Class-Based MDP for improved multimedia transmission over LTE

Madalina Ene, Nesrine Changuel, Bessem Sayadi, Michel Kieffer

CLASS-BASED MDP FOR IMPROVED MULTIMEDIA TRANSMISSION OVER LTE

Nesrine Changuel[1], Madalina Ene[1], Bessem Sayadi[1], Michel Kieffer[2]

1- Alcatel Lucent Bell Labs France, Route de Villejust, 91620 Nozay
2- L2S, CNRS - SUPELEC - Univ Paris-Sud, F-91192 Gif-sur-Yvette

ABSTRACT

This paper proposes an improved cross-layer control mechanism to efficiently stream videos to mobile users over an LTE network. A proxy-based filtering algorithm among scalable layers is considered to decide the number of SVC layers to transmit for each frame according to the communication conditions and to the class to which the video belongs to. The problem is cast in the context of Markov Decision Processes which allow the design of foresighted policies maximizing some long-term accumulated reward. Optimal actions to apply to the system are obtained by reinforcement learning. The proposed solution is implemented in an LTE simulation platform. Experiments show the performance of the proposed class-based layer filtering algorithm for a single video transmission and its robustness to content changes. [1]

Index Terms— Video, 4G mobile communication, optimization

1. INTRODUCTION

Today, we are facing an explosion of the video traffic on wireless network due to the proliferation of multimedia-friendly portable devices [1]. This trend is expected to intensify further in the near future. Even if the rates provided by next-generation mobile networks are increasing, the demand for high-quality multimedia content is growing faster [2].

Network congestion will be the main cause of QoS deterioration in the mobile network [3]. It may occur in the wireless part, due to the nature of the wireless channel and the limited bandwidth, or in the backhaul part. Service providers need to find solutions to deliver multimedia contents in the most efficient way to provide users with the best quality of experience (QoE) while accounting for scarce network resources.

In the current implementation, a multimedia session is controlled either by the client, *e.g.*, via HTTP adaptive streaming [4] or by the server via RTP [5]. In both cases, the network is not aware of the video traffic and the control of any congestion in the network relies only on the way the multimedia session is managed.

Long Term Evolution (LTE) [6] substantially improves end-user throughputs, sector capacity and reduces user plane latency, bringing significantly improved user experience. LTE provides IP connectivity between the User Equipment (UE) and the Packet Data Network (PDN) without any disruption during mobility. LTE supports different types of services including web browsing, video streaming, VoIP, online gaming, real time video, etc. [7].

In order to improve the received video quality, in the context of LTE networks, control may be done in Medium Access Control (MAC) or in PHYsical (PHY) level. In [8] a QoS-aware scheduling algorithm is proposed for real time video delivery over LTE cellular networks. A cross-layer optimized system is proposed to achieve a good perceived video quality under delay constraint by dynamically adapting the resource allocation to the instantaneous channel quality. This requires scheduler adaptation in the eNodeB (eNB). In [9] a quality-oriented video delivery over LTE algorithm based on modulation and coding adaptation is proposed. In this algorithm, PHY layer technologies, MIMO configuration, and MCS are dynamically adjusted.

In [10, 11], video delivery control techniques are proposed for single-video delivery. In [10] a smoothing buffer is used in the eNB to smooth the transmission rates with minimal information about the video to be transmitted. This approach is mapped to be used in an LTE network where a four-state Markov channel model is used to represent the channel variations. In [11], a cross-layer optimization of various parameters of the coding and transmission chain is performed to deliver video to mobile users. The problem is cast in a Markov Decision Process (MDP) framework [12]. Coarse models of the source and of the channel have been considered due to the increase of complexity of the optimization algorithm with the number of state values.

In [13], we proposed an online control mechanism to maximize the quality of the received video, while accounting for the source type (I, or P) of the transmitted frames. The control problem is again cast in the framework of MDP. The solution in [13] was proposed for single-user video delivery with coarse description of the system characteristics. In this paper, we propose to better account for the variability of video characteristics among video streams and within a video stream. Videos are mainly characterized by their Rate-


Distortion (RD) function which depends on the activity in the video, as well as the their spacial and temporal resolution. A cross-layer optimization mechanism should account for these characteristics, which have a significant impact on the user QoE. Various video classes may be defined to to take these characteristics (activity, resolution) into account. Then, the proposed solution consists in learning optimal policies for each class. This allows a finer adaptation of the policy to the video content. Determining the class a video or a frame of it belongs to may be done at the server or in the network using, *e.g,* [14, 15]. The proposed class-based control algorithm is implemented on an LTE simulation platform [16]. Typically, in case of 3GPP mobile network, the last buffer in a multimedia session is the Radio Link Control (RLC) layer. The RLC buffer is monitored and controlled to avoid the underflow or overflow situations which can cause disruptive or low quality of the multimedia content.

