# Joint Optimization of User-Experience and Energy-Efficiency in Wireless Multimedia Broadcast

Chetna Singhal, Swades De, Ramona Trestian, and Gabriel-Miro Muntean

**Abstract**—This paper presents a novel cross-layer optimization framework to improve the quality of user experience (QoE) and energy efficiency of the heterogeneous wireless multimedia broadcast receivers. This joint optimization is achieved by grouping the users based on their device capabilities and estimated channel conditions experienced by them and broadcasting adaptive content to these groups. The adaptive multimedia content is obtained by using scalable video coding (SVC) with optimal source encoding parameters resulted from an innovative cooperative game. Energy saving at user terminals results from using a layer-aware time slicing approach in the transmission stage. A trade-off between energy saving and QoE is observed, and is incorporated in the definition of a utility function of the players in the formulated heterogeneous user composition and physical channel aware game. An adaptive modulation and coding scheme is also optimally incorporated in order to maximize the reception quality of the broadcast receivers, while maximizing the network broadcast capacity. Compared to the conventional broadcast schemes, the proposed framework shows an appreciable improvement in QoE levels for all users, while achieving higher energy-savings for the energy constrained users.

**Index Terms**—Adaptive multimedia broadcast and multicast, scalable video coding, adaptive modulation and coding, heterogeneous users, energy saving, quality of user experience

## 1 Introduction

Rapid advancement in communication technologies in recent years, coupled with the availability of affordable high-end mobile computing devices, such as smartphones, tablets, personal digital assistants, small notebooks, have led to a significant growth in the number of consumers that access multimedia services from various types of devices, while on the move or stationary [1], [2].

The prevalent wireless technologies for multimedia broadcast include Long Term Evolution (LTE) using extended Multimedia Broadcast and Multicast Services (e-MBMS) interface specifications [3] [4], Worldwide Interoperability for Microwave Access (WiMAX) [3], and Digital Video Broadcast (DVB) [5] [6] [7]. Although the latest advances in many wireless network technologies, including broadcast (e.g. DVB-second generation terrestrial (DVB-T2), DVB-hand-held (DVB-H)), broadband (e.g. IEEE 802.11g, IEEE 802.11n [8]), and cellular (e.g. LTE), have enabled the operators to increase network capacity, the demands for popular multimedia content delivery to the mobile devices are growing even faster. Consequently, the overall user experience is still far from optimal, as the rich multimedia content puts pressure on the existing communication resources in terms of their bandwidth requirements and real-time constraints.

*C. Singhal is with Bharti School of Telecommunications, Indian Institute of Technology (IIT) Delhi, New Delhi, India. (e-mail: chetna.singhal@dbst.iitd.ac.in). S. De is with the Electrical Eng. Dept., IIT Delhi, New Delhi, India. (e-mail: swadesd@ee.iitd.ac.in). R. Trestian is with the Computer Communication Dept., Middlesex Univ., London, UK (e-mail: r.trestian@mdx.ac.uk). G.-M. Muntean is with the Performance Eng. Lab., School of Electronic Eng., Dublin City Univ., Ireland (e-mail: gabriel.muntean@dcu.ie).*

Thus, the challenge for the network operators include network resource optimization for popular multimedia content delivery, while ensuring uninterrupted and smooth services over wireless to a diverse customer population with varying degree of user-end constraints.

### 1.1 Motivation and proposed solution

In multimedia broadcast, one challenge is posed by user-end heterogeneity (e.g., different display size, processing capabilities, channel impairments). Another key component that consumers highly care about is the battery lifetime of their high-end mobile device. It is known that, real-time multimedia applications demand strict Quality of Service (QoS), but they are also very power-hungry.

Given the above user-end constraints, a service provider would look for maximizing the number of users served without affecting the Quality of user Experience (QoE). Clearly, attempting to receive a broadcast content irrespective of the device constraints is detrimental to battery resource efficiency, wherein the low-resolution mobile users suffer from redundant processing of high-end data that the device is not even able to use fully.

There have been a few recent studies that address receiver energy constraints [9], [10], display limitations and channel dynamics [11], [12], [13], [14], source and channel rate adaptation [15]. Yet to our best knowledge, a comprehensive look into the optimal broadcast strategy that jointly caters to both user-specific constraints and network dynamics is still missing.

This paper presents *a novel cross-layer optimization framework to improve both user QoE levels and energy efficiency* of wireless multimedia broadcast receivers with varying display and energy constraints. This solution combines user composition-aware source coding rate (SVC) optimization, optimum time slicing for layer coded transmission, and a cross-layer adaptive modulation and coding scheme (MCS).

### 1.2 Key features and findings

The main features of the proposed framework are as follows: 1) user grouping based on individual device capabilities and channel conditions; 2) formulation of a cooperative game to obtain user heterogeneity aware optimized SVC parameters that enable energy saving of the battery constrained users and at the same time maintain high QoE levels for high-end users; 3) optimizing layer-coded time slicing for energy saving and quality trade-off; 4) user heterogeneity and physical channel adaptive MCS allocation to the layered video content that maximizes network capacity.

The main findings of this work are: (a) The proposed user- and channel-aware grouping and cooperative game provide the users options to trade between quality of reception and energy conservation. (b) the usage of time slicing along with user heterogeneity and channel aware MCS significantly reduce energy consumption and increase QoE; the number of users served in the network with a guaranteed minimum quality level is increased.

Specifically, tests in different traffic scenarios reveal that, the proposed adaptive MCS offers about $16.6\%$ higher user serving capacity compared to fixed MCS or simple MCS schemes. With respect to only energy saving based optimization, the proposed joint energy and quality based cross-layer optimizations give about 43% higher video quality, while trading off only about 8% in energy saving and a marginal 0.62% in user serving capacity. Compared to only quality based optimization, the proposed scheme results in about 17% extra energy saving, 3.5% higher quality, and 10.8% higher capacity.

### 1.3 Paper organization

The rest of this paper is organized as follows: Section 2 discusses related works and Section 3 presents the technological details of the system and the proposed framework. This is followed by the analytic system performance model and optimizations in Section 4. Subsequently, Section 5 describes the simulation framework and Section 6 presents the key results of the proposed user-centric optimized multimedia broadcast scheme. Finally, the paper is concluded in Section 7.

## 2 Related Works

Hierarchical video coding [16] is an attractive solution that allows a user to dynamically adapt the video bitstream reception in dynamic wireless channel conditions. This technique encodes the stream into multiple progressively dependent layers. The most important layer is called *base layer* which typically provides an acceptable basic quality. The rest of the layers are known as *enhancement layers* which can be added to the base layer to improve the video quality. To this end, both ITU-T VCEG and ISO / IEC MPEG have standardized the SVC [17], [18] extension of H.264 / AVC [19], [20], [21]. The H.264 / SVC extension achieves a rate-distortion performance comparable to that of H.264 / AVC, where the same visual perceived quality is typically achieved with at most 10% higher bit rate [22].

DVB-H, an European Telecommunications Standards Institute (ETSI) standard [23], provides a built-in function that helps exploiting the video scalability features using Hierarchical Modulation [24] and is an efficient way to broadcast multimedia services over digital terrestrial networks to hand-held terminals. However, it considers transmission level details only, but not the user constraints or video encoding details.

[25] compared group management mechanisms in IP and MBMS models in UMTS networks, but did not discuss group formation criterion and user heterogeneity. An adaptive radio resource allocation scheme for multi-resolution multicast services in orthogonal frequency-division multiplexing (OFDM) systems was proposed in [26], which was shown to achieve an improved system throughput while maintaining fairness among all users. For energy-efficient streaming of scalable video over LTE using e-MBMS, grouping of users based on position and requested video quality was considered in [11]. Discontinuous reception (DRX) and energy saving at the user-end was not considered here; instead energy saving at the base station (BS) was targeted.

A cross-layer adaptive hierarchical video multicast solution in [27] considered jointly application, data link, and physical layers, where channel dependent Auto Rate Selection was proposed. To combat packet losses in multicast, a layered hybrid Automatic Repeat reQuest scheme was proposed in [28], where operating point for the multicast group was selected by a Nash bargaining game. The approach in [29] for video unicast / multicast over wireless proposed to minimize the resource usage while satisfying the diverse QoS requirements. The adaptive multicast in [30] maintains the highest sustainable transmission rate with suitable forward error correction (FEC) to maximize the received video quality. These approaches however did not address channel dependent SVC rate adaptation, MCS, and receiver constraints.

The approach in [9] proposed to enable the heterogeneous receivers render the appropriate sub-streams by time slicing technique in DVB-H for energy saving. This study derived the rate allocation to different layers from uniform, linear, or exponential distribution. But in actuality the rate of the layers depends on the encoding parameters (e.g. frame rate, quantization level, and spatial resolution). Also, the quality of received video and the effect of channel condition were not studied here.

A recent study [15] considered heterogeneous broadcast users, where an objective (temporal-spatial rate) distortion metric was used based on Principal Component Analysis distance between frames, and optimal layer broadcasting policy was obtained to maximize the utility. However, it did not consider channel adaptive scalability of SVC content, dynamic physical resource allocation, and energy saving at the receiver.

Adaptive modulation and coding (AMC) has been widely employed to effectively combat the channel dynamics and maximize physical layer data rate. In the context of video broadcast over wireless there are a few recent works (e.g., [12], [10], [13], [14]) which have used AMC in different forms and with different objectives.

The AMC approach in DVB-H applications in [12], which we call *simple MCS scheme*, decides on adaptation based on the broadcast receiver with an acceptable weakest signal strength and uses the same MCS for all SVC layers. In the AMC approach for DVB-H transmission [10], which we call *fixed MCS scheme*, different layers are assigned a predetermined fixed MCS. This scheme results in saving of power at both data reception and processing. However, in this work, the adaptation is merely on the basis of transmitted frame arrangement which is organized in terms of weaker and incremental codes; it does not incorporate video encoded data rates or the use of SVC to support heterogeneous users.

Unlike in DVB-H, transmission rate optimization in LTE MBMS is not based on time slicing. The adaptive MCS in [13] is in context of orthogonal frequency-division multiple access (OFDMA). The approach in [14] is also for LTE and WiMAX systems, where cooperative reception from multiple BSs is utilized following the Single Frequency Network principle. In all these AMC approaches, device limitations were not considered, thus the application layer encoding rate, and hence MCS is not affected by the heterogeneity of users in the network.

