Computer Networks 185 (2021) 107638

Contents lists available at ScienceDirect

Computer Networks

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

Multipath MMT-based approach for streaming high quality video over multiple wireless access networks

Samira Afzal [a,*], Christian Esteve Rothenberg [b], Vanessa Testoni [c], Prakash Kolan [d], Imed Bouazizi [e]

[a] Department of Electronic Systems Engineering, Polytechnic School, University of São Paulo (USP), São Paulo, SP, Brazil
[b] School of Electrical and Computer Engineering (FEEC), University of Campinas (UNICAMP), Campinas, SP, Brazil
[c] Samsung Research Brazil (SRBR), Campinas, SP, Brazil
[d] Samsung Research America (SRA), Dallas, TX, USA
[e] Qualcomm, San Diego, CA, USA

ARTICLE INFO

Keywords:
Heterogeneous wireless networks
Multipath video streaming
Content-aware and path-aware
Path-awareness
Packet scheduling strategy

ABSTRACT

Demand for video streaming content and multimedia services have increased dramatically. Delivering high throughput and low delay video data are real challenges to provide and sustain high quality mobile video streaming services. One type of solution to provide quality of experience is to exploit multipath strategies by taking advantage of the multiple network interfaces that are currently available in most wireless devices. Our work aims at delivering contributions by exploring the advantages of multipath video streaming using the MPEG Media Transport protocol. MPEG Media Transport is a multimedia application layer protocol with the capability of hybrid media delivery. In this article, we propose a Content-Aware and Path-Aware (CAPA) scheduling strategy for this protocol, considering both video features and path condition. The performance of our CAPA proposal is evaluated using ns-3 DCE with different realistic multipath network scenarios. We evaluate the performance of CAPA over heterogeneous wireless networks under congested network and wireless lossy network conditions, which are common network situations with significant adverse effects on video quality. Our approach yields considerable video quality improvement in both scenarios compared to Path-Aware strategy and a simple scheduling strategy, called Evenly Splitting. For congested network scenario, CAPA could increase PSNR, respectively, by up to 4.25 dB (12.97%), and 7.22 dB (20.58%) compared to Path-Aware and Evenly Splitting strategies. It could also improve SSIM, respectively, by up to 0.033 (3.78%), and 0.102 (12.54%) compared to Path-Aware and Evenly Splitting strategies. For wireless lossy network scenario, CAPA increases PSNR, respectively, by up to 6.84 dB (20.30%), and 9.43 dB (30.32%) compared to Path-Aware and Evenly Splitting strategies. The proposed strategy also provides improvements in terms of SSIM, by up to 0.100 (12.72%), and 0.113 (14.23%) compared to Path-Aware and Evenly Splitting strategies, respectively. We also evaluate how proposed CAPA handles the streaming of different video bitrates under congested network conditions to provide sufficient video quality. Finally, we provide an initial validation of fairness of CAPA to confirm fair access to the resources available across all paths.

1. Introduction

Video streaming is a major part of the current network traffic. Applications such as 4K Ultra HD (UHD) and 8K UHD live or on-demand video, video conferences and online cloud games have become considerably popular over the past few years. High bandwidth demand technologies such as, Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR) as well as Multi-View Video (MVV) have evolved fast and become true in the near future. According to the annual Cisco's report [1], IP video traffic would take 82 percent of all consumers' Internet traffic by 2021. However, the real problem behind these technologies is the huge bitrate they produce [2], since it is insufficient to transmit it through the current available IP based networks. Delivering high-quality video streaming services makes the task of providing real-time wireless transmission of multimedia while ensuring Quality of Experience (QoE) quite challenging due to bandwidth and time constrains [3].

Several solutions have been proposed to provide QoE for video streaming. One type of approach is packet loss resilient methods [4,5]

* Corresponding author.
E-mail address: afzal.samira@usp.br (S. Afzal).

https://doi.org/10.1016/j.comnet.2020.107638
Received 28 April 2020; Received in revised form 2 August 2020; Accepted 19 October 2020
Available online 29 October 2020

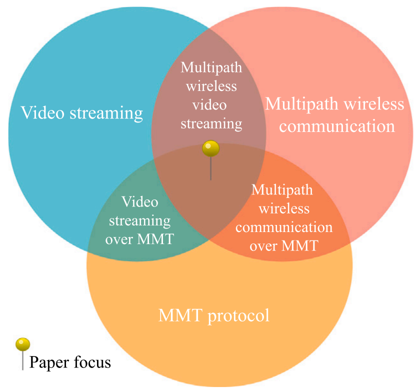


Fig. 1. The focus of this paper is the confluence of MMT protocol for video streaming applications over heterogeneous wireless networks.

such as automatic repeat request (ARQ), Forward Error Correction (FEC), and Error Resilient Coding (ERC), which cope with noisy networks by reducing the effect of data loss. Adaptive streaming mechanisms [6–8] are also remarkable options to dynamically adjust the video delivery data rate to the underlying network conditions.

Another line of research in the literature to improve QoE is Network Coding (NC). NC proves its benefits in network performance, especially in lossy wireless networks and multicast scenarios [9]. In NC, packets from the same flow (intra-session) or different flows (inter-session) can be combined into a single packet and transmitted [10]. One significant feature of NC is that it allows intermediate nodes to perform re-encoding of packets leading to improved throughput [10–12]. Besides that, NC outperforms simple retransmission strategies, specifically for multicasting [13–15], since it reduces the number of packets transmission. However, NC solutions are incompatible with some protocols such as TCP [9], add extra coding and decoding requirements, and impose huge challenges when trying to combine packets from different flows with different transmission rates.

Another type of solution is to exploit multipath strategies [6,16,17] by taking advantage of the multiple network interfaces, which also refers to multihoming, that are currently available in most wireless devices (e.g., laptops, tablets, smartphones). A user connected to more than one network opens the opportunity to leverage different video delivery strategies to provide better coverage and overall more stable network connectivity by circumventing congested network paths and aggregating the bandwidth available over multiple paths. One concept presented in 5G to reach the target of supporting services with extreme bandwidth and ultra-low latency requirements is by exploiting multihoming capabilities.

Multiple efforts have been made regarding multipath data transmission [18,19], and different multipath protocols (e.g., MPTCP and CMT) have been proposed. However, most of the current transport protocols do not match the requirements of video streaming applications or are not designed to address relevant issues, such as delay constraints, network heterogeneity, and head-of-line blocking issues.

At the crossroads, we observe MPEG Media Transport (MMT) [20], as a popular multimedia protocol standardized in 2014 as a part of MPEG-H standard suite [21]. Several other standards have already adopted MMT for major advances in televisual technology worldwide [22,23]. MMT supports UHD, on-demand, live video streaming, and it has also been widely used for VR, AR, and MVV technologies. We regard MMT as an appropriate protocol for exploiting multipath streaming because it derives the capability of the video transmission in heterogeneous network environments. MMT has the principal property of hybrid media delivery, which refers to the combination of delivered media components over different types of networks. For example, one could be a broadcast channel, and another broadband or two simultaneous broadband channels can be combined.

Altogether, as shown in Fig. 1, the focus of this paper is on improving the MMT protocol with multipath strategies to deliver improved video quality of experience for real-time wireless video streaming. In a nutshell, we propose investigating novel multipath scheduling strategies that consider both video features (content-aware) and path characteristics (path-aware). Considering video content features in the scheduling strategy helps to define the priority of each packet, and subsequently, unequal importance packets can be sent through different network paths based on network-level quality estimators. CAPA improves video streaming QoE by increasing goodput, decreasing packet losses, and end-to-end delay.

This article is an extended version of our previous work [16] published in 2018. We would like to point out the main extended contributions to the original work in the following.

- While only congested network situation was considered in our original work, we also study wireless lossy network situation, where channels have high burst losses due to wireless errors (e.g., lossy channels, noise or interference), in this current paper. According to [24], losses due to lossy network situation could even occur more than congestion condition in the real network. Since such burst wireless losses have a high adverse effect on the perceived video quality, it is well worth to consider it. Therefore, the proposed scheduling strategy defined in the original work has been improved considering lossy path condition in the Markov model and related packet distribution strategies in order to guarantee video quality in wireless lossy condition.
- In the original work, CAPA experimental results were only compared with a simple scheduling strategy for the traditional multipath MMT (ES). However, in this current paper, besides ES, we also define a Path-Aware scheduling strategy (PA) to compare with CAPA. PA considers only the paths' condition for packet distribution. Therefore, result comparison with PA makes clear how much content-aware protection contributes to the overall performance gain.
- While we validated the proposed strategy with only two cartoon sequences in the original work, in this current paper, it is validated with a more broad range of video sequences including three cartoons and two natural scene video sequences with motion, details and very tough colors. This is due to this reason that content features of the streamed video sequence have a significant impact on the perceived video quality. Therefore, a proper scheduling strategy should cope with different contents.
- This article provides more scenarios compared to the original work considering different network situations in order to evaluate our proposed strategy, CAPA. First, we consider the congestion network scenario, which could significantly decrease the perceived video quality. We also study our scheduler behavior with different video qualities (bitrates) under our congested network environment. Then, we consider wireless channel lossy networks, where channels have high burst losses. Finally, we evaluate our proposed strategy when it competes with another MMT flow for fairness support.
- As an extension to the original paper, a new section on related work, and also a comprehensive discussion on the experimental results have been added.

The structure of this article is organized as follows. In Section 2, we briefly discuss main aspects of scheduling functions and review the related work to this research. Section 3 details our developed solution and the multipath improved MMT system. Section 4 presents the evaluation of the proposed solution over heterogeneous wireless networks and also provides our contribution to tackle the challenge of scheduling. Finally, in Section 5, we conclude the paper by pointing out the main findings, contributions, and directions.

2. Related work

In this section, we discuss the main aspects of the scheduler functions for multipath video streaming schemes and also provide an overview of notable related work.

Research studies on the wireless multipath video streaming [25] shows that to design a suitable multipath approach for mobile video streaming, several factors have to be considered. The first one is layer dependency. The scheduler should access the accurate measurements and the accuracy level of information, which depends on the layer. For example, the application layer is aware of video features, player buffer, and deadlines. The transport layer is able to calculate the bandwidth and RTT, and it also has a congestion control mechanism. The network layer accesses the IP level, and the link-layer has wireless parameter access. The interaction between layers is known as cross-layer. Mostly, in cross-layer approaches, lower layers (network or below) gather network information and feed them to higher layers (application or transport layers). Therefore, the higher layer could decide the path for data distribution and manage load balancing, or apply a method to save energy. In this work, we focus on the application layer because of its influence in path selection, and we use the application feedback for optimal delivery of the media data.

Table 1 shows some attempts available for the application layer based on the original protocols. In this table, MPRTP [26] is based on the RTP protocol, and therefore, it has no congestion control, and it is unfair to give room to other flows. The work in [27] manages DASH content delivery at the subsegment level and provides multipath delivery at the client-side. Go et al. [28] proposed a hybrid TCP/UDP-based enhanced. HTTP adaptive streaming and a MPEG-DASH-based enhanced system for multihomed mobile devices. The client determines types of transport protocols by analyzing MPD information together with the estimated network condition and buffered video time. When UDP is selected to transfer data, an adaptive Raptor code is also used to provide reliable data transmission. The work in [29] proposes a synchronization scheme for video streams transported over hybrid delivery: a combination of MMT (for broadcasting) and HTTP (for broadband) streaming. In their experiment, the base and first enhanced layers are delivered over the broadcast channel, and the second enhanced layer is delivered over broadband networks. The synchronization scheme is implemented at the receiver side, and the receiver could request the segments that they can deliver on time. However, the approach does not use any scheduling strategy to manage the paths. QUIC-FEC [30] is a FEC extension to QUIC protocol. The proposed scheduler, named HighRB, can perform path interleaving or only using one path when using both paths is harmful. HighRB selects a path randomly by using the number of remaining bytes computed by the congested window as weights for the random selection. The scheme in [31] is an energy-aware bandwidth aggregation middleware for video streaming over HTTP. Middle-ware approaches are designed to enable multipath interfaces to current applications without any application modification. Therefore, middle-ware approaches are easy to deploy. However, they are complex to implement [32]. MP-H2 [33] uses an HTTP-based multipath scheduling solution providing server transparency, middlebox compatibility, and load balancing. It focuses on minimizing the transfer time of a medium to the large size of a Dropbox file, video chunks, mp3 song, and an image. This way, the authors implemented a client-based scheduler on top of HTTP/2, which is responsible for determining when and which chunk should be fetched over which path using file size and network condition for its decision making. Except MPRTP, all aforementioned approaches are not content-aware, and their performance could be improved by considering video content features for path management. In this work, we have improved MMT to make it aware of the content and path situations.

Another important factor in designing a scheduler is network equipment compatibility, to present which part of the network has to be modified (client and/or server and/or network) in order to become compatible with the multipath transmission solution. The most flexible case for implementation is where only client modification is required. Table 1 indicates which parts of the network equipment need to be adjusted to become compatible with related multipath transmission schemes. In this work, our proposed scheduling strategy requires both server and client sides modification.

To optimize the scheduling performance, a scheduler can take into consideration three functions: content awareness, path awareness and protection method. Actually, a key characteristic of video content is that packets have unequal importance. In particular, there is a strong interdependency among video frames in each Group of Pictures (GOP), which causes a compressed video to be very sensitive to data loss. Each GOP consists of one initial Intra (I)-frame, several Predicted (P)-frames, and possibly Bidirectional (B)-frames. While an I frame is encoded without reference to any other video frames, but a P frame is encoded with reference to previous I or P frames, and a B frame is encoded with reference to both immediate previous and forward I or P frames. Therefore, in the decoding process, the loss of some frames may preclude a proper decoding, specifically in the miss of I frames. To ensure receiving of I frames at user, one solution is prioritizing packets and transmitting them over different network paths based on paths' quality. This way, it is possible to meet real-time deadlines, circumventing path heterogeneity issues, and improve QoE. Also, a content-aware scheduling strategy could utilize stronger packet protection for higher priority packets.

Table 1 shows which scheduler functions are used for each of related work. In our work, all those three essential functions are considered. This way, our proposed packet scheduler strategy prioritizes packets by making the scheduler aware of the video content, and protects packets using duplication method and assigning the best-qualified paths for packet distribution.

