lines 2, 3, 4, 5, 6, 7, 8, 9, 10 and 17. This describes the worst-case scenario when the condition in line 8 is not satisfied. The time complexity of our algorithm is linear to the counter k, i.e.

$$T(k) = 10 + 1 + (k + 1) + 3k \quad (6)$$

Please cite this article as: Q.M. Qadir et al., A quality of experience-aware cross-layer architecture for optimizing video streaming services, Computer Networks (2016), http://dx.doi.org/10.1016/j.comnet.2016.02.030

Fig. 2. Snapshots of the video sequences used in this paper, MAD (left) and Grandma (right).

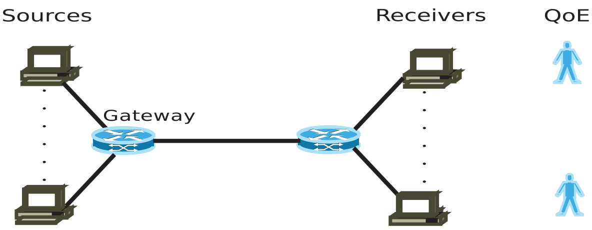


Fig. 3. Topology scenario considered in this paper.

$$T(k) = 12 + 4k \tag{7}$$

that is to say, $T(k) \sim O(k)$. The space complexity of the algorithm such as memory requirement, is insignificant due to the large storage capacity of modern routers.

Each of the on-line rate adaptation and QoE-aware admission control was implemented and investigated separately in [5] and [6], respectively. In this paper, the functionalities of both components are combined and evaluated within our architecture.

4. Evaluation environment

This section describes the settings of the evaluation environment for testing the performance of our architecture. Two video clips with different resolutions were used. The objective of having different video resolutions was to see the impact of video frame size on the performance metrics not to compare these two resolutions. The description of the video contents as well as coding and network parameters are shown in Tables 2 and 3, respectively.

NS-2 [41] was used to simulate the 30 s Common Intermediate Format (CIF) Mother And Daughter (MAD) and 28 s Quarter Common Intermediate Format (QCIF) Grandma video sequences shown in Fig. 2. In this paper, QCIF (176 × 144) and CIF (352 × 288) are specifically chosen as acceptable video formats for most video capable devices such as handsets, mobiles and videoconferencing systems delivered on telephone lines [6,42]. Whereas current mobile devices support bigger sizes, QCIF and CIF make packet level simulation practical. Other resolutions can be applied to the proposed architecture with different values of coefficients α and δ [6].

The topology shown in Fig. 3 with a bottleneck link was considered for evaluating the performance of the architecture. A maximum of (24) video sources were competing for the capacity of the link. As the video sources were always active in this paper, p_i was set to 1. There were also (48) File Transfer Protocol (FTP) sources active. The FTP sessions created background traffic and video sessions started randomly during the first 20–50 s of the simulation. The objective was to have a more realistic scenario where other traffic exists in the same network along with the video traffic. In total 500 s were simulated.

The proposed QoE-aware architecture (referred to as *cross-layer architecture*) was compared in details to an architecture (referred to as *adaptive architecture*) in which video sources adapt their bit rates only, while in the *cross-layer architecture*, the gateway implements the QoE-aware admission control in addition to the rate adaptation by video sources. Both architectures were then compared to a *non-adaptive architecture* in which the video flow is sent without rate adaptation and QoE-aware admission control. Similar simulation parameters and environment were used for the comparison.

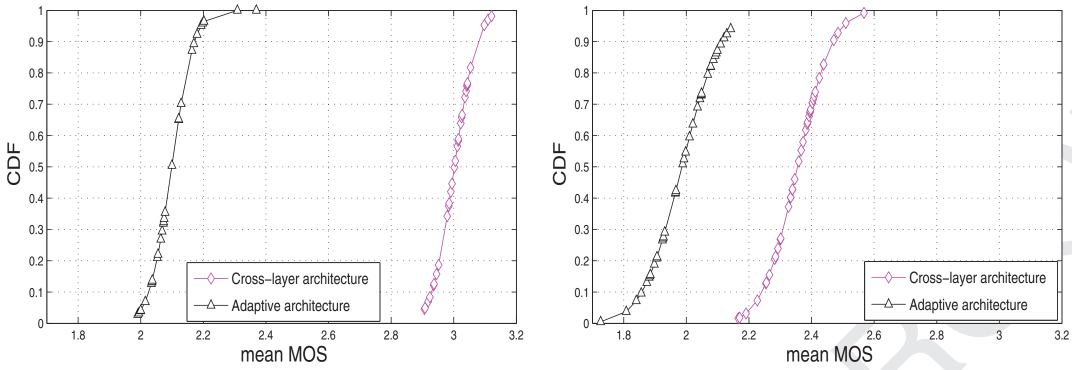
Evalvid-RA [34] was used to implement on-line rate adaptation from different encoded videos each with a valid range (2–31) of Quantization Parameter (QP). A lower QP generates a higher bit rate and better video quality. The MAD and Grandma video sequences were utilized by the NS-2 simulator through video trace files using EvalVid-RA. The *non-adaptive* videos were encoded with QP of 2 whereas the *cross-layer architecture* and *adaptive* videos with QP between 2–31 using ffmpeg encoder [43] (thirty video sequences with different bit rates).

