Improving the performance of wireless H.264 video broadcasting through a cross-layer design
Atici, C.; Sunay, M.O.

Published in:

DOI:
10.1109/ISBMSB.2009.5133837

Published: 01/01/2009

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 09. Jan. 2019
Improving the Performance of Wireless H.264 Video Broadcasting Through a Cross-Layer Design

Çağdaş Atıcı, M. Oğuz Sunay
Electrical and Electronics Engineering
Özyeğin University, Istanbul, Turkey
Email: {cagdas.atici, oguz.sunay}@ozyegin.edu.tr

Abstract—In a wireless video broadcasting service, number of users receiving the service and the average video quality of the received stream have to be maximized for improving the system performance. To solve this problem optimally, one may adjust parameters at the physical and the application layers in a cross-layer fashion while utilizing the characteristics of the video that is being transmitted. In this paper, we propose a multi-objective optimized, cross-layer video broadcasting scheme for a wireless system capable of supporting a multitude of transmission data-rates using the H.264/AVC. The multi-objective, cross-layer optimization aims to find the H.264/AVC as well as the physical layer system parameters jointly to reach the optimal compromise between maximizing the average received video PSNR and minimizing the video broadcast service outage probability. Simulations conducted for the ITU Pedestrian A and Vehicular B channels show that further gains in system performance can be achieved for video broadcasting when such a cross-layer design is used.

I. INTRODUCTION

With growing demand for high quality multimedia content and development of wireless networks, increasing number of users are expected to enjoy services provided by future wireless systems. Considering the share of the multimedia transmission among those services, wireless system solutions for mobile multimedia services are of major importance. Today, widely used wireless systems are based on layered architecture, where each OSI layer is allowed to communicate with one layer above or below. As an evolution, it has been recently realized that designing physical (PHY), MAC, transport and application (APP) layers of wireless multimedia systems jointly under a cross-layer scenario provides gains not attainable with the isolated layer design [1], [2]. At this point, cross-layer optimization seems to be an alternative technique for improving the multimedia transmission performance of the existing wireless networks instead of developing a revolutionary design from scratch.

In the literature, studies focus on improving the design objectives such as QoS guarantees, throughput or video quality maximization, power saving and delay minimization. To achieve a certain QoS level, a joint source rate selection and power management for hybrid ARQ/FEC schemes in wireless multimedia multicast is proposed in [3]. For real-time wireless multimedia transmission, [4] proposes a rate-distortion optimized joint ARQ/FEC scheme where APP layer parameters are adjusted according to parameters in PHY layer. Furthermore, a combined APP-MAC-PHY layer design to achieve the optimum rate adaptation from an information theoretical view for real-time streaming is given in [5]. A similar but more practical approach is taken in [6] for adaptive optimization of cross-layer design parameters. The proposed framework adapts PHY-link-APP layer parameters jointly to achieve QoS requirements for real-time video transmission.

According to the works on this front, the improvement of design objectives is done by exchanging information among various OSI layers adaptively. The adaptation is achieved via the utilization of the feedback information received from the uplink channel. Although various works exist in the literature on cross-layer design, the applicability of these designs to a wireless video broadcasting scenario poses significant difficulties. The main problem is the adaptive tuning of system parameters via the utilization of the uplink channel information when the transmission mode is broadcast (i.e. amount of feedback information is enormous). In a wireless broadcasting scenario, adaptation of the broadcast rate according to the channel conditions of the worst user results in a spectrally inefficient transmission. On the other hand, ARQ for retransmission of lost packets to a specific user is not feasible as number of users grows. As a result, works on cross-layer design for video broadcasting are scarce. There has only been a single cross-layer design proposal for cellular wireless broadcast services in the literature until now [8]. In this proposal, the authors propose a novel cross-layer optimization of the FEC parameters at the PHY, link, transport and APP layers of MBMS platform. Although [8] presented the gains due to optimal FEC selection over different OSI layers, there are no works focusing on the inter-play between the operating SNR region in the PHY layer and the FEC applied in the APP layer for wireless video broadcasting.