Among the related research summarized in Table 1, MPRTP [26] is the most similar solution to our work since it considers video data content and network condition for traffic splitting as well as packet protection technique to decrease data loss rate. However, our approach differs from MPRTP in a series of aspects. For instance, MPRTP monitors and controls network metrics, and as a result, paths are categorized in different conditions based on the packet loss information. Differently, in CAPA, we propose to model the path condition estimation problem as a three-state Markov model, and in order to define the path condition state, several metrics such as one-way delay, packet loss rate, the standard deviation of one-way delay, the weighted moving average of one-way delay, are used. After paths' condition estimation, while MPRTP only prioritize I packets among others to transmit through the best path (non-congested), CAPA has different priority levels of I, NI, and P packets as well as different packet scheduling strategies. For example, I and/or NI packets are transmitted through the best path, and P packets may be discarded to reduce congestion when the path is in bad condition (high congested networks). There is also packet duplication to protect I or/and NI packets when the network path is lossy or congested. Instead of duplication, MPRTP utilizes retransmission technique while prioritizing I packets. For more details and background, we direct the interested reader to a recent holistic literature survey of multipath wireless video streaming [25].

3. Advancing MMT for multipath wireless video streaming

3.1. Multipath MMT system model

A diagram overview of the proposed multipath MMT system considered in this paper is presented in Fig. 2. The goal of this system is to achieve a high-quality video streaming solution for the MMT protocol by adopting scheduling strategies considering path conditions and video content features. This system is completely defined at the application layer. We define a unicast video transmission system considering the multipath data transmission of a single video flow over

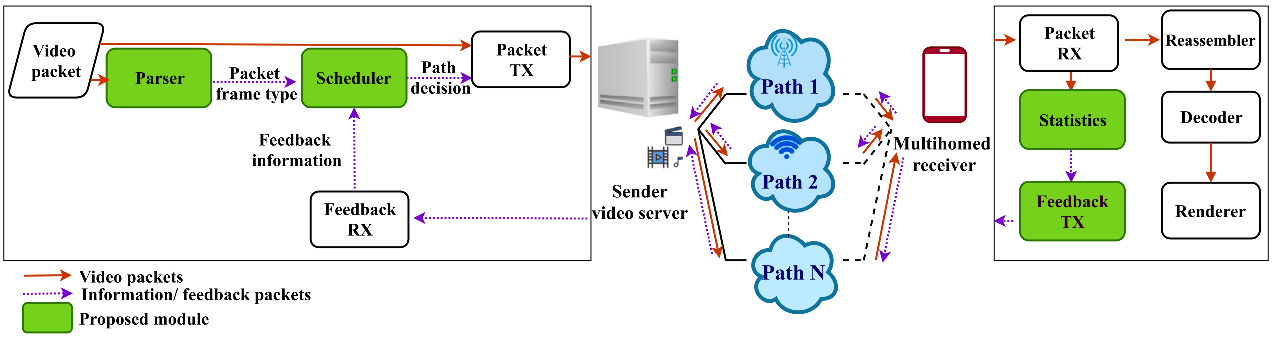


Fig. 2. Diagram overview of the proposed multipath MMT system.

Table 1
State of the art comparison. More details can be found in [25].

Applied protocol	Work	Compatibility	Scheduler function
RTP	MPRTP [26]	Server and Client	Content-aware, path-aware, Protection methods
DASH	Houze et al. [27]	Client	Path-aware, Protection methods
	Go et al. [28]	Server and Client	Path-aware, Protection methods
MMT	Sohn et al. [29]	Server and Client	Protection methods
QUIC	QUIC-FEC [30]	Server and Client	Path-aware, Protection methods
Other adaptive streaming approaches	GreenBag [31]	Client	Path-aware, Protection methods
	MP-H2 [33]	Client	Path-aware, Protection methods

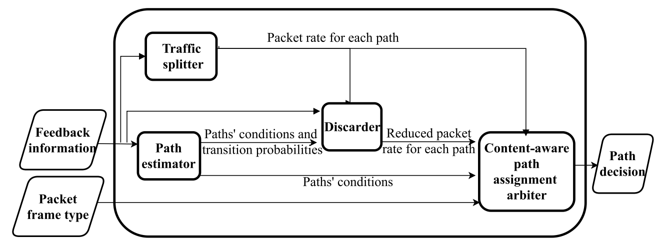


Fig. 3. Overview of the proposed scheduler module architecture.

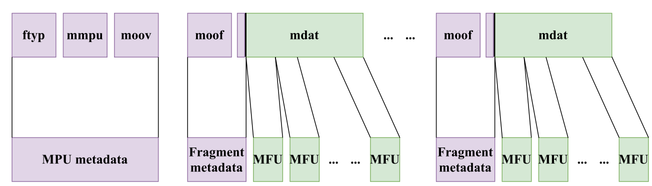


Fig. 4. Structure of ISOBMFF file for a fragmented video sequence.
Source: Adapted from [20].

multiple access networks. Each network is modeled as an independent end-to-end communication path, and UDP is employed to transmit video data.

At the sender side, there are different modules, namely parser, feedback RX, scheduler, and packet TX. First, each video packet is handled by the parser module to extract the frame type of video content. This information would be informed to the scheduler module and later would be used for path decisions by it. Additionally, feedback RX module receives feedback packets periodically from the receiver. This module discards the overdue received feedback packets and informs the updated paths' information to the scheduler. In sequence, the scheduler module evaluates the path metrics and uses them together with the frame type information to assign a proper path for each video packet based on its own strategy. Finally, the packet TX module transmits the video packet through the decided path.

At the receiver side, existing modules include packet RX, statistics, feedback TX, reassemble, decoder, renderer. After receiving packets by the packet RX module, statistics (such as goodput, average delay, number of lost packets, and jitter) should be generated via the statistics module periodically (after each predefined time interval) in order to monitor the quality of the transmission paths. This information should be sent to the sender in a feedback packet by assembling Reception Quality Feedback (RQF) and Network Abstraction for Media Feedback (NAMF) signaling messages defined in MMT standard [20]. Since feedback information is very important and has a high effect on the sender scheduler decision, it is necessary to send it over the best path. For this reason, a scheduler is implemented at the receiver side as well. If the received packets by the packet RX module are not overdue, then they would be sent to the reassembler module to reassemble the encoded video bitstream, and consequently, to the decoder module for error concealment and decoding. Finally, the decoded video is ready to display by the video renderer module.

3.2. Video parsing

As previously explained in Section 2, understanding the frame type of each video packet content and considering this information in scheduling strategy helps to choose the frame packets with higher priority to transmit through more qualified paths or to protect them during transmission. In our multipath MMT system, the parser module is responsible for extracting the frame type information of each MMT video packet content by accessing the ISOBMFF structure of the fragmented video and parsing the first bits of each fragment. The ISOBMFF structure of a fragmented video sequence is shown in Fig. 4. As it is shown in this figure, the MPU metadata consists of 'ftyp', 'mmpu', and 'moov' boxes. It can also contain any other boxes that are applied to the whole MPU. Fragment metadata consists of 'moof' box and 'mdat' box header. The video data is then split into multiple data units of MFU in 'mdat' box. Therefore, the parser module, implemented in this work, looks for the 'moof' box and with this information, it reaches to the content of fragment, which is stored in the 'mdat' box, and then it can access to the frame type information available there. This information is sent to the scheduler to be used for path decision.

3.3. Content-aware and path-aware scheduling proposal

An overview of the proposed scheduler module architecture is presented in Fig. 3. This proposed scheduler module is composed of four

Table 2
Definition of parameters.

Symbol	Definition
N	Total number of transmission paths
bw_p	Maximum bandwidth of path p [kbps]
$delay_p$	Minimum delay of path p [ms]
gp_p	Goodput of path p [Kbps]
d_p	Average one-way delay of path p [ms]
GDD_p	GDD (goodput-division-delay) of path p
λ_p	The packet rate split factor for path p
P	A matrix of transition probabilities among the states
p_{ij}	Transition probability from each state (i) to state (j)
C	A weighted matrix to store c_{ij}
c_{ij}	The number of transitions from each state (i) to state (j)
$d_{p,wma,cur}$	Current weighted moving average of one-way delay of path p
$d_{p,wma,pre}$	Previous weighted moving average of one-way delay of path p
$\sigma_{d_{p,cur}}$	Current standard deviation of one-way delay of path p
$\sigma_{d_{p,pre}}$	Previous standard deviation of one-way delay of path p
T_l	Packet loss rate threshold [%]
T_d	One-way delay threshold [ms]
D_p	Highest one-way delay of path p [ms]
L_p	Packet loss rate of path p

sub-modules: traffic splitter, path estimator, discarder and content-aware path assignment arbiter.

Traffic splitter utilizes feedback information to properly split the video traffic, deal successfully with each path capacity and its current conditions in order to avoid either congestion network situation or resource underutilization, and to perform load balancing. Therefore, for each feedback information received as input, traffic splitter applies the adaptive traffic split scheme proposed in Section 3.3.1 to assign a packet rate to each path. The process of this proposed adaptive traffic split scheme is depicted in the flowchart of Fig. 5.

Path estimator also uses the feedback information to adaptively classifies each path condition as good, mild/lossy, or bad by following the scheme described in Section 3.3.2. A path is classified as good, mild, or bad when it is under congestion, and it is estimated as lossy when the channel has burst wireless losses due to wireless errors (e.g., noise or interface). The process of the proposed path estimation is detailed in the flowchart of Fig. 7.

When the path is in mild or bad condition, the probability of losing packets is high due to network congestion. Therefore, a discard strategy, described in Section 3.3.3, is defined to apply on the input data to find a packet discard rate for each path. This strategy avoids sending packets that would probably be lost, and this way decreases the congestion. However, when a path is in lossy condition, there is no need to apply the discard strategy. The process of the proposed discard strategy is summarized in Fig. 8.

Finally, the content-aware path assignment arbiter receives all the calculated information by modules above; the packet rate, the packet discard rate, and paths' conditions. Then, it applies the defined content-aware strategy to decide the best path for each packet considering its content importance. In this work, I frame packets are the most important ones, and after that, packets of N frames close to the I frame (NI frames) are considered as important packets due to their effect on the video quality [34]. Our proposed content-aware strategy follows the rules described in Section 3.3.4 and illustrated as a flowchart in Fig. 9. The mathematical notations used throughout this work are summarized in Table 2.

3.3.1. Adaptive video traffic split

The proposed video traffic split strategy is illustrated in Fig. 5 and is based on a goodput-division-delay (GDD) metric. Scheduler calculates GDD for each path after receiving each feedback packet as $GDD_p = \frac{gp_p}{d_p}$. The packet rate split factor for each path (λ_p) is then calculated as

$$\lambda_p = \frac{GDD_p}{\sum_{i=1}^{N} GDD_p}, \tag{1}$$

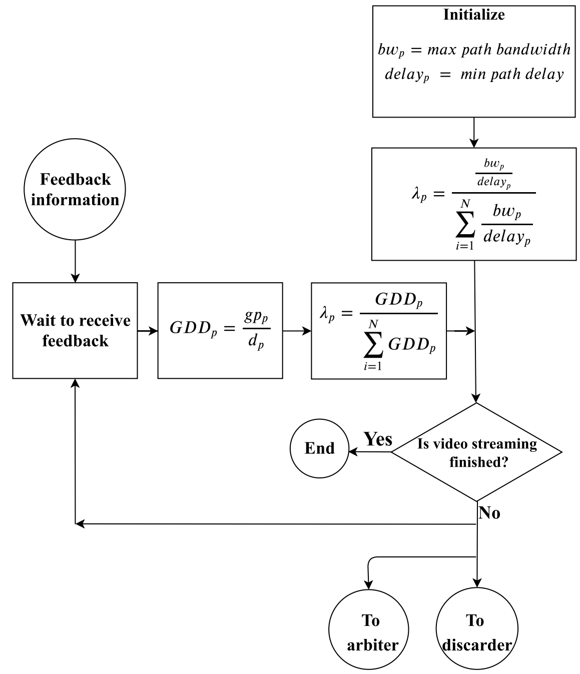


Fig. 5. Flowchart of adaptive video traffic split strategy.

As an initial estimation, when feedback packets were not yet received by the scheduler, λ_p is computed as

$$\lambda_p = \frac{\frac{bw_p}{delay_p}}{\sum_{i=1}^{N} \frac{bw_p}{delay_p}}, \tag{2}$$

Therefore, higher packet rate is assigned to the path with higher goodput and lower one-way delay. Another solution considering resource availability is defined in [35].

3.3.2. Estimation of path condition

We propose to model the path condition estimation problem as a three-state Markov model where each state represents one path condition: Good Condition (**GC**), Mild/Lossy Condition (**MC/Lossy**) and Bad Condition (**BC**), as shown in Fig. 6. In this model, Matrix P and C are dynamically calculated by scheduler after receiving each feedback packet. Matrix P consists of transition probabilities among the three path states and Matrix C is used to store the number of transitions from each state i to state j (c_{ij}). Following [36], the elements of matrix P are computed by the following equation:

$$p_{ij} = \frac{c_{ij} + 1}{\sum_{j=1}^{N} c_{ij} + N}, \tag{3}$$

Two thresholds are defined to determine paths' condition states: T_d for one-way delay and T_l for packet loss rate. In this work, T_d is set as 50 ms following recommendation in [37], which is the maximum delay to reach high video quality. T_l is set as 2% inspired by the work in [38]. Chow et al. [38] proposed a multipath streaming scheme (EMS) combined with FEC (Forward Error Correction) scheme. The work states that the packet loss rate should be less than 1% for H.264 video encoding in order to guarantee high quality real-time live video streaming. Since FEC is not used in this work, this limit was slightly increased to 2%.

The following two metrics, specified in [39], were also computed dynamically after receiving each feedback packet and used in this work:

$$d_{p,wma,cur} = \frac{31}{32} \cdot d_{p,wma,pre} + \frac{1}{32} \cdot d_p \tag{4}$$

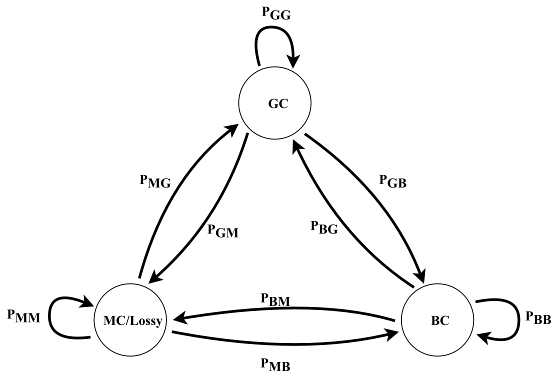


Fig. 6. Three-state Markov model used for estimation of path condition.