The remainder of the paper is organized as follows. Section 2 introduces the considered unicast video streaming. Section 3 details the optimization algorithm and the use of video classification. Section 4 presents the performance of the proposed class-based layer filtering process implemented on an LTE simulation platform.

2. SYSTEM ARCHITECTURE

Consider the video streaming system to a mobile receiver sketched in Figure 1.
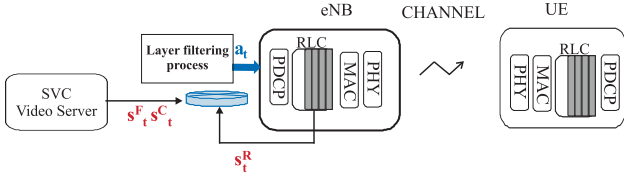
Fig. 1. Proposed class-based layer filtering process using H.264/SVC in LTE network for a unicast connection

The considered system is based on the LTE network topology with an emphasis on the downlink radio access sector between the eNB and the UE. For each unicast connection, a video server hosts a scalable video coder providing L SNR layers and can be situated inside or outside the core network. Frames are generated with a constant period of time T. The video packets travel through a limited delay core LTE network to reach a Post-Encoding (PE) buffer situated in a proxy where a controller performs layer filtering. For each frame, SNR layers may be sent, kept, or dropped. The layer filtering process should maximize the video quality at the receiver side by taking into account most factors impacting it: video class, frame type, number of SNR layers, and lost packets due to possible RLC buffer overflow. Filtered layers are then sent to the eNB. The RLC buffer status in the eNB is monitored to estimate the channel state. In fact, there is one RLC buffer

per radio bearer, and therefore per video service. Moreover, the RLC buffer is the last buffer in the Layer 2 protocol stack that can provide information about the data flow through the stack. The MAC scheduler, as well as its PHY layer, its radio front-end, the wireless channel, the physical layer of the receiver, and the part of the MAC layer at receiver side managing ACK/NACK procedures are considered as the channel.

3. LAYER FILTERING ALGORITHM

3.1. Single user layer filtering

In [13] a layer filtering process for a unicast video delivery is proposed. It is cast in the MDP framework [12]. Time is slotted into discrete-time intervals of length T. The t-th time slot is the time interval $[t, t+1)T$. The states of the system consist of the frame type s^F and the level of the RLC buffer s^R. The system state is then $\mathbf{s} = (s^F, s^R) \in \mathcal{S}$. The actions indicate the number of scalability layers to transmit, keep or drop from the PE buffer. Action chosen at time slot t is described by the vector $a_t = (a_{\ell,f})_t$ where $\ell \in \{1, ..., L\}$ is the layer index and $f \in \{1, ..., F\}$ is the frame index. The possible values of $a_{\ell,f}$ are $0, 1,$ and-1 corresponding respectively to keeping, sending, and dropping a frame. Actions corresponding to sending enhancement layers and keeping or dropping base layer of the same frame are excluded.

The effect of an action is expressed via a reward function $r_t(s_t, a_t)$ depending on the PSNR or any other video quality metric [17, 18, 19], and the level of the RLC buffer to determine the optimal policy. In absence of overflow and frame drop, the reward function provides an estimate of the video quality at the receiver

$$\widehat{r}_t(s_t, a_t) = \sum_{\ell=1}^{L} \max(0, a_{\ell,f,t}) q_t(s_t^F, \ell) \qquad (1)$$

where $q(s^f, \ell)$ is the quality increment measure provided by the transmission of layer ℓ from a frame of type s^F. When there is a frame drop, concealment is performed at receiver side, the video quality of the current frame is assumed to be equal to that of the previous frame reduced by a factor $\lambda(s^f)$ depending on the type of the lost frame (lost I frames will have more impact on the next frames than lost P frames). One then gets

$$\widehat{r}_t(s_t, a_t) = \widehat{r}_{t-1}(s_{t-1}, a_{t-1}) - \lambda(s_t^F). \qquad (2)$$

The optimal *foresighted* policy [12] maximizing the discounted sum of present and future rewards is evaluated through reinforcement learning (RL). In fact, unlike policy or value iteration [12] where state transition probability are needed to find the optimal policy, RL aims at estimating a good policy without requiring knowledge of state transition probabilities which is very practical in the considered situation, where system state transition are partially unknown. RL leads to a two-phase algorithm. In the *exploration* phase,

the system states (frame type, buffer state, channel state) are explored by taking random actions and measuring the resulting rewards. In the *exploitation* phase, the optimal actions learned in the exploration phase are applied as a function of the system state.

