(IEEE WCNC 2019 submitted)

QoE-aware Cross-Layer Adaptation for Delay-Constrained Video Transmission over Wireless Channels

Chaynika Taneja
Defence Research and Development Organization
New Delhi 110054, India
chaynika@hqr.drdo.in

Subhankar Chatterjee, Swades De
Department of Electrical Engineering and
Bharti School of Telecommunication, IIT Delhi
New Delhi 110016, India
subhankar.ece@gmail.com, swadesd@ee.iitd.ac.in

***Abstract*—In this paper a cross-layer adaptive scheme is presented for transmission of H.264 encoded video over wireless channels, under a specified delay constraint. Adaptive modulation and coding (AMC) is considered at the physical layer, accompanied by type-I hybrid automatic repeat request (HARQ) at the link layer. The HARQ scheme is cognizant of the video characteristics. It prioritizes the important independent (I) frames at the cost of dependent (P) frames, maintaining the group of pictures (GOP) deadline boundaries. At the application layer, quantization parameter (QP) variation is carried out to control the rate and quality requirements. This multi-layer adaptive scheme is modelled analytically as a multivariate nested Markov chain (MC). The proposed approach takes into account both fading and shadowing effects of the wireless channel. Numerical results highlight performance benefits of the proposed model. Realistic simulation results indicate improved performance of the proposed approach over similar multi-layer adaptation schemes in terms of video quality metrics.**

I. Introduction

Video streaming applications have gained wide popularity with the recent advancements in encoding techniques and wide proliferation of wireless networking. However due to the highly dynamic nature of the wireless links and stringent delay and throughput requirements of video, wireless streaming is still a challenging task. Different techniques at the distinct layers, namely, adaptive modulation and coding (AMC), forward error correction (FEC), automatic repeat request (ARQ), and source rate control, can be used in combination or in isolation towards this end. Such cross-layer schemes have been of wide interest for improvement in received video quality.

A. Related Works

Cross-layer optimization between physical layer and datalink layer was addressed in [1] considering AMC and fixed limit ARQ. The work in [1] was extended in [2] considering the effects of finite queuing and QoS constraints. However, parameters relevant to the video quality were not considered. The authors in [3] discussed an adaptive truncated Hybrid ARQ (HARQ) scheme that utilizes inter-layer encoding. Similarly, AMC with quantization parameter (QP) adaptation was discussed in [4].

Previous works in literature addressed video-specific ARQ schemes [5], [6], considering differential treatment according to frame type [5] and the use of group of pictures (GOP) as well as packet level deadlines [6] to control retransmissions. Various techniques for video rate adaptation were also discussed in literature, including proxy based adaptation [7] and progressive downloads [8].

Markov models have been extensively used in modelling video transmission schemes. Markov modelling of selective retransmissions in [9] provided a basic framework and motivation for analyzing the multi-layer schemes. A general hierarchical Markov model for characterizing the effect of shadowing was presented in [10]. This was extended for performance analysis of selective-repeat ARQ (SR-ARQ) in [11]. Parameter adaptation at more than two layers for video transmission was also disccussed in literature. A three layer adaptation scheme was studied in [12] considering AMC at physical layer, ARQ at link layer and source coding at application layer.

Recent works also focused on the delay constrained cross-layer designs. The authors in [13] utilized differentiated FEC spreading with delay constrained rate assignment to reduce video distortion in one-to-many scenarios. The cross-layer design of [14], [15] addressed the physical, data link, and application layers, modelling the system as a four dimensional MC. In [16], an interference aware scheme was presented that jointly determined the video encoding rate and link layer transmission policy.

B. Motivation

Main motivation behind this work has been the lack of video content and delay aware parameter adaptation at more than two layers. Previous works in literature considered different encoding parameters such as, frame rate (FR) and spatial resolution (S), in a three-layer based framework. However, QP based rate adaptation was not considered which offers better computational, hardware, and memory efficiency. Literature is also rich in video transmission techniques considering only the fading effects of wireless channel. To this end, it is necessary to design a video transmission scheme taking into account the effects of both fading and shadowing as a more realistic scenario.