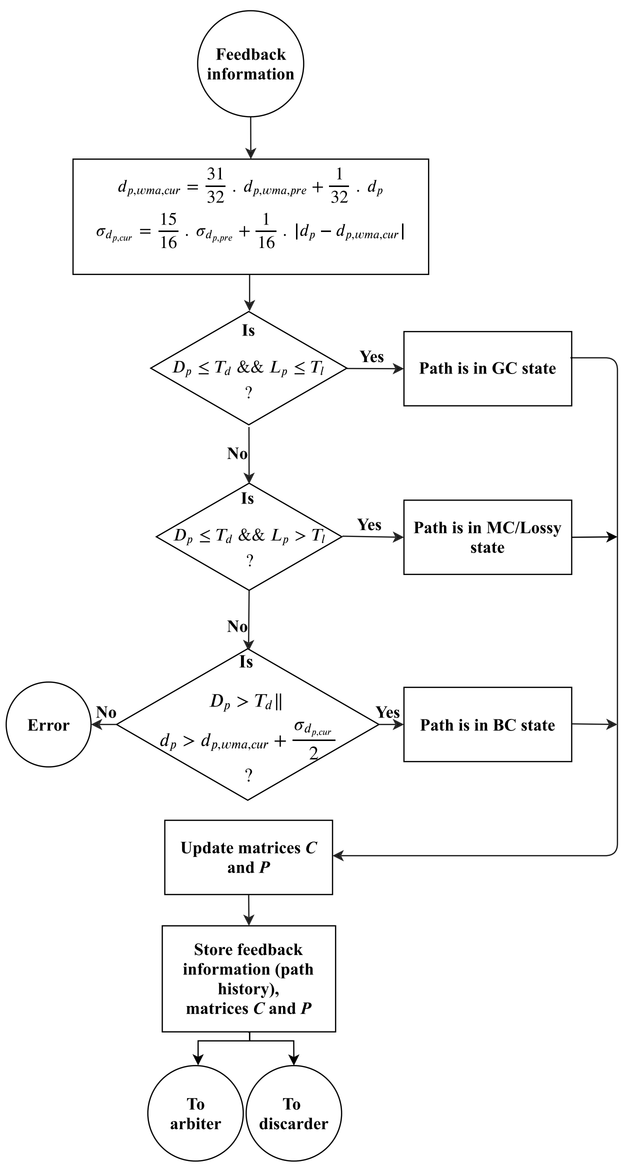


Fig. 7. Flowchart of paths' conditions estimation strategy.

$$\sigma_{d_{p,cur}} = \frac{15}{16} \cdot \sigma_{d_{p,pre}} + \frac{1}{16} \cdot |d_p - d_{p,wma,cur}| \tag{5}$$

Then, following rules are defined to estimate paths' conditions using combination of the two thresholds (T_d and T_l) and two computed metrics ($d_{p,wma,cur}$ and $\sigma_{d_{p,cur}}$):

- Path is in **GC** state if $D_p \leq T_d$ && $L_p \leq T_l$;
- Path is in **MC/Lossy** state if $D_p \leq T_d$ && $L_p > T_l$;
- Path is in **BC** state if $D_p > T_d \;\|\; d_p > d_{p,wma,cur} + \frac{\sigma_{d_{p,cur}}}{2}$.

The flowchart in Fig. 7 illustrates the process of the path conditions estimation strategy.

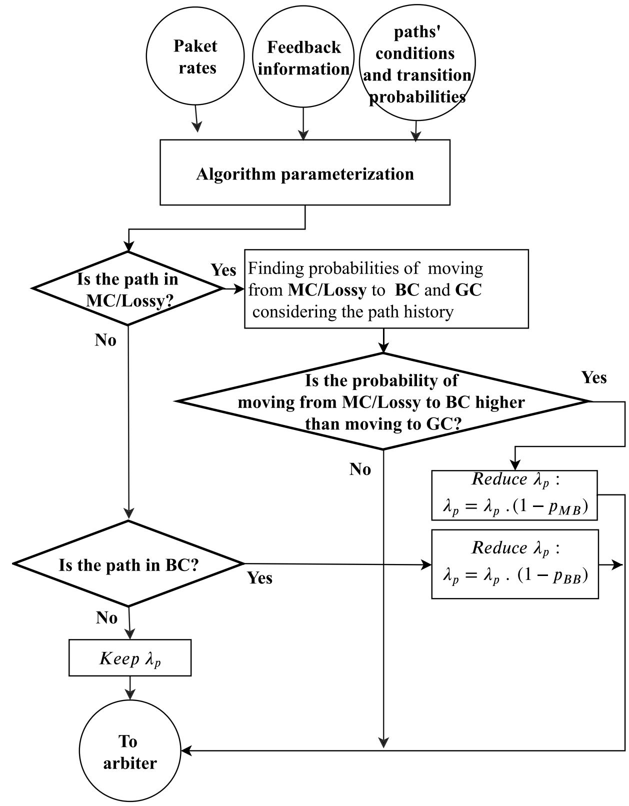


Fig. 8. Flowchart of discard strategy.

3.3.3. Discard strategy

We propose to discard a rate of packets from sending that will probably be either overdue or dropped, as shown in Fig. 8. In the proposed discard strategy, the transition probabilities computed according to Eq. (3) are used and updated the path packet rate split factor λ_p in the following way:

- if the path is in **MC/Lossy**, it is important to consider the path history in order to verify if there is a higher probability of moving to **GC** or to **BC**. The last computed objective metrics are compared with the metrics received in the previous feedback message. If the absolute number of lost packets, values of jitter and delay have increased, then the probability of moving to **BC** is higher and λ_p is updated as $\lambda_p = \lambda_p .(1 - p_{MB})$, where p_{MB} is the probability of transition from **MC** to **BC**. Otherwise, it means that the path congestion condition is improving, or it is only a lossy channel. Therefore, no packet will be discarded;
- if the path is in **BC**, then $\lambda_p = \lambda_p . (1 - p_{BB})$, where p_{BB} is the probability of being in **BC** and staying in **BC** state.

Another solution defined in [40] is estimating of packet display deadline and discarding ones that have no chance to arrive in time.

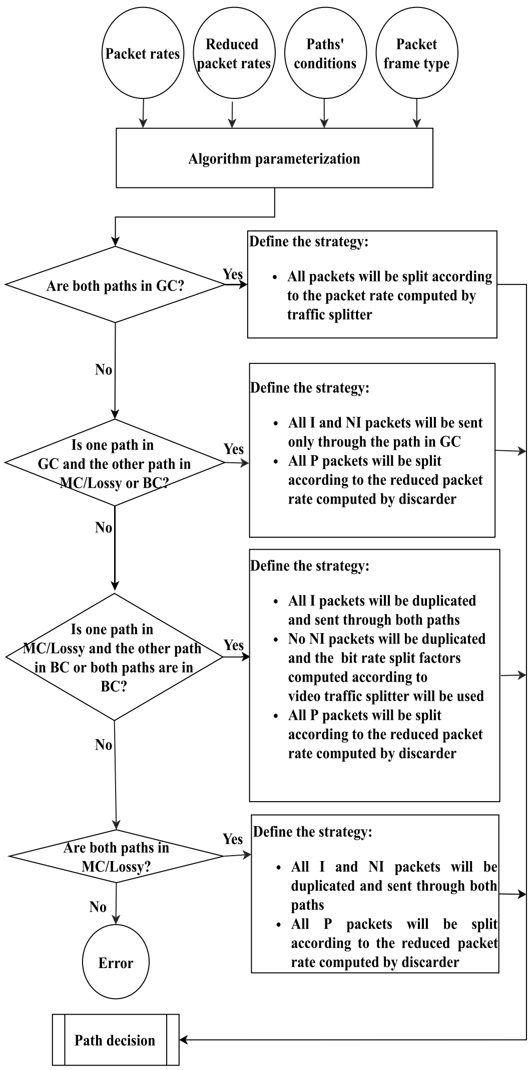


Fig. 9. Flowchart of content-aware strategy.

3.3.4. *Adding content-aware protection*

We here detail the proposed content-aware scheme to show how it protects I and NI frame packets. The flowchart in Fig. 9 presents the process of the proposed content-aware protection strategy. While the protection scheme is not applied to P packets transmitted according to all the rules specified in previous subsections, I or NI packets are protected by duplication and/or rerouting according to the following rules specified for the scenario with two communication paths:

- if both paths are in **GC**, then no packet will be duplicated, and the bitrate split factors computed according to Eq. (1) for each path will be used.
- if one path is in **GC** and the other path is in **MC/Lossy** or **BC**, then all I packets will be sent only through the path in **GC**;
- if one path is in **MC/Lossy** and the other path is in **BC** or both paths are either in **BC** or **MC/Lossy**, then all I packets will be duplicated and sent through both paths;

The same rules are applied for NI frame packets, except they are duplicated only when both paths are in **MC/Lossy**. Note that none of I or NI packets would be discarded, and discard strategy only applies on P packets.

4. Performance evaluation

We provide simulation results to evaluate the efficacy of the proposed CAPA by carrying out ns-3 simulations. We compare the performance of CAPA with the following multipath scheduling strategies:

- **Path-aware scheduling strategy (PA).** PA refers to CAPA, while content-aware is disabled. Such a result comparison makes clear how much content-aware protection contributes to the overall performance gain.
- **Evenly split scheduling strategy (ES).** ES is a simple scheduling strategy that distributes packets evenly through both paths. It should be noted that, although ES is a simple scheduling strategy, it benefits from the network multipath capabilities and improve the total achieved goodput.

In the following, we first explain the evaluation methodology that includes the environment and simulation setup, video sequences, and key performance indicators, which are used to evaluate the user-perceived streaming video quality. Then, we describe different simulation scenarios, such as congested and wireless lossy network situations, together with evaluation result discussion. We also evaluate our proposed strategy when it competes with another MMT flow for fairness support. Finally, we conclude remarks.

4.1. *Methodology*

4.1.1. *Simulation setup*

We simulated our proposed CAPA implementing a ns-3 DCE [41] model. In our evaluation environment, as shown in Fig. 10, our multipath simulation setup comprises LTE and WiFi wireless networks, which are implemented by the LTE and WiFi modules available in the ns-3 simulation library. We chose these two network connections because this is the most common setup for today's mobile devices. The main challenge of this work is different path specifications and heterogeneity to properly split video traffic and provide load balancing. For the LTE path, based on [42], bw_p and $delay_p$ are defined, respectively, as 18.3 Mbps and 15 ms (ms). The 802.11n/5 GHz model is chosen for the WiFi path with bw_p and $delay_p$, respectively, as 54 Mbps and 10 ms [43,44]. In order to turn the simulation setup more real, based on different network scenarios, the ns-3 channel rate (random) error model with choosing "Packet" option for ErrorUnit and Burst Error Model available in the ns-3 simulation library are employed to capture the effects of noisy wireless channels. The packet loss rate values are set later based on different network scenarios. In addition, downlink and uplink background traffics are also added generating, respectively, by the server and network nodes. Background traffic condition detail is also explained later based on the different network scenarios.

4.1.2. *Video sequences*

We have selected a broad range of sequences including cartoon (*Elephants Dream*, *Big Buck Bunny* and *Sintel*) and natural scene (*Meridian*, and *LIVE*) which are the movies with motion, details and very tough colors. The *LIVE* video sequence is a concatenated of nine short videos available in Image & Video Engineering (LIVE) Laboratory[1]; *AirShow*, *AsianFusion*, *Chimera1102347*, *Chimera1102353*, *ElFuenteDance*, *ElFuenteMask*, *Skateboarding*, *Soccer*, and *Sparks*.

All the video sequences are encoded with the same properties. For example, all have 1920 × 1080 resolution and 15,000 total number of frames. The GOP size is of 16 frames, and the employed GOP structure is IPPPP...P. The H.264/AVC JM Reference Software [45] is used as the encoding tool, and the MP4 fragmentation procedures are done by the GPAC MP4BOX [46] tool. Decoding and error concealment are

[1] http://live.ece.utexas.edu/research/Quality/index.htm

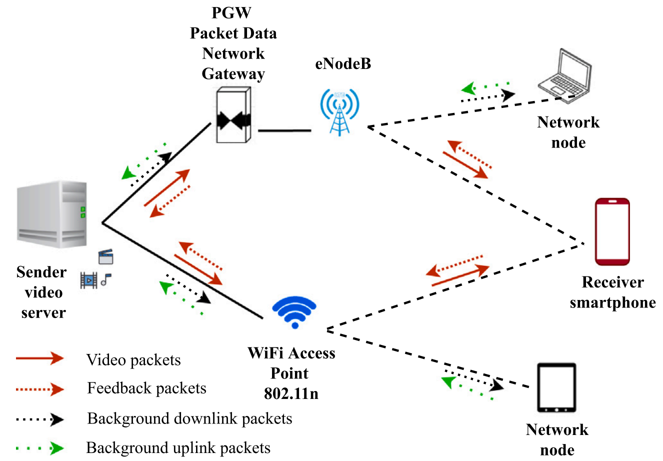


Fig. 10. Evaluation environment where a sender video server uses multiple paths to stream video to a receiver smartphone.

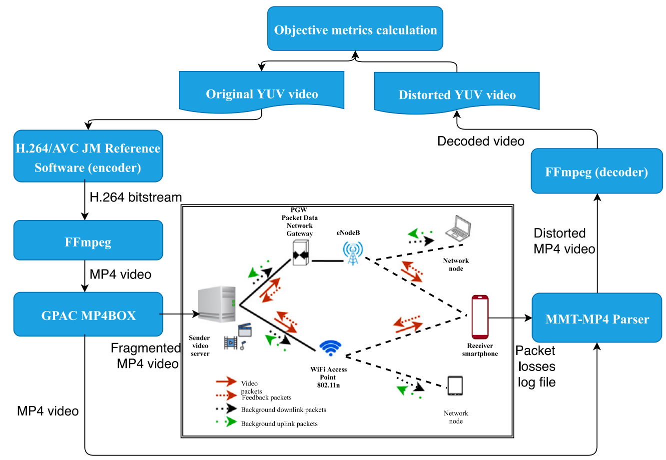


Fig. 11. System architecture for performance evaluation in this paper — objective video quality metrics.

