Reliable Video Streaming with Strict Playout Deadline in Multi-Hop Wireless Networks

Hussein Al-Zubaidy, Viktoria Fodor, György Dán, Markus Flierl
School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden
{hzubaidy, vjfodor, gyuri, mflierl}@kth.se

***Abstract*—Motivated by emerging vision-based intelligent services, we consider the problem of rate adaptation for high quality and low delay visual information delivery over wireless networks using scalable video coding. Rate adaptation in this setting is inherently challenging due to the interplay between the variability of the wireless channels, the queuing at the network nodes and the frame-based decoding and playback of the video content at the receiver at very short time scales. To address the problem, we propose a low-complexity, model-based rate adaptation algorithm for scalable video streaming systems, building on a novel performance model based on stochastic network calculus. We validate the analytic model using extensive simulations. We show that it allows fast, near optimal rate adaptation for fixed transmission paths, as well as cross-layer optimized routing and video rate adaptation in mesh networks, with less than 10% quality degradation compared to the best achievable performance.**

***Keywords*—scalable video coding; wireless multimedia; multihop fading channels; performance analysis; network calculus**

I. Introduction

Low-cost cameras that are able to capture high quality images, combined with increasing wireless transmission rates, and advances in video coding and visual processing are enabling a variety of novel, visual-information-based intelligent services. The services include vision-controlled robotics [1], automated driving applications [1], [2], and telematic surgery [3], and are often safety critical. Their requirements differ significantly from the ones of traditional video content distribution: they require very low latency, high reliability and good video quality for human inspection or for automated visual processing. As an example, use case specifications for eHealth, future factories and automotive [1], [4] require tens of milliseconds of network, and hundreds of milliseconds of application level delay limits, with a 99.99% reliability.

In many of these emerging application areas the cameras are hard to access or are mobile, hence the use of wireless data transmission is inevitable, likely over multiple wireless hops. Multiple wireless hops facilitate the support for multicast or convergecast of the captured streams of images, facilitate the handling of node mobility, may help to cope with hostile wireless environments, and could also allow low power transmission for battery-driven nodes [5]–[7].

Future networks are expected to provide service awareness, and thus support the timely and reliable delivery of video streams [8]. In the wireless domain, softwarization and virtualization will be used to achieve service aware, flexible resource allocation [9]. However, low-latency video streaming in wireless networks is challenged even by the rapid changes of the channel quality due to fading, which may lead to temporal queues at intermediate nodes, and thus makes careful rate adaptation necessary.

Standardization and existing research activities in the area of video coding and streaming use rate adaptation to adjust the video quality to the available network resources in slowly changing dynamic environments. The key enabling technology for rate adaptation is scalable video coding (SVC) [10]–[12], where video frames or groups of pictures (GOP) are encoded into multiple layers. Rate adaptation then involves the selection of an appropriate number of layers to be transmitted. It makes it possible to adjust the transmission rate, and thus the video quality, to the bandwidth available for the transmission. When the bandwidth of the network path deteriorates, less layers are transmitted to reduce the queuing delays at the intermediate nodes, and to safeguard the timely delivery of layers already transmitted. At the same time, the number of transmitted layers is kept as high as possible to minimize the distortion at the receiver.

Recent approaches to rate adaptation with SVC aim at addressing the requirements of video streaming for entertainment purposes, which allow tens of seconds of delay. Therefore, they follow the long-term changes of the transmission rate, either based on the buffer occupancy at the receiver, or by estimating the transmission rate [13]–[20]. Long-term rate adaptation is then combined with buffering at the receiver in order to even out the short-term bandwidth variations. Nonetheless, these rate adaptation solutions cannot be applied for wireless network applications with strict delay limits, as the short-term variability of the wireless channels can not be compensated with in-network and receiver buffering. The rate adaptation problem under strict delay constraints is thus particularly challenging and calls for a novel solution approach.

In this paper we address the problem by proposing model-based rate adaptation for low-latency video streaming in wireless networks and utilizing a stochastic network calculus approach. A significant advantage of this approach is that it provides quantifiable measures of end-to-end quality of service (QoS) as a function of link quality. These measures can then be translated into useful quality of experience (QoE) measures for the video playout. The QoE measures can then be used for rate adaptation and performance optimization, as well as for the evaluation of new coding schemes.

We utilize the wireless extensions of stochastic network calculus to capture both the transmission of video bit streams over the time varying wireless links and the queuing delays at the intermediate nodes [21], [22]. We extend previous results to take the playout process into account and to combine two different time scales, the bit-stream-based transmission and queuing as well as the video-frame-based playout. We validate the analytic model through extensive simulations and demonstrate the efficiency of the model-based source rate adaptation. We show that the analytic model supports the design of new video coding schemes and that it is helpful for cross-layer-optimized delay sensitive routing.

The rest of the paper is organized as follows. Section II discusses recent results on video playout optimization and stochastic network calculus. Section III presents background regarding the methodology used. Section IV describes the considered system. Section V presents the model and provides a lower bound on the playout rate under reliability constraint. Possible modeling extensions are discussed in Section VI. The model is validated in Section VII. In Section VIII we evaluate the efficiency of our model-based rate adaptation, including also transmission path optimization. Section IX concludes the paper.

II. Related Work

As the importance of SVC is widely recognized, scalable extensions of video coding standards are available for H.264/AVC [10], as well as SHVC for H.265/HEVC [11], and new, highly scalable solutions are subject to current research [12]. The objective of SVC is to provide temporal, spatial, and quality scalability by encoding the video stream into multiple layers, a base layer and several enhancement layers, using interlayer processing. As a result, an enhancement layer can be decoded if all previous layers are fully received. The layered structure facilitates a trade-off between video quality (i.e., distortion) and required bandwidth (i.e., rate). This property becomes handy for transmission over wireless and mobile networks [23], where the underlying link quality is subject to the channel variability [24]–[27]. The same property makes SVC attractive for delay-limited applications, since all layers received completely before the playout deadline can be utilized for the decoding. Encoding with a larger number of layers increases the potential of more efficient rate adaptation. However, the actual SVC implementation may limit this potential (i.e. by using limited number of layers) due to complexity and/or efficiency constraints (i.e. due to a high layering overhead bitrate overhead resulting from rate-distortion-inefficient scalable coding). Nevertheless, the rapid technological advances may soon render such limitations obsolete. It is therefore important to quantify the performance gains when using a high number of layers in various application domains.

Recently proposed rate adaptation methods focus on maximizing the quality of experience of video streaming. They consider rebuffering and the frequency of rate switching as quality metrics, and allow client side buffering of tens of seconds. The proposed methods are based on the buffer content [13]–[15], transmission rate estimation [16]–[19], or both [20]. Others build on Markov decision processes [14], Lyapunov optimization [15], control theoretic solutions [18], or Markovian throughput prediction [19]. All these proposed solutions share the advantage that detailed modeling of the network performance is not required.

Low delay applications, with a delay requirement within a second however can not build on buffer-content-based models, where the changes in the buffer content reflect the changes in the networking environment. Results presented in the literature consider second to tens of seconds of playout delays. Similarly for low latency requirements, rate adaptation based on average transmission rate would be overly optimistic; it would result in queuing delays at the network nodes and late arrivals at the playout buffer. Therefore, in this paper we propose rate adaptation based on network performance modeling for low latency wireless applications.

Performance modeling of adaptive video streaming in wireless networks has mostly been considered for a single wireless link. In [26] the effect of an unreliable wireless channel is modelled by an i.i.d packet loss process, and the video coding rate and the packet size are optimized under retransmission-based error correction. In [27] and [28] adaptive media playout and adaptive layered coding is addressed respectively. Both papers define a queuing model on a video frame level, assuming that the wireless channel results in a Poisson frame arrival process at the receiving terminal, a simplification that may be reasonable if the buffering at the receiver side is significant, and therefore packet level delays do not need to be taken into account.

Modeling of video streaming based on network calculus is presented in [29] for the purpose of resource allocation in cellular networks, again, considering a frame level model. Modeling of video transmission over two wireless links is presented in [30]. This work considers the video transmission as a bitstream, but even with this simplifying assumption the results reflect that modeling based on traditional queuing theory quickly becomes intractable as the number of links increases. In [31] a tractable model is derived for the delay violation probability for fluid transmission over multihop wireless links, following the effective capacity concept. This approach however does not lend itself to frame or GOP level modeling.

In this paper we propose model-based rate adaptation using network calculus, advancing previous results through careful modeling of GOP-based video coding and decoding, combined with packet-based transmission over wireless links. Network calculus characterizes the departure process and the network backlog over multihop paths. Together with recent advances on modeling wireless links, this motivates our approach.

Stochastic network calculus has been extended to capture the randomly varying channel capacity of wireless links, following different methods Most of the existing work builds on an abstracted finite-state Markov channel (FSMC) model of the underlying fading channel, e.g., [33], [34] or uses moment generating function based network calculus [36]. However, the complexity of the resulting models limits the applicability of these approaches in multi–hop wireless network analysis with

more than a few states and more than two hops. In this work, we follow the approach proposed by Al-Zubaidy et al [21], where a wireless network calculus based on the $(\min, \times)$ dioid algebra was developed. The main premise for this approach is that the channel capacity, and hence the offered service of fading channels is related to the instantaneous received SNR through the logarithmic function as expressed by the Shannon capacity, $C(\gamma) = \log(1+\gamma)$. Hence, an equivalent representation of the channel capacity in an isomorphic transform domain, obtained using the exponential function, is $e^{C(\gamma)} = 1+\gamma$. This simplifies the otherwise cumbersome computations of the end-to-end performance metrics.

III. Network Calculus for Wireless Networks

Network calculus has been developed to provide an efficient analytic tool for evaluating the quality of service provided by networks with multi-hop transmission path, including the effect of correlated buffering at the network nodes. In network calculus, the generated network traffic at node k in time interval $[\tau, t)$ is characterized by the cumulative arrivals, that is, the real–valued non–negative bivariate process $A_k(\tau, t)$, while the transmission capabilities of node k are described by the process of cumulative services $S_k(\tau, t)$. The resulting departure process, $D_k(\tau, t)$, characterizes the cumulative traffic leaving node k. These processes are non-decreasing in t with $A_k(t,t) = S_k(t,t) = D_k(t,t) = 0$ and $A_k(0,t) \geq D_k(0,t)$ for all t. The objective of stochastic network calculus is to derive the departure processes for complex network topologies, and based on that express the network performance, typically in terms of probabilistic bounds on the end-to-end delay $W(t)$, and the backlog $B(t) = A(0,t) - D(0,t)$, characterizing the amount of traffic delayed in the transmission queues of the network. Network calculus can be used to analyze networks with either packetized or fluid flow traffic and for discrete or continuous time scale. In this work, we consider fluid flow traffic and discrete (slotted) time.

