Analysis of PHY/Application Cross-layer Optimization for Scalable Video Transmission in Cellular Networks

Iffat Ahmed⋆‡, Leonardo Badia†‡, Daniele Munaretto†‡, and Michele Zorzi†‡
⋆ IMT Institute for Advanced Studies, Lucca, Italy
†Department of Information Engineering, University of Padova, Italy
‡ Consorzio Ferrara Ricerche, Ferrara, Italy
E-mail: iffat.ahmed@imtlucca.it, {munaretto,badia,zorzi}@dei.unipd.it

***Abstract*—We investigate the optimization of video transmissions over cellular networks by using the H.264 Scalable Video Coding (SVC) at the application layer and an Adaptive Modulation and Coding (AMC) scheme at the physical layer. We analyze how the cross-layer optimization (XLO) of these two techniques together performs compared to a sequential and independent selection of video packets and Modulation and Coding Schemes (MCS) with no cross-layer optimization (NXLO), in terms of goodput and packet delivery delay. We formulate an analytical model based on a Markov chain representing the wireless channel, where each state is associated to a different channel quality corresponding to a set of possible choices of video layer and MCS. Our numerical results show that XLO significantly outperforms NXLO for video transmissions, thereby pointing out the strong need for cross-layer solutions in video transmission.**

I. Introduction

THE WORLD of wireless communications has faced a tremendous increase in the usage of multimedia applications, such as video conferencing, video on demand, and video streaming. This has led to an increased demand of bandwidth and also has implied a big challenge for network operators to optimally utilize the resources and provide adequate quality to multimedia users in spite of heterogeneous terminals and network infrastructures. Well known documents like the Cisco report [1] offer evidence of such a "video explosion" by impressively foreseeing that the amount of video content crossing global IP networks each month in year 2016 will amount to the equivalent of over 6 million years of video duration. Not only is a policy optimization required, but it is also important to perform the optimization in the proper way.

We advocate the need not only for a generic "cross-layer" optimization, but more specifically a PHY/Application cross-layer optimization (specifically, H.264 Scalable Video Coding (SVC) [2], [3] source's Base Layer (BL) and Enhancement Layers (EL) information from Application layer and user channel condition from PHY layers), which is even more challenging as it has to span through the entire protocol stack. For this purpose, we first optimize the number of transmit opportunities (TXOPs) needed for BL to correctly decode the video [4] using Lagrange Multipliers, and then we design an analytical model, where the channel is represented through a Markov chain whose states represent different channel qualities. Depending on the channel quality, an efficient decision has to be made on which packets to allocate, so as to provide the user with an adequate QoE, which turns into maximizing the number of delivered packets and satisfying the relational requirements between layers.

This work has been partially supported by the European Commission under the FP7 EU project MEDIEVAL, grant agreement no. 258053

II. Analytical Model

We formulate an analytical model based on a Markov chain representing the wireless channel, where each state is associated to a different channel quality corresponding to a set of possible choices of video layer and modulation and coding scheme. For the sake of simplicity, we assume the SVC video source with one base layer and one enhancement layer, for downlink transmission from sender to receiver. The wireless channel is modeled as an N-state Markov chain [5], [6], with the set of transition probabilities. We consider four states, where each of the four states, $\mathcal{S} = \{00, 01, 10, 11\}$, is associated to a channel quality level in increasing order. Denoting the transition from state i to state j as $p_{i\to j}$, the transition probabilities can be collected into a matrix

$$\mathbf{P} = \begin{bmatrix} p_{00\to 00} & p_{00\to 01} & p_{00\to 10} & p_{00\to 11} \\ p_{01\to 00} & p_{01\to 01} & p_{01\to 10} & p_{01\to 11} \\ p_{10\to 00} & p_{10\to 01} & p_{10\to 10} & p_{10\to 11} \\ p_{11\to 00} & p_{11\to 01} & p_{11\to 10} & p_{11\to 11} \end{bmatrix}.$$

