Quality-oriented Video delivery over LTE using Adaptive Modulation and Coding

Amit Pande*, Vishwanath Ramamurthi+ and Prasant Mohapatra*
* Department of Computer Science, University of California, Davis, CA, USA
Email: {amit, prasant}@.cs.ucdavis.edu
+ AT&T Labs, San Ramon CA, USA, Email: vrama@att.com

Abstract—Long Term Evolution (LTE) is emerging as a major candidate for 4G cellular networks to satisfy the increasing demands for mobile broadband services, particularly multimedia delivery. MIMO (Multiple Input Multiple Output) technology combined with OFDMA and more efficient modulation/coding schemes (MCS) are key physical layer technologies in LTE networks. However, in order to fully utilize the benefits of the advances in physical layer technologies MIMO configuration and MCS need to be dynamically adjusted to derive the promised gains of 4G at the application level. This paper provides a performance evaluation of video traffic with variations in the physical layer transmission parameters to suit the varying channel conditions. A quantitative analysis is provided using the perceived video quality (evaluated using no-reference blocking and blurring metrics) along with transmission delay, as video quality measures. Experiments are performed to measure performance with changes in modulation as well as code rates in poor and good channel conditions. We discuss how an adaptive scheme can optimize the performance over a varying channel.

I. INTRODUCTION

Cellular networks are on the verge of a third phase of growth. The first phase was dominated by voice traffic, and the second phase, which we are currently in, is dominated by data traffic. In the third phase it is predict that the traffic will be dominated by video and will require new ways to optimize the network to prevent saturation [1]. The increasing demand for multimedia-based communications is made viable by increased computational resources in mobile phones (advent of GPUs (such as NVIDIA TEGRA) and special purpose video processing chips such as ARM MALI VE6), evolution of video services to mobile segment (such as Youtube and IPTV [2]) and evolution to new mobile broadband standards like WiMAX IEEE 802.16m and 3G LTE and LTE-Advanced. Service and network providers are exploring the opportunity to further enhance their current offerings and to increase revenues by catering for the demand in rich multimedia services to both mobile and fixed users using cellular networks such as LTE.

3GPP Long Term Evolution

Long Term Evolution (LTE) is emerging as a major candidate for 4G cellular networks and is being adopted by various cellular providers (including AT&T and Verizon wireless in the US). The major features that distinguish LTE from 3G technologies at the air-interface are Orthogonal Frequency Multiple Access (OFDMA), advanced MIMO technology, and Hybrid Automatic Repeat Request (HARQ). In addition

LTE uses flat-IP architecture for the core network. LTE uses OFDMA in the downlink (DL) for efficient multiple-access and for countering multipath frequency selective fading. OFDMA divides the available channel into number of sub-carriers and is naturally suitable for scalable bandwidth allocation by varying the FFT (Fast Fourier Transform) size. In this paper, we concentrate only on the DL.

LTE's enhanced UMTS Terrestrial Radio Access (E-UTRA) and evolved packet core network (EPC) also involve the establishment of an all IP network (AIPN) [3]. Any issues that degrade a network's ability to deliver packets will, as a consequence, degrade the quality of any real-time services of customers currently connected to the network. In the case of video services this degradation is likely to result in pausing of playback due to buffer starvation, macroblocking or full loss of picture depending on the video frames.

LTE supports a full range of multiple antenna transmission techniques including transmit diversity (TD) [4], spatial multiplexing (SM) [5] , and closed-loop eigen-beamforming [6] that are suited for different objectives. Transmit diversity is used for obtaining reliable transmissions and is achieved by using Space Frequency Block Codes [7] in LTE. SM is used for obtaining enhanced throughput and is achieved by using layered space time codes [5]. Eigen-beamforming also is used to improve reliability of transmissions when accurate channel state information is available.