Introduced in [21], the $(\min, \times)$ network calculus transforms the problem into an alternative domain, called SNR domain, where the SNR service process ($\mathcal{S}_i$) is obtained by taking the exponent of the original service process[1], i.e., $\mathcal{S}_i = e^{S_i}$. Therefore, we refer to a network element i as *dynamic SNR server*, if it offers a service $\mathcal{S}_i$ that satisfies the input–output inequality [37], $\mathcal{D}(0,t) \geq \mathcal{A} \otimes \mathcal{S}_i(0,t)$, where the $(\min, \times)$ convolution and deconvolution are respectively defined for any two SNR processes $\mathcal{X}_1(\tau, t)$ and $\mathcal{X}_2(\tau, t)$ as

$$\mathcal{X}_1 \otimes \mathcal{X}_2(\tau,t) \stackrel{\triangle}{=} \inf_{\tau \leq u \leq t} \left\{ \mathcal{X}_1(\tau,u) \cdot \mathcal{X}_2(u,t) \right\},$$

$$\mathcal{X}_1 \oslash \mathcal{X}_2(\tau,t) \stackrel{\triangle}{=} \sup_{u \leq \tau} \left\{ \frac{\mathcal{X}_1(u,t)}{\mathcal{X}_2(u,\tau)} \right\}.$$

The key result of network calculus is the possibility to substitute the sequence of service processes on a multi-hop transmission path with a single network service process, S_{net}, by concatenating the service processes for all nodes along a path [38]. In the SNR domain

$$\mathcal{S}_{\text{net}}(\tau,t) = \mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \cdots \otimes \mathcal{S}_N(\tau,t). \tag{1}$$

In addition, network performance bounds, e.g., end-to-end delay and backlog, can be obtained in terms of the $(\min, \times)$ deconvolution of the SNR arrival and service processes [21].

The computation of the $(\min, \times)$ convolution and deconvolution operations are not straight forward as it involves the evaluation of products and quotients of random processes. Thus, an exact solution for (1) may not be feasible. Instead, we may use yet another transform, the Mellin transform, to find bounds on these two operations. The Mellin transform, see [39], is defined for a nonnegative random variable Z as $\mathcal{M}_Z(s) = E[Z^{s-1}]$, for any complex valued s given that the expectation exists. Then, the following holds [21]:

Lemma 1. *Let $\mathcal{S}_1(\tau,t)$ and $\mathcal{S}_2(\tau,t)$ be two independent SNR service processes. The Mellin transform of $\mathcal{S}_1 \otimes \mathcal{S}_2(\tau,t)$, for all $s<1$, is bounded by*

$$\mathcal{M}_{\mathcal{S}_1 \otimes \mathcal{S}_2}(s,\tau,t) \leq \sum_{u=\tau}^{t} \mathcal{M}_{\mathcal{S}_1}(s,\tau,u) \cdot \mathcal{M}_{\mathcal{S}_2}(s,u,t). \tag{2}$$

The Mellin transform of $\mathcal{S}_1 \oslash \mathcal{S}_2(\tau,t)$, for $s>1$, is given by

$$\mathcal{M}_{\mathcal{S}_1 \oslash \mathcal{S}_2}(s,\tau,t) \leq \sum_{u=0}^{\tau} \mathcal{M}_{\mathcal{S}_1}(s,u,t) \cdot \mathcal{M}_{\mathcal{S}_2}(2-s,u,\tau). \tag{3}$$

Lemma 1 above suggests that the Mellin transform of the $(\min, \times)$ convolution/deconvolution of two independent processes is bounded by a function of their Mellin transforms. In the case of wireless networks, the independence follows from the assumption on independent fading on the consecutive wireless links. Consequently, network performance bounds can be obtained in terms of the Mellin transforms of the SNR arrival and service processes of that network.

IV. System Model and Problem Formulation

In this section we describe our model of the wireless network and of video streaming, formulate the rate adaptation and routing problem, and provide the corresponding arrival and service models.

A. Wireless network model

We consider a time slotted multi-hop wireless network with a time slot duration of Δt. We use t to refer to a time slot. For each wireless link we consider a block fading channel [40] with Rayleigh fading distribution, with coherence time larger than Δt. As our focus is not on channel coding, we assume that a channel coding scheme is available at each node, such that each channel provides a service that is equivalent to its instantaneous Shannon channel capacity, $C(\gamma_{k,t}) = W \log_2(1+\gamma_{k,t})$ bits/s, where W is the channel bandwidth, $\gamma_{k,t}$ is the instantaneous SNR at the receiver of channel k at time slot t, and we consider that $\gamma_{k,t} \stackrel{d}{=} \gamma_k, \forall t$, where $\stackrel{d}{=}$ denotes equal in distribution, with average $\bar{\gamma}_k$. We allow the average SNR $\bar{\gamma}_k$ to change over time, but we

[1] We use the calligraphic upper–case letters to represent traffic and service processes in the SNR domain and to distinguish them from their bit domain (where traffic and service are measured in bits) counterparts.

make the reasonable assumption that it is known. Feedback of the channel state information (CSI) over a single link is implemented or can be implemented in modern networks [43], [44], while the CSI or estimated SNR values can be collected in a mesh network with the help of the routing protocol [45].

We consider that video has to be streamed between two nodes in the wireless network, as shown in Fig. 1. We refer to a sequence of links from the sender to the receiver node as a transmission path, and denote it by $\mathcal{P}$. We denote the length of the path by N. We assume that buffers at intermediate nodes are locally FIFO, i.e., frames (belonging to the same flow) and their contents are served according to the order of their arrival. Intermediate nodes do not drop or duplicate traffic.

B. Scalable video streaming

The video is captured at a rate of n frames per second. Depending on the considered coding scheme, a video frame can represent a single image, or a group of pictures (GOP). The captured frames are fed directly to the SVC encoder that generates L layers. We consider that each layer has a size of $m + h$ bits, where an ideal scalable video coder would require m bits, and h is the overhead due to rate-distortion inefficient scalabe coding. This results in a constant rate traffic of $R_E = nr = (m + h)nL$ bits per second, where r is the frame size in bits, and is determined by the number of transmitted layers L. Although, we assume equal size layers and constant rate traffic for ease of presentation, the approach can be extended to handle unequal layer size. We discuss this extension in Section VI. The assumption of equal layer size may be realistic for future scalable video coders, when the number of layers is large and when video streaming needs to be supported in a wide variety of environments. It is also true that the assumption of constant rate traffic is reasonable for GOP-based coding. Even though GOP sizes are varying according to the video content, their variation is much slower when compared to the variation of the image sizes. Nevertheless, the proposed methodology is not limited to constant rate traffic and it can handle variable rate video as well. We provide further discussion regarding this matter in Section VI.

The coded video frames are transmitted over a wireless network of N transmission links. Once transmitted over the wireless network, received bits are stored in a playout buffer. Video frames are decoded and played out regularly with $T_f = 1/n$ time intervals and a fixed playout delay T_D, that is, a video frame i that is generated at time τ_i is played out after a fixed delay T_D at time $\tau_i + T_D$. According to the layered coding, only the completely received layers are used for decoding. Due to the variability of the wireless channels, the number of layers of a video frame i that are received within a deadline T_D is random, leading to a varying per frame playout bitrate R_D at the decoder, and consequently a varying distortion.

Discarding layers with already expired playout deadlines at intermediate nodes would improve the streaming quality. We do not consider this option, since it would require the identification of the layers at the intermediate nodes.

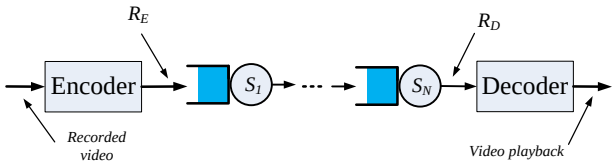

Fig. 1. Video transmission over multi-hop wireless network.

C. Model-based rate adaptation and routing problem

Given the system model presented above, the objective of the model-based rate adaptation and routing problem is to select the optimal number of transmitted layers L^* and the optimal transmission path $\mathcal{P}^*$ that together maximize the lower bound r_D^ε of the playout rate under a reliability constraint ε:

$$\max_{(L,\mathcal{P})} r_D^\varepsilon \tag{4}$$

s.t.

$$Pr(R_D < r_D^\varepsilon) \quad \leq \quad \varepsilon \tag{5}$$

Due to the variability of the wireless channels and the queuing at the intermediate nodes, there is no tractable analytic expression for the distribution of R_D. Therefore, in the following we provide a bound on the tail distribution of R_D, as a function of the video frame size, the average SNR values and the path length N. We then show how to use it for solving the model-based rate adaptation and routing problem.

D. Arrival and Service Models

Recall that R_E is the bitrate of the coded video, and is thus the arrival rate to the wireless network. We can then express the cumulative arrival process as[2]

$$A(\tau, t) = R_E(t - \tau) = (m + h)nL(t - \tau)\,, \tag{6}$$

and the SNR arrival process $\mathcal{A}$ is given by

$$\mathcal{A}(\tau, t) = e^{R_E(t-\tau)} = e^{(m+h)nL(t-\tau)}\,. \tag{7}$$

Hence, the Mellin transform of the arrival process can be expressed as

$$\mathcal{M}_{\mathcal{A}}(s, \tau, t) = e^{(s-1)(m+h)nL(t-\tau)}\,. \tag{8}$$

Similarly, we can define the cumulative service process of a fading channel with SNR $\gamma_{k,u}$ is

$$S(\tau, t) = W \sum_{u=\tau}^{t-1} \log(1 + \gamma_{k,u})\,, \tag{9}$$

Its SNR domain counterpart is given by the log-free form

$$\mathcal{S}(\tau, t) = \prod_{u=\tau}^{t-1} (1 + \gamma_{k,u})^W\,. \tag{10}$$

The Mellin transform of $\mathcal{S}$ depends on the distribution of $\gamma_{k,u}$, i.e., the fading distribution. In Section V-C, we will derive the Mellin transform of the service process for Rayleigh channels.

[2]The process $A(\tau, t)$ can also be obtained from real traces. This requires an extra traffic modelling step. The proposed approach can handle a random arrival process as long as its Mellin transform exists and is obtainable.

V. Performance of Video Communication

In this section we present our main contribution: A system model for adaptive video transmission over a multi-hop wireless network and a bound on the received video quality in terms of the parameters of the transmitted video, as well as the underlying fading channels' parameters. We first derive a general expression on the lower bound of the received rate under playout delay constraint and frame based transmission. Then we give the bound for transmissions over multihop wireless channels, specifically considering Rayleigh fading. Finally we derive the bound on the playout bitrate, considering the layered structure, and address the feasibility of solving the optimal rate adaptation problem.

A. Lower Bound for the Received Rate

We investigate a video decoder which operates as follows: At time $\tau_i + T_D$ it considers the content of the playout buffer. It then drops all content that belongs to video frames $j < i$ (i.e., late arrivals from previous frames), then removes and decodes all frame content that belongs to frame i; arrivals from subsequent frames remain in the playout buffer. The modelling challenge is twofold. First, the received rate should include only data that belongs to a given video frame. Second, we would like to derive a lower bound of the received rate R_D^i, while network calculus usually considers its upper bound to characterize backlog and delay.