The video sessions were competing for the C_l described in Table 3. The *cross-layer architecture* was configured so that new session was requested randomly within every second of the simulation time and accepted if there is enough bandwidth, i.e. the condition $Pro\text{-}IAAR(t) + x_{new} \leq C_l$ is satisfied. The arrival of new sessions in this manner avoids the possibility of having "flash crowd" phenomenon when numerous sessions arrive at the same time [44]. Whereas in the *adaptive architecture* and *non-adaptive architecture*, all sessions were admitted for each simulation run, this paper considered only video sessions that were successfully decoded and played back by the receiver (through ffmpeg decoder) as the metric *number of sessions*. Both alternative architectures allow for more sessions, but only those which are decoded and played back successfully

Table 4
Calculation of β.

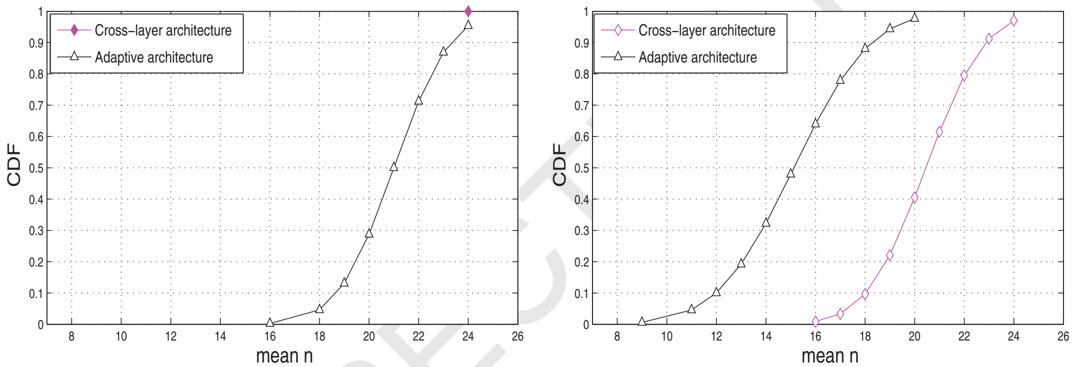
Video sequence	β (Experimental)	β (Eq. (5))	α	δ	C_l(Mbps)	Mean n
MAD (CIF)	0.9	0.84	−0.54	0.96	32	24
Grandma (QCIF)	0.78	0.775	−0.1	0.4	7	20

Please cite this article as: Q.M. Qadir et al., A quality of experience-aware cross-layer architecture for optimizing video streaming services, Computer Networks (2016), http://dx.doi.org/10.1016/j.comnet.2016.02.030



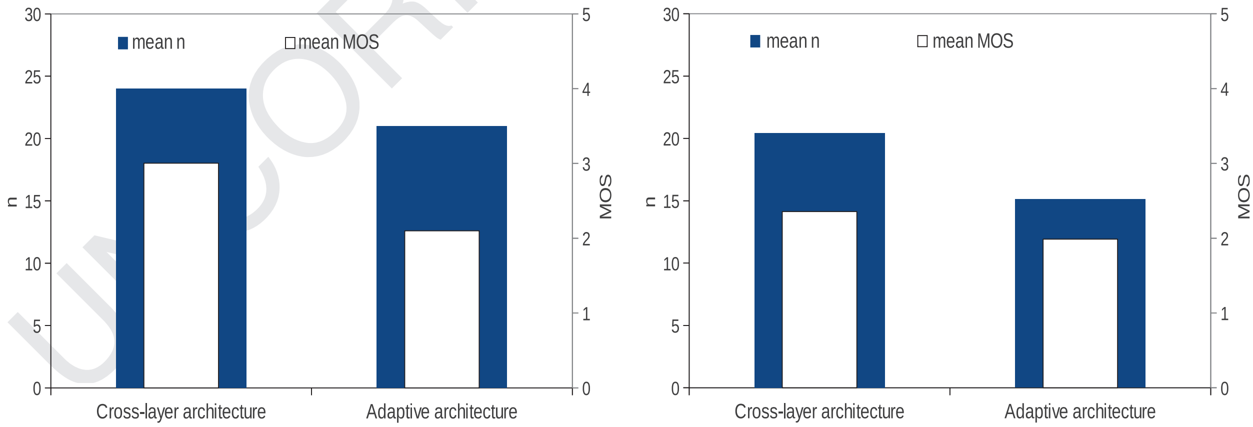
(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 4. CDF of the mean MOS of the video flows in the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences.



(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 5. CDF of the mean *number of sessions* in the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences.