Table 3
Original PSNRs and packets distribution according to frame type for video sequences encoded with 4 Mbps.

Vide sequence	Original PSNR [dB]	I packets [%]	NI packets **[%]**	P packets [%]
Meridian	52.35	32.04	16.93	51.03
Sintel	50.86	18.88	18.70	62.42
Elephants Dream	47.07	25.38	17.15	57.47
LIVE	43.60	23.99	18.75	57.26
Big Buck Bunny	42.95	41.20	10.98	47.82

Table 4
Original PSNRs and packets distribution according to frame type for video sequences encoded with 3 and 5 Mbps.

Video sequence	Encoding rate [Mbps]	Original PSNR [dB]	I packets [%]	NI packets [%]	P packets [%]
Meridian	3	51.1	33.51	16.66	49.83
	5	53.34	30.84	17.09	52.07
Big Buck Bunny	3	41.72	42.20	14.54	43.26
	5	44.02	41.17	15.14	43.69

performed with FFmpeg [47]. These all sequences are encoded with the same source bitrate of 4 Mbps.

The video original PSNR values (PSNR value without losing any packet) are shown in Table 3 ordered from the highest to the lowest value. As expected, we can see that videos have different original qualities due to the different content of each video. However, all of them have enough quality for video transmission above 40 dB. In this work, only the initial 3 P frames in the GOP were considered as NI frames. Therefore, the remaining 13 frames in each GOP are regular P frames. This table also shows the distribution of I, NI, and P packets.

We also encoded *Meridian* and *Big Buck Bunny* with source bitrates of 3 and 5 Mbps for a specific scenario explained in Section 4.2.1. We selected these two video sequences because they have respectively the highest and the lowest PSNRs when encoded with source bitrate of 4 Mbps. Table 4 shows the video original PSNR values related to each of these encoding rates and also distribution of I, NI, and P packets where NI = 3.

Note that video encoded with higher rate has larger PSNR as it is clearly shown in Table 3 together with Table 4. For example, *Meridian* with bitrates of 3, 4, and 5 has 51.1, 52.35, and 53.34 dB respectively. Similarly, *Big Buck Bunny* with bitrates of 3, 4, and 5 has 41.72, 42.95, and 44.02 dB, respectively.

4.1.3. *Key performance indicators*

We evaluate our experimental results in terms of QoS metrics, including delay, packet loss rate, and goodput in order to evaluate performance from the network perspective. Moreover, we also evaluate our experimental results in terms of QoE metrics, including PSNR, and SSIM to evaluate video quality.

The system architecture for performance evaluation considering objective metrics is shown in Fig. 11. In this figure, in order to provide the distorted video to compute PSNR and SSIM, we developed a C++ application, namely MMT-MP4 parser. MMT-MP4 parser has access to the MP4 atom structure and fills the lost packets' positions with zeros in the original MP4 file according to the packet losses information on the receive. After that, the FFmpeg error concealment method is responsible for trying to repair the corrupted or missy bitstream, and consequently, reconstruct the decoded video (distorted YUV video). Finally, PSNR and SSIM are computed between the original and the distorted YUV videos.

4.2. *Different network scenarios*

Here, we first discuss the performance of our proposed CAPA in different network situations; congested network scenario and wireless lossy network scenario. These common network situations cause burst packet losses in the network, and consequently, the receiver may not receive a sufficient number of packets to properly decode the video. Therefore, these network situations could have a severe adverse effect on the perceived video quality. Then, we evaluate the fairness support of proposed CAPA when it competes with another MMT flow.

4.2.1. *Congested network scenario*

In this scenario, we vary background traffic to make congestion in the network in order to observe the response of the proposed CAPA strategy. Besides dynamic background traffic, the ns-3 channel random error is also used in order to capture the effects of noisy wireless channels. Related work [48] assumes loss rate values of 1% for the WiFi path and 0.1% for the LTE path. However, due to the results of our simulations where sequence losses for WiFi become extremely high, we opted to reduce the maximum random loss rate to 0.5%, which is still a realistic assumption.

After explaining the background traffic model in the rest, we discuss the performance of our proposed scheduler under the defined congested network scenario in two different cases; first, where all sequences are encoded with the same bitrate of 4 Mbps. Then, where sequences have different video bitrates.

Background traffic condition. Evaluation considering different background traffic models has been one of the main challenges to be addressed. For a broad and sound evaluation, the background traffic conditions to be explored should cover most of the potential real cases while being able to evaluate the performance of the scheduling strategy.

In this scenario, downlink background traffic is generated by the server and initially set as 70% of full link capacity for both paths. On the other hand, the uplink background traffic is generated by the network nodes and set as 10% of each full link capacity in accordance to real network scenarios where the uplink traffic is smaller than the downlink traffic.

Variable downlink background traffic is illustrated in the top of Fig. 12. As illustrated in this figure, we have three network congestion parts in our simulation; LTE congestion, WiFi congestion, LTE, and WiFi congestion. The first part of the simulation, LTE congestion, is from 40 to 120 s of simulation time. During this period, the background traffic of only the LTE path is increased up to completely (100%) saturating the LTE channel while the background traffic of the WiFi is kept constant. For this purpose, after 40 s of simulation time, the LTE background traffic is first increased to 85% of its full link capacity, and then from 80 to 120 s, it is increased to 100%.

In the second part, WiFi congestion, the opposite behavior is simulated and, from 200 to 400 s of simulation time, the background traffic of the WiFi path is increased up to completely (100%) saturating the WiFi channel while the background traffic of the LTE is kept constant. For this purpose, after 200 s of simulation time, the WiFi background traffic is first increased to 80% of its full link capacity, then from 240 to 280 s, it is increased to 85%, to 90% from 280 to 320 s, to 95% from 320 to 360 s, and finally to 100% from 360 to 400 s.

In the last part, when both paths get congested, from 480 to 600 s of simulation time, background traffic is increased in both paths. For this purpose, the LTE background traffic is increased to 75% from 480 to 520 s, to 80% from 520 to 560 s, to 85% from 560 to 600 s. Simultaneously, the WiFi background traffic is increased to 80% from 480 to 520 s, to 85% from 520 to 560 s, to 90% from 560 to 600 s of simulation time.

In this figure, we also define three background traffic levels; normal, high, and very high background traffic. The path has normal background traffic when its amount is around 65% to 75% of full link capacity. It has high background traffic when its amount is around 75% to 90% of full link capacity, and path has very high background traffic when its amount is more than 90% of full link capacity. Therefore, one can note regarding this proposed background traffic is that we considered extreme cases. Having very high background traffic is a rare network situation and not a condition that happens all the time.

How does the proposed CAPA perform for constant bitrates under congested network scenarios? To evaluate our proposed CAPA in this scenario, all sequences are encoded with the same source bitrate of 4 Mbps and are also streaming with the constant transmission bitrate of 4 Mbps in order to compare results in the same network simulation scenario.

First, we check the fraction of total packets sent through the network for each scheduling strategy, depicted in the third column of Table 5. This information helps to understand how each strategy works and further how they affect the network and video quality performance. One can see is that the percentage of total packets sent according to CAPA strategy is higher than 100% for all sequences, which means it is higher than the number of associated video packets. The reason is mainly duplication. In CAPA, while some packets with low priority are discarded to reduce congestion, the packets with high priority would be rerouted or even duplicated to ensure reliable data delivery during the congestion network. Duplication for protecting packets has the cost of sending more packets through the network. It is important to highlight that high priority frames have a higher number of packets than low priority ones. Note that the content of video always affects on the performance and the amount of duplication.

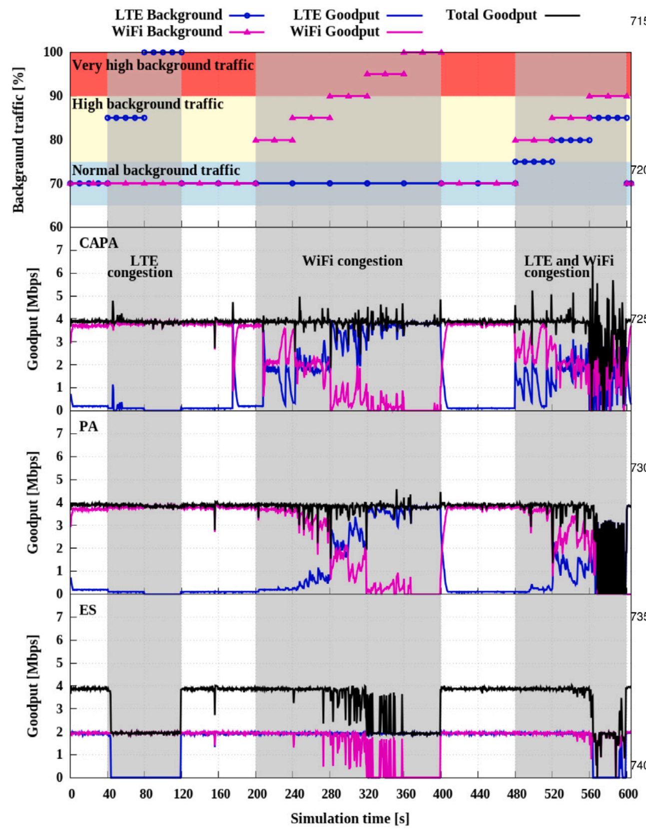


Fig. 12. LTE, WiFi and total (joint) goodput for *Elephants Dream* packets for the defined background traffic according to the different scheduling strategies; the proposed CAPA, PA, and ES, under congested network scenario.

In contrast to CAPA, as shown in Table 5, the percentage of total packets sent according to PA strategy is less than 100% or all sequences, which means less than the number of video packets. It is due to this fact that PA discards some packets at the sender to reduce congestion (and there is no duplication). Regarding ES strategy, the percentage of total packets sent is 100% or all sequences, which is the same as the number of video packets due to there is neither any discard nor any duplication.

Goodput. The goodput performance according to different strategies for all tested video sequences have the same behavior. Here, for instance, Fig. 12 shows the LTE, WiFi and total (joint) goodput for *Elephants Dream* sequence under congested network scenario according to the different scheduling strategies; CAPA, PA, and ES. In order to facilitate readers tracking the network condition, the defined background traffic is also shown in Fig. 12 top of the goodputs.

Fig. 12 represents a more stable and higher total achieved goodput for proposed CAPA compared to PA and ES distribution. One should be noted is that CAPA can clearly provide load balancing, achieve a properly higher inherent capacity of the WiFi path and keep a stable total goodput of 4 Mbps when LTE or WiFi are congested by switching traffic among paths. However, in the last part of the network, where both paths get congested, goodput is decreased. This decrease is not only because of network congestion but also because of the applied packet discard strategy, which discards a rate of packets from sending to avoid further congestion. One important note is that the discard strategy is only applied on the low priority packets. Moreover, the content-aware packet protection method protects the high priority packets by duplicating or rerouting.

Table 5
Comparison results between different scheduling strategies for video sequences encoded with the same bitrate of 4 Mbps under congested network scenario.

Video sequence	Scheduling strategy	Total packets sent [%]	Total packet loss rate [%]	I packet loss rate [%]	NI packet loss rate [%]	Delay [%]	PSNR [dB]	SSIM
Meridian	**CAPA**	**100.97**	**4.30**	**0.91**	**0.73**	–	**42.28 ± 15.15**	**0.961 ± 0.063**
	PA	99.70	4.72	1.23	0.82	−2.7	38.56 ± 16.48	0.938 ± 0.092
	ES	100	17.87	6	3.18	−10	35.08 ± 20.12	0.868 ± 0.177
Sintel	**CAPA**	**100.13**	**3.62**	**0.26**	**0.66**	–	**39.80 ± 16.10**	**0.944 ± 0.096**
	PA	99.58	4.86	0.75	1.02	0	36.31 ± 17.07	0.921 ± 0.116
	ES	100	17.21	3	3.10	−7.6	33.92 ± 19.25	0.874 ± 0.181
Elephants Dream	**CAPA**	**101.20**	**4.52**	**1.24**	**0.67**	–	**37.0 ± 14.10**	**0.922 ± 0.104**
	PA	97.52	6.19	1.85	1.03	0	32.75 ± 14.55	0.892 ± 0.129
	ES	100	17.25	5.23	2.46	−7.5	30.0 ± 17.00	0.819 ± 0.212
LIVE	**CAPA**	**100.97**	**4.33**	**0.59**	**0.98**	–	**34.43 ± 16.10**	**0.919 ± 0.104**
	PA	99.62	5.12	1.18	1.01	0	30.44 ± 17.07	0.885 ± 0.139
	ES	100	17.23	4.28	3.37	−7.7	28.34 ± 14.60	0.839 ± 0.176
Big Buck Bunny	**CAPA**	**109.97**	**4.22**	**1.03**	**0.92**	–	**34.25 ± 11.57**	**0.910 ± 0.107**
	PA	99.03	6.19	2.14	1.03	0	31.44 ± 12.75	0.880 ± 0.131
	ES	100	17.68	7.06	2.56	−5.12	29.15 ± 15.30	0.819 ± 0.202

Fig. 12 also illustrates the total achieved goodput according to PA strategy. Similar to CAPA, PA can also execute load balancing, achieve a higher inherent capacity of the WiFi path and keep a stable total goodput of 4 Mbps through all simulation, except for the last part where both paths get congested. One can note is that the goodput reduction in the last part of the network according to PA strategy is more than goodput reduction at the same time according to CAPA strategy. This is due to this fact that the discard strategy is applied on all types of packets (blindly and without considering the video packet content priority), and there is no content-aware strategy to protect high priority packets.

Regarding ES distribution, one observation is that Fig. 12 is that when there is a normal background, video traffic is equally divided between LTE and WiFi. Therefore, each channel has a goodput of 2 Mbps and the total goodput is 4 Mbps. Then, when LTE gets heavily congested, its goodput sharply decreases due to packet losses, and the total goodput (2 Mbps) achieved in this period is only due to packets transmitted over WiFi. In the second congested part of the simulation, WiFi gets congested, and since inherent WiFi capacity is higher than LTE, congestion is better handled for a while. However, with congestion increasing, then the WiFi goodput decreases to almost zero. Finally, in the last part, where both paths get congested, the goodput reduction is noticeable. This behavior of ES is due to a lack of any path-aware or content-aware strategy.