## 3 System Model

### 3.1 Overview of the system

A single-cell broadcast scenario is considered. Multimedia content delivery is done from the BS and managed jointly with a connected media server. The wireless user equipments (UEs) have varying display resolution and battery capabilities. Based on the users characteristics in the cell and their SNRs, the media server suitably encodes the source content in H.264 / SVC standard of DVB-H. The broadcast over the physical channel is OFDM-based. A UE, depending on its current status, may choose to receive all or part of the broadcast content (layers) by exploiting the time-sliced transmission feature of DVB-H. Fig. 1 illustrates a representative system, where $L$ layers and $\mathcal{T}$ user types are considered. For example, $L = 14$ in the standard 'Harbor' video sequence.

**Definition 1.** *UE type $\tau$ is characterized by the spatial resolution $\mathcal{R}$ of a UE display and battery power, which are device-specific. In a system with $\mathcal{T}$ types ($1 \le \tau \le \mathcal{T}$) of UE, $\mathcal{R}_i > \mathcal{R}_j$ if $j < i$. For example in a system with 3 types of users (i.e., $\mathcal{T} = 3$), $\mathcal{R}_3 > \mathcal{R}_2 > \mathcal{R}_1$.*

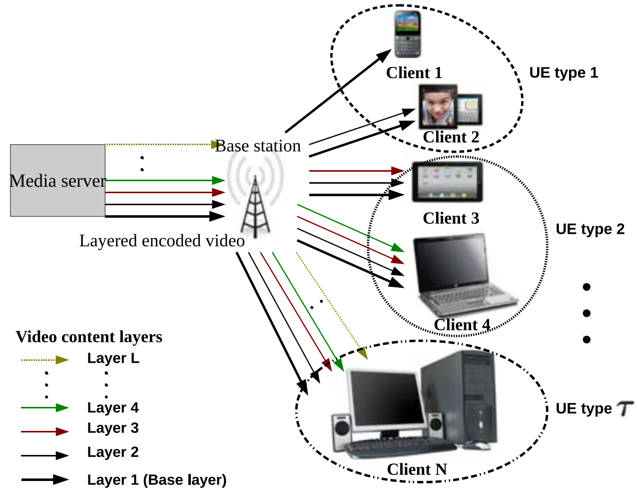



Fig. 1. An example of the DVB-H system, where $L$-layer SVC content and $\mathcal{T}$ types of UEs are considered.

SVC supports three types of scalability: spatial, temporal, and SNR-based. *Spatial scalability* is governed by display resolution of the UE (e.g., QCIF, CIF, D1), *temporal scalability* is related to the frame transmission rate (e.g., $1.875$ fps to $30$ fps), and *SNR scalability* is linked with the SVC coding rate as a function of the SNR experienced by the various UEs. A detailed overview of H.264 / AVC scalable video extension is given in [18].

In our study, the supported spatial resolutions considered are QCIF (quarter common intermediate format, with display resolution $176 \times 144$ pixels), CIF (common intermediate format, resolution $352 \times 288$ pixels), and D1 (D-1 digital recording video standard, resolution $704 \times 576$ pixels) formats, which serve three types of users (i.e., $\mathcal{T} = 3$). Apart from spatial resolution of the individual video frames, variable frame rate is also considered for the transmitted video.

**Definition 2.** *Layer $l$ $(1 \le l \le L)$ of a SVC content with a total of $L$ layers implicitly has its priority $\mathcal{P}_l$ in an inverse order with respect to the other layers, i.e., $\mathcal{P}_i > \mathcal{P}_j$, if $i < j$.*

If $L = 14$, following definition 2, $\mathcal{P}_1 > \mathcal{P}_2 > \cdots > \mathcal{P}_{14}$. If a type $i$ UE finds useful to display content up to the layer $l^{(i)}$ ( $\le L$), then $l^{(i)} < l^{(j)}$ for $i < j$.

SVC encoding generates different layers: base layer (layer 1) and enhancement layers. Layer 1 is the most important that needs to be received by all the UEs for the basic minimum quality. The other layers when received by a UE improve the reception quality by increasing the frame rate and / or resolution at the playback stage.

### 3.2 Proposed DVB-H system framework

The proposed overall system architecture is illustrated in Fig. 2. The server encapsulates the SVC encoded data in real-time transport protocol (RTP) format to IP packets and sends them to the BS. The BS comprises of the IP encapsulator, DVB-H modulator, and the radio transmitter.

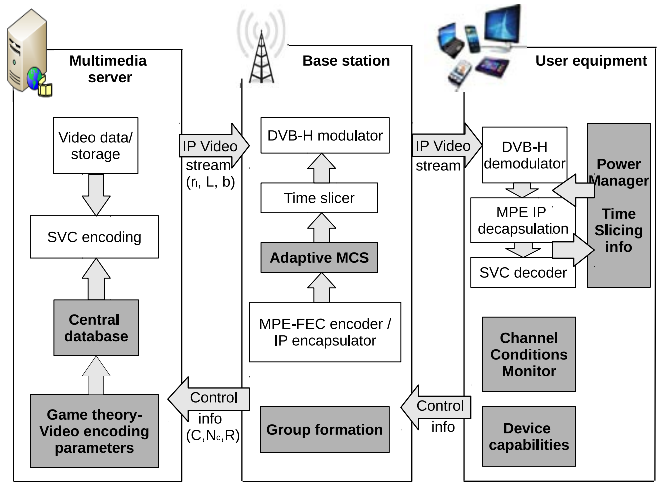



Fig. 2. The proposed DVB-H system architecture components.

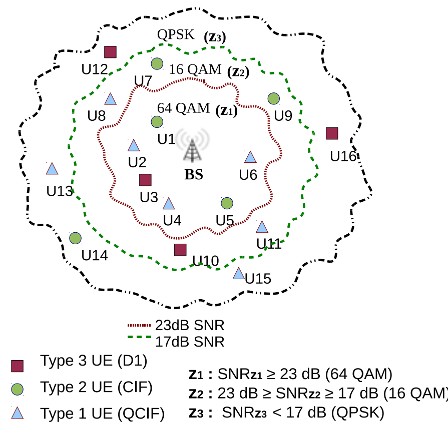



Fig. 3. Signal quality-based grouping example, with 3 UE types.

IP encapsulator puts the IP packets into multiprotocol encapsulation (MPE) frames and forms MPE-FEC for burst transmission as per the *time slicing scheme* (Section 4.2). The DVB-H modulator employs an *adaptive MCS selection* (Section 4.6) for the layered video content and sends it to the radio transmitter for broadcast.

The SVC encoding and MPE-FEC framing operations are inter-dependent and jointly optimized based on some underlying parameters (user, channel, and layer information). The optimized video encoding parameters are obtained through a game theoretic approach and stored in a central database. The UE and channel aware *user grouping* is discussed in Section 4.1, and *SVC parameter optimization game* is detailed in Section 4.5.

The UE informs its capabilities while subscribing to the broadcast service and also time-to-time updates its signal strength to the BS. It also has a power manager that helps to take advantage of the time slicing scheme and save energy based on its remaining power.

**Definition 3.** *A user class $c$ ($1 \leq c \leq C$) defines the capability of receiving the number of layers which is dictated by channel rate constraint experienced by the UE at a given instant of time. C is the total number of user classes.*

If a UE can receive up to $l_s$ useful layers, it belongs to class $c = l_s$. Thus, a user class is dynamically associated to a UE and is upper bounded by its resolution. If the number of useful layers of a type $i$ UE with resolution $\mathcal{R}_i$ is $l^{(i)}$, ($l^{(i)} \leq L$), then it can be in class $c$ such that $1 \leq c \leq l^{(i)}$. For a UE type $k$ with the highest resolution $\mathcal{R}_k$, $l^{(k)} = L$. In that case, $L = C$ and $1 \leq c \leq L$.

The parameters updated by the BS in the database are: $C$, the number of user classes; $N_c$, the number of users in class $c$ ($1 \leq c \leq C$); and $R$, the OFDM channel rate (expressed in bps). The parameters updated by the video server in the database are: $L$, the number of layers in the encoded SVC content; $r_l$, the rate of layer $l$ ($1 \leq l \leq L$); and $b$, the burst size of the base layer (measured in bis).

The proposed system performance optimization involves: i) grouping of users, ii) game theoretic formulation to obtain SVC encoding parameters, iii) time slicing at data-link level transmission, and iv) adaptive MCS allocation to the SVC layers. These are discussed next.

## 4 Performance Modeling and Optimization of the Proposed System

### 4.1 Grouping of users

User grouping is based on the respective UE resolution capabilities and received SNR. A UE capability is determined by the BS at the time of service subscription, when the UE sends its type information, i.e., the number of layers it wants to receive. The UE periodically updated its channel condition to the BS through the uplink channel.

**Definition 4.** *User group $g_z^{l^{(\tau)}}$ refers to the UEs of type $\tau$ ($1 \leq \tau \leq \mathcal{T}$) in zone $z (1 \leq z \leq \mathcal{Z})$ that have requested for $l^{(\tau)}$ layers. $\mathcal{Z}$ is the number of concentric zones around a BS.*

The coverage region of a BS is comprised of concentric zones, as shown in Fig. 3, with the SNR thresholds defining zone boundaries. For a SVC content with $L$ layers, $1 \leq l^{(\tau)} \leq L$, with QCIF resolution $l^{(\tau)} = 4$, and those with CIF and D1 resolutions are respectively 9 and 14. In the user-grouping example of Fig. 3, three UE types and three zones (based on three supported MCS levels in DVB-H [31]) are considered. The groups in this example are: $g_{z_1}^4 = \{U2, U4, U6\}$, $g_{z_1}^9 = \{U1, U5\}$, $g_{z_1}^{14} = \{U3\}$, $g_{z_2}^4 = \{U8, U11\}$, $g_{z_2}^9 = \{U7, U9\}$, $g_{z_2}^{14} = \{U10\}$, $g_{z_3}^4 = \{U13, U15\}$ , $g_{z_3}^9 = \{U14\}$ and $g_{z_3}^{14} = \{U12, U16\}$.

### 4.2 Time slicing as an energy saving measure

Time slicing approach allows discontinuous reception at the UEs, thereby facilitating the UE to turn-off the radio when not receiving data bursts and hence saving energy.

**Definition 5.** *Energy saving (ES) is calculated as the ratio of the time duration for which the UE's radio components are turned-off over the total time of the video transmission cycle.*

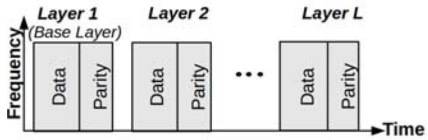



Fig. 4. Time slicing based DVB-H broadcast scheme.