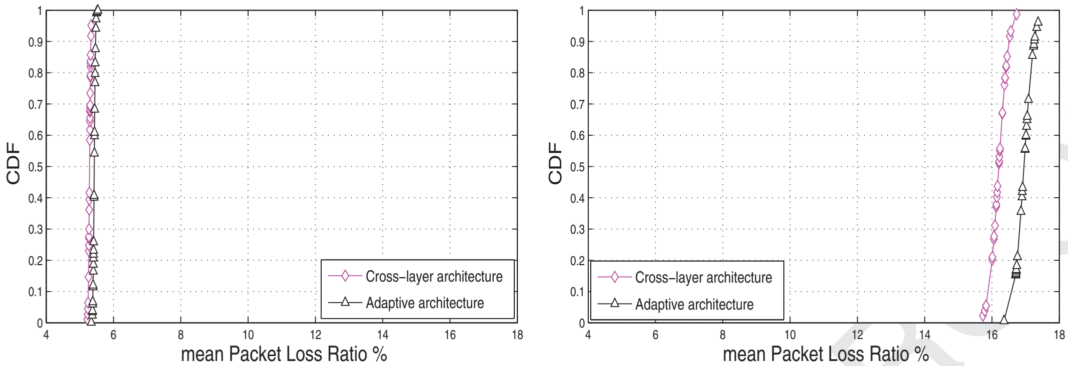
(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 6. Mean MOS of the video flows and mean *number of sessions* in the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.) Q8

by the receiver were taken into account. For simplicity, the maximum number of competing video sessions was limited to 24 sessions.

MOS, *n*, packet loss ratio and delay were measured as performance metrics. There are no significant jitter requirements for streaming video (the target traffic of this paper) [45]. The studied metrics for both resolutions are plotted next to each other for the sake of convenience not comparison. Cumulative Distribution Functions (CDF) of the means were calculated for the video flows for each metric over 30 runs. MOS was measured using Evalvid [46]. Evalvid provides a set of tools to analyze and

Please cite this article as: Q.M. Qadir et al., A quality of experience-aware cross-layer architecture for optimizing video streaming services, Computer Networks (2016), http://dx.doi.org/10.1016/j.comnet.2016.02.030



(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 7. CDF of the mean packet loss ratio of the video flows in the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences.

evaluate video quality by means of PSNR and MOS metrics. The Evalvid MOS metric used in this paper calculates the average MOS value of all frames for the entire video with a number between 1 and 5. This tool has widely been used for the similar purpose [47–50]. Parameter β was experimentally found to be 0.9 for the MAD video sequence and 0.78 for the Grandma video sequence. It was also calculated using Eq. (5). The values of coefficients (α and δ) were adopted from [6] (Table IV for MAD sequence and last paragraph of Section VIII for Grandma sequence). Experimental and calculated β are illustrated in Table 4.

5. Performance evaluation

In this section, the performance of the video flows in the *cross-layer architecture* is compared to the video flows in the *adaptive architecture* in terms of MOS, number of successfully decoded sessions and delay. Finally, a comparison between the video flows in the *non-adaptive architecture*, *adaptive architecture* and *cross-layer architecture* is made.

The CDF of the mean MOS of the video flows in the *cross-layer architecture* and *adaptive architecture* for both resolutions are plotted in Fig. 4. MOS enhancement of the video flows delivered through the proposed *cross-layer architecture* can be seen for both resolutions. The difference between the graphs shows that the result depends on the resolution. The mean MOS of the video flows in the *adaptive architecture* was enhanced by the *cross-layer architecture* from 1.98 to 2.35 for the QCIF resolution and from 2.09 to 3 for the CIF resolution. Although, the enhancement of the QCIF resolution can be considered minor, it is substantial for the CIF resolution as the MOS changes from bad to fair according to the absolute mapping in [51,52]. As the maximum possible MOS for any multimedia service in practice is 4.5 [14], even the slight enhancement of the QCIF MOS by the *cross-layer architecture* can make a difference.

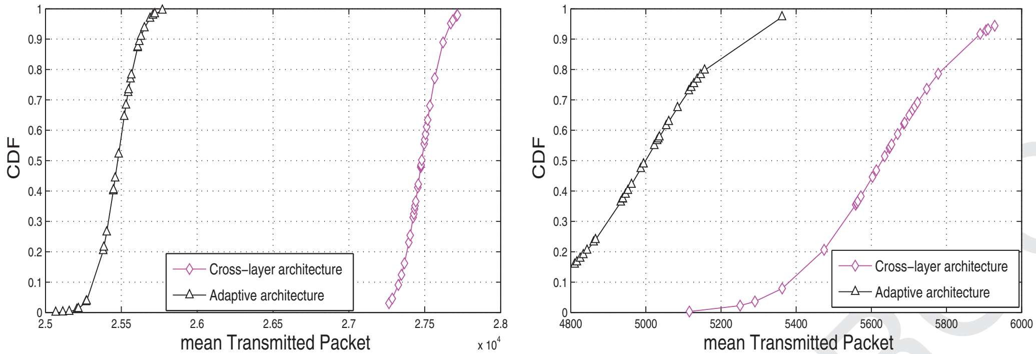
It is worthwhile to mention that the performance of the QoE-aware rate measurement algorithm and associated admission control were more pronounced in terms of MOS when they were evaluated among video flows only in [6]. In this paper, FTP traffic is included as a background traffic. Rate adaptation implemented by the video sources lets the video flows respond to the FTP flows by adapting their sending rates. This resulted in a lower MOS compared to the MOS of the video flows in [6] in which FTP flows were not considered.