Packet loss rate. Table 5 shows the effectiveness of the proposed scheduling strategy to decrease the loss rate in CAPA compared to PA and ES scheduling strategies. The results show that CAPA decreases the total loss rate, respectively, by up to **31.82%**, and **78.96%** compared to the PA and ES. There is also better protection for both of I and NI frame packets according to CAPA compared to PA and ES. CAPA decreases the I frame packet loss rate, respectively, by up to **65.33%**, and **91.3%** compared to the PA and ES, and decreases the NI frame packet loss rate, respectively, by up to **42.59%**, and **78.7%** compared to the PA and ES. Noting that, due to different content of videos previously mentioned in Table 3, *Big Buck Bunny*, the one with the highest percentage of I frame packets, also has the most I packet losses in all compared conditions.

Table 5 also shows the proposed scheduling strategy gains better protection of I and NI frame packets. For all sequences, the I and NI frame packet loss rate over the total packet loss rate decreases. For instance, for *Meridian*, according to our proposed scheduling strategy, total lost packets are 4.30% (of total sent packets), and 1.64% of these losses are I and NI losses (0.91%+0.73%), which corresponds to 38.13% of the total losses. When PA scheduling strategy is applied, total lost packets are 4.72% and 2.06% of these losses are I and NI losses (1.23%+0.82%), which corresponds to 43.64% of the total losses. When ES scheduling strategy is applied, total lost packets are 17.87% and 9.18% of these losses are I and NI losses (6% + 3.18%), which corresponds to 51.37% of the total losses. Therefore, the rate is reduced from 43.64% to 38.13% compared to PA, and from 51.37% to 38.13% compared to ES.

Delay. Seventh column of Table 5 shows CAPA average one-way delay reduction of non-overdue packets compared to the alternative scheduling strategies; PA and ES. In this column, the larger the negative values, the performance is better. The results clearly indicate that our proposed CAPA efficiently enables a better adjustment compared to ES by balancing the bitrate distribution and the discard strategy. However, CAPA does not always have delay reduction compared to PA because actually, the packets are not lost, meaning that we have a priority. The priority is to increase the quality of the perceived quality (measured by PSNR and SSIM). This way, CAPA reduced the number of packets, which would be lost in PA.

Regarding the objective video quality metrics results, the PSNR and SSIM values of Table 5 attest to our objective of improving the QoE of end-users by employing our scheduling strategy.

PSNR. The results in Table 5 show that the proposed CAPA improves significantly the average video PSNR, respectively, by up to **4.25 dB (12.97%)**, and **7.22 dB (20.58%)** compared to PA and ES. This is due to proper load balancing and considering path conditions together with video packet contents to transmit packets and perform packet protection. The differences in measured quality results are caused by the different amount of texture, details, action, etc. Furthermore, the standard deviation of PSNR shows that CAPA has less PSNR variation compared to other strategies for all sequences leading to more stable video quality for users.

To have a more detail view of the PSNR results, the PSNR values from *Elephants Dream* sequence, as an example, for all video frames is depicted in Fig. 13 according to defined network congestion parts, shown in gray color, in terms of frame number. For example, the LTE congestion part, which is from 40 to 120 s of simulation time, corresponds to frame numbers from 960 to 2889 of *Elephants Dream* sequence. One can observe of this figure is that CAPA achieves apparently higher PSNR values compared to PA and ES for all the three network congestion parts. This is due to the higher achieved goodput previously discussed and showed in Fig. 13. Furthermore, while for the first and second parts of the simulation, LTE or WiFi congestion, ES has the lowest values, but in the last part, when both paths get congested, the PSNR values for PA are the lowest, even less than ES. Mainly because of blindly discard strategy applied on the packets by

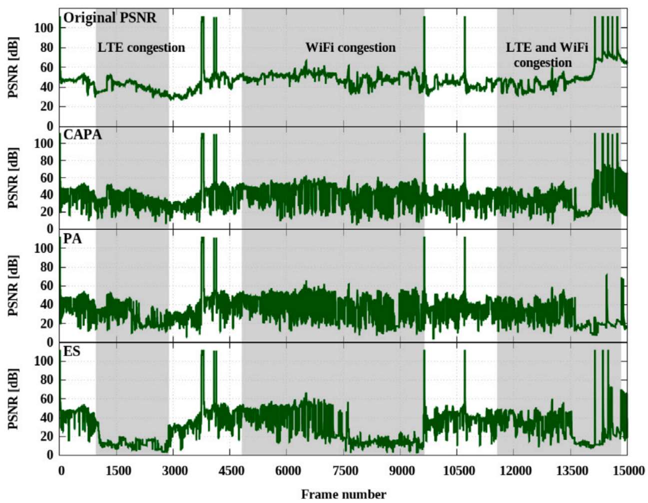


Fig. 13. PSNR values from *Elephants Dream* sequence according to the different scheduling strategies compared to original PSNR; CAPA, PA, and ES, under congested network scenario. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

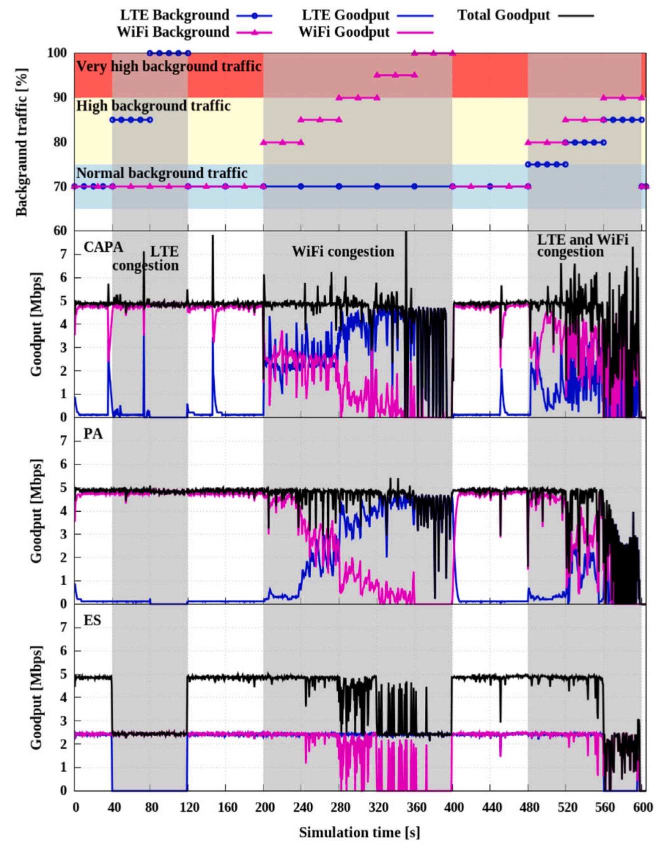


Fig. 14. LTE, WiFi and total (joint) goodput for *Meridian* packets for the defined background traffic according to the different scheduling strategies; the proposed CAPA, PA, and ES, under congested network scenario.

PA, which may discard even high priority packets (I and NI packets). Note that there are PSNR variations for all three strategies even when there is no network congestion compared to the original PSNR. This is due to having random loss through all the simulation time, while FEC is not applied in this work to MMT packets.

SSIM. Table 5 shows that CAPA substantially outperforms other strategies in improving the video SSIM and increases the video SSIM, respectively, by up to **0.033 (3.78%)**, and **0.102 (12.54%)** compared to PA and ES. Also, the standard deviation of SSIM shows that CAPA has less SSIM variation compared to other strategies for all sequences providing more stable video quality for users.

How does the proposed CAPA perform for different video bitrates under congested network scenarios? Here, we have some experiments to check the behavior of our proposed CAPA for different video bitrates in our defined congested network scenario, but it is not the propose of this work to determine for each sequence what is the best. For this scenario, sequences are encoded with the source bitrates of 3 and 5 Mbps and packets are distributed on the network, respectively, with the constant transmission bitrates of 3 and 5 Mbps in order to fit fewer packets or more packets in the same simulation time.

The fourth column of Table 6, similar to what depicted in Table 5, shows the percentage of total packets sent through the network for each scheduling strategy. But differently, it is measured from *Meridian* and *Big Buck Bunny* sequences with different source bitrates of 3 and 5 Mbps. To analyze the measured values in the table, it is important to note that encoded video with higher bitrate has a higher number of packets too. For example, *Meridian* has 163839, 215080, and 2666618 number of packets, respectively, for videos with bitrates of 3, 4 and 5 Mbps. Video qualities are also respectively 51.1, 52.35, and 53.34 dB. Similarly, *Big Buck Bunny* has 166137, 215912, and 265858 number of packets, respectively, for videos with bitrates of 3, 4 and 5 Mbps and video qualities are also respectively 41.77, 42.95, and 44.02 dB. The number of packets matters because having fewer video packets means sending fewer data through the network, having less congestion, and consequently, fewer losses. In contrast, having more video packets means sending more data through the network, having more congestion, and consequently, more losses, which could decrease perceived video quality. Furthermore, CAPA duplicates more packets when there is high congestion in the network to protect them. Therefore, it is possible to see this fact in the table that the total packets sent according to CAPA strategy for 5 Mbps encoded *Big Buck Bunny* and *Meridian* is notably higher than the total packets sent according to CAPA strategy for 3 Mbps encoded.

Goodput. Fig. 14 shows the goodput performance of different scheduling strategies for *Meridian* sequence with bitrate of 5 Mbps through our network environment with such defined extreme cases of background traffic to emphasize the scheduler behavior when transmitting high data. We chose this sequence because it has the highest number of packets (2666618) among our sequences. Generally, the behavior of strategies obtained for *Meridian* sequence with bitrate of 5 Mbps are very similar to what obtained for *Elephants Dream* with bitrate of 4 Mbps, which we explained previously in Section 4.2.1 and illustrated in Fig. 12. The difference here is that the network gets more congested due to higher data transmission. However, yet, CAPA outperforms other scheduling strategies and could achieve higher and more stable goodput compared to PA and ES distribution.

Packet loss rate. Regarding loss results for video streaming with source bitrate of 3, Table 6 shows the effectiveness of the proposed scheduling strategy to decrease loss rate in CAPA compared to PA and ES scheduling strategies by up to **20.65%**, and **87.22%** respectively. Table 6 also shows better protection of I packets for CAPA compared to PA and ES by up to **25.39%**, and **90.44%** respectively, as well as better protect of NI packets for CAPA compared to PA and ES by up to **63.15%**, and **94.26%** respectively.

Regarding loss results for video streaming with source bitrate of 5, CAPA outperforms ES scheduling strategy decreasing the total loss rate, I and NI loss rate, respectively, by up to 86.24%, 25.39%, and 90.44%, but it is not successful to outgo PA strategy in total loss rate. However, it probably outperforms I loss rate and this way could achieve the main objective of our work, which is better-perceived video quality than PA. Actually, CAPA has different content-aware strategies to protect high

Table 6
Comparison results between different scheduling strategies for video sequences encoded with the bitrates of 3 and 5 Mbps under congested network scenario.

Video sequence	Encoding rate [Mbps]	Scheduling strategy	Total packets sent [%]	Total packet loss rate [%]	I packet loss rate [%]	NI packet loss rate [%]	Delay [%]	PSNR [dB]	SSIM
Meridian	3	**CAPA**	**100.07**	**1.88**	**0.47**	**0.14**	–	**43.25 ± 14.57**	**0.962 ± 0.071**
		PA	100	2.04	0.63	0.38	0	40.19 ± 16.66	0.942 ± 0.092
		ES	100	13.67	4.92	2.44	-2.70	38.11 ± 19.43	0.892 ± 0.165
Big Buck Bunny	3	**CAPA**	**100.15**	**1.69**	**0.59**	**0.27**	–	**35.94 ± 10.37**	**0.922 ± 0.099**
		PA	100	2.13	0.79	0.33	0	33.06 ± 11.88	0.891 ± 0.130
		ES	100	13.23	5.01	2.48	-0.66	30.49 ± 19.3	0.830 ± 0.203
Meridian	5	**CAPA**	**102.28**	**10.54**	**2.2**	**2.19**	–	**39.63 ± 15.75**	**0.947 ± 0.081**
		PA	97.94	9.04	2.5	1.47	+8.27	35.09 ± 16.20	0.921 ± 0.105
		ES	100	19.72	6.46	3.44	+6.13	33.81 ± 19.54	0.866 ± 0.167
Big Buck Bunny	5	**CAPA**	**103.79**	**11.45**	**2.22**	**3.43**	–	**32.39 ± 12.77**	**0.883 ± 0.137**
		PA	95.19	8.83	3.13	1.43	+8.18	30.27 ± 13.53	0.853 ± 0.168
		ES	100	19.44	7.51	2.9	+6.22	25.05 ± 19.85	0.803 ± 0.200

priority packets and based on them, it has the strongest protection case for I frame packets. Therefore, having a higher loss rate in this simulation is due to losing more P and NI frame packets. The I frame loss rate for CAPA in this simulation for *Big Buck Bunny* is 2.22% while this value is 3.13% for PA strategy. Similarly, the I frame loss rate for CAPA in this simulation for *Meridian* is 2.2% while this value is 2.5% for PA strategy. Therefore, the content-aware protection method of CAPA properly protects I frame packets, and consequently, results in improved video quality, even under worse network congestion, as shown by the PSNR and SSIM results in Table 6.

Delay. The eighth column of Table 6 shows CAPA average one-way delay reduction of non-overdue packets compared to the alternative scheduling strategies; PA and ES. Reminding this note that video data packets arriving at the destination after decoding deadlines are expired and known as overdue packets. In this table, the larger the negative values, the performance is better. Regarding delay results for videos with a bitrate of 3, the results indicate that our proposed CAPA efficiently enables a better adjustment compared to ES to network conditions by balancing the bitrate distribution and the discard strategy. However, CAPA does not have delay reduction compared to PA. Regarding delay results for videos with a bitrate of 5, CAPA is adding some delay overall. Particularly, CAPA protects I frame packets utilizing duplication. Therefore, compared with PA and ES, which lack of packet duplication, CAPA receives more I frame packets at the cost of higher delay. It is important to note that even if the delay is higher the I frame packets arrive before the deadline. Therefore, the extra delay does not impact the QoE. The resulting trade-off is that while the overall network efficiency may be lower due to the more packets being transmitted due to multipath duplication, the end-user benefits from the increased video quality. For example, while I frame loss rate for CAPA in this simulation for *Big Buck Bunny* is 2.22%, this value is 3.13% and 7.51%, respectively, for PA and ES strategy. Similarly, the I frame loss rate for CAPA in this simulation for *Meridian* is 2.2% while this value is 2.5% and 6.46%, respectively, for PA and ES strategy. Therefore, the QoE does not come free. There is always a trade-off, sending more traffic due to packet duplication adds more delay because of adding more traffic into the buffers and network entities. The increased delay however in below the final deadlines.