C. Contributions

The contributions of this work are summarized as follows:

1) Parameter adaptation at the three layers, particularly, the joint effects of QP based rate control, ARQ at data link

layer, and AMC at the physical layer is studied for video transmission, which has been missing in the literature.
2) An ARQ scheme that is aware of the video characteristics is considered as opposed to the conventional scheme. This is possible as the link layer frames can now be potentially tagged with the originating application layer frame information.
3) Unlike the previous simulation based studies, the fading and shadowing effects of channel on video transmission are jointly accounted analytically. Effect of non instantaneous feedback has also been taken into consideration.
4) Two different time scales are included to account for the different response times of the physical and the application layer. QP variation is carried out at a coarser time scale (GOP level) whereas AMC triggering is done at the closer intervals (frame level).

II. Proposed Scheme

A. System Model

The system model shown in Fig. 1 highlights the transmission of H.264 encoded video over a noisy wireless channel with known statistics. The transmitter implements both AMC and ARQ. Type-I HARQ scheme is considered here as a combination of FEC and ARQ. FEC packets are generated here using a suitable error correcting code, e.g. reed solomon code. The receiver sends positive (ACK) or negative acknowledgement (NACK), respectively to the transmitter depending upon the success or faliure of frame transmission. The feedback channel is assumed to be error-free. Furthermore, the effect of non-instantaneous feedback is modelled by embedding memory. Video flow is comprised here of two types of packets: independently coded I packets and incrementally coded P packets. At the application layer, video rate and quality are controlled by varying QP. H.264 reference software Joint Scalable Video Model (JSVM 9.19.15) [17] is used here as video encoder and decoder. The playback buffer at the receiver end is limited in size. So, each packet in the transmitter buffer must be transmitted before a playback deadline, after which its reception at the receiver is futile. For real-time applications, this delay deadline is given by $D_{fr} = \frac{1}{frame\ rate}$. However, a relaxed requirement is considered here in which the delay constraint is imposed on the GOP as a whole and not on individual frames. The GOP time duration T_{GOP} should be below a specific threshold $D_{threshold}$.

$$T_{GOP} \leqslant D_{threshold} \tag{1}$$

B. Proposed Video-Aware Cross-Layer Adaption

A cross-layer adaptive scheme specifically tailored for video transmission is proposed in this paper. Parameters at the physical, data link, and application layers are adapted to maximize video quality at the receiver under the constraints of delay imposed by the limited buffer at the receiver. Conventional AMC, which divides the signal-to-noise ratio (SNR) range into intervals and assigns a modulation scheme to each interval, is adopted [18]. The proposed ARQ scheme provides preferential

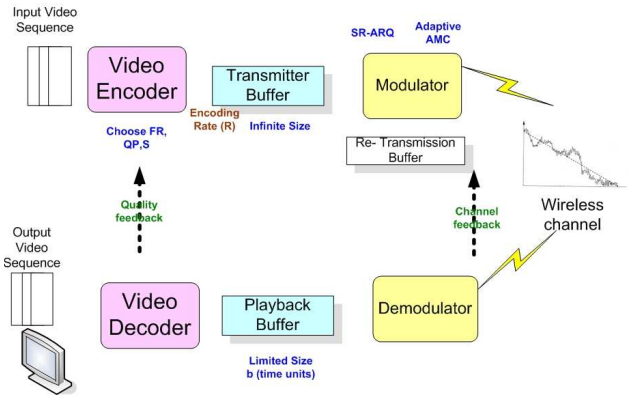


Fig. 1: System model.

treatment to the independent frames at the cost of dependent frames. For application layer adaptation, QP is varied at the GOP level. Specified number of frames are picked up from the raw video data and encoded according to the QP value chosen based on the present shadowing conditions. Unlike other rate control schemes, the entire video sequence is not encoded initially. Rather encoding is done on the fly, thereby saving the computational and memory requirements at the transmitter. Selected QP governs the rate and quality within a GOP. A higher value of QP implies coarse quantization during the encoding process and hence lower rate and quality.