As the main target of this paper is to optimize the QoE-*number of sessions* trade-off, we cannot consider the MOS of the video sessions alone. To account for this, the number of successfully decoded video sessions was measured for the *cross-layer architecture*, *adaptive architecture* and *non-adaptive architecture*. This is plotted for both resolutions in Figs. 5 and 12. Although, all 24 video flows were active in the *adaptive architecture* and *non-adaptive architecture*, an average of 15 QCIF/21 CIF sessions and 5.9 QCIF/19.9 CIF sessions were successfully decoded by the receivers respectively. This is due to the fact that being adaptive, the video sources in the *adaptive architecture* send data in cooperative manners. Thus not all the video frames were sent into the network due to insufficient bandwidth and availability of other traffic (FTP) in the network. In contrast, an average of 20 QCIF and all 24 CIF videos sessions were successfully decoded when delivered on the *cross-layer architecture*. Although the FTP flows again existed, the video sessions were better managed by the QoE-aware admission control and thus more sessions were accommodated.

It can also be noticed in Fig. 5 that the *number of sessions* in the *cross-layer architecture* is not resolution dependent as 5 more QCIF and 3 more CIF sessions are accommodated. As stated in Section 4, due to each resolution's specific simulation settings, the mean MOS and mean *number of sessions* of the two resolutions were not compared to each other.

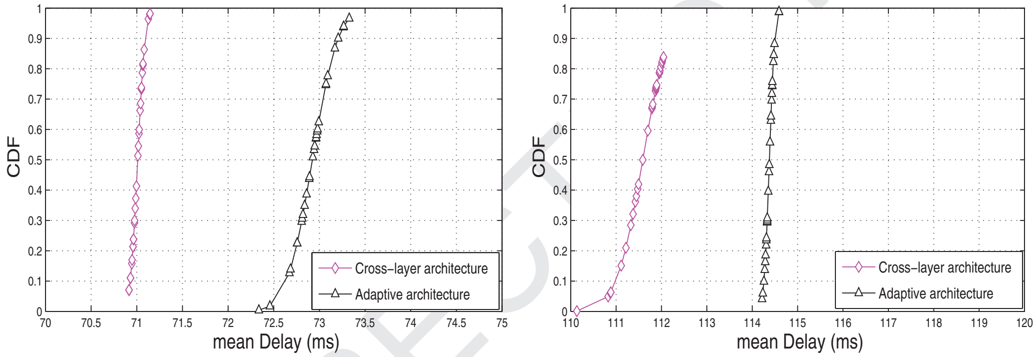
To compare the difference between the mean MOS of the video flows and mean *number of sessions* in the *cross-layer architecture* and *adaptive architecture*, both are plotted in the bar charts in Fig. 6. The white bars represent the mean MOS and blue bars represent the mean *number of sessions*.

Video streaming services are tolerant to packet loss to some extent. Error concealment in the decoder allows video to accept some tolerance of packet loss. We calculated the CDF of the mean packet drop ratio of the video flows in the *cross-layer architecture* and *adaptive*



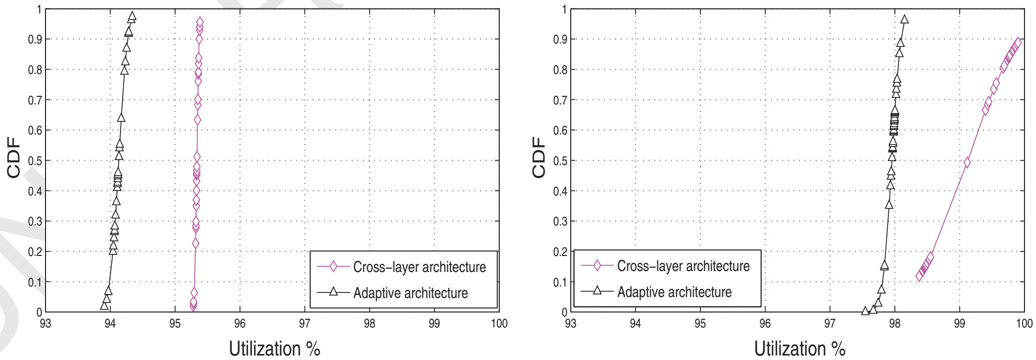
(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 8. CDF of the mean transmitted packet of the video flows in the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences.



(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 9. CDF of the mean delay of the video flows in the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences.

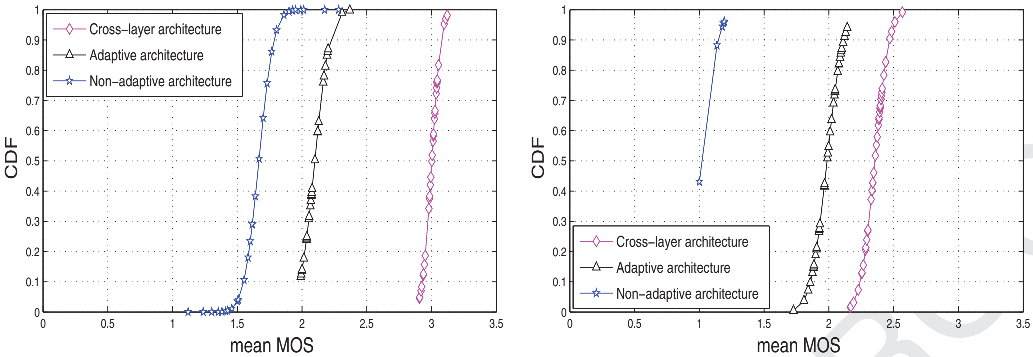


(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 10. Utilization of the *cross-layer architecture* and *adaptive architecture* for MAD and Grandma sequences.