Each possible state is associated to a level of robustness and quality of the channel; this is why it is denoted as a pair of binary digits. The first digit, $\ell \in \{0,1\}$, means that a packet belonging to the base (0) or to the enhancement layer (1) can be sent out without incurring into erasures. The second digit, $\zeta \in \{0,1\}$, represents two orders of MCS in use, with 0 being the lower-order modulation and 1 being the higher-order one; for instance, 0 corresponds to QPSK and 1 to 16-QAM, respectively. Due to the different level of protection of the two video layers (unequal error protection, UEP), typical of H.264/SVC systems, the worse the channel conditions, the more likely the transmission of only a BL packet. On the contrary, packets from the enhancement layer should be sent out only for good channel conditions. Thus, packets of the base layer can be transmitted only in the first two states of the chain ($\mathcal{S}_{00}$ and $\mathcal{S}_{01}$), while the enhancement layer will be considered for transmission in the other two states ($\mathcal{S}_{10}$ and $\mathcal{S}_{11}$). Finally, the selection of MCS (second digit) would add more robustness to the transmission (states $\mathcal{S}_{00}$ and $\mathcal{S}_{10}$) or make the packet delivery quicker (states $\mathcal{S}_{01}$ and $\mathcal{S}_{11}$).

We associate each time slot with an index k. The joint selection of MCS ζ and video layer ℓ in slot k gives the goodput value g_0^k for states $\mathcal{S}_{00}^k$ and $\mathcal{S}_{01}^k$, because of BL transmissions, and g_1^k for states $\mathcal{S}_{10}^k$ and $\mathcal{S}_{11}^k$, because of EL transmissions. On the other hand, the packet delivery delay values, $d_{\zeta\ell}^k$, depend on the MCS and in general on the packet size, i.e., they change if either ζ or ℓ is changed. Without loss of generality, we assume that all packets have the same size, thus only the choice of MCS impacts the delivery delay. Thus, given a state $\mathcal{S}_{\zeta\ell}^k$, at time k, goodput and packet delivery delay can be defined as

$$\mathcal{G} = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{T} g_{\zeta\ell}^k(i)|\mathcal{S}_{\zeta\ell}^k\ , \quad \mathcal{D} = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{T} d_{\zeta\ell}^k(i)|\mathcal{S}_{\zeta\ell}^k$$

Here, N is the total number of users and T is the total number of timeslots before reaching the playback deadline. Since goodput and packet delivery delay are generally contrasting metrics (selecting a higher order modulation will achieve lower delay but also lower goodput), we can define a general quality metric for user i, at time slot k, referred to as quality:

$$\mathcal{Q} = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{T} \Big(\big(\alpha \cdot g_{\zeta\ell}^k(i) - (1-\alpha)\cdot(d_{\zeta\ell}^k(i))^{-1}\big) \mid \mathcal{S}_{\zeta\ell}^k\Big) \quad (1)$$

where $\alpha \in$ [0,1] is a tunable parameter

A. Transmission Optimization

Each user i has BL requirement ($\hat{n}_{BL}$), which is a uniform random distribution and n_{BL}, n_{EL} are the number of BL and EL packets received, respectively. The sender allocates resources according to the need of BL or EL packets as well as the channel conditions of the individual users. Based on individual $\hat{n}_{BL}$, the base station optimizes the number of time slots needed to find the $Opt_{t_1}(i)$ (further details on our optimization module are available in [4]), for each users i, we set θ_i as the threshold value to decode video, defined as

$$\theta_i = Opt_{t_1}(i)\cdot\lambda \quad (2)$$

where $\lambda \in [0,1]$ is a parameter to compute the threshold θ of the base layer, and provides a degree of freedom in the choice between the three allocation policies defined in sec. II-B . Once we have optimized θ_i, we need to allocate resources to send BL and EL packets based on the channel status at each time slot and for each user i, i.e., $alloc_i^k = Opt_i^k(\ell,\zeta)$, such that the BL packet allocation is

$$alloc\_BL_i^k(\theta_i) = Opt_i^k(\ell,\zeta) \quad \forall\,(\zeta,\ell)\in\mathcal{Z}\times\mathcal{L} \quad (3)$$

where, $\mathcal{Z}$ and $\mathcal{L}$ are sets of MCS and SVC Layers, respectively. The EL packet allocation is

$$alloc\_EL_i^k = Opt_i^k(\ell,\zeta) \iff \zeta > \omega \quad (4)$$

where ω is a value depending on the user channel status and the MCS (in this case 16-QAM) that defines the channel rate required to reliably transmit an EL packet. Thus, if the channel quality is below ω we necessarily transmit a BL packet.