Traditional link adaptation techniques used channel quality information only to adapt the MCS level used for transmissions. While it is helpful to adapt modulation or coding rates, lack of application layer feedback leads to wastage or insufficient increase in these modulation rates. An application layer feedback of video-quality can be very helpful to fine-tune the modulation and coding rates in video delivery scenarios.

Contributions

The contributions of this work are as follows:
1) This paper presents trade-offs between perceived video quality and transmission delay with variations in modulation and coding rates.
2) The trade-off can be used to derive an optimal modulation and coding rate for a given network condition.
3) We use no-reference video quality metrics (blocking and blurring) to evaluate video transmission over LTE network.

II. Video quality measurement

There are two primary methods for measuring the perceived video quality, namely subjective and objective methods.

Subjective methods deal with asking a collection of viewers to watch a video stream and then provide that stream with a rating between 1 and 5. For the purpose of a deployed service, this method of measuring video quality is clearly not feasible.

Objective methods are concerned with performing analysis of network and/or video stream data (typically, as close to the user as possible) in order to extract data which can be used as input to an algorithm which is then used to rate the quality of the video sequence. The goal of any proposed objective metric is to provide a rating which would be closely correlated with subjective ratings from a collection of viewers. The input data for these metrics can range from data which analyses a video in a pixel-by-pixel fashion to data from network QoS measurements. Two popular objective metrics are peak signal-to-noise ratio (PSNR) and Structural Similarity Index Metric (SSIM). They are full reference metric, in the sense that they require original (un-transmitted) video for evaluation purposes. However, such testing is not possible in case of real-world deployments. Reduced-reference metrics [8] require transmission of partial information for video quality evaluation. Hence, this overhead quickly becomes quite intolerable in case of cellular networks which have high traffic demands particularly during peak hours.

No-reference quality evaluation metrics such as blocking and blurring are used to evaluate perceived video quality without any reference to original videos. Using such metrics, it is possible for the base station (eNB, as called in LTE) to adapt the delivered stream based on video quality and not the bit-rate [9].

The EvalVid framework [10] allows evaluation of H.264 framework using subjective metrics (such as MOS) and objective metrics (such as PSNR). As such, any derivations of EvalVid, such as EvalSVC [11] also restrict video quality evaluation to PSNR based metrics, which has the disadvantages of not being close to human visual system and being full-reference and thus, not practical in consumer delivery scenario.

III. LTE Downlink Link Level Simulator

In this Section, we briefly describe the LTE DL link level simulator framework (see Figure 1). The main features of the simulator are as follows:

1) MIMO modes (transmit diversity and spatial multiplexing),
2) Forward error correction using turbo codes,
3) OFDMA (Orthogonal Frequency Division Multiplexing Access) with cyclic prefix (CP),
4) Incremental Redundancy (IR) HARQ (Hybrid Automatic Repeat) combining,
5) Transport block error detection using CRC (cyclic redundancy check),
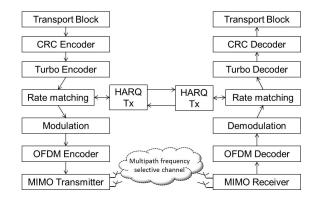6) Rate matching.

Some of them are detailed as follows:

Fig. 1. Block Diagram of Link level LTE Simulator

Turbo Encoder/Decoder: LTE uses a rate-1/3 Parallel Concatenated Convolutional Code consisting of two identical 8-state rate-convolutional encoders connected parallel using an internal interleaver. Viterbi decoding of turbo codes is complex due to the large number of states involved in a concatenated trellis. So we use an iterative MAP (Maximum A Posteriori) detector based algorithm [12] as a practical alternative decoding scheme.

Incremental Redundancy HARQ transmission: In the DL, LTE uses asynchronous and adaptive HARQ mechanism. The schedule of the HARQ transmissions is not pre-declared to the UE. This gives the eNodeB flexibility in scheduling according to priorities and available resource. LTE uses Incremental Redundancy (IR) HARQ as opposed to chase combining. LTE supports up to four redundancy versions for IR HARQ (re)transmissions denoted by $rv_{idx} = 0, 1, 2$ and 3. In each version, a part of rate-$\frac{1}{3}$ turbo-encoded data is transmitted dependent on rv_{idx}.