III. Analytical Model

A. Nested Markov Chain Model

Analytical modelling of the proposed multi-layer adaptive scheme is carried out using MC. At first, the media aware ARQ scheme is modelled as a 1-dimensional (1-D) MC. The following assumptions are made for building this model:

1) Size of a I frame is an integer multiple of the size of a P frame, ratio being denoted by k.
2) Dependencies are assumed only on 1 frame. Bidirectional dependence is ignored.
3) Transmission of a P frame requires m slots and that of a I frame requires mk slots.
4) Frame transmission time in slots is a function of QP.
5) It is assumed that ACK and NACK are never lost.
6) Only two types of frames, i.e. I and P are considered.
7) Size of GOP = N.
8) There is no carry forwarding to the next GOP.
9) I packet is retransmitted at cost of P packets.

A detailed system description for the ARQ process for a GOP size 5 and frame size ratio $k = 2$ is illustrated in Fig. 2(a). Since FEC is combined with retransmissions, only the non-decodable packets are completely retransmitted. The first element in each state represents the frame which has been transmitted and is awaiting acknowledgement from the receiver. The second element is the one who's transmission is

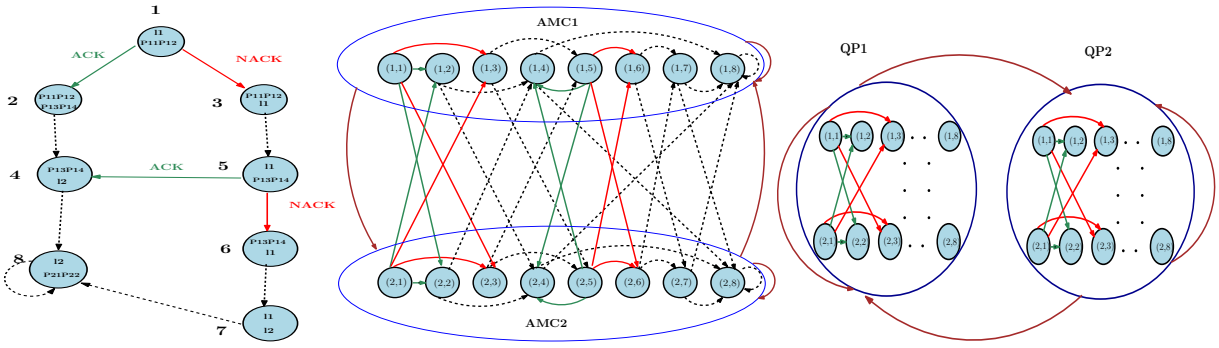


Fig. 2: (a) Video aware ARQ MC, (b) 2-D MC for two modulation schemes + proposed ARQ, and (c) Overall nested MC.

in progress. The transition probability matrix of this MC is represented as:

$$T = \begin{bmatrix} 0 & I_{ack} & I_{nack} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & I_{ack} & 0 & I_{nack} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

where I_{ack} and I_{nack} indicate the probability of success and failure in I frame transmission. Let P_{loss} denote the probability that the transmission in one time slot is in error. So,

$$I_{ack} = \sum_{r=0}^{mk} \binom{mk}{r} (1 - P_{loss})^r P_{loss}^{mk-r} \quad (3)$$

Combining 1-D MC with AMC results in a two dimensional (2-D) chain with the horizontal and the vertical transitions being governed by ARQ and AMC process, respectively. Fig. 2(b) shows a representation for 2D MC with two modulation modes and ARQ. The transition probabilities between the AMC states are calculated following [18]. The detailed analysis is avoided here for brevity. Let P_{AMC} be the transition probability matrix of AMC states. Now, the combined transition probability matrix considering all the micro-states in 2-D MC can be written as