PSNR and SSIM values of Table 4 attest to our objective of improving the QoE of end-users by employing our scheduling strategy.

PSNR. Results in Table 4 show that CAPA improves the average video PSNR for video with bitrate of 3 Mbps, respectively, by up to **3.06 dB (7.61%)**, and **5.45 dB (17.87%)** compared to PA and ES. The results also show that CAPA improves the average video PSNR for video with bitrate of 5 Mbps, respectively, by up to **3.33 dB (9.48%)**, and **7.34 dB (29.30%)** compared to PA and ES. Furthermore, the standard deviation of PSNR in the table shows that CAPA can provide more stable video quality for users.

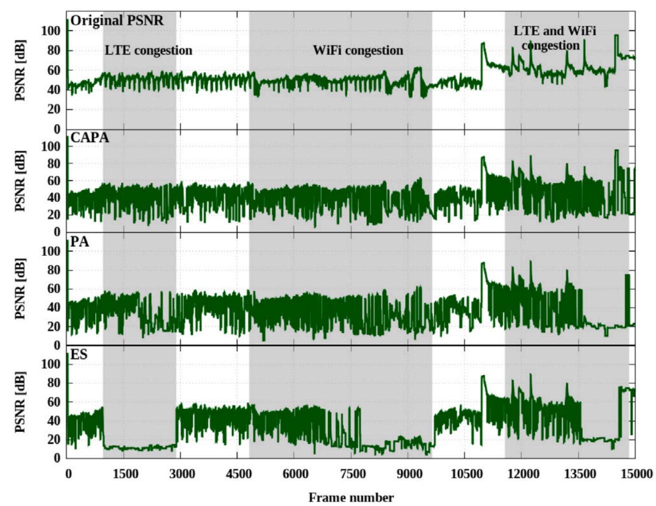


Fig. 15. PSNR values from *Meridian* sequence with bitrate of 5 Mbps according to the different scheduling strategies compared to original PSNR; CAPA, PA, and ES, under congested network scenario.

To have a more detail view of the PSNR results, Fig. 15 shows the PSNR values from *Meridian* sequence, as an example, with source bitrate of 5 Mbps for all video frames. The results behavior obtained for this sequence is very similar to what obtained for the sequence encoded with bitrate of 4 Mbps previously explained and illustrated in Fig. 13.

SSIM. Results show that CAPA improves the video SSIM for video with bitrate of 3 Mbps, respectively, by up to **0.030 (3.367%)**, and **0.091 (10.96%)** compared to PA and ES. The results also show that CAPA improves the video SSIM for video with bitrate of 5, respectively, by up to **0.019 (2.2%)**, and **0.080 (9.96%)** compared to the PA and ES. Additionally, the standard deviation of SSIM shown in the table attests the more stable video quality of CAPA strategy.

In order to conclude our discussion, we form it as reply of two main questions; firstly, **Could our proposed scheduler handle a sequence with more packets?** Yes, results in Table 6 for video with source bitrate of 5 Mbps show that not only CAPA significantly outperforms PA and ES scheduling strategies in improving the video PSNR and SSIM, but also it can achieve PSNR and SSIM, respectively, 38.42 dB and 0.93 for *Meridian*, PSNR and SSIM gain of, respectively, 32.39 dB and 0.883 for *Big Buck Bunny* even in our environment with such defined extreme cases of background traffic, which are sufficient video quality. Secondly, **is it worth to increase the original video quality?** Increasing the original video quality does not always improve the perceived video quality, especially when the network is extremely congested. This is

Table 7
Comparison of original PSNR results with perceived video PSNR results from *Meridian* and *Big Buck Bunny* video sequences with bitrates of 3 and 5 Mbps according to CAPA.

Video sequence	Encoding Rate [Mbps]	Original PSNR [dB]	PSNR results [dB]
Meridian	3	51.1	43.25
	4	52.35	42.28
	5	53.34	38.42
Big Buck Bunny	3	41.77	35.94
	4	42.95	34.25
	5	44.02	32.39

due to the fact that higher quality has more packets, and consequently, increases losses due to higher congestion in our environment with the defined extreme background traffic - Table 7. Therefore, even if the original quality is higher but more quality would be lost. Besides that, the original quality is important because, for example, in our experiment, after adding 1 Mbps, the quality does not change too much. Finally, when the original PSNRs (qualities) are good enough, there is no reason to go higher with the initial quality.

4.2.2. Wireless lossy network scenario

In this scenario, while the background traffic is kept constant in the whole simulation time, different wireless burst loss conditions are defined for paths to observe the response of the scheduling strategy.

In the rest, we first explain the background traffic and wireless loss conditions in detail. Then, we discuss the performance of our proposed scheduler under the defined wireless lossy network scenario in terms of both QoS and QoE metrics.

Background traffic. In this scenario, constant downlink background traffic is set as 70% of full link capacity for both paths. On the other hand, the constant uplink background traffic is set as 10% of each full link capacity in accordance with real network scenarios where the uplink traffic is smaller than the downlink traffic. This explained background traffic is depicted in Fig. 16.

Wireless loss condition. In this scenario, the ns-3 burst error model was employed to capture the effects of burst packet losses caused due to wireless error channels in the network in order to observe the response of the scheduling strategy. Related work [49] assumes loss rates of 5% for the WiFi and 2% for the LTE path. However, for most of our simulations and most of our sequences, the losses for WiFi and LTE were too high. Therefore, we kept the proportion between the loss rates of WiFi and LTE but reduced to 1% for WiFi, and 0.4% for LTE. Finally, we have defined three wireless loss parts in our simulation; LTE lossy channel with a burst loss rate of 0.4%, WiFi lossy channel with a burst loss rate of 1%, LTE and WiFi lossy channels with burst loss rates of 0.4% and 1% respectively.

How does CAPA perform for constant video transmission under wireless lossy network scenarios? To evaluate our proposed CAPA in this scenario, all sequences are encoded with the same source bitrate of 4 Mbps and also are transmitted with a constant bitrate of 4 Mbps.

Table 8 shows the percentage of total packets sent through the network for each scheduling strategy under wireless lossy scenario. One observation is that, similar to congested network scenarios presented in Table 5, the fraction of total packets sent according to CAPA strategy is higher than 100% for all sequences due to the high amount of duplication. However, it is noteworthy that generally no packet is dropped by the discard strategy under lossy network scenarios under the CAPA strategy. In fact, when the network is only lossy (not congested), there is no need to discard any packet and CAPA properly identifies the lossy network situation. Another valuable observation is that the percentage of total packets sent according to PA strategy is 100% for all sequences, meaning that the number of sent packets is equal to the number of video packets. This is due to a similar discard strategy as in CAPA. As

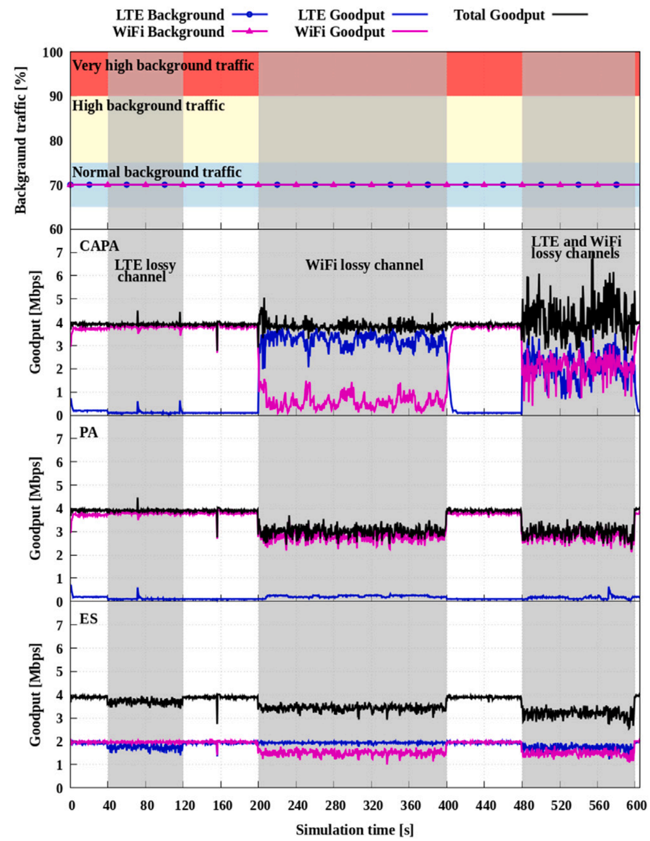


Fig. 16. LTE, WiFi and total (joint) goodput for *Elephants Dream* packets for the constant background traffic with burst wireless loss condition according to the different scheduling strategies; the proposed CAPA, PA, and ES, under wireless lossy network scenario.

a last remark, we recall that the actual video content affects both the performance and the amount of duplication.

Goodput. The goodput performance according to different strategies for all tested video sequences have the same behavior. Here, we show it for *Elephants Dream* sequence in Fig. 16 for instance. This figure, similar to what depicted in Fig. 12, shows the LTE, WiFi and total (joint) goodput according to the different scheduling strategies; CAPA, PA, and ES. However, differently from Fig. 12, it is under the wireless lossy scenario. In order to facilitate readers tracking the network condition, the defined background traffic for this scenario is also shown in Fig. 16 top of the goodputs.

Fig. 16 illustrates that CAPA could gain higher and more total goodput compared with PA and ES distribution, perform load balancing and achieve a properly higher inherent capacity of the WiFi. One can note in this scenario is about where WiFi is lossy. The situation for WiFi lossy channel is so much worse than the situation for LTE in the first lossy part of simulation due to the much higher burst wireless loss rate. Therefore, in the WiFi lossy part or where both paths are lossy, CAPA duplicates many I and NI packets to ensure reliable delivery. Fig. 16 also illustrates the total achieved goodput according to PA scheduling strategy. Although PA can keep a stable total goodput of 4 Mbps for the first part of LTE lossy channel, in part of WiFi lossy channel, and in part of LTE and WiFi lossy channels, it is losing too many packets, and consequently, it has big goodput reduction. The reason is that PA does not have a content-aware method and packet duplication compared to CAPA. Regarding ES, one can observe of Fig. 16 compared to 16(c) is that PA is losing more packets in part of WiFi lossy channel simulation due to higher data rate transmission on WiFi in PA strategy.

Packet loss rate. Table 8 shows the effectiveness of the proposed scheduling strategy to decrease the loss rate in CAPA compared to PA

Table 8
Comparison results between different scheduling strategies for video sequences encoded with the same bitrate of 4 Mbps under wireless lossy network scenario.

Video sequence	Scheduling strategy	Total packets sent [%]	Total packet loss rate [%]	I packet loss rate [%]	NI packet loss rate [%]	PSNR [dB]	SSIM
Meridian	**CAPA**	**105.23**	**4.09**	**0.38**	**0.42**	**40.53** ± **16.11**	**0.943** ± **0.086**
	PA	100	12.65	1.94	3.92	33.69 ± 19.61	0.865 ± 0.157
	ES	100	8.47	1.35	2.57	31.09 ± 19.12	0.857 ± 0.146
Sintel	**CAPA**	**104.63**	**4.5**	**0.25**	**0.46**	**38.17** ± **17.90**	**0.914** ± **0.125**
	PA	100	12.23	2.27	2.42	33.07 ± 20.13	0.846 ± 0.178
	ES	100	8.14	1.56	1.56	30.71 ± 19.37	0.836 ± 0.165
Elephants Dream	**CAPA**	**105.94**	**4.0**	**0.35**	**0.37**	**35.15** ± **14.87**	**0.903** ± **0.121**
	PA	100	12.39	2.94	2.13	29.25 ± 15.88	0.833 ± 0.170
	ES	100	8.23	2.09	1.41	27.37 ± 15.83	0.817 ± 0.167
LIVE	**CAPA**	**105.35**	**4.27**	**0.28**	**0.50**	**33.10** ± **13.38**	**0.886** ± **0.149**
	PA	100	12.31	2.78	2.33	27.49 ± 16.31	0.786 ± 0.232
	ES	100	8.16	1.90	1.49	25.04 ± 15.13	0.773 ± 0.193
Big Buck Bunny	**CAPA**	**108.02**	**3.52**	**0.60**	**0.35**	**33.13** ± **12.60**	**0.889** ± **0.129**
	PA	100	12.65	4.77	1.89	26.59 ± 15.34	0.790 ± 0.193
	ES	100	8.42	3.29	1.19	25.26 ± 14.36	0.749 ± 0.174

and ES scheduling strategies. The results show that CAPA reduces the total loss rate, respectively, by up to **72.09%**, and **58.2%** compared to PA and ES. There is also better protection of I packets according to CAPA compared to PA and ES by up to **89.92%**, and **85.26%** respectively. Moreover, protection of NI packets according to CAPA compared to PA and ES are by up to **89.28%**, and **83.65%** respectively.

The proposed CAPA provides better protection of I and NI frame packets. For all sequences, the I and NI frame packet loss rate over the total packet loss rate decreases. For instance, for *Meridian*, according to our proposed scheduling strategy, total lost packets are 4.09%, where 0.80% (0.38% + 0.42%) of these losses are I and NI losses, which corresponds to 19.55% of the total losses. When PA scheduling strategy is applied, total lost packets are 12.65%, and 5.86% (1.94%+3.92%) of these losses are I and NI losses, which corresponds to 46.32% of the total losses. When ES scheduling strategy is applied, total lost packets are 8.47%, and 3.93% (1.35% + 2.57%) of these losses are I and NI losses, which corresponds to 46.39% of the total losses. Therefore, the rate is reduced from 46.32% to 19.55% compared to PA, and from 46.39% to 19.55% compared to ES.