In this paper, we propose a novel, cross-layer optimized framework for improving the performance of video broadcasting using H.264/AVC [9] in a wireless cellular system supporting multiple data rates. In this framework, we set the parameters at APP and PHY layers jointly so that an optimal compromise is reached among the two goals of the system: minimization of the broadcast service outage and maximization of the average received video quality in PSNR. According to our knowledge, the idea of optimizing the operating SNR in a cellular wireless video broadcasting scenario has not been questioned in the literature until now. Thus, in this work, we focus on the optimization of the system parameters both at...
APP and PHY layers in order to jointly maximize the service coverage and the average video quality at the receiver side.

The remainder of this article is organized as follows. First, we discuss the importance of the operating SNR region in Section II. We formalize the cross-layer problem for wireless video broadcasting in Section III. Then, we present the simulation platform and performance evaluations for the proposed system design in Section V. Finally, conclusions are drawn in the last section.

II. OPERATING SNR REGION

As 2G wireless systems are based on providing only voice service to the users, these systems are designed to improve the quality of this service. With the introduction of 3G systems, the multimedia service become available at very high speeds with the channel packet error rate (PER) inherited from 2G systems. Although advanced error concealment techniques are developed to obtain a satisfactory video quality at the receiver side, 3G wireless systems operate at a high PER of 1% [10]. Therefore, the optimality of the PER for a cellular wireless video broadcasting scenario has not been questioned in the literature until now and accepted as 1% since this value is enough to provide an acceptable quality for voice services in 2G and 3G wireless systems. However, to provide a video broadcasting service through a wireless channel having 1% PER, the transmitter needs to add redundant packets using FEC in order to achieve a satisfactory video quality at the receiver. But, as the applied FEC at the transmitter side increases, it is clear that the effective video data rate, which is directly proportional to the video quality, decreases.

To solve this issue, one may change the operating SNR, e.g., using adaptive modulation and coding, so that the PER of the channel is decreased instead of increasing the applied FEC at the transmitter. However, this approach may cause a significant coverage drop among the users if users experiencing worse channel conditions have low SNR than the operating SNR. The optimal compromise has to be found by searching for the optimal operating PER of the system instead of using an optimized PER for vocoders and endeavoring for concealing the data stream received with packet losses. Therefore, the broadcast system has to be designed in a way that service quality is maximized in the average sense. The key challenge in a wireless broadcast problem is to serve as many users while achieving a maximum average video quality at the receiver side simultaneously.

III. PROPOSED CROSS-LAYER DESIGN

The cross-layer design problem can be formulated as an optimization problem where the objective is to select an optimal compromise operating point including multiple system parameters. In this paper, the cross-layer optimization problem includes system parameters from PHY and APP layers. However, extensions can easily be developed to include extra parameters from other layers with the associated constraints. Let $n$ denote the total number of system parameters considered in the cross-layer optimization problem. To clarify, let $k$ and $m$ be the numbers of system parameters at PHY and APP layers, respectively, satisfying $n = k + m$. The optimization problem searches for the optimal $n$-tuple operating point among possible configurations such as

$$\Delta = \{P_1, P_2, ..., P_k, A_1, A_2, ..., A_m\}$$

where $P_i$ and $A_i$ denote the $i$'th system parameter at PHY and APP layers, respectively. For the rest of the paper, the subscript $\Delta$ refers a wireless system operating according to a configuration set $\Delta$.

The wireless video broadcasting problem aims to maximize the average received video quality while minimizing the service outage probability at the same time. Since wireless channel is prone to variations, each user experiences fluctuations in the SNR and some of the packets are lost during the transmission. At the receiver side, the video decoder uses error concealment techniques to retrieve the transmitted video stream. Therefore, the received video quality is different from one user to another according to the channel conditions of the users.