$$T_{micro} = P_{AMC} \bigotimes T \quad (4)$$

The symbol $\bigotimes$ denotes Kronecker product. The parameters for the internal states of AMC+ARQ 2-D chain depend on the chosen QP and hence on the application layer state. Therefore, the overall system model is a nested MC with the macro-states governed by the chosen QP at the application layer. Each macro-state has a number of micro-states indicating the 2-D MC. Fig. 2(c) shows a representative model for two QP, two modulation schemes and GOP pattern IPPPP. A composite shadowed fading channel model is considered here. Let the local mean of the received signal be Θ. Due to the shadowing effect, Θ varies with time and can be partitioned into B non-overlapping intervals. In case of Fig. 2(c), $B = 2$. At time instance n, the channel is in macro-state i if $\Theta(n)\epsilon[\Theta_{i-1}, \Theta_i)$. Let, p_{ij} be the transition probability from macro-state i to macro-state j. The values of p_{ij} are determined by the shadowing process as follows [10]:

$$\begin{aligned} p_{ij} &= Pr[\Theta(n) \in [\Theta_{j-1}, \Theta_j) | \Theta(n-1) \in [\Theta_{i-1}, \Theta_i)] \\ &= \frac{\int_{\Theta_{i-1}}^{\Theta_i} \int_{\Theta_{j-1}}^{\Theta_j} p_{\Theta_1,\Theta_2}(x,y)dxdy}{\int_{\Theta_{i-1}}^{\Theta_i} p_{\Theta_1}(x)dx} \end{aligned} \quad (5)$$

At two consecutive sampling time instants, the joint probability density function of local mean values is p_{Θ_1,Θ_2}. Now, the micro-state transition probabilities between the different or same macro-states can be determined following [10, eq. (12)].

B. Derivation of Performance Metrics

This section explains the derivation of the performance metrics from the proposed analytical model.

1) Average GOP transmission time: The average number of slots required to transmit a GOP is determined using the theory of absorbing MC [19]. The combined transition matrix of the micro states in each macro-state can be expressed in the following canonical form:

$$T_c = \begin{bmatrix} \mathcal{Q} & \mathcal{R} \\ 0_{m,n} & I_m \end{bmatrix} \quad (6)$$

For m modulation modes in the AMC, T_c will have m absorbing states. Let the number of states in ARQ MC be N_r. Then the number of transient micro-states is n = $(N_r - 1) * m$. Now, $\mathcal{Q}$ is $n \times n$, $\mathcal{R}$ is $n \times m$, $0_{m,n}$ (zero matrix) is $m \times n$ and I_m (Identity matrix) is $m \times m$ matrix. The fundamental matrix is then calculated as:

$$\mathcal{F} = (I_m - \mathcal{Q})^{-1} \quad (7)$$

The expected time to reach the absorbing states, starting from any state i is given by the i^{th} entry of $t = \mathcal{F}c$, where c is a column vector $(n \times 1)$ of all 1s. Thus the average time for transmission of a GOP in terms of number of state transitions is given by

$$T_{GOP} = c^T t \quad (8)$$

TABLE I: Parameters for numerical results

Parameter	Value
Number of AMC	2
Number of States	2 macro-states with 16 micro-states each
I frame transmission slots (mk_1,mk_2)	$mk_1 = 9$ (QP1), $mk_2 = 13$ (QP2)
No of GOPs in the video sequence	10
No of states in ARQ MC (GOP IPPPP)	8
Code correction ratio (HARQ) CR	0.7
Doppler frequency f_d	50 Hz
Area mean μ	0 dB
Standard deviation of shadowing/dB spread σ	10 dB
Threshold for shadowing MC	Θ_0 = 0 dB, Θ_1 = 7 dB, Θ_2 = 22 dB
Spatial correlation coefficient ρ	0.8
QP values	30 and 44 (Sim)
Video Sequence	Bus QCIF (Sim)
GOP Size and structure	5 and IPPPP (Sim)

2) Peak signal-to-noise ratio (PSNR): Peak signal-to-noise ratio (PSNR) can be considered as an objective video quality metric. To consider the effect of both video encoding and transmission scheme, we use PSNR model in [16] that accounts for both encoding and transmission distortion. Here, PSNR of video session i can be expressed as

$$PSNR_i = 10\log_{10}\left(\frac{D_{maximum}}{D_i}\right) \tag{9}$$

$$D_i = D_i^{enc} + D_i^{trans} \tag{10}$$

The maximum possible distortion of a given pixel, i.e. $D_{maximum}$ is given by $2^L - 1$. Here, L is the number of bits required to represent a pixel. The distortions due to packet loss and lossy video compression are represented by D_i^{trans} and D_i^{enc}, respectively.