Another observation of Table 8 is that PA has higher losses and also higher I and NI losses compared to ES. This is mainly because PA has higher losses when WiFi is lossy compared to this part in ES. Goodput achievement of PA previously illustrated in Fig. 16(c) attests these higher losses and lower goodput of PA compared to ES when WiFi is lossy. Yet another point observing of Table 8 is that due to differences of the video sequences previously mentioned and shown in Table 3, *Big Buck Bunny*, the one with the highest percentage of I ad NI frame packets, also has the most I and NI packet losses in all compared conditions.

Delay. In this scenario, since the background traffic is constant and in normal level during all simulation time, delay for most of our simulations and most of our sequence is almost the same.

PSNR. PSNR results shown in Table 8 also attest to our objective of improving the QoE of end-users by employing our scheduling strategy, this time under lossy network scenario. The results show that CAPA improves the average video PSNR, respectively, by up to **6.84 dB (20.30%)**, and **9.43 dB (30.32%)** compared to PA and ES. This is mainly because of the proposed content-aware strategy and packet duplication. Furthermore, the standard deviation of PSNR in the table shows that CAPA provides more stable video quality for users compared to alternative strategies.

Similar to Fig. 13, Fig. 17 shows the PSNR values from *Elephants Dream* sequence, as an example, for all video frames, but differently, under wireless lossy network scenario. One can observe is that PA could only achieve proper PSNR values when LTE is lossy, and it almost completely degrading for the second and third lossy parts due to the high WiFi loss rate. The PSNR resulting QoE for ES would be very low almost completely degrading the sequence in all three defined parts of wireless lossy conditions. Even in this challenging scenario, our proposed strategy was able to keep the QoE in higher levels while optimizing the total network goodput.

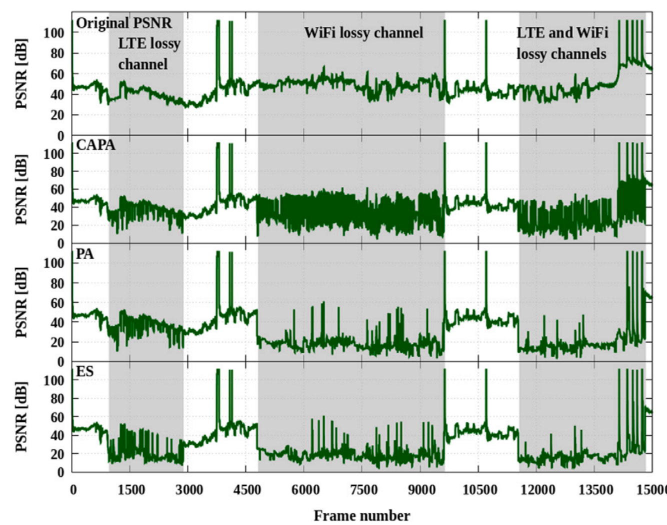


Fig. 17. PSNR values from *Elephants Dream* video sequence according to the different scheduling strategies compared to original PSNR; CAPA, PA, and ES, under wireless lossy network scenario.

SSIM. Besides PSNR, the SSIM values, shown in Table 8, also attest to our objective of improving the QoE of end-users by employing our scheduling strategy. CAPA substantially outperforms other strategies in improving the video SSIM, respectively, by up to **0.100 (12.72%)**, and **0.113 (14.23%)** compared to PA and ES. Furthermore, the standard deviation of SSIM points to more stable video quality of CAPA compared to alternative strategies.

In conclusion, CAPA hugely outperforms PA and ES in improving QoE under wireless lossy network situation by optimizing the total network goodput.

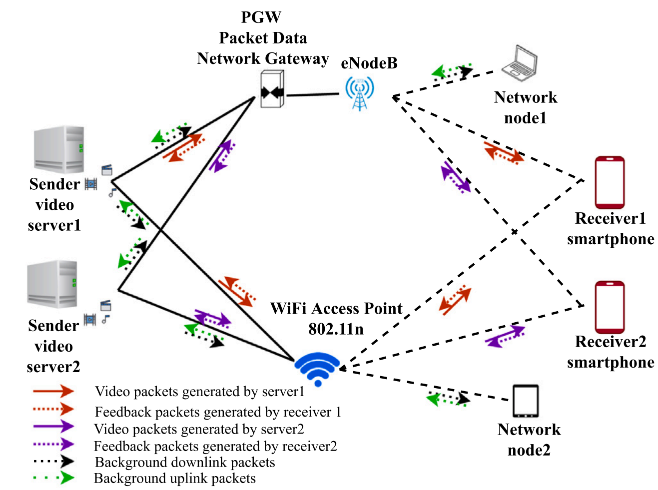


Fig. 18. Evaluation environment where two sender video servers use multiple paths to stream video to receiver smartphones sharing common bottlenecks.

4.3. Support of fairness

We, now, evaluate the initial fairness test of our proposed CAPA when it competes with another MMT flow where all paths between MMT sending and receiving entities share common bottlenecks.

Simulation Setup and Background Traffic Condition. For fairness experiment, as shown in Fig. 18, we have extended the environment explained in Section 4.1 adding one more server named "Sender video server2" and one more receiver named "Receiver smartphone2". In this environment, each video server uses multiple paths to stream video to its assigned receiver. Therefore, video server1 streams video to receiver1, and video server2 streams video to receiver2. Subsequently, receiver1 generates feedbacks to server1, and receiver2 generates feedbacks to server2. Therefore, in this environment, there are bottlenecks for both LTE and WiFi paths.

Similar to what explained in Section 4.1, as depicted in Fig. 18, uplink and downlink background traffics are generated by the network nodes and servers respectively. Differently, TCP packets are generated for background traffic instead of UDP packets for this scenario. These TCP packets are generated by iperf in ns-3 DCE. The reason of choosing TCP background traffic instead of UDP is that there is no congestion control and feedback control for UDP. UDP has constant rate, and that is all. However, we need traffic to react to congestion. This is what TCP delivers. Therefore, we can demonstrate how our protocol behaves.

While in this scenario, WiFi and LTE paths' delay and bandwidth are the same as what was defined in Section 4.1, the wireless loss rate is considered as zero in this scenario because we want to focus on fairness in bandwidth sharing and we do not want to have wireless loss impact on the scheduling strategy. The fairness of CAPA, in this environment, is presented in terms of goodput.

How does CAPA stand in terms of fairness? In this network scenario, *Elephants Dream* sequence, which is encoded with the source bitrate of 4 Mbps is used. We also consider the constant transmission bitrate stream of 4 Mbps.

Fig. 19 depicts fairness results of goodput achievement in WiFi and LTE from 0 to 120 s of simulation time. As it is shown in this figure, at 0 s of simulation time, server1 starts to stream the video to receiver1, then after 50 s of simulation time, the server2 starts to stream the video in parallel. These two MMT streams compete with each other to access the available resource. Goodput results show that both receiver1 and receiver2 have fair access to the available resources in both WiFi and LTE. This result is obtained because the overall actual goodput in the WiFi path, which is smaller than the theoretical 54 Mbps due to MAC layer, is attained when the second video stream starts up. This event affects on the adaptive traffic video split strategy Eq. (1) calculation results.

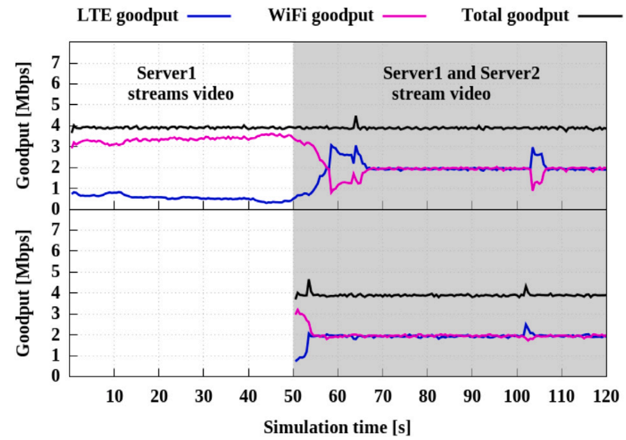


Fig. 19. Fairness results of goodput achievement in WiFi and LTE.

Table 9
Properties of the proposed CAPA adopted for MMT.

Protocol layer		Applied protocol		Compatibility
Application		MMT		Server and Client
Which packet?		**How to protect the packet?**		**Which path?**
Content awareness		Duplication		Path awareness (Delay, PLR, Goodput)
Fairness	**Video compression**	**Error concealment**	**Experimental environment**	**Performance metrics**
Y	H.264	FFMPEG	NS3-DCE, Client interfaces: LTE, WiFi (802.11n)	Goodput, Total packet loss rate, I and NI packet loss rate, Delay, PSNR, SSIM

5. Conclusions and future work

In this work, we explore adding multipath support to MMT proposing a novel Content-Aware and Path-Aware (CAPA) scheduling strategy. CAPA distributes packets through the network according to the wireless network conditions as well as the video packet content characteristics. The proposed CAPA has been fully implemented in the application layer as a compatible module and utilizes the standardized MMT signaling messages to provide feedback information. Some of the key properties of CAPA are summarized in Table 9.

One objective of the proposed approach is increasing the quality of the final sequence, PSNR and SSIM. We could achieve this target by reducing the number of packets that are lost. Table 10 summarizes the achievements by CAPA compared to PA and ES provided in Sections 4.2.1 and 4.2.2 , respectively, under congested network scenario and wireless lossy network scenario, where results are provided for various sequences with bitrate of 4 Mbps. We achieved that our proposing approach has good results for both of these common network scenarios and could improve the QoE of end users. It does not matter if there is congestion or burst wireless losses. But maybe, the proposed solution is better in handling wireless losses than congestion. While, in wireless lossy network scenario, CAPA could increase the average video PSNR, respectively, by up to **6.84** dB **(20.30%)**, and **9.43 dB (30.32%)** compared to PA and ES, in congested network scenario, CAPA could increase the average video PSNR, respectively,

Table 10
Brief description of achievements by CAPA compared to PA and ES provided in Sections 4.2.1 and 4.2.2, respectively, under congested network scenario and wireless lossy network scenario where results are provided for various sequences with bitrate of 4 Mbps.

Performance metrics	Congested network scenario	Wireless lossy network scenario
Goodput	Fig. 12, CAPA achieves higher and more stable goodput compared to PA and ES.	Fig. 16, CAPA achieves higher and more stable goodput compared to PA and ES.
Loss rate	Table 5, CAPA reduces total loss rate compared to PA and ES, respectively, by up to 31.82% and 78.96%. Since CAPA can protect I packets, or I and NI packets based on different levels of packet protection, it can even achieve less I and/or NI packet loss rate compared to alternatives when there is a very high congested network situation.	Table 8, CAPA reduces total loss rate compared to PA and ES, respectively, by up to 72.09% and 58.2%.
Delay	Table 5, CAPA outperforms ES but does not have always delay reduction compared to PA. Even if duplication improves the video quality but it costs of increasing congestion and effect on the delay.	Delay for most of our simulations and most of our sequence is almost the same.
PSNR	Table 5, CAPA outperforms PA and ES, respectively, up to 4.25 dB (12.97%) and 7.22 dB (20.58%).	Table 8, CAPA outperforms PA and ES, respectively, up to 6.84 dB (20.30%) and 9.43 dB (30.32%).
SSIM	Table 5, CAPA outperforms PA and ES, respectively, up to 0.033 (3.78%) and 0.102 (12.54%).	Table 8, CAPA outperforms PA and ES, respectively, up to 100 (12.72%) and 0.113 (14.23%).

by up to **4.25** dB **(12.97%)**, and **7.22 dB (20.58%)** compared to PA and ES. Similarly, while in wireless lossy network scenario, CAPA could increase the SSIM, respectively, by up to **0.100 (12.72%)**, and **0.113 (14.23%)** compared to PA and ES, in congested network scenario, CAPA could increase the average video PSNR, respectively, by up to **0.033 (3.78%)**, and **0.102 (12.54%)** compared to PA and ES.

We have also checked the behavior of the proposed strategy with different video bitrates under congested network scenario. The results show that our scheduler could properly handle these videos. Remark that video with higher bitrate (original quality) also produces more packets, and consequently, more packets would be lost during transmission due to higher congestion. Therefore, higher original quality cannot necessarily improve the perceived quality.

We also had basic validation of fairness for CAPA, when two MMT flows compete with each other to access the available resource, where all paths between MMT sending and receiving entities share common bottlenecks. Even if the full understanding of fairness requires much more work and experience, but these initial experiments could properly show that the second MMT flow does not kill the first one and both receivers have fair access to both paths.

A major target of this work is improving the MMT standard by adding multipath scheduling strategies. We proposed the collaboration document m44902r1 for the MMT IG standardization activity in 2018 and was adopted for the 4th edition of the ISO/IEC 23008-13 standard. The invention has also been protected through patent applications that has been filled out in INPI [50] and US patent [51]. It is important to highlight that, this contribution is valuable because MMT has been already adopted by other standards such as ATSC 3.0 and it is expected to adopt by digital television broadcasting/broadband transmission systems as well as AR, VR, MR devices, smartphones, and tablets. Therefore, there is high potential for our approach to become wildly used.

In future work, we plan to investigate a Forward Error Correction (FEC) stage exploiting what is already specified as the Application Layer FEC of the MMT standard. Exploiting variable bitrate for the video sequence instead of only constant bitrate could improve the QoE by adapting video traffic rate based on the network condition. Extended evaluation plans for CAPA include mobility at different speeds and mobility patterns, including considerations on radio coverage and handoff between multiple networks. Furthermore, we will scale the scenarios in terms of number of servers and receivers, especially critical for taking the fairness evaluation to the next steps. Moreover, we will consider evaluating the use of 802.11e and its capabilities of traffic prioritization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially supported by Samsung Research Brazil and CNPq (Brazilian National Council for Scientific and Technological Development). Dr. Imed bouazizi was at Samsung Research America (SRA) and Dr. Samira Afzal was at University of Campinas (UNICAMP) during the development period of this work.

References

[1] Cisco visual networking Index: Forecast and methodology, 2016–2021, 2017, https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html. (Accessed 15 January 2017).

[2] F.M.d.C.R. Vieira, Systematic Analysis of Super Hi-Vision Systems, FEUP - Faculdade de Engenharia da Universidade do Porto, 2014.

[3] Y. Sani, A. Mauthe, C. Edwards, Adaptive bitrate selection: A survey, IEEE Commun. Surv. Tutor. 19 (4) (2017) 2985–3014.