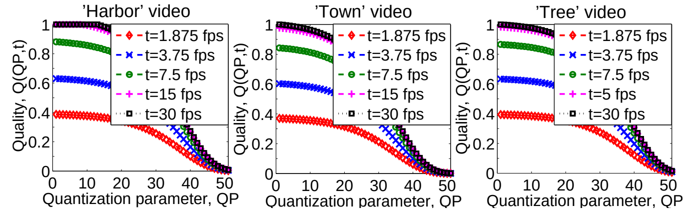



Fig. 5. Quality versus quantization parameter at different frame rates with the three standard video sequences, namely, 'Harbor', 'Town', and 'Tree'.

The broadcast channel rate is considered $R$ (bps). The multimedia content is encoded into $L$ layers. For decoding the layer $l$ ($1 \leq l \leq L$) the UE first needs to correctly receive and decode all layers $\grave{l}$, $1 \leq \grave{l} < l$. Video layer $l$ is allocated rate $r_l$ (bps), such that $\sum_{l=1}^{L} r_l \leq R$.

In time slicing-based layered broadcast, the UEs know apriori the specific layer constituted in a MPE-FEC frame (burst). As shown in Fig. 4, each layer corresponds to a different burst within the recurring window. This allows a UE to safely skip the bursts containing the layers that are irrelevant to it, and thereby save energy. Each MPE-FEC frame consists of two parts: Application Data Table that carries the IP packet, and an R-S (Reed-Solomon coding) Data Table that carries the parity bits.

Given a channel rate $R$ and base layer burst size $b$ bits, the burst size of layer $l$ is proportionally set to $b \cdot r_l/r_1$ bits. The recurring window size is the total burst size of all the layers, given as: $\sum_{l=1}^{L} b \cdot r_l/r_1 = \frac{b \cdot R}{r_1}$ bits. Hence with respect to starting time of the base layer burst, the start time of the layer $l$ burst is: $\frac{b \cdot \sum_{i=1}^{l-1} r_i/r_1}{R}$ sec.

If a user is currently in class $c$, the energy saving factor of that user at that time instant would be:

$$ES_c = 1 - \frac{\sum_{i=1}^{c} r_i}{R} - \frac{\mathcal{H} \cdot c \cdot r_1}{b}. \quad (1)$$

where, in general $1 \leq c \leq l^{(\tau)}$, for a type $\tau$ UE, $\mathcal{H}$ is the overhead duration (typically 100 ms [9]).

### 4.3 Video quality model

The video quality $Q(q,t)$ is a parametric function that best approximates the Mean Opinion Score (MOS). MOS is a subjective measure that indicates the user QoE level. MOS 5 refers to 'excellent' quality, 4 is 'good', 3 is fair, 2 is 'poor', and 1 is 'bad'. The parameters for the quality model are specific to a video based on its inherent features. The quality parametric model in [32] is specified with video specific parameters $\lambda$ and $g$. For a given spatial resolution, $Q(q,t)$ is a function of the quantization parameter $QP$ and frame rate $t$, as follows:

$$Q(q,t) = Q_{max} \cdot Q_{t_c}(t) \cdot Q_q(q), \quad \text{with} \quad (2)$$
$$Q_{t_c}(t) = \frac{1 - e^{(-\lambda \cdot t/t_{max})}}{1 - e^{-\lambda}},$$
$$Q_q(q) = \frac{e^{(-g \cdot q/q_{min})}}{e^{-g}}, \text{ and } q = 2^{(QP-4)/6}.$$

$Q_{max}$ is the maximum quality of the received video at the UE when it is encoded at minimum quantization level $q_{min}$ and at the highest frame rate $t_{max}$. To normalize, we consider $Q_{max}$ to be 100%.

To comprehensively study the video quality in the proposed system framework, we consider three representative video sequences: 'Harbor', 'Town', 'Tree', which cover a wide spatial and temporal perceptual information space [33]. In particular, the 'Harbor' video represents a sequence with sharp edges (high spatial variations) but having a relatively slow motion (low temporal variations), 'Town' has high spatial and temporal variations, whereas 'Tree' has low spatial and temporal variations in first half and high spatial and high temporal changes in the later half. Fig. 5 captures the effect on quality $Q$ of the three different video sequences at different $QP$. The trends of variation of $Q$ (which represents QoE) are observed to be quite similar in all these video sequences. Also, the plots indicate that the quality is a concave function of $QP$.

### 4.4 Energy saving versus quality trade-off

As noted in Section 4.2, the energy saving is a function of rate allocation to the layers. We now consider the scalability factors as parameters in the rate allocation at the source encoding stage. The parametric rate model, as in [34], [35], again is a function of the quantization level $q$, frame rate $t$, and spatial resolution $s$. The parameters $\theta$, $a$ and $d$, here are video specific.

$$R_c(q,t,s) = R_{max} \cdot R_{t_c}(t) \cdot R_q(q) \cdot R_s(s), \quad \text{with} \quad (3)$$
$$R_{t_c}(t) = \frac{1 - e^{(-\theta \cdot t/t_{max})}}{1 - e^{-\theta}}, \quad R_q(q) = e^{a \cdot (1-q)/q_{min}},$$
$$\text{and} \quad R_s(s) = \left(\frac{s}{s_{max}}\right)^d, \ d < 1.$$

Here, $R_{max}$ is the maximum bit rate of the video sequence with minimum quantization level $q_{min}$, maximum frame rate $t_{max}$, and maximum spatial resolution $s_{max}$. By using these rate parametric model equations for energy saving analysis (i.e., in (1)), the energy saving for class $c$ users of type $\tau$ ($1 \leq c \leq l^{(\tau)}$) is given as:

$$ES_c = 1 - \frac{R_c(q,t,s)}{R} - \frac{\mathcal{H} \cdot c \cdot R_1(q, t_{min}, s_{min})}{b}. \quad (4)$$

To study the impacts of SVC quantization parameter and time slicing scheme on the energy saving, we again consider the three representative video sequences: 'Harbor', 'Town', and 'Tree'. Fig. 6(a) shows the variation of normalized average energy saving with the change in

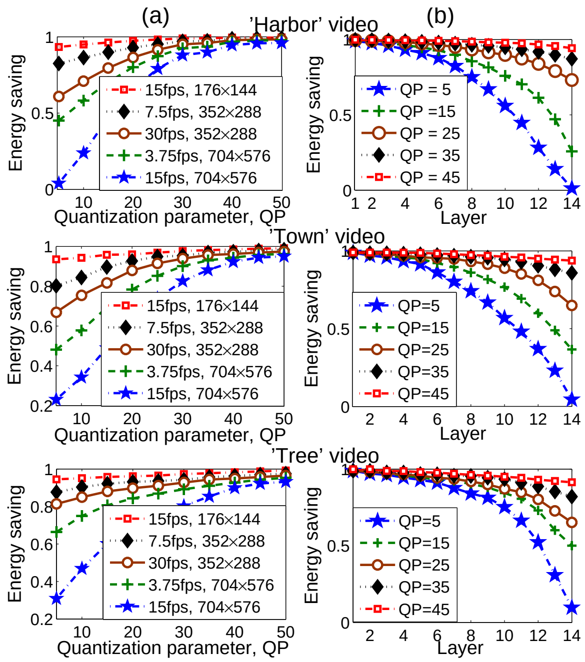



Fig. 6. Energy saving performance at different SVC quantizations and time slicing schemes for the three video sequences: 'Harbor', 'Town', and 'Tree': (a) effect of quantization parameter $QP$; (b) effect of number of layers transmitted.

quantization parameter $QP$. It is notable that, in all the three cases, an increase in $QP$ results in a higher energy saving. Also a decrease in $t$ (by DRX mechanism) and smaller spatial resolution of the video sequence results in more energy saving for the UEs. It is also observed that the nature of the energy saving is concave with respect to $QP$. For the three video sequences, Fig. 6(b) shows the normalized energy saving with respect to the layers received by the UEs. Also, a higher $QP$ (i.e., higher $q$ and hence a lower allocated rate) corresponds to a higher energy saving. Hence, there is a clearly evident trade-off between the energy saving and quality for a specific value of quantization level $q$. Also different type of users have different energy savings and QoE requirements.

It is observed from Figs. 5 and 6 that, the quality and energy saving performances of the three representative test sequences ('Harbor', 'Tree', 'Town') follow similar trends with respect to the variations of quantization parameter, frame rate, spatial resolution, and the number of layers transmitted. Thus, the proposed framework in the paper and the optimizations (discussed subsequently in this section) should generically hold true for any possible SVC video sequences. Therefore, our remaining performance results are discussed with respect to only one representative test sequence ('Harbor' video).

### 4.5 Energy saving and quality optimization game

Based on the energy saving and quality trade-off that depends on the quantization level $q$, we now formulate a cooperative game to obtain the optimal video encoding parameters. Note that, the development in Section 4.4 demonstrates the possibility of an optimal SVC encoding from the individual user's perspective. However it does not provide an insight to the encoding optimality for broadcast when there are different user class in different proportions. Here, we address this optimization aspect.

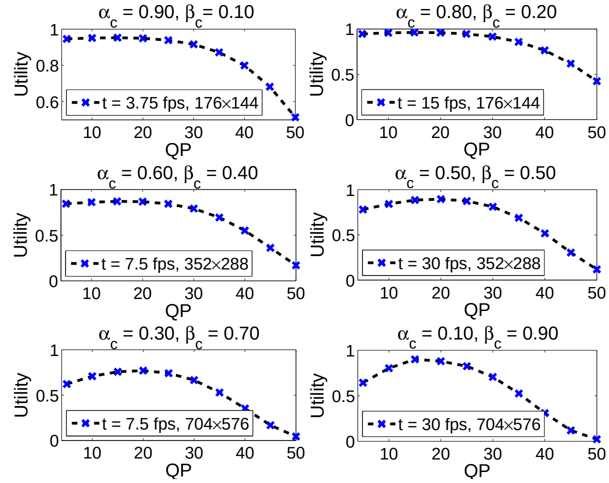



Fig. 7. Examples of utility plots of individual class of users for the standard 'Harbor' video sequence.

This optimization game jointly accounts for the users of different classes (Definition 3) as well as the fraction of users in each class. The game is defined below:

**Players:** Class $c$ comprising of a set of users who can be served up to $l = c$ layers, where $1 \leq c \leq l^{(\tau)}$, $\tau = 1, 2, 3$. (Recall that, $c$ is dynamic, computed at the BS, depending on the UE type $\tau$ and their individual SNRs.)