3) Video quality utility: The incremental nature of video encoding leads to loss in packets and frames. This results in PSNR degradation and undecodable frames. Frame copying as a measure of error concealment helps in the case of abrupt drops in PSNR. Still, the the effects of undecodable frames need to be accounted for. The term decodable rate can be used [20] to include the effects of such frames. Decodable rate is calculated as the ratio of successfully decoded frames over total number of transmitted frames. Utility is further defined as a product of video PSNR and decodable rate.

$$\text{Utility} = \text{PSNR} \times \text{Decodable rate} \tag{11}$$

4) Lost frame rate: Percentage of the frames lost is another parameter to measure the performance. This is indicative of the delivery rate of the scheme. Hence, it is different from PSNR which is calculated after suitable frame concealment.

5) Structural Similarity Index (SSIM): SSIM provides a better measure of perceptual quality as compared to PSNR and takes into account the similarities of luminance, contrast and structure [21]. The index is defined as:

$$SSIM(i,j) = \frac{(2\mu_i\mu_j + a_1)(2\sigma_{ij} + a_2)}{(\mu_i^2 + \mu_j^2 + a_1)(\sigma_i^2 + \sigma_j^2 + a_2)} \tag{12}$$

where μ_i and σ_i represent the mean and the standard deviation of i^{th} signal. The symbol σ_{ij} represents the cross correlation between i^{th} and j^{th} signal. The symbols a_1 and a_2 are determined following [21]. The Video Quality Measurement tool in [22] is used here for measurement purpose.

TABLE II: Realistic simulation parameters

Parameter	Value
Video sequence	Foreman QCIF
Resolution	288×352
QP values	$10, 15, 20, 25, 30, 40$
GOP size and structure	15 IPPPPPPPPPPPPP
No of frames	200
Fading	Rayleigh
Distance between Tx and Rx	10m
Path loss model	Friss
Doppler frequency	50 Hz
Slot duration T_s	2 ms
No of multi-paths	100
Frame length	1080 bytes
Shadowing	Log Normal
Standard deviation of shadowing	4
Frequency	2 GHz
Frame rate	25 fps
Modulation	QAM 4,16,64

TABLE III: Analysis versus simulation

Parameter	NRMSE	Coefficient of Determination (R-squared)	Adjusted R-squared
GOP time QP set1	0.0352	0.972	0.969
GOP time QP set2	0.0632	0.922	0.913

IV. Results and Discussion

A. Verification of Analysis by Numerical results

For numerical analysis, we consider the two state MC channel model widely used in the literature. The states are indicated by G and B, representing good and bad channel conditions. The bad and good channel coditions can be translated to modulation modes BPSK and QPSK, respectively. Rayleigh flat fading channel is considered in this paper. The transition probabilities of 2-D MC can be calculated following [23]. The detailed analysis is omitted here due to page constraints. Now, the transition probability between any two macro-states considering log-normal shadowing is [10]

$$p_{\Theta_1,\Theta_2}(x,y) = \frac{\frac{10}{\ln 10}}{xy\sigma_\Theta^2 2\pi}\exp\left\{-\frac{1}{1-\rho^2}\left[\frac{(10\log_{10}x - \mu_\Theta)^2}{2\sigma_\Theta^2} + \frac{(10\log_{10}y - \mu_\Theta)^2}{2\sigma_\Theta^2} - 2\rho\frac{(10\log_{10}x - \mu_\Theta)(10\log_{10}y - \mu_\Theta)}{2\sigma_\Theta^2}\right]\right\} \tag{13}$$

Here, ρ indicates spatial correlation coefficient. To verify the validity of the proposed model, simulation is carried out using the parameters listed in Table I. The analytical and the simulation results of the average GOP transmission times are compared for two sets of QP values in Fig. 3(a). The model predicts the GOP times satisfactorily, as indicated by the normalized root-mean-square error (NRMSE) values in Table III. The coefficient of determination (R-squared) and the adjusted R-squared values are also calculated for both QP values.