System states can be chosen with a fine granularity leading to a fine description of the system (*e.g,* number of PDU in the RLC buffer) and finely tuned actions. This may result in a very long exploration phase before getting the optimal action to take in each system state, see Figure 2. On the other hand,
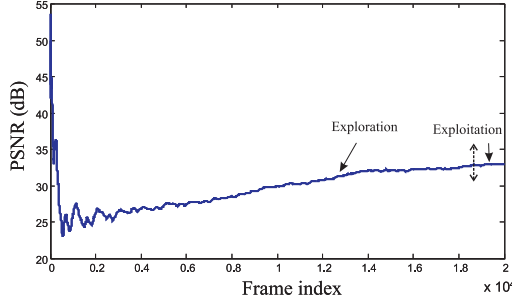
Fig. 2. Cumulative PSNR reward for foresighted policies obtained using the technique in [13].

system states can be chosen with a coarse granularity leading to an averaged description of the system (using, *e.g*, a coarse quantization of the number of PDU in the RLC buffer) which may decrease the learning time. The second choice, considered in [13], is more robust to small system changes since it corresponds to an average behavior of the system.

With this approach, video characteristics are not accurately taken into account at the learning process, and only policy optimal *in average* are obtained. Better tuned policies may be obtained by taking more characteristics into account.

3.2. Class-based layer filtering

In order to improve the accuracy of the solution in [13], we propose to add a new state in the system model. Consider $s^C \in \{1, \ldots, C\}$ corresponding to index of the class the frame, Group of Picture (GoP), or video belongs to. A video class can be considered at different granularity levels, it can be related to an entire video, or it can be updated at each frame, GoP, or set of GoPs. When considering classes at a finer granularity than the video level, the evolution of the class state may be distributed by a Markov chain as in [20].

Information related to the video class can be derived in the proxy according the R-D characteristics of the stored video packets in the PE buffer or it can be defined by the server and marked in the header of the compressed flow at each time the class changes. Depending on the class s_t^C at time t, different policies may be used. Due to the unknown class state transition probabilities, RL is used which allows estimating a good policy without requiring an accurate knowledge of the class

transition probability $P(s_{t+1}^C | s_t^C)$.

4. EXPERIMENTAL RESULTS

The proposed solution consists of on the following steps: First, action-state learning matrices are derived either dynamically using the delivered flows or offline for different video classes and stored in the layer filtering block, see Figure 1. Second, for each frame to be filtered, the video class s_t^C is identified. Third, the optimal policy corresponding to video class s_t^C is used while filtering the transmitted frame. One has to verify the robustness of the algorithm with respect to the video class.

4.1. LTE SIMULATOR

In this section we briefly describe the layer 2 of the LTE downlink link level simulator framework and the implementation of the proposed algorithm.

Open Air Interface (OAI) [16] provides an open source wireless technology platform with both hardware and software components. It hosts a prototype representative of the LTE infrastructure offering the possibility of a large scale simulation. From the packet delivery point of view, the simulation is soft real-time with an emulation which permits the existence of several instances (eNB and UE) on the same machine.

Video delivery begins at the server side where video packets are transmitted to the proxy and stored in the PE buffer. Video class as well as the frame type are assumed delivered by the server and marked in the header of the transmitted video packets. Decision making is then performed by the layer filtering control block using information about the frame type, the video class from the stored video packets and the level of the RLC buffer from the eNB. The filtered layers travel to the eNB via OAI. Once transmitted, video packets are segmented and placed in the RLC buffer. The new RLC buffer status is then fed back to the layer filtering block and the system state is updated. Based on the new system state a PSNR reward is generated and used to update the policy during the exploration phase of the RL. Then, the algorithm continues its loop.