A statistical description of the rate at the decoder, R_D^i, can be obtained by observing the departure process of the wireless network, $D(\tau, t)$. Specifically, R_D^i can be obtained by considering all departures during the time period from video frame generation until playout, that is, $D(\tau_i, \tau_i + T_D)$, and then counting only $D^i(\tau_i, \tau_i + T_D)$, the part of the traffic that belongs to video frame i. Since the instantaneous received rate at the decoder, R_D^i, includes only traffic that belongs to fully received layers by the frame's playout deadline, we can write

$$R_D^i = \frac{\left\lfloor \frac{D^i(\tau_i, \tau_i + T_D)}{m+h} \right\rfloor \cdot m}{T_f}. \tag{11}$$

A probabilistic lower bound on the departures belonging to video frame i during the period $[\tau_i, \tau_i + T_D)$, $D^i(\tau_i, \tau_i + T_D), i = 1, 2, \cdots$, is given by the following lemma.

Lemma 2. *Given a video frame i generated at time τ_i and destined to a decoder with playout deadline T_D, the departure process $D^i(\tau_i, \tau_i + T_D), i = 1, 2, \cdots$, is characterized as follows*

$$\begin{aligned} &Pr(D^i(\tau_i, \tau_i + T_D) \le d) \le \varepsilon \\ &\Leftrightarrow Pr(D(0, \tau_i + T_D) \le d + (m+h)nL\tau_i) \le \varepsilon\,, \end{aligned} \tag{12}$$

for all $d \le (m+h)L$ and all $\varepsilon \in [0, 1]$.

It is worth noting that this probability is equal to 1 for all $d > (m+h)L$, i.e., departures belonging to a video frame i can never exceed the frame size.

Proof. Fig. 2 shows the encoding, transmission, and decoding of the consecutive $i-1, i, i+1$ frames and can be used to derive $D^i(\tau_i, \tau_i + T_D)$. We express $D^i(\tau_i, \tau_i + T_D)$ by first considering all departures $D(\tau_i, \tau_i + T_D)$ and then removing traffic that does not belong to frame i. As shown in Fig. 2, the departures belonging to previous frames within the interval $[\tau_i, \tau_i + T_D)$ are equal to the backlog $B(\tau_i)$ at time τ_i, i.e., the traffic from all previous frames that is still in the network when the i^{th} frame arrives. Once we remove $B(\tau_i)$ the remaining departures belong to frame i, up to a size of $(m+h)L$, followed by traffic from subsequent frames. Using this argument we arrive at the following equivalence statement

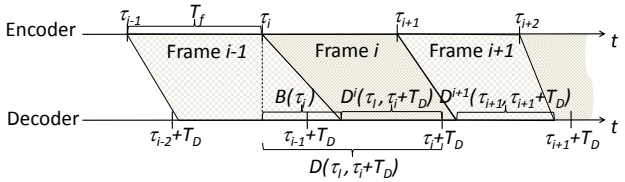

Fig. 2. Determination of $D^i(\tau_i, \tau_i + T_D)$.

$$\begin{aligned} &Pr(D^i(\tau_i, \tau_i + T_D) \le d) \le \varepsilon \\ &\Leftrightarrow Pr(D(\tau_i, \tau_i + T_D) - B(\tau_i) \le d) \le \varepsilon\,, \end{aligned} \tag{13}$$

for all $d \le (m+h)L$.

Then using the fact that the backlog at any time τ is given by the difference of all arrivals and all departures from time $t = 0$, where $B(0) = 0$, until time $t = \tau$, the right hand side of (13) can be evaluated as follows

$$\begin{aligned} &Pr(D(\tau_i, \tau_i + T_D) - B(\tau_i) \le d) \\ &\quad = Pr(D(\tau_i, \tau_i + T_D) - A(0, \tau_i) + D(0, \tau_i) \le d) \\ &\quad = Pr(D(0, \tau_i + T_D) - A(0, \tau_i) \le d) \\ &\quad = Pr(D(0, \tau_i + T_D) \le d + A(0, \tau_i)) \\ &\quad = Pr(D(0, \tau_i + T_D) \le d + (m+h)nL\tau_i)\,. \end{aligned} \tag{14}$$

Substituting (14) in (13), the lemma follows. □

Lemma 2 states that a probabilistic lower bound for D^i can be obtained in terms of the probabilistic lower bound on the departure process D given by (14), for the specific arrival process described in (6). When the arrival and service processes have stationary increments, which is the case here, then D^i is identically distributed for all i. The next step is to derive this bound, which we can accomplish using network calculus.

Lemma 3. *For any work-conserving server with dynamic bivariate service process $S(\tau, t)$ and an arrival process $A(\tau, t)$, the departure process $D(\tau, t)$ is bounded as*

$$D(\tau, t) \ge A \oplus S(\tau, t) = \inf_{\tau \le u \le t} \left\{ A(\tau, u) + S(u, t) \right\}, \tag{15}$$

where $\oplus$ denotes the $(\min, +)$ convolution.

Proof. Using Reich's recursive backlog formula we have

$$\begin{aligned} B(t) &= [B(t-1) + a(t) - s(t)]^+ \\ &\le \sup_{0 \le u \le t} \left\{ A(u, t) - S(u, t) \right\}, \end{aligned}$$

where $a(t), s(t), B(t)$ are the instantaneous arrival, service and backlog at time slot t respectively. Hence,

$$\begin{aligned} D(0,t) &= A(0,t) - B(t) \\ &\geq \inf_{0\leq u\leq t}\left\{A(0,u)+S(u,t)\right\} \quad = A \oplus S(0,t)\,, \end{aligned}$$

where in the second step we used the fact that $A(0,u) = A(0,t) - A(u,t)$.

But due to causality we also have

$$D(0,\tau) \leq A(0,\tau)\,.$$

Then for work-conserving server and for $0 \leq \tau \leq t$ we get

$$\begin{aligned} D(\tau,t) &= D(0,t) - D(0,\tau) \\ &\geq A \oplus S(0,t) - A(0,\tau) \\ &= \inf_{0\leq u\leq t}\left\{A(0,u) - A(0,\tau) + S(u,t)\right\}. \end{aligned}$$

Since A is non-decreasing, $A(0,u) - A(0,\tau) = 0\,, \forall u < \tau$, and since

$$\exists u \geq \tau \quad \text{s.t.} \quad A(\tau,u) + S(u,t) \leq S(\tau,t)\,.$$

Then, $D(\tau,t) \geq \inf_{\tau\leq u\leq t}\left\{A(\tau,u)+S(u,t)\right\}$ and the lemma follows. □

B. Departure Process Lower Bound for Wireless Channels

To use Lemma 3, we must evaluate the right hand side of (15) which is not an easy task for a wireless channel where $S(\tau,t)$ is a randomly varying process due to random fading. Therefore, with the following theorem we provide a probabilistic bound on $D(\tau,t)$ in terms of the Mellin transform of the arrival and service process by using the $(\min,\times)$ network calculus approach [21].

Theorem 1. *Let $\mathcal{A}$ be the SNR arrival process to a work-conserving queuing system with SNR service process $\mathcal{S}$, then for any $0 \leq \tau \leq t$ and any $s < 1$, a lower bound d ($d \geq 0$) for the departure process $D(\tau,t)$ must satisfy the following inequality*

$$Pr(D(\tau,t) \leq d) \leq e^{(1-s)d} \sum_{u=\tau}^{t} \mathcal{M}_{\mathcal{A}}(s,\tau,u) \cdot \mathcal{M}_{\mathcal{S}}(s,u,t)\,. \tag{16}$$

Proof. We start by formulating a probabilistic lower bound on the departures in terms of the SNR departure process $\mathcal{D}$, $\forall s < 1$, as follows

$$\begin{aligned} Pr(D(\tau,t) \leq d) &= Pr(\mathcal{D}(\tau,t) \leq e^d) \\ &= Pr(\mathcal{D}^{s-1}(\tau,t) \geq e^{(s-1)d}) \\ &\leq e^{(1-s)d}\mathcal{M}_{\mathcal{D}}(s,\tau,t)\,, \end{aligned} \tag{17}$$

where we used the assumption $s < 1$ to obtain the second line and then we applied Markov's inequality and used the definition of the Mellin transform to arrive at the last step.

Using Lemma 3, the SNR departure process $\mathcal{D}$ can be bounded as follows

$$\begin{aligned} \mathcal{D}(\tau,t) = e^{D(\tau,t)} &\geq e^{\inf_{\tau\leq u\leq t}\left\{A(\tau,u)+S(u,t)\right\}} \\ &= \inf_{\tau\leq u\leq t}\left\{e^{A(\tau,u)+S(u,t)}\right\} \\ &= \inf_{\tau\leq u\leq t}\left\{\mathcal{A}(\tau,u)\cdot\mathcal{S}(u,t)\right\} \\ &= \mathcal{A}\otimes\mathcal{S}(\tau,t)\,, \end{aligned} \tag{18}$$

where we used the definition of the $(\min,\times)$ convolution in the last step.

Note that when $s > 1$, the Mellin transform is order-preserving. On the other hand, when $s < 1$, the order is reversed [21]. Hence, the Mellin transform for the SNR departure process for any $s < 1$ is computed using (18) as follows

$$\begin{aligned} \mathcal{M}_{\mathcal{D}}(s,\tau,t) &\leq \mathcal{M}_{\mathcal{A}\otimes\mathcal{S}}(s,\tau,t) \\ &= E\left[\left(\inf_{\tau\leq u\leq t}\left\{\mathcal{A}(\tau,u)\cdot\mathcal{S}(u,t)\right\}\right)^{s-1}\right] \\ &= E\left[\sup_{\tau\leq u\leq t}\left\{\left(\mathcal{A}(\tau,u)\cdot\mathcal{S}(u,t)\right)^{s-1}\right\}\right] \\ &\leq \sum_{u=\tau}^{t}\mathcal{M}_{\mathcal{A}}(s,\tau,u)\cdot\mathcal{M}_{\mathcal{S}}(s,u,t)\,, \end{aligned} \tag{19}$$

where we used the non-negativity of $\mathcal{A}$ and $\mathcal{S}$ and their independence and then we applied the union bound in the last step.

Substituting (19) into (17), the theorem follows. □

C. Lower Bound for Multihop Rayleigh Channels

We will now use Theorem 1 to obtain a probabilistic lower bound on the departure process for an N-hop wireless network subject to Rayleigh fading. For simplicity, we present results for the case when for $\bar{\gamma}_k = \bar{\gamma}$ for the N hops, but the methodology works for non-identically distributed channel fading using a more complex representation of the network service process as we show later in Section VI.