B. Numerical Comparison with Related Schemes

The performance of the proposed approach is compared with two similar three layer schemes utilizing different retransmission strategies. The first scheme uses the fixed ARQ scheme where each frame has a fixed retransmission limit,

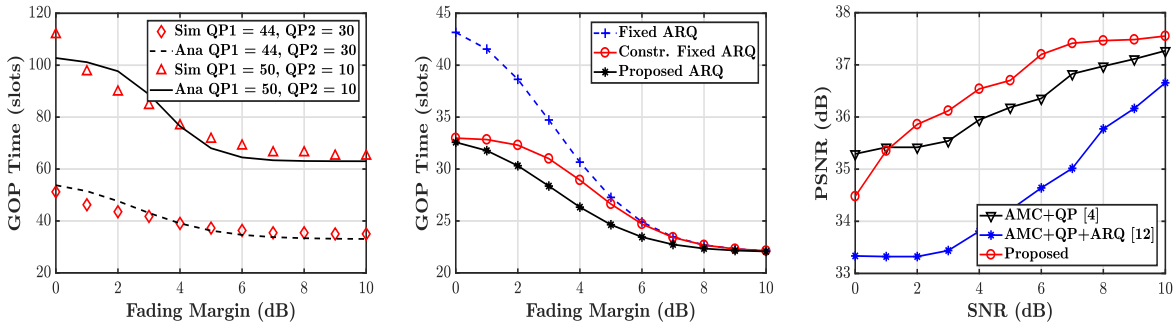


Fig. 3: (a) Average GOP time versus fading margin, (b) variation of GOP times for different schemes; and (c) variation of PSNR with SNR for different schemes.

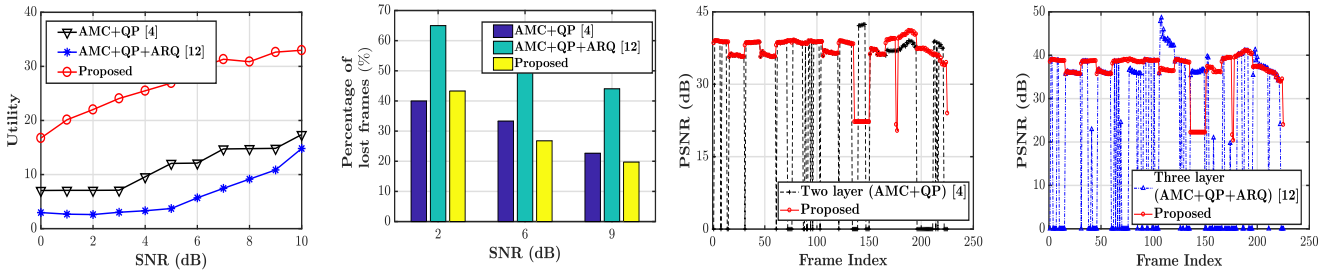


Fig. 4: (a) Utility versus SNR, (b) frame loss percentage versus SNR; (c) proposed approach versus two layer approach [4]; and (d) proposed approach versus three layer approach [12].

exceeding which the frame is dropped. This scheme is denoted by AMC+Fixed_ARQ+QP. To provide a fair comparison in terms of equal delay constraint, a variant of this scheme is also considered for comparison. There, a limit is placed on the total number of slots/transitions a GOP can consume. We denote this scheme as AMC+Constrained Fixed_ARQ+QP .

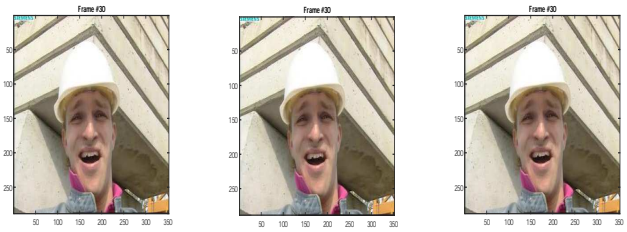
Fig. 3(b) shows the variation of GOP transmission time with fading margin for the proposed approach and the two compared schemes. As the conventional fixed ARQ scheme operates unconstrained, the GOP times are much higher as compared to the other two schemes. For higher values of fading margin (above 8 dB), the conventional ARQ schemes become comparable to the proposed scheme. The increase in the GOP transmission times compared to the proposed scheme is due to the fixed retransmission of P frames. However in the proposed approach, only initial transmissions are permitted for P frames, that too conditioned on the successful transmission of the preceding I frame.