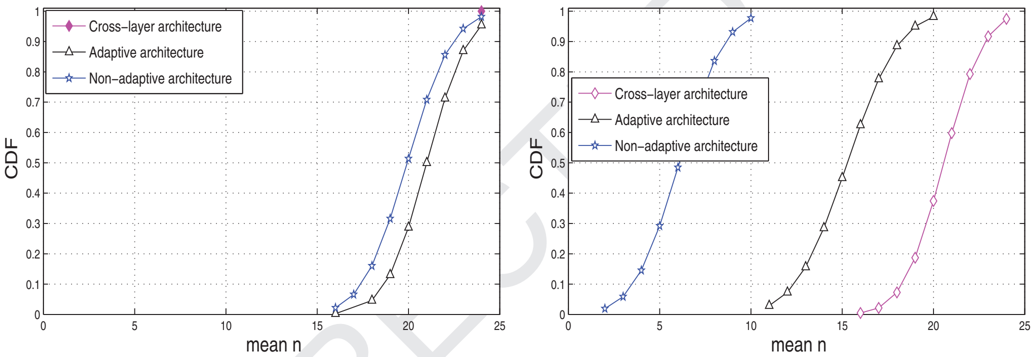
architecture and plotted them in Fig. 7. The video flows delivered over the *cross-layer architecture* experienced less packet drop compared to the video flows in the *adaptive architecture*.

In contrast to the substantial difference in the mean MOS as shown in Fig. 4, there is a small difference between the packet drop ratio of the video flows in the *cross-layer architecture* and *adaptive architecture* as can be seen in Fig. 7. However, these packets were dropped out of the total number of the transmitted packets. The CDF of the mean transmitted packet are shown in Fig. 8 in which the difference between the number of packets transmitted by the video sources in each of the *cross-layer architecture* and *adaptive architecture* is evident. Therefore, a smaller ratio of the packet loss of the video flows out of a higher number of transmitted packets of the same video content in



(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 11. CDF of the mean MOS of the video flows in the *cross-layer architecture*, *adaptive architecture* and *non-adaptive architecture* for MAD and Grandma sequences.



(a) MAD (CIF) (b) Grandma (QCIF)

Fig. 12. CDF of the mean *number of sessions* in the *cross-layer architecture*, *adaptive architecture* and *non-adaptive architecture* for MAD and Grandma sequences.

the *cross-layer architecture* compared to the *adaptive architecture* ensured a better quality (in terms of MOS) as discussed earlier in this section.

From Eq. (1), it is evident that a higher bitrate provides a better MOS for the same packet drop ratio. Sending a higher number of video packets by the *cross-layer architecture* compared to *adaptive architecture* as shown in Fig. 8 and a lower packet drop ratio as shown in Fig. 7 over the same simulation time (500 s), indicates that the video content was sent with a higher bitrate, thus a better MOS was provided by the *cross-layer architecture*.

Video streaming applications have a lenient delay requirement. Depending on the application's buffering capabilities, 4–5 s delay is acceptable [45]. The CDF of the mean delay of the video flows for each of the *cross-layer architecture* and *adaptive architecture* was measured. Fig. 9 show that the video flows in the *cross-layer architecture* experienced less delay compared to the video flows in the *adaptive architecture*.

The *adaptive architecture* utilizes the capacity of the bottleneck link less efficiently than the *cross-layer architecture* as can be observed in Fig. 10. Note that the utilization measure includes the FTP flows also. It is calculated as the number of transmitted bits over the capacity of the link for the simulation period. Thus, the *adaptive architecture* led to a high link utilization; 94% for MAD sequence and 98% for Grandma sequence. The utilization of the *cross-layer architecture* however, increased to 95% for MAD sequence and 99% for Grandma sequence. We can conclude that the utilization figures cannot decide the performance of the two architectures for the video flows as it is calculated for video and FTP flows.

Finally, the video flows delivered over the proposed *cross-layer architecture* are compared to the video flows transmitted by each of the *adaptive architecture* and *non-adaptive architecture*. Fig. 11 shows the CDF of the mean MOS of the video flows in the three architectures for both video resolutions. While, there is an improvement of the mean MOS of the video flows in the *adaptive architecture* through the adaptation of the sender rate compared to the video flows in the *non-adaptive architecture*, a higher value of the mean MOS of the video flows in the *cross-layer architecture* is observed.