**Strategy:** Quantization level $q$ used by the SVC encoder for encoding the source video. Optimum $q$ determines the rate distribution (i.e. the minimum bit rate) $r_l$ for the different layers $l$ of the SVC content, that satisfy the users' ES and quality requirements.

**Utility:** For class $c$ the utility is defined as: $u_c = (ES_c(q,t))^{\alpha_c} \cdot (Q_c(q,t))^{\beta_c}$, where $\alpha_c$, $\beta_c$ are the parameters for a particular class of users based on their emphasis on energy saving or quality, with $\alpha_c + \beta_c = 1$. The higher the $\alpha_c$ value is, the higher is the emphasis on energy saving by the users in that class. On the other hand, the higher the value of $\beta_c$ is, the more will be the emphasis on receiving higher quality video. Here, for class $c$, energy saving $ES_c(q,t)$ is given in (4) and the quality value $Q_c(q,t)$ is given in (2).

We use multiplicative exponent weighting (MEW) in defining the utility $u_c$ instead of simple additive weighting (SAW), because in SAW based optimality poor value of a parameter can be outweighed by a very good value of another parameter. Instead, MEW penalizes alternatives with poor parameter values more heavily. For example, if energy utility is near zero (which means the UE consumes a lot of energy), the MEW based utility function avoids selecting this because it is multiplicative, whereas the SAW utility may end up choosing this case of near-zero energy but very high quality.

In Fig. 7, some examples of utility function with $QP$ variation are shown. Quantization level $q$ is can be obtained from $QP$ (see (2)). The ($\alpha_c$, $\beta_c$) combination shown are the optimum values that achieve the maximum possible utility for the individual user class $c$. The plots indicate that for each class, the considered utility function is concave in nature in terms of the $QP$.

Since the number of users in a class impacts the overall system utility, we define a modified utility function. If there are $N_c$ users in class $c$, the modified utility is:

$$\grave{u}_c = N_c(ES_c(q,t))^{\alpha_c} \cdot (Q_c(q,t))^{\beta_c} \tag{5}$$

The objective is to maximize the total average utility for the system,

$$U_{total} = \max \left( \frac{\sum_{\tau=1}^{\mathcal{T}} \sum_{c=1}^{l^{(\tau)}} \grave{u}_c}{\sum_{\tau=1}^{\mathcal{T}} \sum_{c=1}^{l^{(\tau)}} N_c} \right) \tag{6}$$

Before we proceed further, we prove the concavity of the utility functions in (5) and (6).

**Proposition 1.** *The utility function of a class $c$, defined in* (5)*, and the system utility, defined in* (6)*, are strictly concave functions of $QP$ in the range $[QP_{min}, QP_{max}]$.*

**Proof**: Given two functions $f_1(x)$ and $f_2(x)$, a function $\phi(x) = f_1(x) \cdot f_2(x)$ is said to be strictly concave and has a unique maxima in $[x_{min}, x_{max}]$ if the following conditions hold [36]:

(1) $f_1''(x) < 0$ and $f_2''(x) < 0$, i.e., $f_1(x)$ and $f_2(x)$ are concave functions of $x \in [x_{min}, x_{max}]$, (2) $f_1(x), f_2(x)$ are non-negative, and (3) $f_1'(x) \cdot f_2'(x) < 0$ in $[x_{min}, x_{max}]$.

Since $\phi''(x) = f_1''(x) \cdot f_2(x) + 2 \cdot f_1'(x) \cdot f_2'(x) + f_1(x) \cdot f_2''(x)$, by the above conditions $\phi''(x)$ is negative in $[x_{min}, x_{max}]$. Hence it is concave down with a maxima at $k \in [x_{min}, x_{max}],\ s.t.\ \phi'(k) = 0$.

In our context the two functions are: $ES_c(q,t)^{\alpha_c}$ and $Q_c(q,t)^{\beta_c}$. Since, the proof is generic and holds true $\forall$ $c \in [1, l^{(\tau)}]$, and $t$, $s$ are constant values for any class $c$, the variable over which the optimization is carried out is $QP$, which is related to $q$ by (2). Thus the two functions can be written as $f_1(x) = ES_c(QP)^{\alpha_c}$ and $f_2(x) = Q_c(QP)^{\beta_c}$, and the interval of concavity is $[QP_{min}, QP_{max}]$. We want to show the concavity of the utility function and joint optimization of the system in terms the best suited QP for the video encoding.

Firstly, we prove $Q_c(QP)^{\beta_c}$ is concave in $[QP_{min}, QP_{max}]$. The proof is as follows:

$Q_c(QP)^{\beta_c} = (Q_q(q) \cdot Q_{max} \cdot Q_{t_c})^{\beta_c}$, where $QP \in [QP_{min}, QP_{max}]$.

Then, from (2):

$$Q_c(QP)^{\beta_c} = \left( \frac{e^{(-g \cdot 2^{(QP-4)/6})/2^{(QP_{min}-4)/6}}}{e^{-g}} \cdot D \right)^{\beta_c}, \text{ where}$$

$D$ is a constant with respect to the variable $QP$ and is given as $D = Q_{max} \cdot Q_{t_c}(t)$.

For obtaining the derivative, denote $QP = x$, $QP_{min} = x_{min}$, $QP_{max} = x_{max}$. Also let $w = 2^{((x-4)/6)}/2^{((x_{min}-4)/6)}$. Then we have:

$$\frac{dQ_c(x)^{\beta_c}}{dx} = \frac{dQ_c(w)^{\beta_c}}{dw} \cdot \frac{dw}{dx}, \text{ where} \tag{7}$$

$$Q_c(w)^{\beta_c} = \left( \frac{e^{-g \cdot w}}{e^{-g}} \cdot D \right)^{\beta_c},$$

$$\frac{dQ_c(w)^{\beta_c}}{dw} = (-\beta_c \cdot g) \cdot e^{-\beta_c \cdot g \cdot (w-1)} \cdot D, \text{ and } \frac{dw}{dx} = \frac{w}{6}.$$

It is evident from (7) that, for $x_{min} = QP_{min} = 1$, $x_{max} = QP_{max} = 51$:

$$\frac{dQ_c(x)^{\beta_c}}{dx} < 0, \forall\ x \in [x_{min}, x_{max}] \tag{8}$$

Differentiation (7) again with respect to $x$ we have,

$$\frac{d^2 Q_c(x)^{\beta_c}}{dx^2} = \left( \beta_c \cdot g - \frac{1}{6} \right) \times \Big( (\beta_c \cdot g) \cdot e^{-\beta_c \cdot g \cdot (w-1)} \cdot D \cdot 2^{((x-4)/6)}/(6 \cdot 2^{((x_{min}-4)/6)}) \Big) \tag{9}$$

Since $\beta_c \cdot g < \frac{1}{6}$ ($\beta_c(\max) = 1$, $g = 0.06$), from (9), $\frac{d^2 Q_c(x)^{\beta_c}}{dx^2} < 0$. Thus $Q_c(QP)^{\beta_c}$ is concave in $[QP_{min}, QP_{max}]$.

We now prove that $ES(QP)^{\alpha_c}$ is concave in $[QP_{min}, QP_{max}]$.

From (4), $ES_c(QP)^{\alpha_c} = \left( 1 - \frac{R_c(QP)}{R} - \frac{\mathcal{H} \cdot c \cdot R_1(QP)}{b} \right)^{\alpha_c}$, where $QP \in [QP_{min}, QP_{max}]$.

From (3), it implies that $ES_c(QP)^{\alpha_c} = \left( (1 - e^{a(1-2^{(QP-4)/6})/(2^{(QP_{min}-4)/6})}) \cdot P \right)^{\alpha_c}$, where $P$ is a constant with respect to variable $QP$. Using (3), $P$ is obtained as: $P = \left( \frac{R_{max} \cdot R_{t_c}(t) \cdot R_{s_c}(s)}{R} \right) - \left( \frac{\mathcal{H} \cdot R_{max} \cdot R_{t_c}(t_{min}) \cdot R_s(s_{min})}{b} \right)$.

Again, for the derivative we denote $QP = x$, $QP_{min} = x_{min}$, $QP_{max} = x_{max}$. Let $v = e^{a(1-2^{(x-4)/6)})/2^{(x_{min}-4)/6}} \cdot P$. By simplifying we have,

$$\frac{dES_c(x)^{\alpha_c}}{dx} = \frac{dES_c(v)^{\alpha_c}}{dv} \cdot \frac{dv}{dx}, \text{ where} \tag{10}$$

$$ES_c(v)^{\alpha_c} = (1-v)^{\alpha_c} \text{ and}$$

$$\frac{dv}{dx} = -\frac{a e^{a(1-2^{(x-4)/6)})/2^{(x_{min}-4)/6}}}{(2^{(x_{min}-4)/6})} \cdot P \cdot \frac{2^{(x-4)/6}}{6} \cdot \ln 2.$$

Since $v < 1\ \forall\ x \in [x_{min}, x_{max}]$, $\frac{dv}{dx} < 0$ and $\frac{dES_c(v)^{\alpha_c}}{dv} < 0$. This implies that,

$$\frac{dES_c(x)^{\alpha_c}}{dx} > 0,\ \forall\ x \in [x_{min}, x_{max}]. \tag{11}$$

On similar lines as in (7)-(9), it is observed that $\frac{d^2}{dx^2}(ES_c(x)^{\alpha_c}) < 0,\ \forall\ x \in [x_{min}, x_{max}]$. This implies that $ES(QP)^{\alpha_c}$ is concave in $[QP_{min}, QP_{max}]$.

Thus, Condition (1) is shown to be true. Condition (2) holds in the proposed scheme as per the basic design of the system, since $ES_c(QP)^{\alpha_c}$ and $Q_c(QP)^{\beta_c}$ are always positive. From (8) and (11), the product $ES_c(QP)^{\alpha_c} \cdot Q_c(QP)^{\beta_c}$ is negative. So Condition (3) also holds true. Hence, the utility of class $c$, $u_c$ is proven to be strictly concave and has a maxima in the given range.

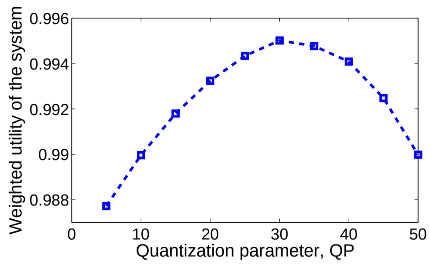



Fig. 8. Total average system utility with the 'Harbor' sequence.

Note that, although the exact value of the maxima for a class of utility function and the corresponding value of $QP$ can be easily obtained from the above development, these are not of our interest here.