C. Comparison with Similar Cross-Layer Designs

The performance of the proposed approach is compared with similar cross-layer schemes using realistic simulation environment. The parameters necessary for the realistic simulation are listed in Table II. The first scheme chosen for comparison purpose is the two-layer adaptation scheme using AMC and QP denoted by AMC+QP. Similar scheme was addressed in [4]. The second scheme used for comparison is the three-layer adaptation using AMC, QP as well as conventional ARQ dentoted by AMC+QP+ARQ. Similar scheme was considered in [12].

Fig. 3(c) shows the variations of video PSNR as a function of SNR for the proposed scheme and the compared schemes. AMC+QP+ARQ performs worst as there is no error correction during frame reconstruction. Recovery is only through retransmission and transmission is pre-empted to satisfy GOP deadline. AMC+QP with FEC performs better than only retransmission at the link layer. However, the measured PSNR values are lower than the proposed scheme by about 3 dB.

The proposed scheme considers the effects of the lost frames as well as frames which are not decoded at the receiver. Therefore, the utility measure of the proposed scheme is significantly higher than the compared schemes. The variations of utility with SNR is shown in Fig. 4(a). Fig 4(b) shows the variations of the percentage of lost frames for the three schemes at three different SNR values. At low channel conditions, FEC coupled with AMC can rectify the channel impairments, hence results in low percentage of lost frames. At moderate and high SNR values, the media aware HARQ strategy of the proposed scheme is beneficial. At high SNR region, percentage loss of less than 10% can be observed. Fig 4(c) compares PSNR of the proposed scheme with the two layer scheme [4] and Fig 4(d) shows the corresponding comparison with the three-layer



(a) Two layer [4] (b) Three layer [12] (c) Proposed

Fig. 5: Subjective comparison of video quality of Foreman frame number 30.

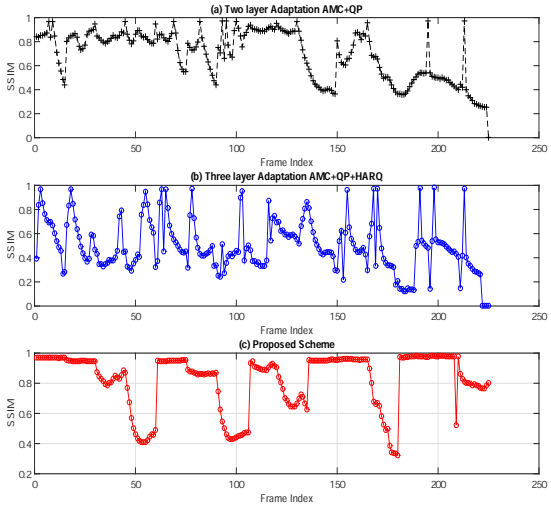


Fig. 6: Variation of SSIM with frame index.

scheme [12]. PSNR variation with frame index is much higher in the other schemes compared to the proposed scheme. It results in significant degradation of the video quality.

The screen shots of the received sequences for the three schemes are shown in Fig. 5 for Foreman frame number 30. The proposed scheme also achieves higher perceptual quality as indicated by the higher values of SSIM index in Fig. 6.

V. Conclusion and Scope of Future Work

In this paper, we have proposed a cross-layer adaptation scheme for video transmission over wireless channels experiencing fading as well as shadowing. The scheme combines AMC, video aware HARQ and QP based rate control, under a GOP level delay constraint. Thereafter a Markov framework has been proposed to quantify the GOP transmission time. Applicability of the model has been validated by comparing with the simulation results on actual encoded video samples, and the results have been found to be in agreement. Realistic simulation results have shown improved performance of the proposed scheme over the closest competitive ones.