Moreover, the proposed *cross-layer architecture* accepts and delivers a higher *number of sessions* compared to the other two architectures (*adaptive architecture* and *non-adaptive architecture*). This can be observed in Fig. 12. The bar chart in Fig. 13 illustrates the difference in the mean

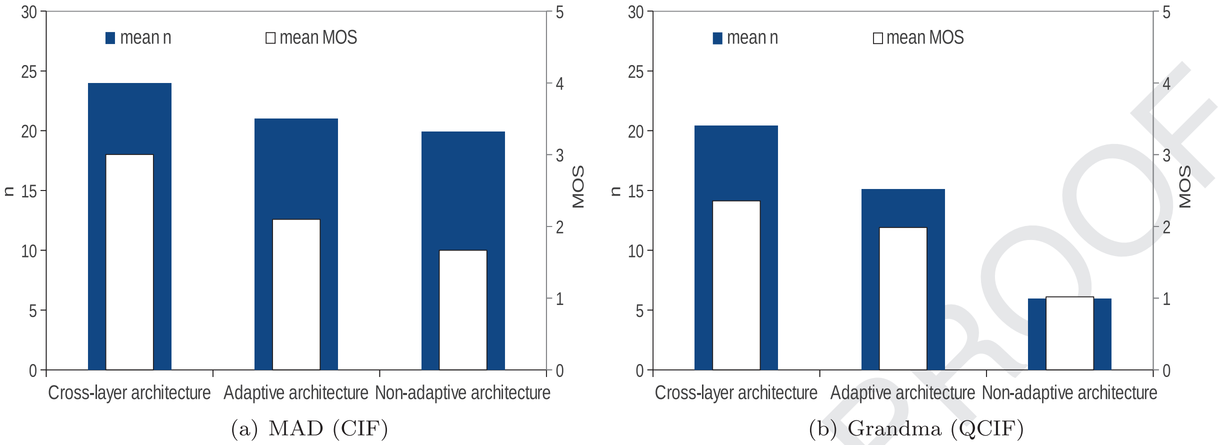


Fig. 13. Mean MOS of the video flows and mean *number of sessions* in the *cross-layer architecture*, *adaptive architecture* and *non-adaptive architecture* for MAD and Grandma sequences.

MOS of the video flows and mean *number of sessions* between all three architectures for both resolutions.

6. Conclusion

A QoE-aware *cross-layer architecture* for video streaming services was proposed in this paper. A combination of the rate adaptation and QoE-aware admission control are two main components of the architecture. The performance of the *cross-layer architecture* was analyzed and compared to two other architectures (*adaptive architecture* and *non-adaptive architecture*). The extensive simulation results have shown that the *cross-layer architecture* can provide an improvement in the mean MOS, considerably higher number of successful decoded video session, less mean delay and packet loss. At the same time it utilizes the link more efficiently. Evaluating the architecture with a greater variety of video contents and developing Algorithm 1 to include post-acceptance bit rate switching are interesting areas of future research. Future studies may also consider higher resolutions for the evaluation of the *cross-layer architecture*. Another interesting area of research is to have a dynamic number of video variants instead of transcoding each content into a fixed number of video files.

References

[1] Cisco documentation, Cisco visual networking index: Forecast and methodology 2013-2018, Cisco white paper, Jun. 2014. 2014.
[2] ITU-T Document FG IPTV-IL-0050, Definition of quality of experience (QoE) (Jan. 2007). 2007
[3] P. Brooks, B. Hestnes, User measures of quality of experience: Why being objective and quantitative is important, IEEE Netw. 24 (2) (2010) 8–13.
[4] B. Fu, D. Munaretto, T. Melia, B. Sayadi, W. Kellerer, Analyzing the combination of different approaches for video transport optimization for next generation cellular networks, IEEE Netw. 27 (2) (2013) 8–14.
[5] S. Qadir, A. Kist, Quality of experience enhancement through adapting sender bit rate, in: Proceedings of the 2013 IEEE TENCON Conference, 2013, pp. 490–494.
[6] Q.M. Qadir, A.A. Kist, Z. Zhang, A novel traffic rate measurement algorithm for quality of experience-aware video admission control, IEEE Trans. Multimed. 17 (5) (2015a) 711–722.
[7] Q.M. Qadir, A.A. Kist, Z. Zhang, Mechanisms for QoE optimisation of video traffic: a review paper, Australas. J. Inf. Commun. Technol. Appl. 1 (2) (2015b) 1–18.
[8] J.B. Ernst, S.C. Kremer, J.J. Rodrigues, A survey of QoS/QoE mechanisms in heterogeneous wireless networks, Phys. Commun. Part B (0) (2014) 61–72.Special Issue on Heterogeneous and Small Cell Networks Q9
[9] R. Maallawi, N. Agoulmine, B. Radier, T.b. Meriem, A comprehensive survey on offload techniques and management in wireless access and core networks, IEEE Commun. Surv. Tutor. 17 (3) (2014) 1582–1604. Q10
[10] P. Goudarzi, Scalable video transmission over multi-hop wireless networks with enhanced quality of experience using Swarm intelligence, Signal Process. Image Commun. 27 (7) (2012) 722–736.
[11] P. Goudarzi, M. Hosseinpour, Video transmission over MANETs with enhanced quality of experience, IEEE Trans. Consum. Electron. 56 (4) (2010) 2217–2225.
[12] I. Politis, L. Dounis, T. Dagiuklas, H.264/SVC vs. H.264/AVC video quality comparison under QoE-driven seamless handoff, Signal Process. Image Commun. 27 (8) (2012) 814–826.
[13] S. Khan, Y. Peng, E. Steinbach, M. Sgroi, W. Kellerer, Application-driven cross-layer optimization for video streaming over wireless networks, IEEE Commun. Mag. 44 (1) (2006) 122–130.
[14] S. Thakolsri, S. Khan, E. Steinbach, W. Kellerer, Qoe-driven cross-layer optimization for high speed downlink packet access, J. Commun. 4 (9) (2009) 669–680.
[15] A. Khalek, C. Caramanis, R. Heath, A cross-layer design for perceptual optimization of H.264/SVC with unequal error protection, IEEE J. Sel. Areas Commun. 30 (7) (2012) 1157–1171.
[16] X. Chen, J.-N. Hwang, C.-N. Lee, S.-I. Chen, A near optimal QoE-driven power allocation scheme for scalable video transmissions over MIMO systems, IEEE J. PP Sel. Top. Signal Process. (99) (2014a). 1–1
[17] S. Latré, Autonomic QoE Management of Multimedia Networks, Universiteit Gent, 2011 Ph.D dissertation.
[18] S. Latré, F. De Turck, Joint in-network video rate adaptation and measurement-based admission control: Algorithm design and evaluation, J. Netw. Syst. Manag. 21 (4) (2013) 588–622.
[19] C. Chen, X. Zhu, G. de Veciana, A. Bovik, R. Heath, Rate adaptation and admission control for video transmission with subjective quality constraints, IEEE J. Sel. Top. Signal Process. 9 (1) (2015) 22–36.
[20] C. Debono, B. Micallef, N. Philip, A. Alinejad, R. Istepanian, N. Amso, Cross-layer design for optimized region of interest of ultrasound video data over mobile WiMAX, IEEE Trans. Inf. Technol. Biomed. 16 (6) (2012) 1007–1014.
[21] C. Singhal, S. De, R. Trestian, G.-M. Muntean, Joint optimization of user-experience and energy-efficiency in wireless multimedia broadcast, IEEE Trans. Mob. Comput. 13 (7) (2014) 1522–1535.
[22] M. Rugelj, U. Sedlar, M. Volk, J. Sterle, M. Hajdinjak, A. Kos, Novel cross-layer QoE-aware radio resource allocation algorithms in multiuser OFDMA systems, IEEE Trans. Commun. 62 (9) (2014) 3196–3208.