As shown in [37], [38], a non-negative linear combination of strictly concave functions is also strictly concave. Since the system utility in (6) is a non-negative linear combination of the utilities of all the individual classes that are already proven to be strictly concave, the system utility in (6) is a strictly concave in the given range. This implies the existence of a unique solution that maximizes the utility in the joint optimization formulation. □

In terms of algorithmic complexity, the joint energy-saving and video quality optimization has a complexity of O($l^{\mathcal{T}}$), where $l^{\mathcal{T}}$ is the number of SVC layers being broadcast for the highest resolution type $\mathcal{T}$.

Fig. 8 shows that, the sum total weighted average utility for all classes is a concave function of $QP$ with a unique maxima. This plot corresponds to 60% type 1 users, 30% type 2 users, and 10% type 3 users in the system, with their corresponding random location dependent SNR feedbacks accounted at the BS to determine the user classes $c$ and the corresponding $N_c$ values. (Different traffic scenarios and the SNR-channel rate relationships are given in Tables 1 and 2, respectively.) The maxima of this scenario corresponds to $QP = 30$.

### 4.6 Adaptive modulation and coding scheme

As noted in Section 3, in our approach, besides user-and-channel aware SVC rate optimization at the application layer and time slicing at the link layer, at the physical layer adaptive MCS is applied which is optimized for enhanced energy efficiency and network capacity. Clearly, this adaptation is a function of the heterogeneous users composition in a cell and the dynamic physical channel rate constraint. Physical channel dynamics is accounted in a slow (shadow fading) scale to avoid high bandwidth overhead of frequent channel state feedback and computation of coding and MCS optimizations at the BS as well as the video server.

The total number of users in the cell that are subscribed to the broadcast service is taken to be $N$. The different MCSs are considered to be $m = 1, \cdots, M$ (for example, $m = 1$ represents QPSK with code rate = 1/2, and so on). The SVC encoded video is considered to have $L$ layers. In our formulation, $R_m$ represents the data rate provided by MCS $m$, $r_l$ is the rate allocated for layer $l$ ($l = 1, \cdots, L$). If a layer $l$, can be served by the BS to the users, then we set $l_{served} = 1$, else it is set to 0. We have used an indicator function $\chi_{l,m}$ that takes a value 1 if layer $l$ is modulated with MCS $m$ and takes a value 0 otherwise. The value of $m_l$, $l = (1, \cdots, L)$ specifies the MCS used for layer $l$ subject to $l_{served} = 1$. For a user to be able to decode any layer, it is necessary to have received all the layers lower than the current layer. Only then the layer is said to be valid for the user. The number of valid layers for any user is denoted by:

$$l_{s_j} = \max\{l \mid \forall\ \grave{l} \leq l \leq l^{(\tau_j)}, \sum_{\grave{m}=1}^{M_j} \chi_{\grave{l},\grave{m}} = 1\}.$$

$l_{s_j}$ is the maximum number of continuous layers modulated with 1 to $M_j$ starting from base layer, where $M_j$ is the highest possible modulation level that $j^{th}$ user can receive, such that these layers are either equal to or less than the requested number of layers by the user ($l^{(\tau_j)}$) based on its type $\tau_j$, $1 \leq \tau_j \leq \mathcal{T}$. The received rate for user $j$ is given by: $r_{\sum_j} = \sum_{l=1}^{l_{s_j}} r_l$.

The utility for the user $j$ is defined as a general function of its received rates, requested quality rates and maximum possible feasible received rate based on its channel conditions, i.e. SNR with shadowing at any given time: $U_j(r_{\sum_j}, r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}})$. Here, $r_{\sum}^{l^{(\tau_j)}}$ corresponds to the rate requested by the user for its desired maximum quality level. So it is based on the maximum number of layers $l^{(\tau_j)}$ of the SVC content requested by the user of type $\tau_j$. $r_{\sum_{j(SNR)}}$ corresponds to the maximum rate that the user would be able to receive if the user alone was present in the network and optimization of the MCS was to be just based on this user's channel conditions (i.e., its experienced SNR).

The user $j$'s utility is defined as:

$$U_j(r_{\sum_j}, r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}}) = Q(r_{\sum_j}) - Q(\min\{r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}}\}) \tag{12}$$

where $Q(r_{\sum_j})$ is the quality value based on the parametric model given in (2). Since the possible rates that a user $j$ can receive is from a set of possible layer rates, if a user is able to receive $l_{s_j}$ layers, such that $r_{\sum_j} \geq r_{l_{s_j}}$, with layer $l_{s_j}$ having a frame rate $t_{l_{s_j}}$ and a quantization level $q_{l_{s_j}}$, then, $Q(r_{\sum_j}) = Q(q_{l_{s_j}}, t_{l_{s_j}})$.

The objective is to maximize the total system utility, i.e., $\max\{\sum_{j \in N} U_j(r_{\sum_j}, r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}})\}$ subject to (5), (6), and the following constraints:

$$\begin{aligned} & r_{\sum_j} \geq r_1, \text{ given that, } r_{\sum_{j(SNR)}} \geq r_1 \quad \forall j \in N \\ & \sum_{m=1}^{M} \chi_{l,m} \leq 1, \quad l = 1, \cdots, L \end{aligned} \tag{13}$$

The first set of constraints mentioned in (13) states that for every user $j \in N$ the rate received should be

**Algorithm 1** Adaptive MCS selection for SVC layers

**Input**: $L,\ \gamma_m,\ R_m,\ r_l,\ r_{\sum_{j(SNR)}},\ r_{\sum}^{l^{(\tau_j)}},\ \forall\ m = 1,\cdots,M,\ l = 1,\cdots,L,$ and $j = 1,\cdots,N$

**Pseudocode**:

1) **Initialize variables:** $U_{total} = 0$ and $\chi_{l,m} = 0, \forall\ l = 1,\cdots,L, \forall\ m = 1,\cdots,M$
2) **for each** $l = 1$ to $L$
   - **if** $R_M < r_l$ **then** Set $l_{served} = 0$
   - **else**
     - **for each** $i = 1$ to $M$
       - **if** $r_l < R_i$ and $m_{l-1} \leq i$ **then**
         - Set $m_l = i$, $\chi_{l,i} = 1$, and $l_{served} = 1$
         - **go to** 3
3) **for each** user $j = 1$ to $N$
   - Using $m_l(l = 1,\cdots,l^{(\tau_j)})$ and $l_{served}$, find $r_{\sum_j}$
   - Compute $U_j(r_{\sum_j}, r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}})$ using (12)
   - $U_{total} = U_{total} + U_j(r_{\sum_j}, r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}})$

**Output**: $\chi_{l,m},\ m_l,\ l_{served},\ \forall\ l = 1,\cdots,L$ and $m = 1,\cdots,M$.

at least greater than or equal to the rate of the base layer $r_1$, for the condition that $r_{\sum_{j(SNR)}} > r_1$, i.e. the channel condition of the user $j$ supports a rate greater than that required for the base layer. The second set of constraint in (13) uses integer relaxation, which states that for a given video layer $l$, it can be modulated and coded with at most one MCS. It is important to note that a layer $l$ may not even be modulated with any MCS, i.e., $\sum_{m=1}^{M} \chi_{l,m} = 0$. In such a case the layer $l$ is not transmitted. The users experiencing extremely bad channel conditions with $r_{\sum_{j(SNR)}} < r_1$ will not be able to receive any layer, since they are experiencing the SNR below the minimum SNR threshold for the most basic MCS (e.g., QPSK with code rate of 1/2).

The different supported MCS have a minimum SNR threshold ($\gamma_m$, for MCS $m = 1,\cdots,M$) under the given DVB-H standard specifications, based on the quasi error free reception and MPE-FEC error rate of 5% with a BER value of $10^{-12}$ [31]. The rates corresponding to SNR threshold of different MCS are given by:

$$\begin{aligned} &R_m = B.\log_2(1+\gamma_m), \quad 1 \leq m \leq M, \\ &\gamma_{\grave{m}} > \gamma_m, \text{ where MCS } \grave{m} > m, \quad 1 < \grave{m} \leq M \\ &\text{Hence, } R_{\grave{m}} > R_m \quad \forall \quad m, \grave{m} \in [1, M], \grave{m} > m. \end{aligned} \tag{14}$$

The proposed MCS assignment algorithm is summarized in Algorithm 1.

The adaptive MCS algorithm has O($L \cdot M + N$) complexity, where $L$ is the number of SVC layers broadcast, $M$ is the number of MCS levels, and $N$ is the total number of users. The proposed approach ensures that, with optimal MCS allocation for all the SVC layers, $Q(r_{\sum_j}) \geq Q(\min\{r_{\sum}^{l^{(\tau_j)}}, r_{\sum_{j(SNR)}}\})$, $\forall\ j \in [1, N]$, leading to maximum system utility $U_{total}$.

### 4.7 Video reception quality measure

For a fair comparison of the quality of reception performance of the different competitive strategies, we define a video reception quality measure.

**Definition 6.** *In a system having $\mathcal{T}$ types of users, with the highest number of layers $l^{(\tau)}$ that a type $\tau$ user $(1 \leq \tau \leq \mathcal{T})$ is capable of receiving and the corresponding reception quality denoted by $Q\left(l^{(\tau)}\right)$, the* weighted average video reception quality, *or the $\mathcal{Q}$ measure is expressed as:*

$$\mathcal{Q} = 1 - \frac{1}{\sum\limits_{\tau=1}^{\mathcal{T}} \sum\limits_{l_s=0}^{l^{(\tau)}} N_{l_s}^{(\tau)}} \sum_{\tau=1}^{\mathcal{T}} \sum_{l_s=0}^{l^{(\tau)}} \left[Q\left(l^{(\tau)}\right) - Q\left(l_s\right)\right] N_{l_s}^{(\tau)} \tag{15}$$

*where $N_{l_s}^{(\tau)}$ is the number of type $\tau$ users actually receiving $l_s$ number of layers, with the corresponding quality measure $Q\left(l_s\right)$. $Q\left(l_s\right) = 0$ if $l_s = 0$, i.e. when no layers are received.*

$Q(l^{(\tau)})$ and $Q(l_s)$ are obtained based on the parametric model in (2), as a function of quantization level $q$ and frame rate $t$ ($Q\left(l_s\right) = Q_{l_s}(q,t)$, i.e., the quality value corresponding to layer $l_s$ ($1 \leq l_s \leq l^{(\tau)}$) of SVC content).