Rate matching: The rate matching converts the rate-1/3 output from the turbo encoder into the target coding rate. This is done by a block consisting of a three sub-block interleavers, a circular buffer, and a bit-selection block [13]. The number of

TABLE I
LTE OFDMA PARAMETERS

Channel BW (MHz)		1.25	2.5	5	10	15	20
Sampling Frequency		1.92	3.84	7.68	15.36	23.04	30.72
# subcarriers		128	256	512	1024	1536	2048
Frame Duration		307200 x T_s = 10ms					
Oversampling factor		192/125= 1.536					
Sub-carrier spacing		15 KHz					
# Sub-carriers		12					
T_μ		2048 x T_s= 66.667					
Short	T_s	71.875, 71.354 (μs)					
cyclic	CP time	5.21, 4.69 (μs)					
prefix	Sym./frame	160					
Extended	T_s	83.33 (μs)					
cyclic	CP time	16.667 (μs)					
prefix	Sym./frame	120					

bits selected depends on the target coding rate. The start point (or offset) of the selected bits is determined by the HARQ redundancy version rvidx.

OFDMA: LTE uses OFDMA for DL access. The available frequency is divided into sub-carriers of 15 kHz bandwidth. LTE specific OFDMA parameters are listed are listed in Table I[14].

MIMO (Multiple Input Multiple Output): One of the main features of LTE is the use of multiple antennas or MIMO technology to enhance the throughput in an unreliable wireless channel. A $Nt \times Nr$ MIMO system consists of Nt transmitter antennas and Nr receiver antennas. LTE supports MIMO configurations of 4×2, 2×2, 2×1 in the DL. Transmission is done in blocks. Multiple antennas can be used either to obtain more reliable transmissions using Transmit Diversity (TD) or to obtain higher transmission rates through Spatial Multiplexing (SM). In this paper open loop MIMO with TD and SM modes associated with 2×2 antennas are evaluated in the simulation.

In LTE TD is obtained by use of Space Frequency Block Codes (SFBC) as opposed to Space Time Block Codes (STBC). SFBC obtains TD using redundancy in spatial and frequency domains. For a 2×2 MIMO system the optimal diversity scheme is the "Alamouti Code" [4]. The transmission matrix for Alamouti code is given by:

$$X = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}$$

where s_1 and s_2 are symbols from the constellation set of the digital modulation used and s^* denotes the complex conjugate of symbol s. The rows of the matrix correspond to antennas and the columns represent sub-carriers when SFBC is used. Because of the orthogonal structure of X, Maximum Likelihood (ML) decoding of SFBC has remarkably low complexity.

Channel Modeling: A tapped delay line model is used to model multipath frequency selective channel $h_{i,j}$ between transmit antenna j and receive antenna i as follows:

$$h_{i,j} = \sum_{i=1}^{N_{taps}} c_i(t)\delta(t - \tau_i)$$

where N_{taps} represents the number of significant paths from the transmitter to the receiver, $c_i(t)$ represents the gain of path i at time t, and τ_i represents the relative delay of path i. $c_i(t)$ has a Rayleigh distribution and the overall gain of the gain vector $[c_1(t), c2(t)...c_{N_{taps}}(t)]$ is normalized to 0 dB. This power delay profile gives the statistical power distribution of the channel at a particular instant. We model a frequency selective channel in which the delay spread is larger than the symbol duration. But this is compensated by using OFDMA with CP duration longer than the delay spread. The multipath coefficients have Normal distribution with mean chosen according to ITU channel models [10]. The MIMO channel model is then given by an $Nr \times Nt$ matrix consisting of elements $h_{i,j}(t, \tau)$. For a 2×2 MIMO system, the channel