First objective, maximization of the average received video quality, can be represented by:

$$\max \left( \frac{1}{M} \sum_u RVQ_{\Delta}(u) \right)$$

where $RVQ_{\Delta}(u)$ defines the video quality experienced by user $u$. To clarify, the average received video quality can be expressed by the concatenated functions below:

$$RVQ_{\Delta}(u) = VD_{\Delta} (WC_{\Delta} (SC_{\Delta}, u))$$

where $VD_{\Delta}$ is the video decoder function at the receiver side, which takes the received video packets as the input and returns the video quality as the output. $WC_{\Delta}$ is the function that performs packet losses on the transmitted video stream according to the wireless channel conditions of a given user. $SC_{\Delta}$ returns the video stream coded by the video coder at the transmitter side. For a feasible video broadcasting session, the transmission rate is subject to:

$$R_{SC_{\Delta}} + R_{FEC_{\Delta}} \leq \overline{R}_{\Delta}$$

where $R_{SC_{\Delta}}$ and $R_{FEC_{\Delta}}$ are the rates required by source coder and by the FEC technique in use, respectively, and $\overline{R}_{\Delta}$ is the video broadcasting data rate.

Secondly, if user $u$ experiences fading at time slot $t$, the SNR value, $\gamma_{\Delta}(u, t)$, may drop under a threshold SNR, $\overline{\gamma}_{\Delta}$, for $\overline{R}_{\Delta}$, which is a service outage effective data rate of user $u$ at time slot $t$, $R_{\delta}(u, t)$, is equal to zero. Second objective, minimization of the service outage probability, can be represented by:

$$\min \left( \sum_u \sum_t \delta_{R_{\delta}(u, t), 0} \right)$$
\[
\delta_{i,j} = \begin{cases} 
1, & \text{if } i = j \\
0, & \text{if } i \neq j 
\end{cases}
\]

(6)

\[
R_{\Delta}(u,t) = \begin{cases} 
\bar{R}_{\Delta}, & \text{if } \gamma_{\Delta}(u,t) \geq \bar{\gamma}_{\Delta} \\
0, & \text{if } \gamma_{\Delta}(u,t) < \bar{\gamma}_{\Delta} 
\end{cases}
\]

(7)

where \( \bar{\gamma}_{\Delta} \) is the associated threshold SNR value at transmission rate of \( \bar{R}_{\Delta} \).

An illustration of the cross-layer optimizer is shown in Fig. 1. The optimizer receives the original video stream from the video server and finds the optimal values of parameters in both the physical layer and the application layer. For cross-layer optimization, multi-objective optimization (MOO) is used since the problem includes two conflicting objectives, 2 and 5, which cannot be satisfied simultaneously.

IV. MULTIPLE-OBJECTIVE OPTIMIZATION

MOO aims to find the solution of an optimization problem with a set of multiple objectives. A solution is called globally Pareto-optimal if any one of the objectives cannot be improved without degrading the other objectives for this solution [11]. Assume that the optimization problem under investigation consists of a number of distinct and possibly conflicting objective functions. Without any loss of generality, assume further that the problem in hand requires all of the objective functions to be minimized. Then, a Pareto-optimal solution exists if there is no other feasible solution that is at least as good as this Pareto-optimal solution in all of the objective functions and also is strictly better in one or more objective functions. For single objective optimization problems, it is possible to have multiple optimal solutions resulting in a unique optimal functional value. It is also possible to have multiple Pareto-optimal solutions in multi-objective optimization problems. However, unlike the single objective problems, the multiple Pareto-optimal solutions do not necessarily result in a unique functional value. In many instances, as different objective functions represent different system aspects on a specific scale, variance, and units of measurement, it is difficult to discriminate between these Pareto-optimal points and determine which one is better than the others. However, if relative importance weights for each of the objective functions are specified, a so-called best compromise solution may be determined.