With respect to a given UE type, $\mathcal{Q}$ is indicative of the difference in actually experienced video reception quality with respect to its highest reception capability. It is a performance metric that indicates the QoE of the broadcast users in a given heterogeneous users distribution. Being a weighted average, it also indicates the proportion of total number of users that are served with a specified video quality level. Hence, a higher value of $\mathcal{Q}$ measure signifies that a higher proportion of total number of users are being served in the cell with a higher video reception quality. In our system example, $\mathcal{T} = 3$, and $l^{(\tau)} = 4, 9, 14$ respectively, for $\tau = 1, 2, 3$.

Further, it may be noted that the parametric quality measure $Q(q,t)$and hence the weighted average quality measure $\mathcal{Q}$, that we use to characterize the transmission strategies, have a direct relationship with the subjective measure MOS [32], given as: MOS $= 4 \times Q(q,t) + 1$. Thus, numerically, $Q(q,t) = 0$ corresponds to MOS = 1, $Q(q,t) = (0.0 - 0.25]$ corresponds to MOS = 2, $Q(q,t) = (0.25 - 0.5]$ corresponds to MOS = 3, $Q(q,t) = (0.5 - 0.75]$ corresponds to MOS = 4, and $Q(q,t) = (0.75 - 1.0]$ corresponds to MOS = 5. This mapping between QoE measure of video quality (MOS) and parametric video quality $Q(q,t)$ is shown in Fig. 9.

## 5 Simulation Settings

For the simulation purpose and in order to encode the SVC streams, we have used the SVC encoder reference software JSVM_9_19_12 [39]. In the considered scenario, scalable video covers three levels of spatial resolution formats: QCIF, CIF, and D1, serving three type of users, and five temporal resolutions: 1.875, 3.75, 7.5, 15, 30 fps (cf. Fig. 10), which serve the users in variable channel conditions. The sample ‘Harbor’ video sequence with

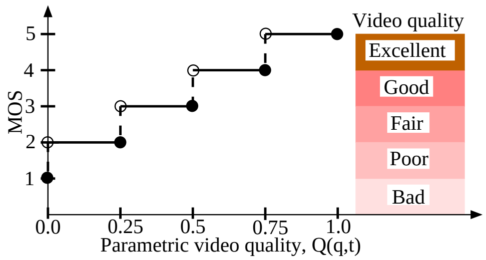



Fig. 9. Mapping between MOS and parametric video quality.

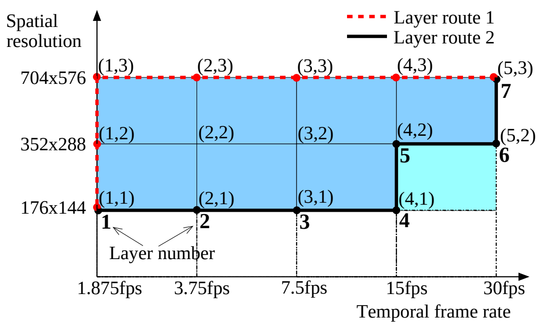



Fig. 10. Spatial-temporal scalable layer structure used in system simulations.

TABLE 1
Simulation scenarios, with variable ratios of user type

| Scenario | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Type 1 | 33.3% | 10% | 45% | 60% | 90% | 5% |
| Type 2 | 33.3% | 30% | 10% | 30% | 5% | 5% |
| Type 3 | 33.3% | 60% | 45% | 10% | 5% | 90% |

300 frames was taken for evaluating the proposed framework. For this video, the parameters $\lambda$, $g$, $\theta$, $a$, and $d$ in (2)-(3) are respectively found to be 7.38, 0.06, 1.429, 1.551, and 0.845. Fig. 10 shows the flexible layer structure with each coordinate representing different spatial and temporal resolutions. Note that there are two possible layer routes for the hierarchical broadcast reception.

We consider a single-cell video broadcast network with 500 randomly distributed users belonging to three user types. Six simulation scenarios are considered with different ratios of user type distributions as listed in Table 1. The users belonging to type 1 require QCIF format video, those belonging to type 2 require CIF format, and the ones of type 3 need D1 format.

The first step is to obtain the optimized SVC encoding parameters as a function of the user types distribution in the system using the proposed cooperative game. The outcome of the game is the optimized adaptive video coded sequence with the optimal rate allocation for each layer, such that it aids the energy saving for type 1 users and the quality for type 3 users.

The rate allocation is followed by the adaptive MCS allocation for the different SVC layers that are transmitted in a time-sliced arrangement. Note that, since the optimal rate allocation to different SVC layers $r_l$, $1 \leq l \leq L$ is a function of the user type distribution ratios, the time-slice allocation in (1) is also accordingly a function of user type distribution ratios. The adaptive MCS in our approach is additionally governed by the SNR experienced by the various user groups.

TABLE 3
Simulation parameters

| Parameter | Value |
|---|---|
| OFDM channel bandwidth | 8MHz |
| Frequency | 800MHz |
| Carrier spacing | 4KHz |
| DVB-H transmission mode | 2K |
| Number of data carriers | 1705 |
| Receiver noise figure | 5.2dB |
| Transmitter output power | 63.8dBm |
| Transmitter cable and connector loss | 3.0dB |
| Transmitter power splitter loss | 3.0dB |
| Transmitter antenna gain | 13.1dBi |
| Receiver antenna gain | −7.3dBi |
| Building loss | 14.0dB |
| Location variation loss for 95% area probability | 5.3dB |
| Receiver noise input power | −99dBm |
| Shadowing standard deviation | 8dB |
| Guard interval | 1/4 |
| Wireless channel model | Gaussian |
| Shadowing model | Log-normal |
| Path loss model | Free space |

As per the DVB-H specifications [23], the minimum SNR threshold for each MCS for a given wireless channel and the corresponding channel rates with the relevant guard interval (GI $= 1/4$) are listed in Table 2. The overall system simulation parameters considered are listed in Table 3. The performance results are presented below.

## 6 Results and Discussions

### 6.1 Energy-quality trade-off performance with time slicing technique

Considering the different traffic scenarios as listed in Table 1, we first analyze the energy–SVC quantization-dependent quality trade-off. Fig. 11, illustrates the energy-quality ($Q$, given in (2)) trade-off for the three types of users. It can be seen that, when the video encoding parameters are so chosen that the quality of the video at the UEs is higher, the corresponding energy saving for the users of all the three types is lower. However, under all the six scenarios the energy saving for the type 1 users is the highest among the three type of users. This is primarily due to the time slicing approach of transmission. Considerable variation in the energy saving and quality values is evident when there is a remarkable change in proportion of any particular type of user in the network. For instance, in scenario 5 with 90% of type 1 users, the joint optimization approach results in energy saving of more than 90% for the UEs, with approximately 20% quality. This is because, more than 90% users are energy-constrained and the objective is to satisfy these users in terms of their energy-saving.

TABLE 2
MCS parameters with Gaussian channel model and guard interval GI $= 1/4$ in DVB-H standard [31]

| Modulation | QPSK | | | | | 16QAM | | | | | 64QAM | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Code rate | 1/2 | 2/3 | 3/4 | 5/6 | 7/8 | 1/2 | 2/3 | 3/4 | 5/6 | 7/8 | 1/2 | 2/3 | 3/4 | 5/6 | 7/8 |
| SNR Threshold (dB) | 3.1 | 4.9 | 5.9 | 6.9 | 7.7 | 8.8 | 11.1 | 12.5 | 13.5 | 13.9 | 14.4 | 16.5 | 18.0 | 19.3 | 20.1 |
| Channel rate (Mbps) | 4.98 | 6.64 | 7.46 | 8.29 | 8.71 | 9.95 | 13.27 | 14.93 | 16.59 | 17.42 | 14.93 | 19.91 | 22.39 | 24.88 | 26.13 |

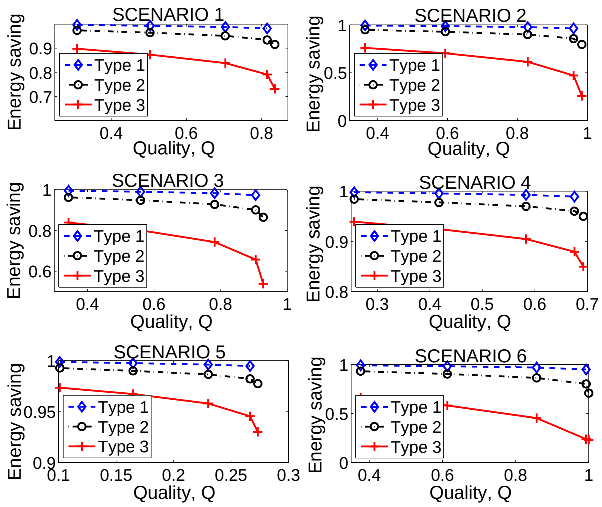



Fig. 11. Normalized energy saving versus reception quality of the different types of users.

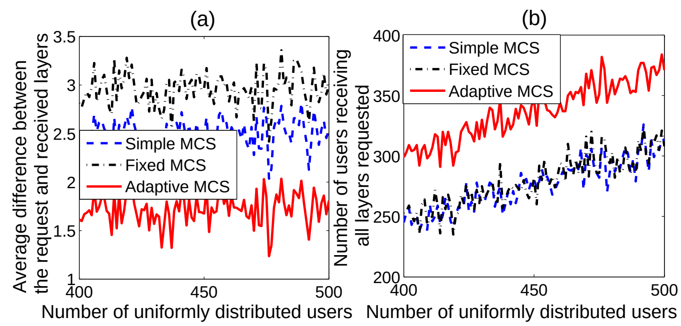



Fig. 12. MCS-dependent quality and capacity performance: (a) average difference between requested and received layers; (b) fully serviced users count under different MCS.

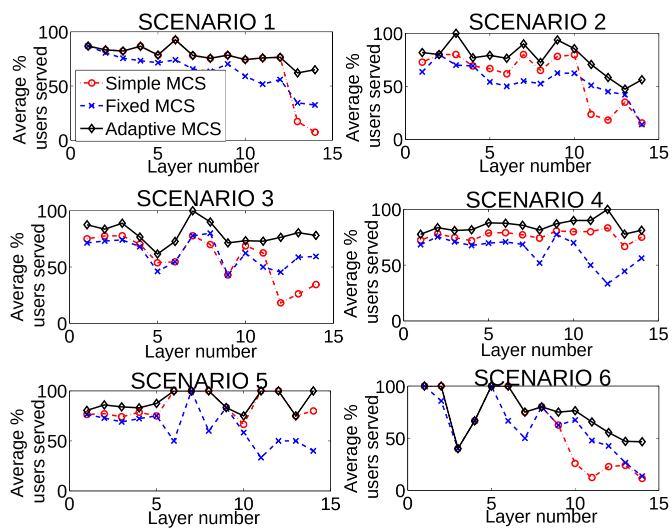



Fig. 13. Effect of MCS and number of layers on the average percentage of users served.