Fig. 2. (a) Original Image, (b) Blurring and (c) Blocking artifacts

matrix is given as:

$$H_{2\times2}(t, \tau) = \begin{bmatrix} h_{1,1}(t,\tau)h_{1,2}(t,\tau) \\ h_{2,1}(t,\tau)h_{2,2}(t,\tau) \end{bmatrix}$$

While the power delay profile is caused by multipath effect, motion of objects causes Doppler spectrum which gives the statistical distribution of the channel at a particular frequency. In addition zero-mean Additive White Gaussian Noise (AWGN) is added at each receiver antenna, the variance of which is varied to get different SNR realizations. The received signal vector is finally given by:

$$y(t, \tau) = H(t, \tau) * X + N_0$$

where $y(t, \tau)$ is the $Nr \times T$ receive signal matrix, $H(t, \tau)$ is the $N_r \times N_t$ channel matrix, X is the $N_t \times T$ transmission matrix, N_0 is the $N_r \times T$ AWGN matrix, and T is the transmission block size.

IV. EXPERIMENTAL SETUP

We integrated video coding subsystem into our LTE link level simulation developed in Simulink. We used Matlab v7.11 with Simulink for this set of experiments. The encoded video (motion vectors and bits) were transmitted using LTE simulator. The receiver decodes the received video stream which is used for no-reference evaluation. We also used delay metric in our experiment to account for the transmission delay (and delay variations) caused by the use of different channel modulation schemes and Turbo code rates. The HARQ retransmission value was kept to a moderate level of 2.

Delay measurement is difficult in Matlab because the simulation is actually slower than real-time and thus we can't use real values directly as indicative of delay, as it is done conventionally using real-time workshop features in Matlab. Instead, we implemented this feature by obtaining the difference between displayed frames in encoder and decoder, normalized by frame-per-second of the sample video.

Blocking artifacts arise from the appearance of vertical and horizontal edges along a regular blocking gris resulting from block based processing in image and video coding standard. Typically, in a wireless scenario, packet loss may result in increased blockiness. The goal of blocking effect measurement algorithm is to detect and estimate the power of blocky signal along horizontal and vertical blocking dimensions [15]. The algorithm estimates the power of blocky signal and then reduces the DC value to remove the power contribution of natural image, making it not subjective to the type of image. This metric is good to model the channel packet drops and its
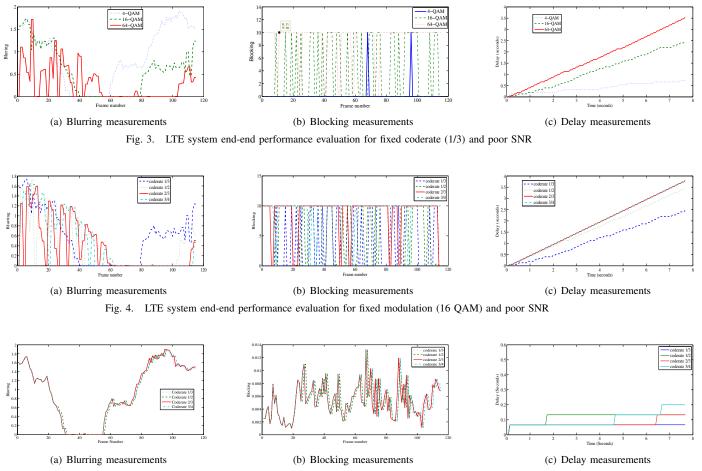
(a) Blurring measurements   (b) Blocking measurements   (c) Delay measurements

Fig. 3.   LTE system end-end performance evaluation for fixed coderate (1/3) and poor SNR

(a) Blurring measurements   (b) Blocking measurements   (c) Delay measurements

Fig. 4.   LTE system end-end performance evaluation for fixed modulation (16 QAM) and poor SNR

(a) Blurring measurements   (b) Blocking measurements   (c) Delay measurements

Fig. 5.   LTE system end-end performance evaluation for fixed modulation (16 QAM) and SNR 10

effect on video quality. The measure of this value is between 0-10 (the higher values are truncated to 10).