In order to find the best compromise solution among the objective functions one has to first re-scale their range of values to lie in the intervals \([0; w_p]\); where \(w_p\) is the importance weight of the \(p\)’th objective function. Once scaling is done, all feasible operating points are mapped onto the \(P\)-dimensional space where each dimension represents one of the objectives. In multi-objective optimization, an infeasible operating point that optimizes all of the objective functions simultaneously is called the utopia point. When \(P = 2\) and both of the scaled objectives need to be minimized, the utopia point corresponds to the \((0, 0)\) point in the two-dimensional objective space as illustrated in Fig. 2. The best compromise solution is then found as the feasible point that is closest to the utopia point in the Euclidean-distance sense.

V. SIMULATIONS

A. Simulation Platform

We use the 3G packet data standard, IS-856 [12] in the simulations to provide realistic results. For H.264/AVC codec, we use the most recent software from JVT available freely [13]. To analyze the performance gains of the cross-layer design for wireless video broadcasting applications, we focus on a variety of system parameters. All the PHY and APP layer parameters used in the simulations are given in Table I. The simulations can be divided into four main categories, as system level, PHY layer, video coding and cross-layer simulations.

System level simulations provide a detailed wireless channel model, with path loss, shadow fading, and multi-path fading to imitate the mobility characteristics. Path loss is modeled with the COST231-Walfish-Ikegami [15], shadow fading is modeled as described in [15] using the Gudmundson model [16] with parameters \(\zeta_D = 0.3\) for a distance of \(D = 10\) m, zero-mean Gaussian variable with \(\sigma = 4.3\) dB and a sample rate \(T = 0.5\) sec. Multi-path fading is modeled with a Rayleigh distribution as in [15] by filtering two independent white Gaussian noise sources. The effect of the wide-band channel is modeled using two of the ITU-IMT 2000 channel models, namely the Pedestrian A and Vehicular B models [17].
System level, we model a 3-tier cellular layout with hexagonal cells in order to observe all the diversity and interference effects. Mobile users are uniformly dropped into the center cell and the resulting SNR values of each terminal are computed for 60 seconds. Then, this procedure is iterated for 50 times to have 3000 seconds of broadcast simulation time. Since the system under consideration is for broadcast service, all of the base stations in the cellular layout are transmitting the same signal at the same time. Then, macro-diversity techniques can be employed to enhance the received signal-to-noise ratio of the mobile terminals [14]. In the simulations, a ten-fingered RAKE receiver with maximal ratio combining is assumed.

PHY layer simulations determine the minimum SNR levels required to maintain each packet error rate value for the various PHY layer modes corresponding to different transmission data-rates. For this purpose, we use Agilent’s Advanced Design System for the PHY layer simulations. The minimum required SNR values to maintain each of the packet error rates are computed for each of the transmission data-rates supported by the 1xEV-DO. Once the PHY layer simulations are complete, the results can be incorporated to system level simulations to determine the transmission data-rate that can be supported with the corresponding service outage percentage for all users in the cell as a function of time.

Video coding simulations provide encoded video packets from the network abstraction layer (NAL) for a given video stream. For a given setting of the video resolution, frame rate, GOP size, intra period as well as QP, an encoded video stream is generated by the H.264/AVC codec [13]. The H.264/AVC codec tags each encoded video packet as nondiscardable or discardable according to its position in the hierarchical GOP structure. It is clear that losing a nondiscardable video packet instead of a discardable video packet causes a larger PSNR degradation at the decoded video sequence. We encoded two different reference video sequences, “Foreman” and “Mobile” at the CIF resolution of 352x288.

Cross-layer simulation is a tool for combining the results of the three simulations described above. It is assumed that the video packets (VPs) obtained as the output of the video coding simulation are to be broadcasted to the users. The NAL generates video frames (or packets) whose sizes are not constant and may vary significantly over the course of the video stream. While one frame may fill only a small portion of the PHY layer transmission packet, another may occupy multiple packets. On the other hand, when a VP is divided into multiple PLPs, it becomes more vulnerable to packet losses since losing one of these PLPs results in losing the whole VP. During the course of transmission over the wireless channel, we observe that the loss of even a single nondiscardable video frame hurts the video quality, and thus the PSNR value, significantly. Therefore, we apply simple block codes with different strengths to nondiscardable and discardable VPs as FEC and assume that nondiscardable VPs are not divided between multiple PLPs unless necessary. As a result, some frame fill inefficiencies are unavoidable. Also, the header packets at the start of each video sequence are required for successful decoding at the receiver side, so they are protected with a powerful repetition code that almost prevents any erroneous transmission for header files.