It is also notable that, since each user has the independent control of time-sliced reception, even though the high-end (e.g. type 3) users may not achieve the maximum desired quality due to the system optimization for large proportion of low-end (e.g. type 1) users, they can improve the QoE by the time slicing flexibility.

### 6.2 Adaptive MCS performance

We now study the MCS-dependent broadcast performance. The proposed adaptive MCS is compared against the two other schemes: simple MCS and fixed MCS.

Fig. 12(a) shows the average difference between the user request for a certain number of layers and what they actually receive under the three MCS schemes. Fig. 12(b) shows that the total number of users receiving exactly the requested quality is much higher with adaptive MCS as compared to the two other schemes. The results are averaged over several iterations with the number of users varying between 400 and 500. It can be noticed that the adaptive MCS outperforms the other two MCS schemes in terms of the number of served users. Moreover by using the adaptive MCS the received number of layers are very close to the requested number of layers, reflecting a higher amount of user satisfaction.

Fig. 13 captures the layer based values for the average percentage of users being served in the broadcast network for the six scenarios listed in Table 1 under the three MCS allocation schemes. The results show that the adaptive MCS allocation scheme outperforms the other schemes, by ensuring a higher percentage of users that are getting served under all scenarios.

The composite gain achieved by the adaptive MCS under the six different scenarios is illustrated in Fig. 14. The ratio of the total number of users served by the proposed adaptive MCS is compared against the simple MCS and fixed MCS. It can be noted that, among the three schemes, the adaptive MCS ensures more number of users served. In particular, the gain of adaptive MCS in terms of number of additional users served over the

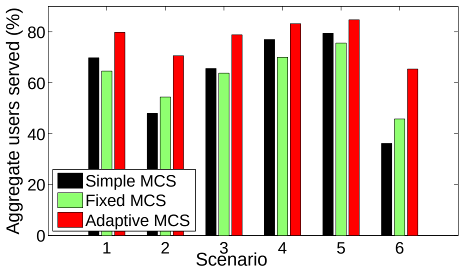



Fig. 14. Comparison of MCS allocation schemes in terms of service capacity.

fixed or simple MCS scheme under the six scenarios with 'ES+Q' strategy are 12.6, 19.4, 14.1, 9.7, 7.2, and 24.4%, respectively. The average number of additional users served with adaptive MCS under the six scenarios is 16.57% with respect to the simple MCS and 16.63% with respect to the fixed MCS.

### 6.3 Energy-Quality trade-off with optimized SVC and adaptive MCS

Fig. 15 presents a comparison of the three MCS along with the three SVC optimization measures: joint energy saving and quality (ES+Q), energy saving only (ES only), and quality only (Q only). The line plots indicate the number of users served versus the number of SVC layers transmitted. The adaptive MCS with 'ES+Q' is shown to perform better than the 'Q only' case. Although the 'ES only' serves a higher number of users (cf. Table 4), the average reception quality is very low (e.g., $\mathcal{Q}$ is 18.19% in scenario 1). This means, in 'ES only' case a large proportion of users would experience a low QoE. The uneven trend in the plots may be due to random distribution of heterogeneous UEs in the network.

The bar plots on the right in Fig. 15 capture the composite gain on number of users served in the three schemes ('ES+Q', 'ES only', and 'Q only') with the three MCS strategies. 'ES+Q' with adaptive MCS serves a lesser number of users in comparison with the 'ES only' case, but performs better with respect to 'Q only' case. The number of users served in 'Q only' case is generally low (e.g., 35% only in scenario 2). The results demonstrate that, with the proposed adaptive MCS, the average reduction of number of users served with 'ES+Q' is only 0.62% lower than that in 'ES only' scenario. On the other hand, the number of users served with 'ES+Q' is about 10.8% higher than that with 'Q only' scenario.

Table 4 further quantifies the energy-quality trade-off with the three optimization schemes: 'ES+Q', 'ES only', and 'Q only' where the weighted quality measure $\mathcal{Q}$ in (15) is used. The table also includes the optimum quantization parameters for the ES and Q trade-off game. Under the six user-heterogeneity scenarios with adaptive MCS, when compared with 'ES only' strategy, the 'ES+Q' strategy offers on average, about 43% higher quality. The corresponding trade-off on the amount of energy saving is only about 8%. With respect to 'Q only' scenario, the 'ES+Q' scheme offers about 17% extra energy saving as well as about 3.5% higher quality performance.

## 7 Conclusion

This paper has introduced a novel cross-layer optimization solution to improve both the quality of user experience (QoE) and energy efficiency of wireless multimedia broadcast receivers with varying display and energy constraints. This joint optimization is achieved by grouping the users based on their device capabilities and estimated channel conditions experienced by them and broadcasting adaptive content to these groups. The optimization is a game theoretic approach which performs energy saving versus reception quality trade-off, and obtains optimum video encoding rates of the different users. This optimization is a function of the proportion of users in a cell with different capabilities, which in turn determines the time slicing proportions for different video content layers for maximized energy saving of low-end users, while maximizing the quality of reception of the high-end users. The optimized layered coding rate, coupled with the receiver groups' SNRs, adaptation of the MCS for transmission of different layers, ensure higher number of users are served while also improving users' average reception quality. Thorough testing has shown how the proposed optimization solution supports better performance for multimedia broadcast over wireless in comparison with the existing techniques.

## References


[1] *The New Multi-screen World: Understanding Cross-platform Consumer Behavior*, Google, Aug. 2012.

[2] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012–2017*, Cisco, 2013.

[3] K. Rao, Z. S. Bojkovic, and D. A. Milovanoic, *Multimedia Communication Systems: Techniques, Standards, Networks*. Prentice Hall, 1st Ed., 2002.

[4] S. Sesia, I. Toufik, and M. Baker, *LTE UMTS Long Term Evolution, from Theory to Practice*. John Wiley & Sons, 2009.

[5] G. Faria, J. Henriksson, E. Stare, and P. Talmola, "DVB-H: Digital broadcast services to handheld devices," *Proc. of IEEE*, vol. 94, no. 1, pp. 194–209, Jan. 2006.

[6] *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television*, ETSI EN 300 744, 2009.

[7] *Digital Video Broadcasting (DVB); DVB-H Transmission System for Handheld terminals*, ETSI EN 302 304, 2011.

[8] *IEEE 802.11n-2009- Amendment 5: Enhancements for Higher Throughput. IEEE-SA.*, 29 Oct. 2009.

[9] C.-H. Hsu and M. M. Hefeeda, "Flexible broadcasting of scalable video streams to heterogeneous mobile devices," *IEEE Trans. Mob. Computing*, vol. 10, no. 3, pp. 406–418, Mar. 2011.

[10] T. E. Shabrawy and S. Wahed, "Adaptive modulation and coding for broadcast DVB-H systems," in *Proc. IEEE PIMRC*, Tokyo, Japan, Sept. 2009, pp. 1292–1296.

[11] O. Karimi and L. Jiangchuan, "Power efficient high quality multimedia multicast in LTE wireless networks," in *Proc. IEEE MASS*, Valencia, Spain, Oct. 2011.

[12] C. Atici and M. Sunay, "High data-rate video broadcasting over 3G wireless systems," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 212–223, Mar. 2007.

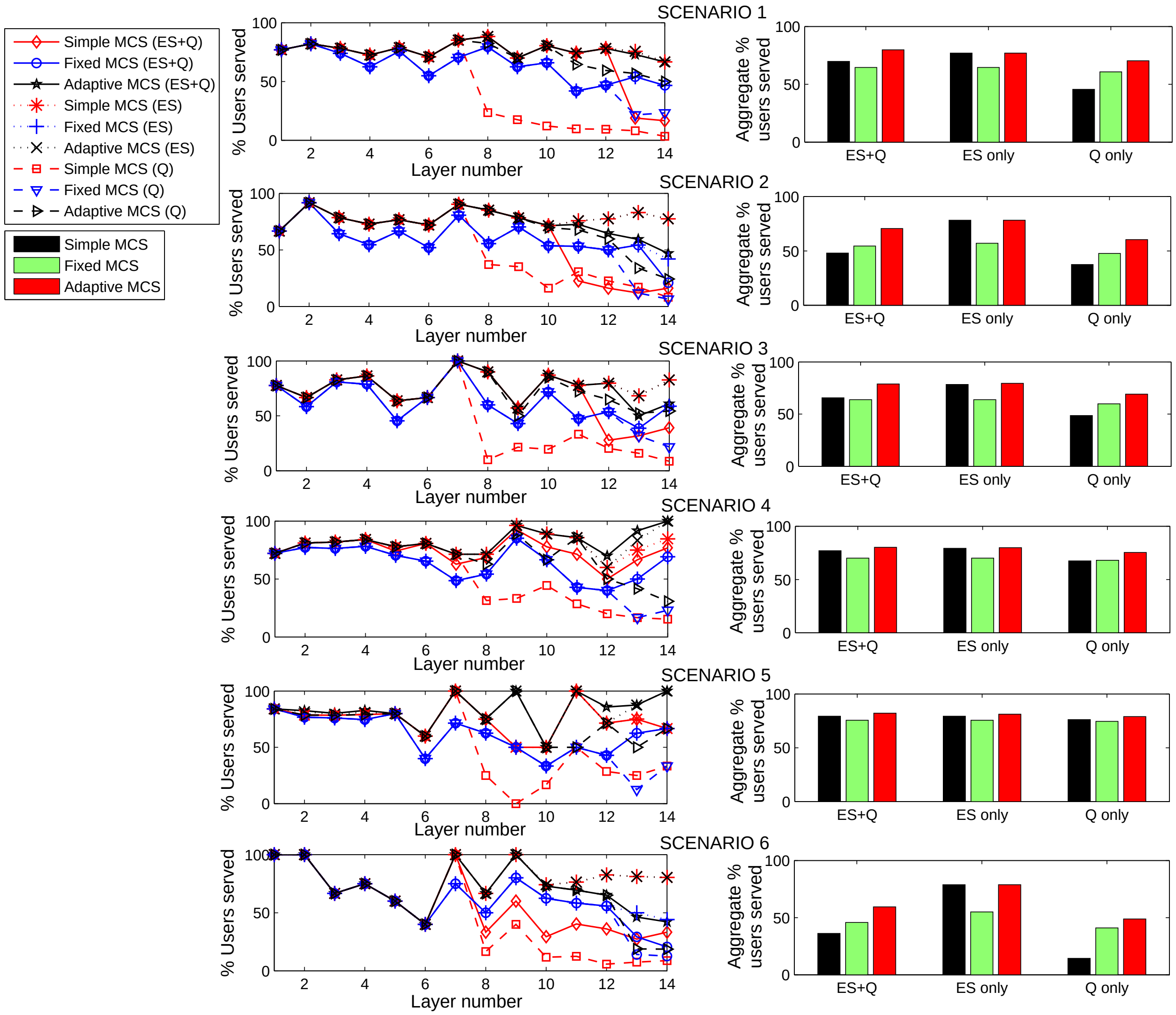



Fig. 15. Average percentage of users served with different MCS versus number of layers, and the aggregate performance.