Blurring is caused by removal of high frequency content from the original video signal. This metric is useful to see the variation in frame quality with changing quantization parameter of video codec or loss in enhancement bistreams in a video. The blur detection scheme is based on histogram computation of non-zero DCT coefficients [16].

Figure 2(a) shows a sample frame from test video while Figure 2(b) shows a sample frame with blurring while Figure 2(c) shows a sample frame with blocking.
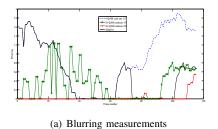
V. EXPERIMENTS

Measurement of blocking artifacts indicate rapid decline in performance with increasing modulation. With 4-QAM, almost all the frames experience zero-blocking, for 16-QAM more number of frames experience high blocking, and with 64-QAM almost all the frames experience high blocking.
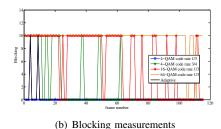
Figure 3 illustrates performance evaluation using 'rhino.avi' sample video clip (115 frames, 7.67 seconds at 15 fps, resolution: 320×240). The channel conditions were poor and SNR was set to 0 to illustrate the variations with modulation. The code rate for Turbo encoder was set to 1/3. Figure 3(a)

illustrates the effect of modulation change in blurring metric. The variations in blurring values are small, because blurring metric usually captures the subtle rate-distortions due to changing quanitzation parameter of the codec. However, the measurements of blocking artifacts (see Figure 3(a)) suggest a rapid decline in performance with increasing modulation.

In Figure 3(c), we plot the video reception delay (the delay between the first transmission of a frame and the passing of received frame to the video display by the LTE receiver) in poor channel conditions for different modulation rates. Note that this delay is cumulative i.e., once a frame experiences a certain delay, subsequent frames experience higher delays. Individual frame delays are limited by the maximum number of HARQ re-transmissions allowed. The slope of any delay curve represents the delay performance for the corresponding modulation scheme. We observe that higher modulation rates experience higher delays when the channel condition is poor. This is due to the higher packet loss, and subsequently higher number of HARQ re-transmissions involved when using higher modulation rates in poor channel conditions. Figures 3 and 4 indicate that when the channel conditions are poor, it is best to use low modulation and coding rates.

Figure 4 gives performance evaluation for the same scenario

(a) Blurring measurements

(b) Blocking measurements

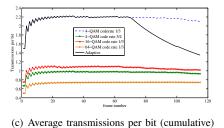(c) Average transmissions per bit (cumulative)

Fig. 6. LTE system end-end performance evaluation for changing network profile. The channel conditions are varying from poor (0dB) to good (25dB)

with changes in coderates, while keeping the modulation constant to 16-QAM. Code rates of 2/3 and 3/4 give poor performance, as evident by high blocking in these cases. Figure 4(c) indicates the large delay incurred due to retransmissions at MAC level (HARQ) at low code rates (3/4). The above mentioned results considered a scenario with low SNR and showed extreme results. Figure 5 shows the performance with different code rates when we have good channel conditions (SNR=10). It can be observed that different code rates follow the same trend in blocking and blurring values, for good channel conditions.

We observe that the blocking and delay measurements observed in Figure 5 are much smaller than what we observe in Figures 3 and 4. It can also be observed that different code rates follow the same trend in blocking and blurring values, for good channel conditions, suggesting that using a higher coding rate would be adequate to get good video reception in good channel conditions.