During physical layer transmission, some of the physical layer packets are lost according to the operating PER value. Then, VPs are extracted from the received PLPs and video stream is decoded using error concealment. Error concealment technique used is frame copy, which simply copies the previous successfully decoded frame. After the decoding, PSNR calculations are done by averaging over the PSNR value of each frame. The corresponding average service outage is measured according to the physical layer SNR of each user and averaged over users in order to compute the outage for the specific mobility model. In this manner, average service outage and average received PSNR are calculated after incorporating the results of the system level, PHY layer and video coding simulations into the experiment. This procedure is repeated for all settings of the design parameters listed in Table I. As a result, the values of the two objectives are obtained for all possible values of the design parameters considered. Then, multiple-objective optimization is invoked to determine the best compromise operating point for the proposed cross-layer video broadcast.

### B. Simulation Results

In order to understand the performance gains achieved by the cross-layer design, we need to compute a traditional system performance for comparison purposes. For this manner, each video sequence is encoded with heuristic parameters, i.e. non-cross-layer mode. The GOP size, and the intra frame period are selected to be 16 and 128 respectively, since these values are widely used in the literature due to their good performance in general. We employ a simple, rate 1/2 block code to protect both non-discardable and discardable video packets since the wireless system operates at 1% PER. Then, this configuration is used to encode the video sequences at each data-rate supported by the wireless system.

For both video sequences, the optimal data rate is found to be 614.4 kbps since it has almost zero outage probability for the Pedestrian A channel while having an acceptable outage probability for the Vehicular B channel among different PER regions. Fig. 3 compares the PSNR values achieved by the cross-layer design and the traditional approach for both video sequences. We can observe that a significant increase in video

<table>
<thead>
<tr>
<th>Table I</th>
<th>CROSS-LAYER SIMULATION PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Parameters</td>
<td>Set</td>
</tr>
<tr>
<td>Physical Layer</td>
<td>Data Rate (kbps)</td>
</tr>
<tr>
<td></td>
<td>Packet Error Rate</td>
</tr>
<tr>
<td>Application Layer</td>
<td>Frame Rate (Hz)</td>
</tr>
<tr>
<td></td>
<td>QP</td>
</tr>
<tr>
<td></td>
<td>GOP</td>
</tr>
<tr>
<td></td>
<td>Intra Frame Period</td>
</tr>
<tr>
<td></td>
<td>FEC</td>
</tr>
</tbody>
</table>
quality is achieved with the cross-layer framework as expected. Furthermore, the optimal video configurations for each PER values can be seen in Table II.

We observe that as the operating SNR increases, the optimal video configurations need weaker FEC, meaning that more video information can be sent at the same data rate instead of redundant bits. This can be observed from the QP values since a low QP value at the encoder creates a better quality video as an output. On the contrary, as PER decreases, intra frames become unnecessary since they are not optimized in the rate-distortion sense, they allocate more space and decrease the performance. In “Foreman” sequence, PSNR increases by 1.6 dB at the high PER region, while this increase becomes 2.3 dB at the low PER region. Similarly, in “Mobile” sequence these values are 1.5 dB and 2.6 dB. It is clear that using a cross-layer design that simultaneously optimizes PHY and APP layer parameters, the average video quality at the receiver side is increased substantially.

To observe the actual performance gain achieved via the cross-layer design, we need to invoke MOO framework. For MOO, we first determine the utopia points for each video sequence used in the simulations. The utopia point has a PSNR value of a video sequence encoded at the maximum data-rate supported by the wireless system. In addition, it is assumed that the video sequence is transmitted through an error-free channel so that the outage probability for the utopia point is zero. The maximum transmission data-rate for the IS-856 rev. 0 system is 2.4576 Mbps. Since PSNR values above 40 dB do not improve the visual video quality further, 40 dB is accepted as an upper bound for the utopia point. As a result, the utopia point for “Foreman” has a PSNR value of 40 dB while “Mobile” has a PSNR value of 38.9701 dB.