4.2. Considered tests

The performance of the proposed layer filtering process has been evaluated with two classes ($C = 2$) at the video level. *Foreman.qcif* (Average Encoding Rate (AER) 3.1 Mbps) and *Carphone.qcif* (AER 3 Mbps) are considered belonging to video class $s^C = 1$ due to their similar R-D characteristics and low activity level. Then, *Football.cif* (AER 5.2 Mbps), and *Basket.cif* (AER 5 Mbps) 100 s long both extracted from real TV programs belong to video class $s^C = 2$, characterized by very high activity level. Video classification is performed offline; it can also be realised using [20, 21, 14, 15].

The Packet Data Units (PDUs) in the RLC buffer are of variable size, depending on the the condition of the channel assigned by the eNB to the UE. The maximum number of PDU in the RLC buffer is $R^{\text{MAX}} = 128$. Five states for the RLC are considered corresponding to a quantization of the number of PDU in the buffer. The intervals are $[0, 0.6) R^{\text{MAX}}$, $[0.6, 0.75) R^{\text{MAX}}$, $[0.75, 0.85) R^{\text{MAX}}$, $[0.85, 0.95) R^{\text{MAX}}$, and $[0.95, 1] R^{\text{MAX}}$. The fifth interval is smaller than the others to anticipate overflow and prevent PDUs from being dropped. All videos are encoded at 30 fps using the H.264/SVC encoder (JSVM 9.19) [22] with three MGS scalability layers per frame ($L = 3$). The period at which the controller is operating is $T = 1/30$ s. IPPP GoPs of 16 frames are considered. A *base layer only* control scheme [23] for the encoder is adopted to minimize the drift when refinement layers are lost. Thus, two possible frame type s^{F} are considered: I or P frame. OAI emulates the LTE channel variation and the MAC scheduler selects the appropriate Modulation Coding Scheme (MCS) so that the channel rate values range from 0.5 to 4 Mbps, the optimal action for each possible state is determined for the foresighted policy. In our simulations, no handover is considered and the receiver buffer size is assumed infinite. The robustness of the proposed approach to variations of the characteristics of the system is evaluated.

4.2.1. Intra-class robustness

To make sure that an optimal filtering policy learned for a video is satisfying for another sequence within the same class, intra-class robustness is evaluated. The performance of the proposed algorithm is evaluated based on the decoded PSNR received at the UE. The policy is learned using *Carphone* or *Foreman* at a channel rate of 2Mbps. These two policies are then used to filter SVC layers while transmitting *Carphone* sequence. Results are shown in Figure 3. The dotted
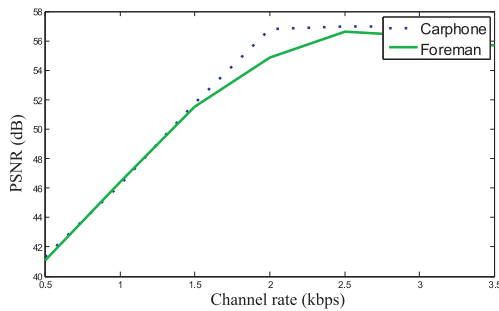

Fig. 3. Decoded PSNR using *Carphone* and *Foreman* foresighted policies for different rates.

curve represents the average PSNR of *Carphone* as a function of the channel rate. The *Carphone*-tuned policy applied on *Carphone* (no mismatch). Solid curve represents the average PSNR of *Foreman* - tuned policy applied on *Carphone*

(mismatch). In case of mismatch in the same class the PSNR is slightly decreased compared to case of no mismatch.

4.2.2. Inter-class robustness

Inter-class robustness is now evaluated. The class of each sequence is assumed available while transmitting the video. The experiment consists in learning the three optimal foresighted policies using: *Basket*, *Football*, and then *Foreman*. The three policies are then applied separately on the *Basket* sequence. Results are shown in Figure 4. One can see that curves refer-
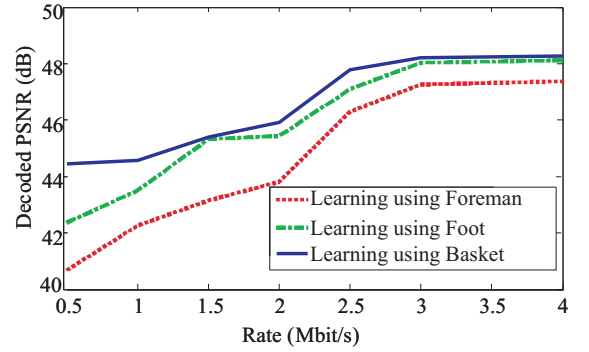
Fig. 4. Decoded PSNR of *Basket* video using optimal foresighted policies of different video class.

ring to videos which fall in the same video class (*Basket* and *Football*) show a comparable behavior. On the other hand, applying the optimal actions learned with *Foreman* (not the same class as *Basket*) gives a PSNR degraded in average of 2 dB compared to using the optimal actions learned with *Basket* sequence. This result shows the importance of using video classes in the system state in order to consider variations of the video characteristics. Later, more videos classes will be considered to test the robustness with various video contents.

