Multimedia Tools and Applications (2019) 78:26807–26831
https://doi.org/10.1007/s11042-019-07868-8

Joint rate adaptation and resource allocation for real-time H.265/HEVC video transmission over uplink OFDMA systems

Fan Li[1] · Taiyu Wang[1] · Pamela C. Cosman[2]

Received: 6 September 2018 / Revised: 30 May 2019 / Accepted: 5 June 2019 /
Published online: 14 June 2019

Abstract
We consider multiuser video communication over uplink orthogonal frequency-division multiple access (OFDMA) systems. A cross-layer algorithm of joint bit allocation, packet scheduling and wireless resource assignment are proposed to minimize the end-to-end expected video distortion. Video rate adaptation is performed under the wireless resource constraints. The target number of encoding bits for each video packet is obtained to minimize the estimated distortion based on the online content-based rate-distortion function. Due to the inaccuracy of the rate control algorithm in H.265/HEVC encoding, the actual number of bits may differ from the target. Accordingly, the actual encoder distortion may deviate from the estimated distortion. Then, we propose an iterative algorithm to re-assign wireless resources based on the actual number of encoded bits to obtain the final resource allocation policy and packet scheduling decision. Numerical simulation results show that our proposed approach significantly outperforms the baseline algorithms in terms of received video quality.

Keywords Video communication · Wireless resource allocation · Encoding rate adaptation · Video packet scheduling · OFDMA

✉ Fan Li
lifan@mail.xjtu.edu.cn

Taiyu Wang
wangtaiyu@stu.xjtu.edu.cn

Pamela C. Cosman
pcosman@eng.ucsd.edu

[1] School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China

[2] Department of Electrical and Computer Engineering, University of California, San Diego, San Diego, CA, USA

1 Introduction

Orthogonal frequency-division multiplexing (OFDM) is an advanced communications technique to provide high-bit-rate transmission and quality of service (QoS) guarantees. In an OFDM system, a user can dynamically be assigned a set of subcarriers with good channel states due to the frequency-selective fading across different subcarriers [22]. Orthogonal frequency-division multiple access (OFDMA) provides flexibility for subcarrier assignment to maximize system capacity and spectral efficiency. The scheduler can dynamically allocate power to subcarriers to satisfy different QoS requirements.

However, it is challenging to improve the end-to-end video quality in uplink OFDMA systems. Users who face severe channel fading may need more subcarriers, which might degrade other users' video quality. Due to the temporal correlation of video frames, video quality may be degraded when packet loss occurs during transmission. Therefore, the problems of video transmission over OFDMA lie in determining video packet encoding and scheduling, and subcarrier and power assignment based on users' diverse video contents and channel states.

Much research has focused on wireless resource allocation for multi-user systems to improve end-to-end video quality. These methods aim to maximize the overall system throughput and spectrum efficiency for downlink [1, 2, 29, 31] or uplink [8, 15, 20, 26] transmission. However, for video transmission, the increase of throughput does not always enhance the received video quality, because quality depends on encoding rate and video content among other factors [19]. Jointly considering wireless resources and video content information is useful.

For video communication over OFDMA systems, optimal resource allocation and packet scheduling has been studied in [4–6, 9, 13, 14, 17, 18, 30]. The video content along with the channel state are jointly considered to optimize the transmission. In [9, 14, 17, 18, 30], downlink video transmission is studied, in which the videos are precoded. In [9, 17, 18], cross-layer algorithms are proposed to optimize wireless resource allocation and packet scheduling by evaluating the importance of video packets in the pixel or compressed domain. The authors in [14] proposed adaptive modulation and coding allocation and adaptive energy allocation for wireless video transmission by utilizing only the information in the slice headers. In these methods, however, videos are precoded, and the data size is fixed which cannot adapt to the time-varying characteristic of wireless channels. Transmission for layered video has also been studied. Layered video, encoded by H.264/AVC and SHVC, can dynamically adjust the transmission rate to adapt to the channel state [4–6, 13]. However, it has a high computational complexity.

To overcome these problems, some research takes the encoding process into consideration. In [10, 23, 28], real-time video encoding optimization based on the channel state is studied. They predict the packet loss rate through feedback from the receive side, and estimate the overall transmission distortion caused by quantization, error propagation and error concealment for video packets. However, these studies are only applicable to the one-user scenario, and wireless resource allocation is not included in the optimization. Research on joint encoding and transmission resource allocation for multi-user video transmission is also a critically important research area. In [25], based on the rate-distortion model developed at the level of group of pictures (GOPs), the authors proposed a joint encoding rate and wireless resource allocation optimization algorithm to minimize the predicted distortion. Similarly, the authors of [27] presented a near optimal solutions to both long-term bit rate assignment and packet scheduling for the uplink wireless network. Research in [21]

proposed a QoE model which is related to quantization parameter (QP) and packet loss rate during transmission. For each algorithm period, the receiver needs to send back the QoE values, which are evaluated after decoding the data received in the last period. With these historical QoE values, the algorithm chooses QP to maximize the current QoE value under the channel state. In [16], the authors proposed a optimized cross-layer allocation acheme for multi-user uplink transmission by designing an objective function to maximize the average PSNR. However, there are two problems with the methods in [16, 21, 25, 27]. First, the rate-distortion (R-D) function is fitted before transmission, and developed off-line. It can not adapt to various contents of video frames. Thus, the coding distortion is not accurately predicted. Second, rate control in the video encoders is not precise. The actual number of encoding bits usually deviates from the target one. Thus, the methods of joint video coding and wireless resource allocation are not optimal.