Before using MOO, we need to define upper and lower limits for each objective function space since these spaces will be used in the rescaling as mentioned in Section IV. Therefore, 10% is accepted as the maximum allowable percentage outage probability while a lower-bound for the PSNR value, i.e. the minimum satisfactory video quality, is set to 29 dB. Then, we compute the distance of each configuration according to the goals of the system and conduct MOO using equal weights for each of the design objectives. Fig. 6 shows the distances from the utopia point for optimal modes in each PER region for the cross-layer and traditional approaches for both sequences. For the vehicular case, the results show that the optimal PER

**Table II**

<table>
<thead>
<tr>
<th>Video Sequence</th>
<th>PER</th>
<th>Frame Rate (Hz)</th>
<th>QP</th>
<th>GOP Size</th>
<th>I Frame Period</th>
<th>NFEC</th>
<th>DFEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreman</td>
<td>10^-4</td>
<td>30</td>
<td>28</td>
<td>32</td>
<td>64</td>
<td>2/3</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>5.10^-3</td>
<td>30</td>
<td>28</td>
<td>32</td>
<td>128</td>
<td>2/3</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>2.10^-3</td>
<td>30</td>
<td>28</td>
<td>32</td>
<td>128</td>
<td>2/3</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>10^-4</td>
<td>30</td>
<td>27</td>
<td>16</td>
<td>64</td>
<td>5/6</td>
<td>1/1</td>
</tr>
<tr>
<td></td>
<td>5.10^-4</td>
<td>30</td>
<td>27</td>
<td>16</td>
<td>128</td>
<td>5/6</td>
<td>1/1</td>
</tr>
<tr>
<td>Mobile</td>
<td>10^-4</td>
<td>30</td>
<td>34</td>
<td>32</td>
<td>Null</td>
<td>2/3</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>5.10^-3</td>
<td>30</td>
<td>34</td>
<td>32</td>
<td>Null</td>
<td>2/3</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>2.10^-3</td>
<td>30</td>
<td>34</td>
<td>32</td>
<td>Null</td>
<td>3/4</td>
<td>4/5</td>
</tr>
<tr>
<td></td>
<td>10^-4</td>
<td>30</td>
<td>34</td>
<td>16</td>
<td>128</td>
<td>5/6</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>5.10^-4</td>
<td>30</td>
<td>33</td>
<td>16</td>
<td>Null</td>
<td>5/6</td>
<td>1/1</td>
</tr>
</tbody>
</table>
region for “Mobile” sequence shifts towards left. This is due to the significant variations in the outage probability among all the PER values simulated. The optimal PER for vehicular channel is $10^{-3}$ while it is $10^{-2}$ for the pedestrian channel. Thus, the optimal operating PER decreases when the mobility of the users increases. This indicates that the increase in the average video quality suppresses the increase in the service outage and a best compromise operating point is found at a lower PER when users are traveling at a higher speed. The change in the optimal PER for different mobility models shows the benefit of optimizing PER and the additional increase on the overall system performance.

VI. CONCLUSIONS

In this paper, we propose a novel, content-adaptive, MOO framework for wireless H.264/AVC video broadcasting using a cross-layer design. The proposed scheme improves the overall performance of the video broadcasting service by finding the best compromise system parameter tuples among PHY and APP layers by maximizing the average received video quality and the geographical coverage. The key contribution of this work is the extensive analysis among system parameters including the wireless system operating SNR region. The simulations show that using a cross-layer design for wireless video broadcasting is necessary and beneficial. In addition, the proposed cross-layer design framework can be used to compute a best compromise mode for any wireless system supporting multiple data-rates in a broadcasting scenario.

REFERENCES