References

[1] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.

[2] X. Wang, Q. Liu, and G. B. Giannakis, "Analyzing and optimizing adaptive modulation coding jointly with ARQ for QoS-guaranteed traffic," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 710–720, Mar. 2007.

[3] C. Zhu, Y. Huo, B. Zhang, R. Zhang, M. El-Hajjar, and L. Hanzo, "Adaptive-truncated-HARQ-aided layered video streaming relying on interlayer FEC coding," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1506–1521, Mar. 2016.

[4] D. Wu, S. Ci, H. Luo, W. Zhang, and J. Zhang, "Cross-layer rate adaptation for video communications over LTE networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 4834–4839.

[5] C. Shu, N. Ma, T. Wu, Y. Wang, and P. Zhang, "QoS differentiation adaptive retransmission limits ARQ for ieee 802.16e BWA system," in *Proc. IEEE 66th Veh. Technol. Conf.*, Sep. 2007, pp. 1533–1538.

[6] C. M. Chen, C. W. Lin, and Y. C. Chen, "Cross-layer packet retry limit adaptation for video transport over wireless LANs," *IEEE Trans. Circuits Syst. Video Technol*, vol. 20, no. 11, pp. 1448–1461, Nov. 2010.

[7] H. Hu, X. Zhu, Y. Wang, R. Pan, J. Zhu, and F. Bonomi, "Proxy-based multi-stream scalable video adaptation over wireless networks using subjective quality and rate models," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1638–1652, Nov. 2013.

[8] K. J. Ma, M. Li, A. Huang, and R. Bartos, "Video rate adaptation in mobile devices via HTTP progressive download of stitched media files," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 320–322, Mar. 2011.

[9] 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, Apr. 2010.

[10] H. Yang and M. S. Alouini, "A hierarchical markov model for wireless shadowed fading channels," in *Proc. IEEE 55th Veh. Technol. Conf.*, vol. 2, May 2002, pp. 640–644.

[11] F. Chiti, R. Fantacci, and R. Marchiani, "Performance analysis of a ARQ-SR protocol over a wireless packet network channel," in *Proc. IEEE Int. Conf. Commun.*, vol. 11, June 2006, pp. 5117–5122.

[12] D. Wu, S. Ci, and H. Wang, "Cross-layer optimization for video summary transmission over wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 4, pp. 841–850, May 2007.

[13] J. Wu, B. Cheng, C. Yuen, N. M. Cheung, and J. Chen, "Trading delay for distortion in one-way video communication over the internet," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 26, no. 4, pp. 711–723, Apr. 2016.

[14] C. Singhal, S. De, R. Trestian, and G. Muntean, "Joint optimization of user-experience and energy-efficiency in wireless multimedia broadcast," *IEEE Trans. Mobile Computing*, vol. 13, no. 7, pp. 1522–1535, July 2014.

[15] C. Singhal and S. De, "UE-tv: User-centric energy-efficient HDTV broadcast over LTE and Wi-Fi," *IEEE Trans. Mobile Computing*, July 2018 (accepted).

[16] J. Tian, H. Zhang, D. Wu, and D. Yuan, "Interference-aware cross-layer design for distributed video transmission in wireless networks," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 26, no. 5, pp. 978–991, May 2016.

[17] "Joint scalable video model," https://www.hhi.fraunhofer.de/en/departments/vca/research-groups/image-video-coding/research-topics/svc-extension-of-h264avc/jsvm-reference-software.html, accessed: 2017-05-17.

[18] Q. Liu, S. Zhou, and G. B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: cross-layer analysis and design," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1142–1153, May 2005.

[19] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Springer Verlag, 1960.

[20] L. Cai, S. Xiang, Y. Luo, and J. Pan, "Scalable modulation for video transmission in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4314–4323, Nov. 2011.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[22] "Video Quality Measurement Tool," http://mmspg.epfl.ch/vqmt, accessed: 2017-06-8.

[23] M. Zorzi, R. R. Rao, and L. B. Milstein, "ARQ error control for fading mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 445–455, May 1997.