B. Allocation Policies

We consider three allocation policies: (i) a theoretical upper bound on the performance obtained using a genie-like channel knowledge and offline optimization, denoted as *Offline*; (ii) a sequential selection with optimization performed separately at each layer, denoted *NXLO*; (iii) a joint cross-layer optimization of both layers, denoted as *XLO*.

1) Offline Policy: The complete evolution of the channel is known, thus this is an upper bound used for comparison. This policy takes into consideration the slots with best available channel conditions amongst $\mathcal{Z}$ above the threshold ω, as defined in (4) for EL and remaining slots for BL, upon the condition that the required BL packets have been transmitted. If the remaining slots are below a certain threshold ω, then only BL packets can be transmitted with either high or low MCS, while EL packets cannot, as presented in Algo. 1. Furthermore, if the BL is completed and the only available slots are below ω, then the EL packets are dropped and no more transmissions are performed.

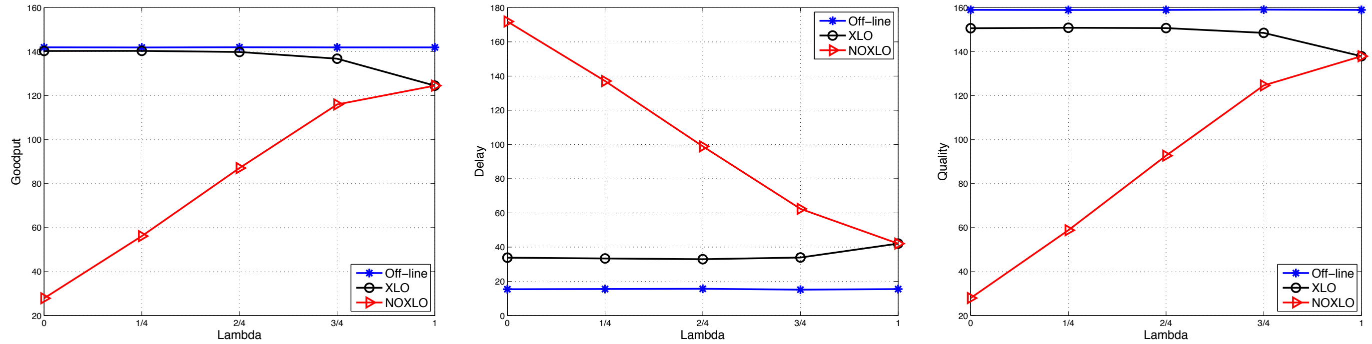
```
Input: (L, Z) and Output: Optimum assignment of BL/EL
for each user i ∈ N do
    for each time slot k ∈ T do
        if ζ_i ≥ ω and n_EL(i) not finished then
            transmit EL with high MCS ;
        else
            transmit BL with high MCS;
    for each time slot T do
        if ζ_i < ω and n_BL(i) not finished then
            transmit BL with low MCS;
```

Algorithm 1: Offline Policy

```
Input: (L, Z) and Output: Optimum assignment of BL/EL
for each user i ∈ N do
    for each time slot k ∈ T do
        if n_BL(i) ≥ θ_i then
            BL done ;
        else
            BL not done ;
        if BL done then
            if n_EL(i) < n_BL(i) then
                if ζ_i ≥ ω and n_EL(i) not finished then
                    transmit EL with high MCS ;
            else if ζ_i < ω and n_BL(i) not finished then
                transmit BL with low MCS ;
        else if ζ_i < ω and n_BL(i) not finished then
            transmit BL with low MCS ;
```

Algorithm 2: NXLO Policy

2) NXLO: In the non-cross-layer policy the base station first picks the SVC layer packet to be sent, that is, BL or EL packet based on θ_i and then checks the channel (and the possible MCS) of the user in the current time slot. Before reaching θ_i, the system is forced to transmit the BL packets, so that the video can be correctly decoded at the receiver's side with the minimum number of BL packets. The procedure is illustrated in Algo. 2.