[23] Y. Ju, Z. Lu, W. Zheng, X. Wen, D. Ling, A cross-layer design for video applications based on QoE prediction, in: Proceedings of the 2012 Fifteenth International Symposium on Wireless Personal Multimedia Communications (WPMC), 2012, pp. 534–538.
[24] K. Ivesic, L. Skorin-Kapov, M. Matijasevic, Cross-layer QoE-driven admission control and resource allocation for adaptive multimedia services in LTE, J. Netw. Comput. Appl. 46 (0) (2014) 336–351.
[25] L. Zhou, Z. Yang, Y. Wen, H. Wang, M. Guizani, Resource allocation with incomplete information for QoE-driven multimedia communications, IEEE Trans. Wirel. Commun. 12 (8) (2013) 3733–3745.
[26] A. Khan, L. Sun, E. Jammeh, E. Ifeachor, Quality of experience-driven adaptation scheme for video applications over wireless networks, IET Commun. 4 (11) (2010a) 1337–1347.
[27] M. Zhao, X. Gong, J. Liang, W. Wang, X. Que, S. Cheng, Qoe-driven cross-layer optimization for wireless dynamic adaptive streaming of scalable videos over HTTP, IEEE Trans. Circuits Syst. Video Technol. (99) (2014). 1–1
[28] A.E. Essaili, D. Schroeder, E. Steinbach, D. Staehle, M. Shehada, Qoe-based traffic and resource management for adaptive HTTP video delivery in LTE, IEEE Trans. Circuits Syst. Video Technol. (99) (2014).1–1
[29] M. Fiedler, H.-J. Zepernick, L. Lundberg, P. Arlos, M. Pettersson, Qoe-based cross-layer design of mobile video systems: Challenges and concepts, in: Proceedings of the International Conference on Computing and Communication Technologies RIVF '09, 2009, pp. 1–4.
[30] S. Latré, P. Simoens, B. De Vleeschauwer, W. Meersscche, F. Turck, B. Dhoedt, et al., An autonomic architecture for optimizing QoE in multimedia access networks, Comput. Netw. 53 (10) (2009) 1587–1602.
[31] O. Oyman, S. Singh, Quality of experience for HTTP adaptive streaming services, IEEE Commun. Mag. 50 (4) (2012) 20–27.
[32] J. Zhang, N. Ansari, On assuring end-to-end QoE in next generation networks: challenges and a possible solution, IEEE Commun. Mag. 49 (7) (2011) 185–191.
[33] B. Mathieu, S. Ellouze, N. Schwan, D. Griffin, E. Mykoniati, T. Ahmed, et al., Improving end-to-end QoE via close cooperation between applications and ISPs, IEEE Commun. Mag. 49 (3) (2011) 136–143.
[34] A. Lie, J. Klaue, Evalvid-RA: trace driven simulation of rate adaptive MPEG-4 VBR video, Multimed. Syst. 14 (2008) 33–50.
[35] S. Latré, R. Klaas, T. Wauters, F. DeTurck, Protecting video service quality in multimedia access networks through PCN, IEEE Commun. Mag. 49 (12) (2011) 94–101.
[36] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Am. Stat. Assoc. 58 (301) (1963) 13–30.
[37] S. Floyd, M. Handley, J. Padhye, J. Widmer, TCP friendly rate control (TFRC): protocol specification, RFC 5348, IETF, [Oline] <https://www.ietf.org/rfc/rfc3448.txt>, 2008. (accessed 07.11.15).
[38] H. Hamdi, J. Roberts, P. Rolin, Rate control for VBR video coders in broad-band networks, IEEE J. Sel. Areas Commun. 15 (6) (1997) 1040–1051.
[39] Rosenberg J., Schulzrinne H., Camarillo G., Johnston A., Peterson J., Sparks R., et al. SIP: Session initiation protocol. 2002.
[40] W. Zhang, Y. Wen, Z. Chen, A. Khisti, Qoe-driven cache management for HTTP adaptive bit rate streaming over wireless networks, IEEE Trans. Multimed. 15 (6) (2013) 1431–1445.
[41] NS-2, The Network Simulator, [Online] <http://www.isi.edu/nsnam/ns/⟩, 2015. (accessed 28.05.15).
[42] A. Khan, L. Sun, E. Ifeachor, Qoe prediction model and its application in video quality adaptation over UMTS networks, IEEE Trans. Multimed. 14 (2) (2012) 431–442.
[43] FFMPEG Multimedia System, [Online] <http://ffmpeg.mplayerhq.hu/>, 2004. (accessed 05.11.15).
[44] Y. Chen, B. Zhang, C. Chen, D.M. Chiu, Performance modeling and evaluation of peer-to-peer live streaming systems under flash crowds, IEEE/ACM Trans. Netw. 22 (4) (2014b) 1106–1120.
[45] T. Szigeti, C. Hattingh, End-to-End QoS Network Design: Quality of Service in LANs, WANs, and VPNs (Networking Technology), Cisco Press, Indianapolis, IN, USA, 2004.
[46] J. Gross, J. Klaue, H. Karl, A. Wolisz, Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming, Comput. Commun. 27 (11) (2004) 1044–1055.
[47] P. Papadimitriou, V. Tsaoussidis, SSVP: A congestion control scheme for real-time video streaming, Comput. Netw. 51 (15) (2007) 4377–4395.
[48] D. Li, J. Pan, Performance evaluation of video streaming over multi-hop wireless local area networks, IEEE Trans. Wirel. Commun. 9 (1) (2010) 338–347.
[49] K. Zheng, X. Zhang, Q. Zheng, W. Xiang, L. Hanzo, Quality-of-experience assessment and its application to video services in LTE networks, IEEE Wirel. Commun. 22 (1) (2015) 70–78.
[50] A. Khan, L. Sun, E. Ifeachor, Learning models for video quality prediction over wireless local area network and universal mobile telecommunication system networks, IET Commun. 4 (12) (2010b) 1389–1403.
[51] J.-R. Ohm, Multimedia Communication Technology, vol. 1, Springer, 2004.
[52] R. Stankiewicz, P. Cholda, A. Jajszczyk, Qox: what is it really? IEEE Commun. Mag. 49 (4) (2011) 148–158.