The instantaneous SNR γ_t of a Rayleigh fading channel is exponentially distributed with average $\bar{\gamma}$. Then the Mellin transform for the cumulative service process of a Rayleigh fading channel defined in (10) is given by [21]

$$\mathcal{M}_{\mathcal{S}}(s,\tau,t) \leq \left(e^{\frac{1}{\bar{\gamma}}}\bar{\gamma}^{s-1}\Gamma(s,\bar{\gamma}^{-1})\right)^{t-\tau}\,, \tag{20}$$

where $\Gamma(s,a) = \int_a^\infty x^{s-1}e^{-x}dx$ is the incomplete Gamma function.

Theorem 2. *A probabilistic lower bound on the departure of N-hop i.i.d. Rayleigh channels with average SNR $\bar{\gamma}$, when the arrival process is given by* (6)*, and for $0 \leq \tau \leq t$ is*

$$Pr(D(\tau,t) \leq d(t-\tau)) \leq \inf_{s<1}\left\{\frac{e^{(s-1)((m+h)nL(t-\tau)-d(t-\tau))}}{\left(1-V(1-s)\right)^N}\right\} \tag{21}$$

whenever the stability condition

$$V(1-s) \stackrel{\triangle}{=} e^{(1-s)(m+h)nL}e^{\frac{1}{\bar{\gamma}}}\bar{\gamma}^{s-1}\Gamma(s,\frac{1}{\bar{\gamma}}) < 1 \tag{22}$$

is satisfied.

Proof. Let the service offered by a network of N store and forward nodes be characterized by the SNR service process $S(s,\tau,t) = \mathcal{S}_{\text{net}}(s,\tau,t)$. Then using Theorem 1 we obtain for all $s<1$

$$Pr(D(\tau,t)<d) \le e^{(1-s)d} \sum_{u=\tau}^{t} \mathcal{M}_{\mathcal{A}}(s,\tau,u)\cdot \mathcal{M}_{\mathcal{S}_{\text{net}}}(s,u,t)\,. \tag{23}$$

A bound on $\mathcal{M}_{\mathcal{S}_{\text{net}}}(s,\tau,t)$ for N i.i.d. Rayleigh fading channels is obtained by using the server concatenation property (1) and then applying the convolution bound in Lemma 1 repeatedly $N-1$ times,

$$\mathcal{M}_{\mathcal{S}_{\text{net}}}(s,\tau,t) \le \sum_{u_1,\ldots,u_{N-1}} \prod_{n=1}^{N} \mathcal{M}_{\mathcal{S}_n}(s,u_{n-1},u_n)\,, \tag{24}$$

where the sum runs over all sequences $u_0 \le u_1 \le \cdots \le u_N$ with $u_0=\tau$ and $u_N=t$, $\mathcal{S}_n(\tau,t)$ is the SNR service process for node $n\in\{1,\ldots,N\}$.

In a homogeneous service processes setting (due to the assumption of i.i.d. channel gain) we can use the binomial identity

$$\sum_{u_1=\tau,u_2,\ldots,u_{N-1}=t} 1 = \binom{N-1+t-\tau}{t-\tau},$$

and can substitute into (20) to obtain for all $s<1$

$$\begin{aligned}\mathcal{M}_{\mathcal{S}_{\text{net}}}(s,\tau,t) &\le \binom{N-1+t-\tau}{t-\tau}\mathcal{M}_{\mathcal{S}}(s,\tau,t)\\ &\le \binom{N-1+t-\tau}{t-\tau}\left(e^{\frac{1}{\bar\gamma}}\bar\gamma^{s-1}\Gamma(s,\frac{1}{\bar\gamma})\right)^{t-\tau}. \end{aligned}\tag{25}$$

The binomial coefficient is the result of expanding the $N-1$ sums and then collecting all terms for the i.i.d channels case.

Substituting (6) and (25) in (23) we obtain the following probabilistic lower bound

$$\begin{aligned}Pr(D(\tau,t)\le d) &\le e^{(1-s)d}\sum_{u=\tau}^{t} e^{(s-1)(m+h)nL(u-\tau)}\\ &\quad\cdot\binom{N-1+t-u}{t-u}\left(e^{\frac{1}{\bar\gamma}}\bar\gamma^{s-1}\Gamma(s,\frac{1}{\bar\gamma})\right)^{t-u}\\ &\le e^{(s-1)((m+h)nL(t-\tau)-d)}\sum_{v=0}^{\infty}\binom{N-1+v}{v}\left(V(1-s)\right)^{v},\end{aligned}\tag{26}$$

where, we use the change of variables $v=t-u$, let $t\to\infty$ and define $V(1-s)\stackrel{\triangle}{=}e^{(1-s)(m+h)nL}e^{\frac{1}{\bar\gamma}}\bar\gamma^{s-1}\Gamma(s,\frac{1}{\bar\gamma})$.

Using the binomial identity

$$\sum_{v=0}^{\infty}\binom{N-1+v}{v}x^v = \frac{1}{(1-x)^N},$$

for all $N\ge 1$ and $|x|<1$, the sum in (26) converges to the following

$$Pr(D(\tau,t)\le d(t-\tau)) \le \frac{e^{(s-1)((m+h)nL(t-\tau)-d(t-\tau))}}{\left(1-V(1-s)\right)^N},$$

for all $s<1$, whenever the condition $V(1-s)<1$ is satisfied. Optimizing over s results in the best possible bound and concludes the proof. □

Note that the function $V(1-s)$ in (22) is defined in terms of the ratio of the Mellin transforms of the SNR arrival and the SNR service processes. This function approaches 1 when the traffic intensity increases towards the service capacity of the system and vice versa. Furthermore, the sum in (26) converges only if $V(1-s)<1$. Otherwise, the sum will not converge, the violation probability will always be 1, and the system enters unstable operation mode.

D. A Bound on Playout Bitrate R_D

Combining the results obtained in Lemma 2 and Theorem 2 for stable system operation, we can compute a lower bound on the departures $D^i(\tau,\tau+T_D)$ for all $s<1$ as follows

$$\begin{aligned}Pr(D^i(\tau,\tau+T_D)\le d) &= Pr(D(0,\tau+T_D)\le d+A(0,\tau))\\ &= Pr(D(0,\tau+T_D)\le d+(m+h)nL\tau)\\ &\le \inf_{s<1}\left\{\frac{e^{(s-1)((m+h)nLT_D-d)}}{\left(1-V(1-s)\right)^N}\right\},\end{aligned}\tag{27}$$

if $d\le (m+h)L$, otherwise, i.e., if $d>(m+h)L$, $Pr(D^i(\tau,\tau+T_D)\le d)=1$.

To obtain the lower bound on the departures such that $Pr(D^i(\tau,t)\le d^{\varepsilon})\le\varepsilon$, we equate the right hand side of (27) to ε and solve for d^{ε} to get

$$\begin{aligned}d^{\varepsilon}(T_D) \ge \min\Bigg[&(m+h)L,\ \sup_{s<1}\Big\{(m+h)nLT_D\\ &+\frac{1}{1-s}\left[N\log(1-V(1-s))+\log\varepsilon\right]\Big\}\Bigg].\end{aligned}\tag{28}$$

The first expression in the min operation in (28) is there to insure that what is received is no more than what was transmitted. The second expression shows that the departure per frame during the period T_D is governed by the amount of traffic transmitted during that time, the number of hops N, the target reliability of the bound ε, and the network utilization through the function $V(1-s)$.

Using (28), the distribution of the number of usable bits (i.e., bits received within the video frame's playback deadline, T_D) per second is bounded by

$$Pr(R_D < r_D^{\varepsilon}) \le \varepsilon\,,$$

where

$$r_D^{\varepsilon} \ge \frac{\left\lfloor \frac{d^{\varepsilon}(T_D)}{m+h}\right\rfloor\cdot m}{T_f}\,. \tag{29}$$

For steady state operation, this corresponds to the decodable rate per frame.

Note that the right hand side of (29) reduces to $\frac{mL}{T_f}$ when $d^{\varepsilon}(T_D)=(m+h)L$, i.e., all layers of the video frame are received within the playout deadline T_D. This can happen when the underlying wireless links have high channel quality during the video frame transmission, and it represents the best distortion performance that can be achieved for the given coding scheme.

E. Effect of T_D on Received Video Quality

The allowable playout delay T_D has a noticeable effect on the received video quality, as stated in the following corollary.

Corollary 1. *The lower bound on the departures per video frame $d^{\varepsilon}(T_D)$ increases linearly in the playout deadline T_D, independent from the network and channel conditions, and of the violation probability requirement.*

Proof. Rewriting (28) as follows

$$d^{\varepsilon}(T_D) \geq (m+h)nLT_D + \sup_{s<1}\left\{\frac{1}{1-s}\left[N\log(1-V(1-s))+\log\varepsilon\right]\right\}, \tag{30}$$

the corollary follows since the second term of (30) does not depend on T_D. □

VI. Model extensions

For ease of presentation and to be able to derive key conclusions, we made a number of simplifying assumptions in the model presented in Section V. In the following, we show how to relax those assumptions. We are able to address important realistic scenarios and to discuss further use cases for the model.

A. Video Coding with Unequal Layer Size

The arrival model in (6) assumes scalable coding with equal layer sizes. Although this may be a reasonable assumption when using arbitrarily many layers, it is useful to extend the methodology to allow for scalable coding with layers that have different sizes. To do that, we modify (6) as follows

$$A(\tau,t) = \sum_{j=1}^{L}(m_j+h_j)\cdot n\cdot(t-\tau)\,, \tag{31}$$

where m_j, h_j are the j^{th} layer message and overhead portions. Furthermore, the departure bound in (27) becomes

$$Pr(D^i(\tau,\tau+T_D)\leq d)\leq \inf_{s<1}\left\{\frac{e^{(s-1)(\sum_{j=1}^{L}(m_j+h_j)\cdot n\cdot T_D-d)}}{(1-V'(1-s))^N}\right\} \tag{32}$$

if $d \leq \sum_{j=1}^{L}(m_j+h_j)$, and $Pr(D^i(\tau,\tau+T_D)\leq d) = 1$ otherwise, where

$$V'(1-s) \triangleq e^{(1-s)sum_{j=1}^{L}(m_j+h_j)\cdot n} e^{\frac{1}{\bar{\gamma}}}\bar{\gamma}^{s-1}\Gamma(s,\frac{1}{\bar{\gamma}})\,.$$

The instantaneous received rate at the decoder, R_D^i, is then given by

$$R_D^i = \frac{1}{T_f}\sum_{j=1}^{J} m_j\,, \tag{33}$$

where,

$$J = \max\left\{l\leq L : D^i(\tau_i,\tau_i+T_D)\geq \sum_{j=1}^{l}(m_j+h_j)\right\}.$$

Then a probabilistic bound on R_D^i similar to that in (29) can be obtained using (32).