In this paper, we consider a scenario of multiuser video communication over uplink OFDMA systems. The algorithms of joint encoding rate adaptation, packet scheduling and wireless resource allocation are proposed to minimize the end-to-end expected distortion for video users. The main contributions of this paper can be summarized as follows.

1) *Quality-driven cross-layer framework.* We propose a framework that synthetically considers content-based video rate adaption, multiuser packet scheduling, subcarrier assignment and power distribution. The objective is to maximize the received video quality under application layer, media access control layer and physical layer constraints in uplink OFDMA systems. Therefore, the proposed architecture is a quality-driven cross-layer design.
2) *Online content-based rate-distortion optimization for video rate adaption.* Video rate-distortion (R-D) functions depend on video contents, and are different for various video sequences, frames and even slices. Therefore, video R-D functions with constant R-D parameters are usually inaccurate. In this paper, we develop content-based R-D functions at the level of video slices. Therefore, it is inadvisable to develop the R-D function of videos by using constant R-D parameters. The R-D parameters are dynamically updated based on the temporal correlations of the video sequences.
3) *Iterative algorithm for wireless resource re-assignment based on the actual number of encoded bits.* After users encode videos by H.265/HEVC, the actual number of bits for the encoded video may differ from the target due to the inaccuracy of the rate control in the encoder. Accordingly, the actual encoder distortion may deviate from the estimated distortion. Then, an iterative algorithm is proposed for wireless resource re-assignment based on the actual number of encoded bits.

The rest of this paper is organized as follows. In Section 2 we describe our system architecture and formulate the optimization problem. A cross-layer optimization solution of joint video rate adaptation and wireless resource allocation is proposed in Section 3. Then, the baseline algorithms are introduced in Section 4. The performance evaluations are given in Section 5. Section 6 concludes the paper. Notation used in this paper is summarized in Table 1.

2 System overview and problem formulation

We consider an application scenario, such as live video and video surveillance, which is described in Fig. 1. Videos are captured by users and sent to the base station (BS) with

Table 1 Summary of key notation

Notation	Meaning
K	Number of users
N	Number of subcarriers
M_k	Number of packets of user k
B	Bandwidth of the network
N_0	Power spectral density of additive white Gaussian noise
P_k	Power constraint of user k
Ω_k, Ω_k^*, Ω_k^{**}	Subcarrier set of user k, subcarrier set after rate adaptation, subcarrier set after resource re-allocation
$h_{k,n}$	Channel gain of subcarrier n when assigned to user k
$a_{k,m}$, $b_{k,m}$	$R-D$ parameters of packet m of user k
$R_{k,m}$, $R_{k,m}^*$, $R_{k,m}^{act}$	Number of packet bits, target number of encoding bits after rate adaptation algorithm, actual number of bits after encoding k
$Q_{k,m}$	Quality contribution of packet m of user k
$p_{k,n}$, $p_{k,n}^*$, $p_{k,n}^{**}$	Power allocated to subcarrier n of user k, power allocation after rate adaptation, power allocation after resource re-allocation
$\rho_{k,n}$, $\rho_{k,n}^*$, $\rho_{k,n}^{**}$	Subcarrier assignment indicator, subcarrier assignment indicator after rate adaptation algorithm, subcarrier assignment indicator after resource re-allocation algorithm
$u_{k,m}$, $u_{k,m}^{**}$	Indicator of packet m of user k is scheduled or dropped, packet scheduling indicator after resource re-allocation
$D_{k,m}^{display}$	Distortion of packet m when stored in the play buffer of user k
$D_{k,m}^{act}$	Actual distortion of packet m of user k caused by encoding
$D_{k,m}^{est}$	Predicted distortion of packet m of user k caused by encoding
$D_{k,m}^{drp}$	Distortion of packet m of user k caused by error concealment if the packet is lost

uplink channels, sharing the same network resources. Users and the BS are located within a single cell. The connections between them use hybrid time-division multiple-access and OFDMA. Then the BS transmits video packets to the receive side, such as a cloud media server through the backbone network, which can be considered as high bandwidth and lossless.