Qahhar Muhammad Qadir received the B.Sc. from Salahaddin University, Kurdistan Region, Iraq, in 1993 and the M.Sc. From IIUM, Kuala Lumpur, Malaysia in 2001. He is currently pursuing the Ph.D. degree at the University of Southern Queensland, Toowoomba, Australia. He is the CCNP, CCNA, MCSE and MCSA certified. His current research interests include green communication, wireless/mobile networks, QoS/QoE enhancement and multimedia quality assessment.

Alexander A. Kist received the Ph.D. degree in Communication and Electronic Engineering from RMIT University, Melbourne, Australia in 2004. He is an Assoc Professor (Telecommunications) and the School Coordinator (Learning and Teaching) in the School of Mechanical and Electrical Engineering in the Faculty of Health, Engineering and Science at the University of Southern Queensland, Toowoomba, Australia. His research interests include teletraffic engineering, performance modeling, remote access laboratories and engineering education. He has authored more than 90 scientific articles and from 2009 until 2011 he has lead the Faculty of Engineering and Surveying Remote Access Laboratory initiative. He is an elected executive member of the International Association of Online Engineering (IAOE) and the Global Online Laboratory Consortium (GOLC). He is a Steering Committee member of the annual Australasian Networks and Applications Conference (ATNAC).

Zhongwei Zhang received a B.S. degree from the Harbin Institute of Technology, China in 1986, and received the Ph.D. from Monash University, Australia, in 1998. His research interests include Computer Communications and networks modeling, Wireless networking and Simulations. He joined the University of Southern Queensland, Australia in 1999, and taught various courses on Computer communications and System Security ever since. He was involved in many research projects in application of Wireless Sensor networks in medical and health care, and in irrigation control systems for agriculture. He has published 64 articles in international journals and conferences, mainly in the field of computer communications network modelling, and wireless technology simulation. From 2008–2010, he has participated in the project "Innovative Application of Wireless Sensor Networks in Healthcare". From 2011 to 2014, he led a research group "Wireless Sensor Underground Networks and Its Application in Irrigation Control".