B. Service Model with Interfering Flows

The service model presented above for a single flow can be extended to capacity sharing between flows. On the one hand, when there is no information about the scheduling algorithm, this can be achieved by using the concept of leftover service process [22]. In this case, the interfering (cross) flow is assumed to have static priority over the evaluated (through) flow, the video streaming flow in this case, which results in that the interfering flow has maximum impact on video streaming. On the other hand, when the scheduling algorithm is known, the performance of the through flow can be enhanced. Network calculus provides service bounds for many important schedulers such as EDF and FIFO schedulers [46].

In what follows, we provide a leftover service process characterization which can be used to analyze the performance of scalable video streaming with background (interfering) traffic. Let us denote by $\mathcal{A}_o$ the SNR arrival process of the tagged (through) flow, and by $\mathcal{A}_c$ that of the other (cross) flows. We can then describe the service offered to the through flow by the leftover service process, which can be characterized as shown in Lemma 4 in [22].

Lemma 4. *Consider a network element with a through flow $\mathcal{A}_o$ and cross traffic flow $\mathcal{A}_c$. Assume that the network element provides a dynamic SNR server to the aggregate of the two flows, with service process $\mathcal{S}(\tau,t)$ then*

$$\mathcal{S}_o(\tau,t) = \max\left\{1, \frac{\mathcal{S}(\tau,t)}{\mathcal{A}_c(\tau,t)}\right\}$$

is a dynamic SNR server satisfying for all $t\geq 0$ that

$$\mathcal{D}_o(0,t) \geq \mathcal{A}_o \otimes \mathcal{S}_o(0,t)$$

The proof of Lemma 2 can be found in [22].

The Mellin transform for the leftover service is given by

$$\mathcal{M}_{\mathcal{S}_o}(s,\tau,t) = \mathcal{M}_{\mathcal{S}}(s,\tau,t)\cdot\mathcal{M}_{\mathcal{A}_c}(2-s,\tau,t)\,.$$

Inserting the above in (25) we have

$$\mathcal{M}_{\mathcal{S}_{\text{net}}}(s,\tau,t)\leq \binom{N-1+t-\tau}{t-\tau}\mathcal{M}_{\mathcal{S}}(s,\tau,t)\mathcal{M}_{\mathcal{A}_c}(2-s,\tau,t)\,. \tag{34}$$

Assuming a constant rate cross traffic with rate ρ_c, a bound on the departure process, $D^i(\tau,\tau+T_D)$, can still be obtained from (27) if the function V is replaced with the function V'', defined as follows

$$V''(1-s) = e^{(1-s)(m+h)nL}\, e^{(1-s)\rho_c}\, e^{\frac{1}{\bar{\gamma}}}\bar{\gamma}^{s-1}\Gamma(s,\frac{1}{\bar{\gamma}})\,.$$

Using the above, a probabilistic bound on the decodable rate per frame, and hence the playback quality, is obtainable using (28) and (29).

C. Optimal Rate Adaptation

The bounds (27) – (29) characterize the effects of the system parameters on the overall system performance, under the considered Rayleigh fading process with given $\bar{\gamma}$. We propose to utilize these bounds to approximate the optimal operating point of the scalable video coder.

For a given transmission path $\mathcal{P}$ and for known average $\bar{\gamma}_k$, the selection of the optimal number of layers per video frame L^* requires the evaluation of the right hand side (RHS) of (28). As $V(1-s)$, $s<1$, is convex in L, the number of layers per frame, the RHS of (28) can be shown to be concave in L. It can also be easily shown that $V(1-s)$, $s<1$, is convex in s whenever $V(1-s)<1$ (see [42]) and hence, $N\log(1-V(1-s))$ in the RHS of (28) is concave in s. Therefore, the optimal L^* and the corresponding bound $d^{\varepsilon}(T_D)$ can be efficiently obtained via binary search.

Observe that the above optimization is based on a performance bound, and hence the obtained solution does not necessarily coincide with the true optimum. Therefore we will evaluate the accuracy of the approach via simulations.

D. Heterogeneous Networks and Routing

The model presented in Section V builds on the closed form bound of $\mathcal{M}_{\mathcal{S}_{\rm net}}$ in (25). Such a closed form bound does not exist for heterogeneous network paths. Nevertheless, a bound on the Mellin transform of the network service process can still be obtained by computing (24) directly. A possible alternative modeling approach for heterogeneous networks is proposed in [41], providing a recursive equation for computing a lower bound on the service quality over an n hop path $\mathbb{L}$, based on the service quality of the $n-1$ path before adding the $n^{\rm th}$ hop (i.e., $\mathbb{L}\setminus\{n\}$) and that of the $n-1$ path when replacing the $m^{\rm th}$ hop, $m<n$, with the $n^{\rm th}$ hop (i.e., $\mathbb{L}\setminus\{m\}$), as follows.

Let, $\mathcal{M}^{\mathbb{L}}_{\mathcal{S}}(s,\tau,t)$ be the Mellin transform for the SNR network service process for the path defined by the ordered set of nodes $\mathbb{L}$. Then, $\forall s<1$,

$$\mathcal{M}^{\mathbb{L}}_{\mathcal{S}}(s,\tau,t)\le\left(\frac{\mathcal{M}_{g(\gamma_n)}(s)}{\mathcal{M}_{g(\gamma_n)}(s)-\mathcal{M}_{g(\gamma_m)}(s)}\mathcal{M}^{\mathbb{L}\setminus\{m\}}_{\mathcal{S}}(s,\tau,t)\right)+\left(\frac{\mathcal{M}_{g(\gamma_m)}(s)}{\mathcal{M}_{g(\gamma_m)}(s)-\mathcal{M}_{g(\gamma_n)}(s)}\mathcal{M}^{\mathbb{L}\setminus\{n\}}_{\mathcal{S}}(s,\tau,t)\right),$$

for $\mathcal{M}^{\{1\}}_{\mathcal{S}}(s,\tau,t)=\big(\mathcal{M}_{g(\gamma_1)}(s)\big)^{t-\tau}$ and for any $m\in\{1,2,\ldots,n-1\}$. Applying the above recursively $n-1$ times, starting from $|\mathbb{L}|=1$, a path with $|\mathbb{L}|=n$ can be obtained, where, $|\mathbb{L}|$ is the cardinality of the set $\mathbb{L}$.

This result suggests a solution for finding the transmission path that provides the best video streaming quality in the wireless network. The solution would build on the combination of the calculation of the streaming quality over path segments based on [41], the modeling approach in Section V, and an appropriate routing decision at alternative segments. We leave the design and the evaluation of such a routing algorithm for future work.

E. Variable-Rate Arrival Process

Although we assume a constant-rate arrival process as defined in (6), the proposed methodology can handle a randomly-varying arrival process as long as its Mellin transform exists and is obtainable. It has been suggested in the literature (e.g., in [47] and references within) that a GoP-based video traffic can be accurately modeled as a Markov process. We derive next the Mellin transform for a Markov modulated arrival process.

Lets define the M-state Markov modulated arrival process with rate r_i when in state $i\in\{1,\ldots,M\}$ and a transition matrix $\mathbf{P}$. Define $\phi_i(s)=Ee^{sr_i}$ and let $\phi(s)=diag(\phi_1(s),\ldots,\phi_M(s))$. Then a bound on the Mellin transform of this arrival process is given by $sp\big(\phi(\mathbf{s}-\mathbf{1})\mathbf{P}\big)$, where $sp(B)$ is the "spectral radius" of the matrix B defined as the maximum of the absolute values of the eigenvalues of that matrix [37].

As a demonstrating example, for $M=2$, the spectral radius of $\phi(\mathbf{s})\mathbf{P}$ is given by

$$sp\big(\phi(\mathbf{s})\mathbf{P}\big)=\frac{1}{2}\Bigg(P_{11}\phi_1(s)+P_{22}\phi_2(s)+\sqrt{(P_{11}\phi_1(s)-P_{22}\phi_2(s))^2+4P_{12}P_{21}\phi_1(s)\phi_2(s)}\Bigg)$$

and the Mellin transform for the SNR arrival process is then given by

$$\mathcal{M}_{\mathcal{A}}(s,\tau,t)\le\Big[sp\big(\phi(\mathbf{s}-\mathbf{1})\mathbf{P}\big)\Big]^{t-\tau}$$

VII. Model validation and performance evaluation

The analytic model described in Section V provides a lower bound on the departures per frame d within the playout deadline T_D. Therefore, we first validate the bounds via simulation. Then, we evaluate the effect of the network and video streaming parameters on the received quality, based on the results in (27) and (28). The effect of T_D have been addressed by Corollary 1.

We consider an SVC scheme that encodes group of pictures (GOP), that is, a frame in the analytic model represents a GOP. One GOP consists of 10 video images. 25 images are generated per second, which results in $n=2.5$ frames per second. The video is coded with $L=4$ to 24 layers of size $m=100$ kbits of video payload each, resulting in a payload of $r=0.4-2.4$ Mbits per frame, and a video transmission rate of $1-6$ Mbps, typical rates from standard to HD content [13], [15]. For default, we consider $h=0$. The playout deadline is $T_D=450$ msec, which corresponds to a strict delay constraint for real-time machine-to-machine video delivery. We consider transmission paths of $N=1,3,5$ links, a channel of bandwidth $W=2.2$ MHz and average SNRs of the fading channels in the range of $\bar{\gamma}=6-10$ dB. This corresponds to average channel capacities of $C_{avg}=4.24-6.39$ Mbps. We select a small range of channel capacities on purpose to show that rate adaptation is important already at low rate variations. We choose a slot duration of $\Delta t=10$ msec.

We ran the simulations for a period of 10^{10} time slots, which allows an empirical evaluation of the system performance up to a violation probability of $\varepsilon=10^{-8}$.

Fig. 3 shows the CDF of the departures per frame d, not yet considering the effect of the layering at the decoding, for $N=3$ and for various transmitted frame size r and average channel SNR values $\bar{\gamma}$. For reference, the channel utilization for the case $r=2.08$ Mb is 0.8 under $\bar{\gamma}=10$ dB, and is 0.99 for $\bar{\gamma}=8$ dB.