[13] A. Alexiou, C. Bouras, V. Kokkinos, A. Papazois, and G. Tsichritzis, "Modulation and coding scheme selection in multimedia broadcast over a single frequency network-enabled long-term evolution networks," *Intl. J. Commun. Sys.*, vol. 25, no. 12, pp. 1603–1619, Nov. 2012.

[14] P. Li, H. Zhang, B.-H. H. Zhao, and S. Rangarajan, "Scalable video multicast with adaptive modulation and coding in broadband wireless data systems," *IEEE/ACM Trans. Networking*, vol. 20, no. 1, pp. 57–68, Feb. 2012.

[15] W. Ji, Z. Li, and Y. Chen, "Joint source-channel coding and optimization for layered video broadcasting to heterogeneous devices," *IEEE Trans. Multimedia*, vol. 14, no. 2, pp. 443–455, Apr. 2012.

[16] Y. Wang, J. Ostermann, and Y.-Q. Zhang, *Video Processing and Communications*. Prentice Hall, 2001.

[17] *Overview of the MPEG-4 Standard*, ISO / IEC JTC1 / SC29 / WG11 N4668, 2002.

[18] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264 / AVC standard," *IEEE Trans. Circuits, System and Video Tech.*, vol. 17, no. 9, pp. 1103–1120, Sept. 2007.

[19] T. Wiegand, G. J. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264 / AVC video coding standard," *IEEE Trans. Circuits, System and Video Tech.*, vol. 13, no. 7, July 2003.

[20] T. Schierl, T. Stockhammer, and T. Wiegand, "Mobile video transmission using scalable video coding (SVC)," *IEEE Trans. Circuits, System and Video Tech.*, vol. 17, no. 9, pp. 1204–1217, Sept. 2007.

[21] G. Iwacz, A. Jajszczyk, and M. Czkowski, *Multimedia Broadcasting and Multicasting in Mobile Networks*. UK: Wiley, 2008.

[22] *Scalable video coding*, ISO / IEC JTC1 / SC29 / WG11. N9560, 2008.

[23] *Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines*, ETSI TR 102 377 V1.3.1, 2009.

[24] M. Ghandi and M. Ghanbari, "Layered H.264 video transmission with hierarchical QAM," *J. Visual Commun. Image Representation*, vol. 17, no. 2, pp. 451–466, Apr. 2006.

[25] G. Xylomenos, "Group management for the multimedia broadcast / multicast service," in *Proc. IST Mobile Summit*, Dresden, Germany, June 2005.

[26] Y.-C. Chen and Y.-R. Tsai, "Adaptive resource allocation for multi-resolution multicast services with diversity in OFDM systems," in *Proc. IEEE VTC*, Barcelona, Spain, Apr. 2009, pp. 1–5.

[27] J. Villalon, P. Cuenca, L. Orozco-Barbosa, Y. Seok, and T. Turletti, "Cross-layer architecture for adaptive video multicast streaming over multirate wireless LANs," *IEEE J. Sel. Areas in Commun.*, vol. 25, no. 4, pp. 699–711, May 2007.

[28] Z. Liu, Z. Wu, P. Liu, H. Liu, and Y. Wang, "Layer bargaining: multicast layered video over wireless networks," *IEEE J. Sel. Areas in Commun.*, vol. 28, no. 3, pp. 445–455, Apr. 2010.

[29] Q. Du and X. Zhang, "Statistical QoS provisionings for wireless unicast / multicast of multi-layer video streams," *IEEE J. Sel. Areas in Commun.*, vol. 28, no. 3, pp. 420–433, Apr. 2010.

[30] O. Alay, T. Korakis, Y. Wang, and S. Panwar, "Dynamic rate and FEC adaptation for video multicast in multi-rate wireless networks," *Mob. Netw. Appl.*, vol. 15, no. 3, pp. 425–434, June 2010.

TABLE 4
Average energy saving and quality performance with adaptive MCS under different traffic scenarios

| Scenario (ref. Table 1) | Energy saving (ES only) | $\mathcal{Q}$ measure (ES only) | Energy saving (Q only) | $\mathcal{Q}$ measure (Q only) | Optimum $QP$ (ES+Q) | Energy saving (ES+Q) | $\mathcal{Q}$ measure (ES+Q) |
|---|---|---|---|---|---|---|---|
| 1 | 92.45% | 18.19% | 62.78% | 65.64% | 25 | 89.13% | 67.68% |
| 2 | 72.41% | 24.56% | 45.92% | 62.17% | 15 | 60.36% | 77.63% |
| 3 | 90.10% | 21.40% | 59.18% | 67.96% | 20 | 81.62% | 73.23% |
| 4 | 96.21% | 29.89% | 80.67% | 63.58% | 30 | 92.31% | 56.64% |
| 5 | 98.21% | 32.26% | 82.32% | 71.12% | 40 | 97.59% | 64.96% |
| 6 | 62.14% | 26.59% | 31.49% | 60.82% | 10 | 43.32% | 72.03% |


[31] J. T. J. Penttinen, P. Jolma, E. Aaltonen, and J. Väre, *DVB-H Radio Network*. John Wiley & Sons, Ltd, 2009.

[32] Y. Wang, Z. Ma, and Y.-F. Qu, "Modeling rate and perceptual quality of scalable video as functions of quantization and frame rate and its application in scalable video adaptation," in *Proc. Intl. Packet Video Wksp.*, Seattle, USA, May 2009.

[33] *Subjective video quality assessment methods for multimedia applications*, ITU-T Recommendation, P.910, Apr. 2008.

[34] S. Azad, W. Song, and D. Tjondronegoro, "Bitrate modeling of scalable videos using quantization parameter, frame rate and spatial resolution," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010.

[35] S. Parakh and A. Jagannatham, "Game theory based dynamic bit-rate adaptation for H.264 scalable video transmission in 4G wireless systems," in *Proc. SPCOM*, Bangalore, India, July 2012.

[36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[37] A. Takayama, *Analytical Methods in Economics*. University of Michigan Press, 1993.

[38] A. Cambini and L. Martein, *Generalized Convexity and Optimization: Theory and Applications*. Springer-Verlag, 2009.

[39] J. Reichel, H. Schwarz, and M. Wien, "Joint Scalable Video Model JSVM-12 text," Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Shenzhen, China, Doc. JVT-Y202, Oct. 2007.

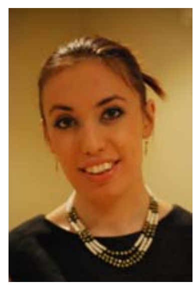

## ACKNOWLEDGMENT


This work has been supported in parts by the Department of Science and Technology (DST) under the grant no. SR/S3/EECE/0122/2010 and Indo-Ireland cooperative science program. The authors are thankful to the anonymous reviewers for the insightful comments and valuable suggestions, which have significantly improved the quality of presentation.


## BIOGRAPHIES

**Chetna Singhal (S'13)** received her B.Eng. in Electronics and Telecommunications from Univ. of Pune in 2008 and M.Tech. in Computer Technology from Electrical Eng. Dept., IIT Delhi in 2010. She worked in IBM Software Lab, New Delhi, as a Software Engineer, from June 2010 to July 2011. She is currently pursuing Ph.D. from Bharti School of Telecommunications, IIT Delhi, since July 2011. Her research interests include handoff schemes and cross-layer optimization in wireless networks, adaptive multimedia multicast and broadcast schemes and technologies. She is a student member of ACM, IEEE, and IEEE Computer and Communications Societies.

**Swades De (S'02-M'04)** received his PhD in Electrical Eng. from the State Univ. of New York at Buffalo in 2004. He is currently an Associate Professor of Electrical Eng. at IIT Delhi. His research interests include performance study, resource efficiency in wireless networks, broadband wireless access, and communication and systems issues in optical networks. Dr. De currently serves as an Associate Editor of IEEE Communications Letters and Springer Photonic Network Communications journal. He is a member of IEEE, IEEE ComSoc, and IEICE.



**Ramona Trestian (S'08-M'12)** is a Lecturer with the Computer and Communications Eng. Dept., School of Science and Technology, Middlesex Univ., London, UK. She received her Ph.D. from Dublin City Univ., Ireland in March 2012 and B.Eng. in Telecommunications from the Electronics, Telecommunications and the Technology of Information Dept., Technical Univ. of Cluj-Napoca, Romania in 2007. She has published in prestigious international conferences and journals and has two edited books. She is a reviewer for international journals and conferences and an IEEE member. Her research interests include mobile and wireless communications, multimedia streaming, handover and network selection strategies, and software-defined networks.

**Gabriel-Miro Muntean (S'02-M'04)** received the B.Eng. and M.Sc. degrees in software eng. from the Computer Sc. Dept., Politehnica Univ. of Timisoara, Timisoara, Romania, in 1996 and 1997, respectively, and the Ph.D. degree from the School of Electronic Eng., Dublin City Univ. (DCU), Dublin, Ireland, in 2003 for his research on quality-oriented adaptive multimedia streaming over wired networks. He is currently a Senior Lecturer with the School of Electronic Eng., DCU. He is Co-Director of the Performance Eng. Lab. Research Group, DCU, and Director of the Network Innovations Centre, part of the Rince Institute Ireland. He has published over 150 papers in prestigious international journals and conferences, has authored three books and 12 book chapters, and has edited five other books. His current research interests include quality-oriented and performance-related issues of adaptive multimedia delivery, performance of wired and wireless communications, energy-aware networking, and personalized e-learning. Dr. Muntean is an Associate Editor of the IEEE Transactions on Broadcasting, the IEEE Communication Surveys and Tutorials, and a reviewer for other important international journals, conferences, and funding agencies. He is a member of the ACM, IEEE and the IEEE Broadcast Technology Society.