[4] M. Kazemi, S. Shirmohammadi, K.H. Sadeghi, A review of multiple description coding techniques for error-resilient video delivery, Multimedia Syst. 20 (3) (2014) 283–309.

[5] Y. Huo, C. Hellge, T. Wiegand, L. Hanzo, A tutorial and review on inter-layer FEC coded layered video streaming, IEEE Commun. Surv. Tutor. 17 (2) (2015) 1166–1207, http://dx.doi.org/10.1109/COMST.2015.2392378.

[6] R. Trestian, I.-S. Comsa, M.F. Tuysuz, Seamless multimedia delivery within a heterogeneous wireless networks environment: Are we there yet? IEEE Commun. Surv. Tutor. 20 (2) (2018) 945–977.

[7] S. Kim, D. Yun, K. Chung, Video quality adaptation scheme for improving QoE in HTTP adaptive streaming, in: 2016 International Conference on Information Networking, ICOIN, 2016, pp. 201–205, http://dx.doi.org/10.1109/ICOIN.2016.7427114.

[8] A.H. Zahran, J.J. Quinlan, K. Ramakrishnan, C.J. Sreenan, ASAP: Adaptive stall-aware pacing for improved DASH video experience in cellular networks, ACM Trans. Multimedia Comput. Commun. Appl. (TOMM) 14 (3s) (2018) 61.

[9] J.K. Sundararajan, D. Shah, M. Médard, S. Jakubczak, M. Mitzenmacher, J. Barros, Network coding meets TCP: Theory and implementation, Proc. IEEE 99 (3) (2011) 490–512.

[10] G. Bhat, J. McNair, Delivering low latency video using TCP with network coding over wireless network, Int. J. Comput. Appl. 87 (15) (2014) 1–8.

[11] H. Cui, Q. Sun, B. Zhang, X. Zhang, W. Wu, Video-aware multipath network coding over multi-hop wireless networks, in: 2019 International Conference on Information Networking, ICOIN, IEEE, 2019, pp. 171–175.

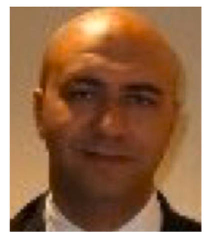
[12] H. Seferoglu, A. Markopoulou, Opportunistic network coding for video streaming over wireless, in: Packet Video 2007, IEEE, 2007, pp. 191–200.
[13] A. Švigelj, M. Junuzović, Network coding-assisted retransmission scheme for video-streaming services over wireless access networks, in: Broadband Communications Networks-Recent Advances and Lessons from Practice, IntechOpen, 2017.
[14] D. Nguyen, T. Nguyen, X. Yang, Multimedia wireless transmission with network coding, in: Packet Video 2007, IEEE, 2007, pp. 326–335.
[15] Y. Keshtkarjahromi, H. Seferoglu, R. Ansari, A. Khokhar, Device-to-device networking meets cellular via network coding, IEEE/ACM Trans. Netw. 26 (1) (2018) 370–383.
[16] S. Afzal, V. Testoni, J.F.F. de Oliveira, C.E. Rothenberg, P. Kolan, I. Bouazizif, A novel scheduling strategy for MMT-based multipath video streaming, in: 2018 IEEE Global Communications Conference, GLOBECOM, IEEE, 2018, pp. 206–212.
[17] J. Wu, C. Yuen, B. Cheng, Y. Yang, M. Wang, J. Chen, Bandwidth-efficient multipath transport protocol for quality-guaranteed real-time video over heterogeneous wireless networks, IEEE Trans. Commun. 64 (6) (2016) 2477–2493.
[18] J. Chakareski, S. Han, B. Girod, Layered coding vs. multiple descriptions for video streaming over multiple paths, Multimedia Syst. 10 (4) (2005) 275–285.
[19] J. Chakareski, B. Girod, Rate-distortion optimized packet scheduling and routing for media streaming with path diversity, in: Data Compression Conference, IEEE, 2003, pp. 203–212.
[20] MPEG, Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 1: MPEG media transport (MMT), 2014, ISO/IEC 23008-1.
[21] Y. Lim, S. Aoki, I. Bouazizi, J. Song, New MPEG transport standard for next generation hybrid broadcasting system with IP, IEEE Trans. Broadcast. 60 (2) (2014) 160–169.
[22] S. Bae, Method for configuring and transmitting M-unit, Google Patents, US Patent App. 13/651, 191, 2013, URL http://www.google.sr/patents/US20130094594, (Apr.).
[23] S. Aoki, Emerging 8K services and their applications towards 2020, in: ITU-T 2nd Mini-Workshop on Immersive Live Experience, 2017.
[24] C. Xu, T. Liu, J. Guan, H. Zhang, G.-M. Muntean, CMT-QA: Quality-aware adaptive concurrent multipath data transfer in heterogeneous wireless networks, IEEE Trans. Mob. Comput. 12 (11) (2013) 2193–2205.
[25] S. Afzal, V. Testoni, C.E. Rothenberg, P. Kolan, I. Bouazizi, A holistic survey of wireless multipath video streaming, 2019, arXiv preprint arXiv:1906.06184.
[26] V. Singh, S. Ahsan, J. Ott, MPRTP: multipath considerations for real-time media, in: Proceedings of the 4th ACM Multimedia Systems Conference, ACM, 2013, pp. 190–201.
[27] P. Houzé, E. Mory, G. Texier, G. Simon, Applicative-layer multipath for low-latency adaptive live streaming, in: IEEE International Conference on Communications, ICC, 2016, pp. 1–7.
[28] Y. Go, H. Noh, G. Park, H. Song, Hybrid TCP/UDP-based enhanced HTTP adaptive streaming system with multi-homed mobile terminal, IEEE Trans. Veh. Technol. 68 (5) (2019) 5114–5128.
[29] Y. Sohn, M. Cho, M. Seo, J. Paik, A synchronization scheme for hierarchical video streams over heterogeneous networks, KSII Trans. Internet Inf. Syst. 9 (8) (2015).
[30] F. Michel, Q. De Coninck, O. Bonaventure, Adding forward erasure correction to quic, 2018, arXiv preprint arXiv:1809.04822.
[31] D.H. Bui, K. Lee, S. Oh, I. Shin, H. Shin, H. Woo, D. Ban, Greenbag: Energy-efficient bandwidth aggregation for real-time streaming in heterogeneous mobile wireless networks, in: Real-Time Systems Symposium (RTSS), 2013 IEEE 34th, IEEE, 2013, pp. 57–67.
[32] K. Habak, K.A. Harras, M. Youssef, Bandwidth aggregation techniques in heterogeneous multi-homed devices: A survey, Comput. Netw. 92 (2015) 168–188.
[33] A. Nikravesh, Y. Guo, X. Zhu, F. Qian, Z.M. Mao, MP-H2: A client-only multipath solution for HTTP/2, in: MobiCom, ACM, 2019.
[34] Á. Huszák, S. Imre, Analysing GOP structure and packet loss effects on error propagation in MPEG-4 video streams, in: Proceedings of the 4th International Symposium on Communications, Control and Signal Processing, ISCCSP, IEEE, 2010, pp. 1–5.
[35] J. Wu, C. Yuen, B. Cheng, M. Wang, J. Chen, Streaming high-quality mobile video with multipath TCP in heterogeneous wireless networks, IEEE Trans. Mob. Comput. 15 (9) (2016) 2345–2361.
[36] M. Xing, S. Xiang, L. Cai, A real-time adaptive algorithm for video streaming over multiple wireless access networks, IEEE J. Sel. Areas Commun. 32 (4) (2014) 795–805.
[37] J. Wu, C. Yuen, N.-M. Cheung, J. Chen, Delay-constrained high definition video transmission in heterogeneous wireless networks with multi-homed terminals, IEEE Trans. Mob. Comput. 15 (3) (2016) 641–655.
[38] A.L. Chow, H. Yang, C.H. Xia, M. Kim, Z. Liu, H. Lei, EMS: Encoded multipath streaming for real-time live streaming applications, in: 17th IEEE International Conference on Network Protocols, ICNP, 2009, pp. 233–243.
[39] S. Cen, P.C. Cosman, G.M. Voelker, End-to-end differentiation of congestion and wireless losses, IEEE/ACM Trans. Netw. 11 (5) (2003) 703–717.
[40] X. Corbillon, R. Aparicio-Pardo, N. Kuhn, G. Texier, G. Simon, Cross-layer scheduler for video streaming over MPTCP, in: Proceedings of the 7th International Conference on Multimedia Systems, ACM, 2016, p. 7.
[41] ns3, Direct code execution (DCE) for ns-3, version 1.9, 2016.
[42] OpenSignal, 2017, https://opensignal.com/reports/2017/02/global-state-of-the-mobile-network. (Accessed 15 June 2017).
[43] Y. Kim, S. Choi, K. Jang, H. Hwang, Throughput enhancement of IEEE 802.11 WLAN via frame aggregation, in: IEEE 60th Vehicular Technology Conference, 2004, Vol. 4, VTC2004-Fall. 2004, IEEE, 2004, pp. 3030–3034.
[44] B.-H. Oh, J. Lee, Constraint-based proactive scheduling for MPTCP in wireless networks, Comput. Netw. 91 (2015) 548–563.
[45] JM, H.264/AVC JM reference software, version 19.0, 2015.
[46] GPAC, GPAC MP4BOX, version 0.5.1, 2015.
[47] FFmpeg, 2017, https://ffmpeg.org. (Accessed 10 July 2018).
[48] Y.-C. Chen, Y.-s. Lim, R.J. Gibbens, E.M. Nahum, R. Khalili, D. Towsley, A measurement-based study of multipath tcp performance over wireless networks, in: Proceedings of the 2013 Conference on Internet Measurement Conference, ACM, 2013, pp. 455–468.
[49] J. Wu, Y. Shang, J. Huang, X. Zhang, B. Cheng, J. Chen, Joint source-channel coding and optimization for mobile video streaming in heterogeneous wireless networks, Eur. Assoc. Signal Process. (EURASIP) J. Wirel. Commun. Netw. 2013 (1) (2013) 283.
[50] J.F. Oliveira, S. Afzal, V. Testoni, Método para transmissão de vídeo digital adotando estratégias de escalonamento com reconhecimento de percurso e conteúdo em redes heterogêneas usando protocolo MMT, 2018, INPI BR 10 2018 070605 5.
[51] J.F. Oliveira, S. Afzal, V. Testoni, Method for digital video transmission adopting packaging forwarding strategies with path and content monitoring in heterogeneous networks using MMT protocol, method for reception and communication system, US Patent 16387249, 2019.

Samira Afzal received her master's degree in IT engineering at Sharif University of Technology (SUT) in 2011 and graduated with her Ph.D. degree in Electrical Engineering from the University of Campinas (UNICAMP), in 2019. Currently, she is a post doctoral researcher at the University of Sao Paulo researching cloud and edge computing, Internet of Things, swarm, and smart cities. Her research interests cover Internet of Things, machine learning, software defined networking, mobile video streaming and video standards, including codecs and network transport protocols (e.g., MMT, DASH, MPTCP, QUIC).

Christian Esteve Rothenberg holds the Telecommunication Engineering degree from the Technical University of Madrid (ETSIT - UPM), Spain, and the M.Sc. (Dipl. Ing.) degree in Electrical Engineering and Information Technology from the Darmstadt University of Technology (TUD), Germany, 2006. From 2010 to 2013, he worked as Senior Research Scientist in the areas of IP systems and networking at CPqD R&D Center in Telecommunications, Campinas, Brazil. Since August 2013, Christian Esteve Rothenberg is Head of the INTRIG (Information & Networking Technologies Research & Innovation Group) and Assistant Professor at the University of Campinas (UNICAMP), where he received his Ph.D. in Electrical and Computer Engineering in 2010.

Vanessa Testoni eceived the B.S. degree in computer science from the Pontifical Catholic University of Paraná (PUCPR), Brazil, in 2002, the B.S. degree in electrical engineer from the Federal University of Paraná (UFPR), Brazil, in 2004, and the M.Sc. and the Ph.D. degrees in electrical engineering from the University of Campinas (UNICAMP), Brazil, respectively, in 2006 and 2011. Right after her graduation, she joined the University of California, San Diego (UCSD), USA, as a postdoctoral employee. Since 2013 she has been the research leader in Media Standards at SRBR (Samsung Research Institute Brazil) where her research interests include digital signal processing, computer vision, machine learning, and specially image and video coding. She has authored/co-authored several scientific publications and patents. Dr. Testoni was the Brazilian PhD Student awarded the Microsoft Research Ph.D. Fellowship Award in 2009. In 2014 she received the MIT TR35 (Young Innovators under 35) in the first Brazilian edition of the award. In 2016, she was elected an affiliate member of the Brazilian Academy of Sciences (ABC) and also the chair of the IEEE SPS (Signal Processing Society) S ã o Paulo chapter. In 2018, she was elevated to IEEE senior member and nominated the first national head of the ISO/IEC JTC 001/SC 29 (JPEG/MPEG) Brazilian delegation.

Prakash Kolan is Standards Engineer, Senior Staff at Samsung Mobility Innovation (SMI) Lab in Samsung Research America, US. His primary research areas include end-to-end network slicing in 5G networks, cloud media processing, next generation media services and delivery using multi-access edge computing (MEC). Dr.Kolan holds several patents in the areas of cloud media processing and 5G media delivery, and authored articles that are published in leading IEEE and ACM journals and conferences. Dr. Kolan is currently a standards delegate for Samsung in 3GPP SA4 SWG that specifies standards for media processing and delivery architecture in 5G networks. Prior to 3GPP SA4, he was a standards delegate for Samsung in 3GPP SA2 SWG (5G core network architecture) and MPEG (media processing technologies).



Imed Bouazizi received a M.Sc. degree in computer sciences from Technical University of Braunschweig, Braunschweig, Germany, and a Ph.D. degree from the RWTH Aachen University, Aachen, Germany, in 2000 and 2004, respectively. From 2015 to 2019, he was a principal research engineer at Samsung Research America, and currently, he is a principal standards engineer at Qualcomm. His research interests cover immersive multimedia distribution and communication, including VR/AR, 6DoF media, and 5G architecture for media distribution, with a strong focus on standardization activities. He has contributed to the development of several standardization efforts in the most relevant organizations, such as MPEG, 3GPP, ATSC, DVB, and the IETF.