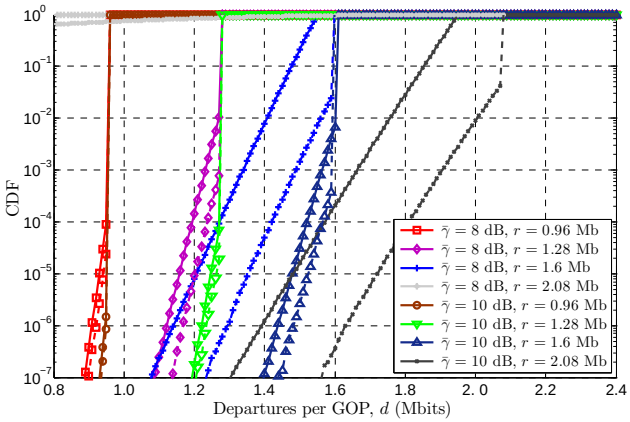

Fig. 3. Violation probability (ε^d) (computed and simulated) vs. departure bound d for SVC over multi-hop wireless network for different GOP size r and for $\bar{\gamma} = 8, 10$ dB, with $T_D = 450$ ms, $N = 3$, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

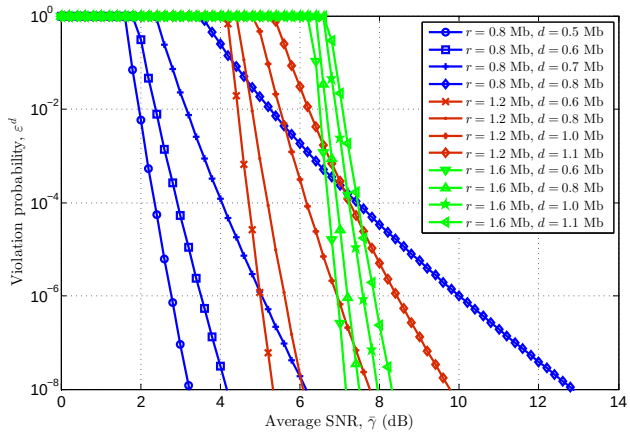

Fig. 4. Violation probability (ε^d) vs. average SNR ($\bar{\gamma}$) for SVC over multi-hop wireless network for three different GOP sizes $r = 0.8, 1.2$ and 1.6 Mb and for different departure within T_D per frame d, with $T_D = 450$ ms, $N = 3$, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

The figure confirms that the model provides a lower bound on the number of bits received per frame, and shows that the empirical CDF shows the same exponential increase as the model-based lower bound. This exponential growth in d can clearly be observed from (27). The bound is tight for low and moderate load, but acceptable even for high utilization of 0.99, specifically, the gradient for the model and simulation based results are equal which means that the error diminishes as ε grows smaller.

We notice that reducing utilization, e.g., by reducing frame size for a given SNR, results in sharper curves, which means that the channel impairments have smaller effect on the video quality. On the other hand, the figure shows that high utilization may lead to overload and low received quality, see for example the 0.99 utilization case of $\bar{\gamma} = 8$ and $r = 2.08$, where the probability of receiving even $d = 0.8$ is close to zero. These results reflect well that allowing transmission rates close to the average channel capacity would lead to overload and low quality streaming for latency critical applications.

Figs. 4 and 5 evaluate the effect of the channel quality on

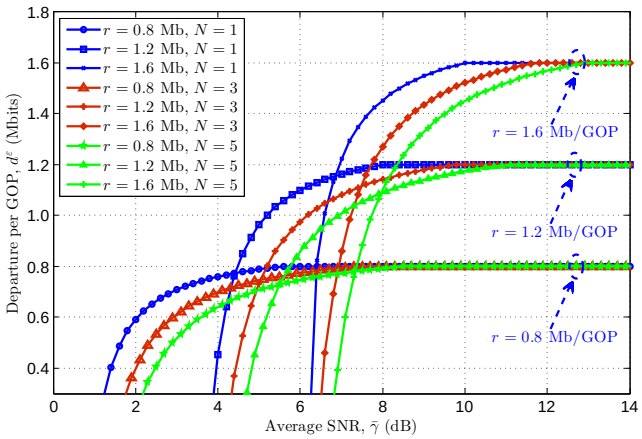

Fig. 5. Departure per frame (d^ε) vs. average SNR ($\bar{\gamma}$) for SVC over multi-hop wireless network for $\varepsilon = 10^{-4}$, for three different GOP sizes $r = 0.8, 1.2$ and 1.6 Mb and for $N = 1, 3$ and 5 hops, with $T_D = 450$ ms, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

the received video performance. Fig. 4 shows for $N = 3$ and for various r and d, that the violation probability ε for given d decreases almost exponentially when increasing the average SNR, as soon as the system becomes stable. Fig. 5 shows how the per hop average SNR affects the per frame departures d^ε for a violation probability $\varepsilon = 10^{-4}$, for different transmitted frame sizes r and number of hops N. We can see that the SNR has significant effect on the optimal transmission scheme, for example, at $\bar{\gamma} = 6$, $r = 1.2$ Mbits provides the best performance among the considered video frame sizes, $r = 0.8$ Mbits does not fully utilize the network, while $r = 1.6$ Mbits leads to low quality due to network congestion. We also observe that the effect of number of hops, N, is significant at high utilization, but diminishes as $\bar{\gamma}$, and thus the channel capacity increases.

Finally, we evaluate the effect of the overhead ratio, r_0, on the playback quality in Fig. 6. We compare an ideal scalable video coding scheme with no overhead due to layering, i.e., $r_0 = 0$ to an overhead-burdened scalable coding scheme with layering overhead ratio $r_0 = h/m > 0$. We investigate how the violation probability of the departure process for a given bound $d = d_0(1 + r_0)$ changes as the overhead ratio is increased. The figure shows a significant cost, in terms of the violation probability, when adding overhead. Furthermore, the lower the channel quality, the more critical is the effect of the overhead on the streaming quality. This highlights the importance of designing more efficient scalable coding schemes for delay-sensitive video streaming over future multi-hop wireless networks.

VIII. Adaptive Video Transmission and Routing

Our analysis exposes the effect of two extreme network behaviours that influence received video quality, namely, network congestion (at high utilization) due to bad channel quality and/or high frame rate, and network underutilization due to small transmitted video frame size. It also shows that the optimal operating point, where the transmitted video frame size maximizes the received video quality depends on the

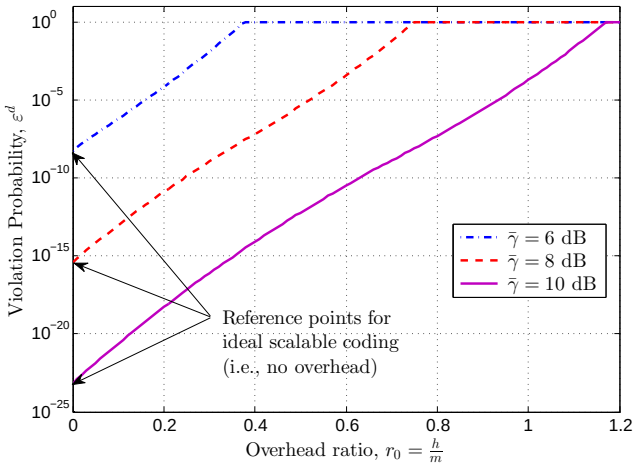

Fig. 6. Violation probability (ε^d) vs. overhead ratio (r_0) for SVC over multi-hop wireless network for $d = d_0 * (1 + r_0)$ where $d_0 = 0.8$ Mb and for different average SNR per hop ($\bar{\gamma} = 6, 8, 10$ dB) and $N = 3$ hop, $T_D = 450$ ms, $W = 2.2$ MHz and $n = 2.5$ GOP/s. The GOP size, r, is selected optimally for each point on the traces.

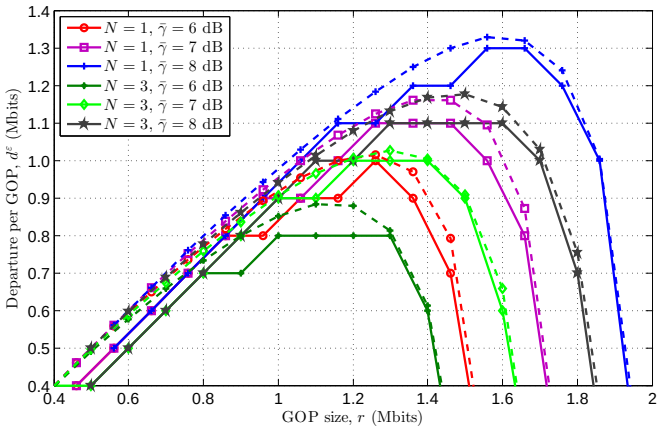

Fig. 7. Departure per frame (d^ε) vs. GOP size (r) for SVC over multi-hop wireless network (solid line for layered video frames and dashed line for fluid traffic model) with layer size $m = 100$ kbits, for different average SNR ($\bar{\gamma} = 6, 7, 8$ dB) and for $N = 1, 3$ hop, $\varepsilon = 10^{-6}$, $T_D = 450$ ms, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

channel conditions and on the length of the transmission path. Since the wireless channel quality may vary with time, the optimal performance can be achieved by adapting the transmitted video frame size to the SNR of the corresponding channels. It may also be beneficial to adapt the routing to the underlying channel quality. In this section, we examine both scenarios and provide examples to illustrate the benefits of such adaptation.

To evaluate the effect of under utilization as well as system overload, Fig. 7 shows the departures per frame, d, that fulfill the violation probability limit $\varepsilon = 10^{-6}$, as a function of the transmitted frame size r, for different SNR values and number of hops. The figure shows that the video frame size leading to maximum departures per frame depends on both of the network parameters. Increasing the frame size above this maximizing value leads to fast quality degradation as

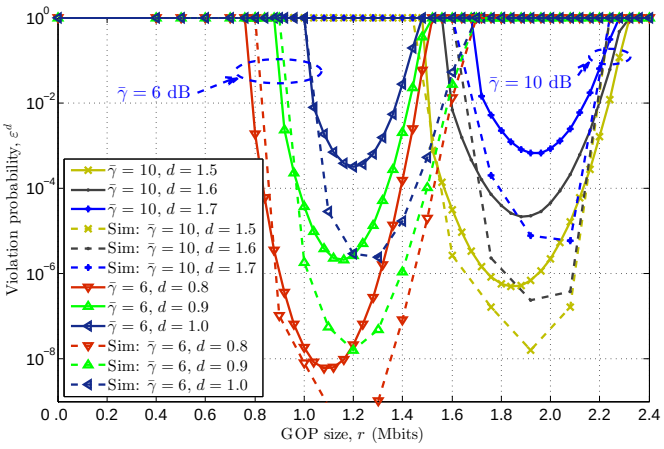

Fig. 8. Violation probability (ε^d) (computed and simulated) vs. GOP size (r) for SVC over multi-hop wireless network (solid line for bounds and dashed line for simulated) for $\bar{\gamma} = 6, 10$ dB and for different target departure per GOP d, with $T_D = 450$ ms, $N = 3$, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

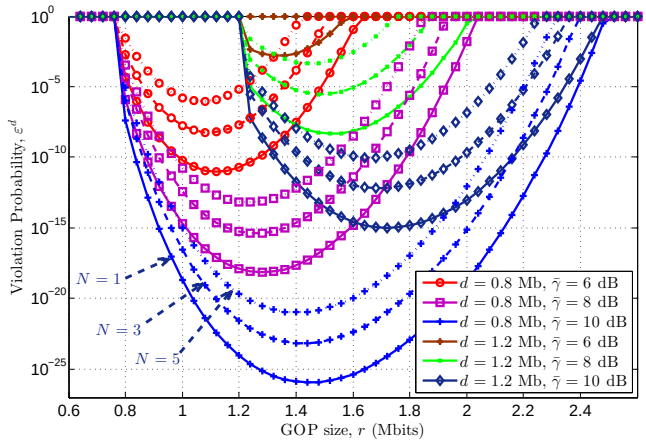

Fig. 9. Violation probability (ε^d) vs. GOP size (r) for SVC over multi-hop wireless network for $d = 0.8, 1.2$ Mb and for different average SNR per hop ($\bar{\gamma} = 6, 8, 10$ dB) and $N = 1, 3, 5$ hop, $T_D = 450$ ms, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

the network becomes more saturated. In this figure we also show the effect of layered transmission compared to its fluid counterpart, considering a layer size of $m = 100$ kbits. As layering affects both the possible transmitted and received video frame sizes, we can see performance degradation of a maximum of one layer size. Moreover, we can see that the same performance can be achieved under a range of transmitted video frame sizes, which means that an adaptation algorithm would have to find the smallest value to maximize the performance under the lowest transmission rate, and thus lowering the energy consumption.