(a) Goodput vs. lambda (b) Packet delivery delay vs. lambda (c) Quality vs. lambda, $\alpha = 0.6$

Fig. 1. Performance evaluation versus the reservation parameter λ

3) XLO: In this policy the channel status is first checked to select the best MCS, then the base station jointly selects the BL/EL and MCS based on θ_i. Similarly, before reaching θ_i, the system is forced to transmit the BL packets, so that video can be correctly decoded at the receiver with the minimum number of BL packets, as described in Algo. 3.

Input: $(\mathcal{L}, \mathcal{Z})$
Output: Optimum assignment of BL/EL
for *each user* $i \in N$ **do**
  for *each time slot* $k \in T$ **do**
    if $n_{BL}(i) \geq \theta_i$ **then**
      BL done ;
    else
      BL not done ;
    if *BL done* **then**
      if $\zeta_i \geq \omega$ *and* $n_{EL}(i)$ *not finished* **then**
        transmit EL with high MCS ;
      else
        transmit BL with high MCS;
      if $\zeta_i < \omega$ *and* $n_{BL}(i)$ *not finished* **then**
        transmit BL with low MCS ;
    else
      if $\zeta_i \geq \omega$ *and* $n_{BL}(i)$ *not finished* **then**
        transmit BL with high MCS ;
      else
        transmit BL with low MCS ;

Algorithm 3: XLO Policy

III. Simulation Results

The performance of the aforementioned policies is assessed by means of simulation in Matlab. Each user receives an H.264/SVC video within T time slots before reaching the playback deadline. Based on the channel conditions and the adopted scheduling policy, the user will get a certain amount of BL and EL packets. Multiple users receive unicast video streams in parallel. For the sake of simplicity, in the numerical evaluations we use a matrix $\mathbf{P}$ where transitions only happen to adjacent states, that is $p_{00\rightarrow01} = p_{01\rightarrow10} = p_{10\rightarrow11} = 0.6$ and $p_{01\rightarrow00} = p_{10\rightarrow01} = p_{11\rightarrow10} = 0.4$.

From Fig. 1(a) we can see that XLO performs almost optimally, i.e., close to the *Offline* solution. The XLO approach takes into consideration the channel conditions and jointly selects the BL/EL packets to be sent with a given MCS, whereas the NXLO checks the need for BL/EL packets first and then checks the channel conditions. The performance is reported in terms of goodput, in Fig. 1(a). On the other hand, if we look at the results for packet delivery delay, as in Fig. 1(b), it can be seen that XLO has lower delay values compared to NXLO and is almost as efficient as *Offline*. The quality in terms of goodput and packet delivery delay combination, which is defined in Section II, is illustrated in Fig. 1(c). Moreover, it is worth noting that the lower λ (i.e., more degrees of freedom), the better the performance of XLO.

IV. Conclusions

To investigate the performance of PHY/Application cross-layer optimization we considered an analytical model, where the channel is modeled by means of a Markov chain, whose states represent different channel qualities. We proposed a Cross-Layer (PHY/Application) solution with respect to adaptive transmission rates for SVC layers. We formulated a model to assess the performance of a Cross-layer solution as opposed to a separate sequential optimization of the two layers. We further evaluated via simulation the performance in terms of goodput, packet delivery delay and quality. We observed that joint selection of both video layer and modulation scheme can improve the quality compared to a sequential selection. Further, the joint selection solution provides provides performance very close to the optimal/theoretical best (as provided in the *Offline* policy).

References

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

[2] 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*, vol. 17, no. 9, pp. 1103 –1120, 2007.

[3] H. Sohn, H. Yoo, W. De Neve, C. S. Kim, and Y. M. Ro, "Full-reference video quality metric for fully scalable and mobile SVC content," *IEEE Trans. on Broadcasting*, vol. 56, no. 3, pp. 269–280, 2010.

[4] I. Ahmed and L. Badia, "Analysis of management policies for multicast transmission of scalable video content in next generation networks," in *Proc. of IEEE IWCMC*, 2013.

[5] M. Rossi, L. Badia, and M. Zorzi, "SR ARQ delay statistics on N-state Markov channels with non-instantaneous feedback," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1526–1536, Jun. 2006.

[6] L. Badia, N. Baldo, M. Levorato, and M. Zorzi, "A Markov framework for error control techniques based on selective retransmission in video transmission over wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 488–500, 2010.