In Figure 6, we vary the channel SNR from poor to good scenario with time. This represents a situation where a user starts a video download in a bad channel condition and is moving towards better channel conditions with progress in video download. Specifically, we increase the SNR by 1 dB for every 4 frames transmitted. An adaptive profile tried to use the highest modulation and least code rate which can achieve least blocking. Figure 6 shows how an adaptive profile can reduce the blocking artifacts. For first 70 frames of the simulation, the adaptive profile chooses 4-QAM with code rate 1/3, and then shifts to 4-QAM with code rate 3/4 folllowed by 16-QAM for high SNR (above 100 frames). The advantage of such adaptation is illustrated in Figure 6(c) which measures the number of transmissions per bit, when HARQ is set to 2. For the case of high blocking, the value saturates to 2 times the original value (for all profiles). For 4-QAM at code rate 1/3, even before saturation (at low SNR) there is improvement in this value. For higher modulations, this value is saturated quickly and doesn't recover (as in case of 64-QAM). For high SNR, there is a slight dip in the curves. The adaptive profile changes the code rate and modulation with changes in SNR, thus shows significant improvement over 4-QAM at code rate 1/3 with similar image quality (in terms of blocking).

VI. CONCLUSION

We present a cross-layer approach to adaptive modulation and coding (AMC) in LTE scenario using no-reference block-

ing and blurring metrics. No-reference metrics can be used in practical deployment scenarios for link-adaptation based on blocking and blurring values, instead of using channel conditions (CQI - Channel Quality Index metric) as a reference. In our future works, we would like to make an in-depth analysis of AMC using these metric and their gain over using CQI values.

REFERENCES

[1] O. Oyman, J. Foerster, Y. joo Tcha, and S.-C. Lee, "Toward enhanced mobile video services over wimax and lte [wimax/lte update]," *Communications Magazine, IEEE*, vol. 48, no. 8, pp. 68 –76, 2010.
[2] S. Wenger, "H. 264/AVC Over IP," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 645–656, 2003.
[3] K. Nishida, I. Tanaka, and T. Koshimizu, "Basic SAE Management Technology for Realizing All-IP Network," *NTT DOCOMO Technical Journal*, vol. 11, no. 3, pp. 4–12, 2009.
[4] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on selected areas in communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
[5] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell labs technical journal*, vol. 1, no. 2, pp. 41–59, 1996.
[6] E. Sengul, E. Akay, and E. Ayanoglu, "Diversity analysis of single and multiple beamforming," *Communications, IEEE Transactions on*, vol. 54, no. 6, pp. 990–993, 2006.
[7] H. Bolcskei and A. Paulraj, "Space-frequency coded broadband OFDM systems," in *Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE*, vol. 1. IEEE, 2002, pp. 1–6.
[8] Z. Wang and E. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," in *Proc. of SPIE Human Vision and Electronic Imaging*, vol. 5666. Citeseer, 2005, pp. 149–159.
[9] G.-M. Muntean, P. Perry, and L. Murphy, "A new adaptive multimedia streaming system for all-ip multi-service networks," *Broadcasting, IEEE Transactions on*, vol. 50, no. 1, pp. 1 – 10, 2004.
[10] J. Klaue, B. Rathke, and A. Wolisz, "Evalvid–A framework for video transmission and quality evaluation," *Computer Performance*, pp. 255–272, 2003.
[11] T. A. Le, H. Nguyen, and H. Zhang, "Evalsvc: An evaluation platform for scalable video coding transmission," in *Consumer Electronics (ISCE), 2010 IEEE 14th International Symposium on*, 2010, pp. 1 –6.
[12] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes," *Communications Letters, IEEE*, vol. 1, no. 1, pp. 22–24, 2002.
[13] G. T. 36.212:, "Evolved universal terrestrial radio access (e-ultra); multiplexing and channel coding," 2008.
[14] G. T. 36.211:, "Evolved universal terrestrial radio access (e-ultra); physical channels and modulation," 2008.
[15] Z. Wang, A. Bovik, and B. Evan, "Blind measurement of blocking artifacts in images," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 3. IEEE, 2000, pp. 981–984.
[16] X. Marichal, W. Ma, and H. Zhang, "Blur determination in the compressed domain using DCT information," in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol. 2. IEEE, 1999, pp. 386–390.