As very small layer sizes may not be viable due to the introduced overhead, Fig. 7 can provide some insight on the desirable layering for video streaming in a wireless network. Comparing results for $\bar{\gamma} = 6, 7$ and 8 dB we see that the GOP size that maximizes the departure per GOP increases linearly with the $\bar{\gamma}$ values. That is, layering in future SVC solutions should be designed according to the expected wireless environment.

Fig. 8 compares the achieved violation probabilities as a function of the transmitted video frame size r, for different departure per video frame values d and SNR values $\bar{\gamma}$, showing the analytic upper bounds as well as the simulation results. Again, we see that there is an optimum r that minimizes ε. This optimum depends significantly on $\bar{\gamma}$, and slightly also on the aimed received quality d. Since the model provides only bounds on the achievable video quality, it is not expected that the model-based optimization gives the optimal r values. The simulation results reflect, however, that even though the model overestimates the violation probability itself, the suggested r values are reasonably close to the real optimum, found via simulation. Consider for example $\bar{\gamma} = 6$ and $\varepsilon = 10^{-6}$. The model predicts that $d = 0.9$ Mbits can be achieved with the required reliability with $r = 1.1$ Mbits, while according to the simulation results, the combination $d = 1$ Mbits, $r = 1.2$ Mbits is achievable as well. That is, the model-based parameter selection leads to 10% bitrate loss only, despite the slacknesss of the violation probability bound. Fig. 8 also shows a rapid increase in the violation probability when moving away from the optimum frame size. Therefore, the availability of a large number of enhancement layers is critical for fine-grained rate adaptation to channel conditions, subject to reliability constraints.

Fig. 9 summarizes the achievable performance for different expected received video frame size values d, SNR and number of hops. We see that the range of transmitted video frame sizes that yield acceptable violation probability depends on d on one side, and on $\bar{\gamma}$ on the other side. The optimal frame size is determined by these two parameters, while the number of hops, N, affects significantly the achievable violation probability, but not the optimal value of the video frame size.

In order to examine the efficiency of model-based frame size adaptation, we consider adaptation over a fixed transmission path and cross-layer optimized routing and rate adaptation. We compare the proposed model-based adaptation (MOD) to the optimal adaptation (OPT), where the optimum transmission video frame sizes, and the resulting departures per frame are obtained by conducting extensive simulations. In Figs. 10 and 11, we show the transmitted and received frame size r and d for OPT. For MOD we show the transmitted frame size that is suggested by the model, the bound on the received frame size, and the actual received frame size where the reliability constraint holds, derived through simulations.

Fig. 10 considers fixed routing with $N = 3$, and layered coding with 100 kbits layer sizes. We consider a scenario where the SNR $\bar{\gamma}$ changes from 10 dB to 6 dB and back to 10 dB at times $t = 30$ seconds and $t = 80$ seconds respectively. We use results similar to the ones reported in Fig. 8 to demonstrate the frame size adaptation in time. We assume that both the OPT and the MOD based schemes have stabilized at $t = 0$. OPT transmits with a frame size of $r = 1.9$ Mbits, and receives a frame size of $d = 1.6$ Mbits with violation probability $\varepsilon = 10^{-5}$. The model-based scheme slightly underestimates both r and d, but due to the layering, it reaches the same actual per frame departures as the OPT solution. After the channel quality degradation, the

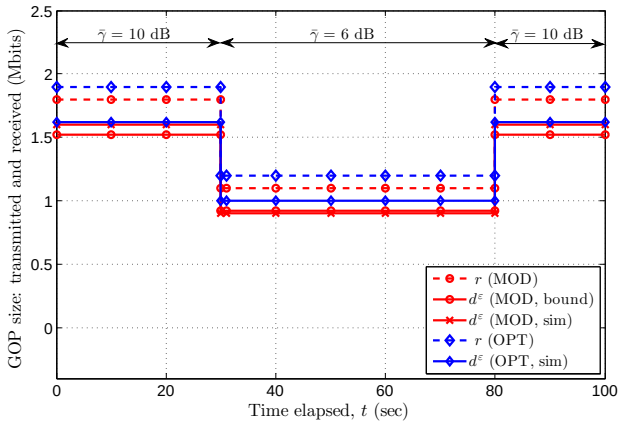

Fig. 10. Video frame size (r) adaptation for SVC over 3-hop wireless network for the model-based adaptation (MOD) and for violation probability $\varepsilon = 10^{-5}$ compared to the optimal adaptation (OPT) when average SNR, $\bar{\gamma} = 10$ dB then it drops to $\bar{\gamma} = 6$ dB and get back to $\bar{\gamma} = 10$ dB again, for $T_D = 450$ ms, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

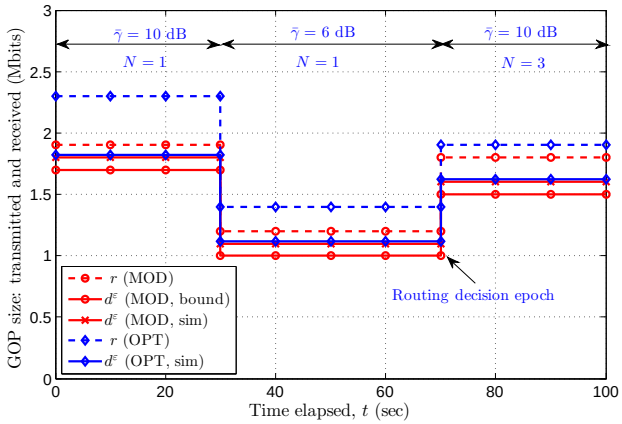

Fig. 11. Video frame size (r) adaptation with routing for SVC over wireless network for the model-based adaptation (MOD) and for violation probability $\varepsilon = 10^{-5}$ compared to the optimal adaptation (OPT) for a single hop link with average SNR $\bar{\gamma} = 10$ dB then it drops to $\bar{\gamma} = 6$ dB while an alternative 3-hop link with $\bar{\gamma} = 10$ dB per hop exist, for $T_D = 450$ ms, $W = 2.2$ MHz and $n = 2.5$ GOP/s.

MOD scheme decreases r, maintaining the system stability, again operating slightly below the OPT scheme. These results demonstrate that albeit the proposed network calculus based model provides only a lower bound on the per frame departures under some quality constraints, it enables the determination of a near optimal transmission frame size as it was suggested by Fig. 8.

In a real implementation of the model-based scheme, the channel quality change would be followed by a transient phase, where the average SNR value is gradually updated, leading to a period with lower than optimal performance. The characterization of this transient phase is out of the scope of the paper.

Finally, Fig. 11 demonstrates an example of rate adaptation combined with routing. We assume that the source node receives routing information, including the per link SNR values periodically, for example every 30 seconds as suggested for

the RPL standard [45]. Between routing updates, the source performs rate adaptation based on the SNR feedback on the actual path. We consider the case when the quality of the single hop path deteriorates from $\bar{\gamma} = 10$ dB to $\bar{\gamma} = 6$ dB at $t = 30$ seconds, and new routing information is received at $t = 70$ seconds, about an $N = 3$ path with 10 dB per link SNR. In this case, the longer path provides better service for the delay constrained transmission, as it has also been shown in Fig. 7. As a result, the MOD scheme first adapts to the poor channel quality on the single hop path, it then selects the three-hop path, and increases the transmitted frame size according to the better channel conditions. As the reporting of per hop SNR values, or the minimum SNR perceived on a path can be easily accommodated in routing protocols like RPL, routing combined with the model-based rate adaptation provides an excellent approach to ensure reliable, high quality, delay sensitive video transmission in wireless networks.

IX. Conclusion

In this paper we propose a network-calculus-based rate adaptation for delay-sensitive scalable video transmission over multi-hop wireless transmission paths. We derive new network calculus results that provide a probabilistic lower bound on the received video quality while considering the variability of the wireless channels, the effect of the queuing delays at the network nodes, and the frame-based playout at the receiver. Our evaluation shows that the channel quality has a more significant effect on the playout performance than the number of hops in the traversed path under low and moderate loads. Nonetheless, the effect of the hop count becomes significant as the network load increases. We show that even if the lower-bound-based model underestimates the achievable reliability, the transmission rate suggested by the model is close to the real optimum. Our results also show that the performance degradation due to the layering effect, compared to the perfect adaptation using the fluid model, depends significantly on the layer size, and hence, the number of enhancement layers per frame. That is, reliable, low latency video streaming over wireless links benefits greatly from adding more layers in layered coding.

The proposed model provides a tool for low-complexity and fast adaptation of the number of transmitted layers to the underlying channel conditions, the playout delay limit, and the desired reliability constraints. Our results show that the streaming performance under the model-based rate adaptation is very close to the achievable optimum for various network parameters (within 10% in the considered numerical examples). This suggests that the proposed network-calculus-based approach is an efficient tool for channel-aware rate control and routing for adaptive layered video transmission under strict playout delay limits.

References

[1] 5G PPP, "5G and the Factories of the Future," *White paper*, Oct. 2015.

[2] M. Gerla, E.-K. Lee, G. Pau, and U. Lee, "Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds," in *Proc. IEEE World Forum on Internet of Things (WF-IoT)*, March 2014.

[3] M. Ghodoussi, S. Butner, and Y. Wang, "Robotic surgery - the transatlantic case," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, May 2002.

[4] 5G PPP, "5G Automotive Vision," *White paper*, Oct. 2015.

[5] L. Baroffio, et al., "Enabling visual analysis in wireless sensor networks," in *Proc. IEEE International Conference on Image Processing (ICIP)*, Oct. 2014.

[6] S. Movassaghi, et al., "Wireless Body Area Networks: A Survey," *IEEE Communications Surveys & Tutorials*, vol.16, no.3, pp.1658-1686, 2014.

[7] H. Nishiyama, M. Ito, N. Kato, "Relay-by-smartphone: realizing multihop device-to-device communications," *IEEE Communications Magazine*, vol.52, no.4, pp.56-65, April 2014.