5. CONCLUSION

This paper proposes a cross-layer optimization solution to efficiently deliver scalable video over LTE networks. Optimal SVC layer filtering decisions are learned for different video classes in order to take into consideration the variability of the video characteristics mainly related to their rate-distortion properties, as well as their spacial and temporal resolution. The proposed algorithm is implemented on an LTE simulation platform. Intra-class robustness shows that the solution is robust for video sequences belonging to the same class. When applying a policy learned for one class to a video belonging to an other class, a more significant quality decrease is observed. This illustrates the importance of accounting for the class a video belong to in such layer filtering algorithm. More classes and more videos in each class will be considered in the final version of the paper.

6. REFERENCES

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

[2] N. Amram, B. Fu, G. Kunzmann, T. Melia, D. Munaretto, S. Randriamasy, B. Sayadi, J. Widmer, and M. Zorzi, "QoE-based transport optimization for video delivery over next generation cellular networks," in *Proc. IEEE SCC*, 2011, pp. 19 – 24.

[3] "Beyond LTE: New solutions for easing escalating network congestion," Tech. Rep., Nokia Siemens Networks, 2012.

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

[5] V. Singh, "Rate-control for RTP-based multimedia applications," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2011, pp. 1 – 4.

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

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

[8] H. Luo, S. Ci, D. Wu, J. Wu, and H. Tang, "Quality-driven cross-layer optimized video delivery over LTE," *IEEE Communications Magazine,*, vol. 48, no. 2, pp. 102 – 109, 2010.

[9] A. Pande, V. Ramamurthi, and P. Mohapatra, "Quality-oriented video delivery over LTE using adaptive modulation and coding," in *IEEE Globecom*, 2011, pp. 1 – 5.

[10] K. Shuaib and F. Sallabi, "Smoothing of video transmission rates for an LTE network," in *IEEE WiMob*, 2010, pp. 713 – 719.

[11] A. Combernoux, C. Delestre, N. Changuel, B. Sayadi, and M. Kieffer, "Cross-layer optimization of a multimedia streaming system via dynamic programming," in *Proc. ICIP*, 2012, pp. 1 – 4.

[12] R.S Sutton and A.G Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.

[13] N. Changuel, S. Sayadi, and M. Kieffer, "Online learning for QoE-based video streaming to mobile receivers," in *Proc. Globecom*, 2012, pp. 1 – 5.

[14] L. Xu and W. Li, "Video classification using spatial-temporal features and PCA," in *Proc. ICME*, 2003, vol. 3, pp. 485 – 492.

[15] G. Toderici, H. Aradhye, M. Pasca, L. Sbaiz, and J. Yagnik, "Finding meaning on youtube: Tag recommendation, category discovery and their implications in browsing and search," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1 – 8. 2010.

[16] Mobile Communications Department, Institut Eurecom, http://www.openairinterface.org/, *Open Air Interface*.

[17] ITU-T, "Objective perceptual multimedia video quality measurement in the presence of a full reference," Tech. Rep. Rec. J.247, ITU-T Rec. J.247 (08/08), 2008.

[18] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74 – 90, November 1998.

[19] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction*, vol. 27 of Applications of Mathematics, Springer, second edition, 2005.

[20] L. Xie, S. Chang, A. Divakaran, and H. Sun, "Structure analysis of soccer video with hidden markov models," in *Proc. ICASSP*, 2002, vol. 4, pp. 4096 – 4099.

[21] W. Zeng, W. Gao, and D. Zhao, "Video indexing by motion activity maps," in *Proc. ICIP*, 2002, vol. 1, pp. 912 – 915.

[22] J. Vieron, M. Wien, and H Schwarz, *JSVM9 software,*, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG,, JVT-V203, January 2007.

[23] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. on Circuits and Systems for Video Technology In Circuits and Systems for Video Technology,*, vol. 17, no. 9, pp. 1103 – 1120, September 2007.