[8] D. Matsubara, et al., "Open the Way to Future Networks – A Viewpoint Framework from ITU-T," in *Proc. Future Internet: Future Internet Assembly,* 2013.

[9] M. Yang, et al., "Software-Defined and Virtualized Future Mobile and Wireless Networks: A Survey," *Mob. Netw. Appl.* vol.20, no.1, pp.4-18, February 2015.

[10] H. Schwarz, D. Marpe and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 17, no. 9, pp. 1103-1120, Sept. 2007.

[11] J. M. Boyce, Y. Ye, J. Chen and A. K. Ramasubramonian, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 26, no. 1, pp. 20-34, Jan. 2016.

[12] D. Rüfenacht, R. Mathew and D. Taubman, "A Novel Motion Field Anchoring Paradigm for Highly Scalable Wavelet-Based Video Coding," *IEEE Transactions on Image Processing,* vol. 25, no. 1, pp. 39-52, Jan. 2016.

[13] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson, "A buffer-based approach to rate adaptation: evidence from a large video streaming service," in *Proc. ACM SIGCOMM,* August, 2014

[14] C. Zhou,C-W. Lin and Z. Guo, "mDASH: A Markov Decision-Based Rate Adaptation Approach for Dynamic HTTP Streaming." *IEEE Trans. Multimedia* vol.18, no.4, pp.738-751, April 2016.

[15] K. Spiteri, R. Urgaonkar, R. K. Sitaraman, "BOLA: Near-Optimal Bitrate Adaptation for Online Videos," in *Proc. IEEE Infocom*, April, 2016.

[16] L. De Cicco, V. Caldaralo, V. Palmisano and S. Mascolo, "ELASTIC: A Client-Side Controller for Dynamic Adaptive Streaming over HTTP (DASH)," in *Proc. International Packet Video Workshop*, December 2013.

[17] Z. Li, X. Zhu, J. Gahm, R. Pan, H. Hu, A.C. Begen and D. Oran, "Probe and Adapt: Rate Adaptation for HTTP Video Streaming At Scale," *IEEE Journal on Selected Areas in Communications,* vol.32, no.4, pp.719-733, April 2014.

[18] S. Meng, J. Sun, Y. Duan and Z. Guo, "Adaptive Video Streaming With Optimized Bitstream Extraction and PID-Based Quality Control," *IEEE Transactions on Multimedia,* vol. 18, no. 6, pp. 1124-1137, June 2016.

[19] Y. Sun, X. Yin, J. Jiang, V. Sekar, F. Lin, N. Wang, T. Liu and B. Sinopoli. "CS2P: Improving Video Bitrate Selection and Adaptation with Data-Driven Throughput Prediction." in *Proc. ACM SIGCOMM,* August 2016.

[20] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli, "A Control-Theoretic Approach for Dynamic Adaptive Video Streaming over HTTP," *SIGCOMM Comput. Commun. Rev.* vol.45 no.4, pp.325-338, August 2015.

[21] H. Al-Zubaidy, J. Liebeherr, and A. Burchard. "A (min, x) network calculus for multi-hop fading channels," in *Proc. IEEE Infocom*, April 2013.

[22] H. Al-Zubaidy, J. Liebeherr, and A. Burchard, "Network-layer performance analysis of multihop fading channels," *IEEE/ACM Transactions on Networking*, vol. 24, no. 1, pp. 204–217, February 2016.

[23] J. Nightingale, Qi Wang, C. Grecos, "Scalable HEVC (SHVC)-Based video stream adaptation in wireless networks," in *Proc. IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2013.

[24] T. Schierl, T. Stockhammer, and T. Wiegand, "Mobile video transmission using scalable video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1204–1217, Sept 2007.

[25] S. Chen, J. Yang, E. Yang, and H. Xi, "Receiver-driven adaptive layer switching algorithm for scalable video streaming over wireless networks," in *Proc. IEEE International Conference on Networking, Sensing and Control (ICNSC)*, April 2014.

[26] H.-L. Lin, T.-Y. Wu, and C.-Y. Huang, "Cross layer adaptation with QoS guarantees for wireless scalable video streaming," *IEEE Communications Letters*, vol. 16, no. 9, pp. 1349–1352, Sept, 2012.

[27] S. Chen, J. Yang, Y. Ran, and E. Yang, "Adaptive layer switching algorithm based on buffer underflow probability for scalable video streaming over wireless networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol 26, no. 6, pp.1146–1160, June 2016.

[28] J. Yang, H. Hu, H. Xi, and L. Hanzo, "Online buffer fullness estimation aided adaptive media playout for video streaming," *IEEE Transactions on Multimedia*, vol. 13, no. 5, pp. 1141–1153, Oct. 2011.

[29] A. Rizk and M. Fidler, "Queue-aware uplink scheduling: Analysis, implementation, and evaluation," in *Proc. IFIP Networking Conference*, May 2015.

[30] W. Song, "Delay analysis for compressed video traffic over two-hop wireless moving networks," in *Proc. IEEE Globecom*, Dec 2011.

[31] D. Wu and R. Negi, "Effective Capacity-Based Quality of Service Measures for Wireless Networks," *Mobile Networks and Applications,* vol. 11, no. 1, pp. 91–99, February 2006.

[32] F. Ciucu, "Non-asymptotic capacity and delay analysis of mobile wireless networks," in *Proc. ACM Sigmetrics*, June 2011.

[33] M. Fidler, "A network calculus approach to probabilistic quality of service analysis of fading channels," in *Proc. IEEE Globecom*, Nov. 2006.

[34] K. Mahmood, A. Rizk, and Y. Jiang, "On the flow-level delay of a spatial multiplexing MIMO wireless channel," in *Proc. IEEE ICC*, June 2011.

[35] G. Verticale and P. Giacomazzi, "An analytical expression for service curves of fading channels," in *Proc. IEEE Globecom*, Nov. 2009.

[36] M. Fidler, "An end-to-end probabilistic network calculus with moment generating functions," in *Proc. IEEE IWQoS*, June 2006.

[37] C.-S. Chang. *Performance guarantees in communication networks*. Springer Verlag, 2000.

[38] Y. Jiang and Y. Liu. *Stochastic network calculus*. Springer, 2008.

[39] B. Davies, *Integral transforms and their applications*, Springer-Verlag, NY, 1978.

[40] R. McEliece and W.E. Stark. "Channels with block interference," *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp.44–53, Jan 1984.

[41] N. Petreska, H. Zubaidy, R. Knorr, and J. Gross, "On the recursive nature of end-to-end delay bound for heterogeneous wireless networks," in *Proc. IEEE ICC*, June 2015.

[42] N. Petreska, H. Zubaidy, R. Knorr, and J. Gross, "Power-minimization under statistical delay constraints for multi-hop wireless industrial networks," arXiv:1608.02191v2 [cs.PF], Aug. 2016.

[43] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: gathering 802.11n traces with channel state information," *SIGCOMM Comput. Commun. Rev.* vol.41, no.1 Jan. 2011.

[44] E. Dahlman, S. Parkvall and J. Skld, "4G LTE/LTE-Advaned for Mobile Broadband." Academic Press, 2011.

[45] N. Accettura, L. Grieco, G. Boggia, and P. Camarda, "Performance analysis of the RPL routing protocol," in *Proc. IEEE International Conference on Mechatronics (ICM)*, April 2011.

[46] J. Liebeherr, Y. Ghiassi-Farrokhfal, and A. Burchard, "On the Impact of Link Scheduling on End-to-End Delays in Large Networks," IEEE Journal on Selected Areas in Communications, 29(5):1009–1020, May 2011.

[47] W. Abbessi and H. Nabli, "GoP-based fluid Markovian modelling of video traffic," The Second International Conference on Communications and Networking, Tozeur, 2010, pp. 1-8.

Hussein Al-Zubaidy (S07M'11SM'16) received the Ph.D. degree in electrical and computer engineering from Carleton University, Ottawa, ON, Canada, in 2010. He was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, from 2011 to 2013. In the Fall of 2013, he joined the School of Electrical Engineering (EES) at the Royal Institute of Technology (KTH), Stockholm, Sweden, as a Post-Doctoral Fellow. Since Fall 2015, he has been a Senior Researcher with EES at the Royal Institute of Technology (KTH), Stockholm, Sweden. Dr. Al-Zubaidy is the recipient of many honors and awards, including the Ontario Graduate Scholarship (OGS), NSERC Visiting Fellowship, NSERC Summer Program in Taiwan, OGSST, and NSERC Post-Doctoral Fellowship.

ViktoriaFodor is professor at KTH Royal Institute of Technology, Stockholm, Sweden. She received her M.Sc. and Ph.D. degrees in computer engineering from the Budapest University of Technology and Economics in 1992 and 1999, respectively. She worked at the Hungarian Telecommunication Company in 1998 and joined KTH in 1999. She is associate editor at IEEE Transactions on Network and Service Management, and the Transactions on Emerging Telecommunications Technologies. In 2017 she acted as co-chair of IFIP Networking. Her current research interests include network performance evaluation, protocol design, wireless and multimedia networking.

György Dán is an Associate Professor at KTH Royal Institute of Technology, Stockholm, Sweden. He received the M.Sc. in Computer Engineering from the Budapest University of Technology and Economics, Hungary in 1999, the M.Sc. in Business Administration from the Corvinus University of Budapest, Hungary in 2003, and the Ph.D. in Telecommunications from KTH in 2006. He worked as a Consultant in the field of access networks, streaming media and videoconferencing from 1999 to 2001. He was a Visiting Researcher at the Swedish Institute of Computer Science in 2008, a Fulbright research scholar at University of Illinois at Urbana-Champaign in 2012-2013, and an invited Professor at EPFL in 2014-2015. He was co-chair of the Cyber Security and Privacy Symposium at IEEE SmartGridComm 2014, and is an Area Editor of Elsevier Computer Communications. His research interests include the design and analysis of content management and computing systems, game theoretical models of networked systems, and cyber-physical system security in power systems.

Markus Flierl (S'01-M'04) is Associate Professor of Electrical Engineering at KTH Royal Institute of Technology, Stockholm. He received the Doctorate in Engineering from Friedrich Alexander University, Germany, in 2003. From 2000 to 2002, he visited the Information Systems Laboratory at Stanford University. From 2003 to 2005, he was a senior researcher with the Signal Processing Institute at the Swiss Federal Institute of Technology Lausanne, Switzerland. From 2005 to 2008, he was Visiting Assistant Professor at the Max Planck Center for Visual Computing and Communication at Stanford University, California. He has authored the book "Video Coding with Superimposed Motion-Compensated Signals: Applications to H.264 and Beyond. He was the recipient of the SPIE VCIP 2007 Young Investigator Award. Currently, he serves as an Associate Editor for the IEEE Transactions on Circuits and Systems for Video Technology. His research interests include visual computing and communication, mobile visual search, and video representations.