2.1 System architecture

The block diagram is shown in Fig. 2. Once a video frame is captured by an end user, the encoder can predict the R-D parameters of each video packet in the frame. The R-D parameters are estimated by using the video content information in the previous frames. After that, users send the BS their estimated R-D parameters of each packet along with the channel-state information (CSI), which in practice can be estimated by users but is assumed here to be known accurately. Based on estimated R-D parameters and CSI, the rate adaptation algorithm will be performed in the BS to get the target number of encoding bits

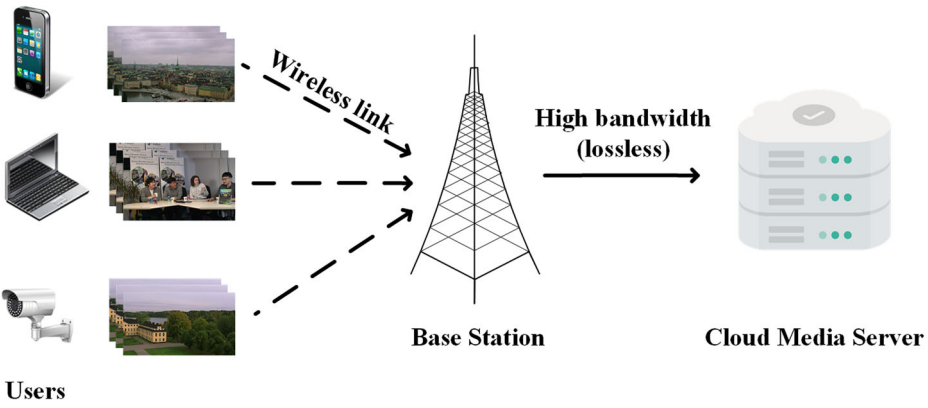


Fig. 1 Application scenario

$R^*_{k,m}$ for packet m of user k, power distribution $p^*_{k,n}$ for user k in subcarrier n and subcarrier assignment $\rho^*_{k,n}$.

Then, each user encodes the video by H.265/HEVC based on the target number of encoding bits. After that, the actual number of bits $R^{act}_{k,m}$ is obtained by each user, which may differ from $R^*_{k,m}$ due to rate control inaccuracy. Meanwhile, the actual value of encoding distortion $D^{act}_{k,m}$ can also be obtained from the encoded video frames. There also exists a deviation between the predicted value of encoding distortion $D^{est}_{k,m}$ and the actual one $D^{act}_{k,m}$.

Using the values of $R^{act}_{k,m}$ and $D^{act}_{k,m}$, the BS re-allocates wireless resources to obtain the optimal subcarrier assignment $\rho^{**}_{k,n}$, power distribution $p^{**}_{k,n}$ and packet scheduling decision $u^{**}_{k,m}$.

Video decoders use temporal replacement (TR) as the error concealment strategy to combat packet loss during transmission. TR reconstructs the lost blocks using the collocated blocks in the previous frame. It is widely used for its simplicity.

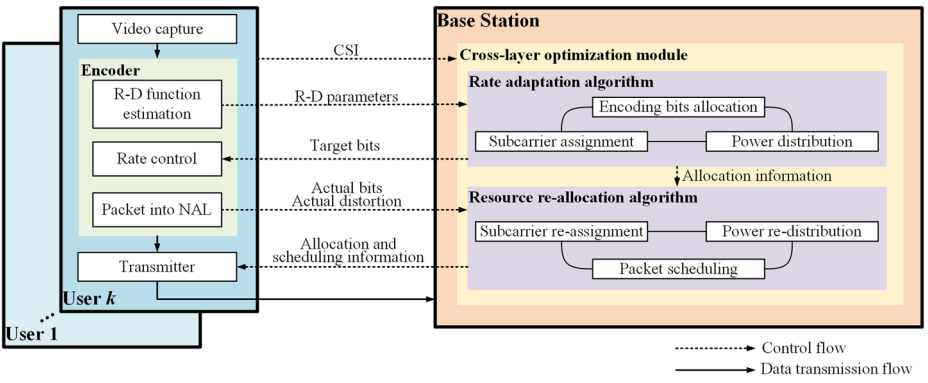


Fig. 2 Block diagram

2.2 Problem formulation

For the uplink system, the objective is to minimize the expected end-to-end distortion for all users. The end-to-end expected distortion $E\{D^{display}\}$ can be expressed as

$$E\left\{D^{display}\right\} = \sum_{k=1}^{K}\sum_{m=1}^{M_k} E\left\{D_{k,m}^{display}\left(u_{k,m}, R_{k,m}\right)\right\} \tag{1}$$

where $R_{k,m}$ is the number of encoding bits allocated for packet m of user k, and $u_{k,m}$ represents whether packet m of user k will be scheduled ($u_{k,m} = 1$) or dropped ($u_{k,m} = 0$) during the current period. K represents the number of users and M_k is the number of video packets to be transmitted by user k.

Define $D_{k,m}^{act}$ as the distortion caused by the encoding for a received packet, and $D_{k,m}^{drp}$ is the distortion caused by the error concealment method for a lost packet. Thus

$$\begin{aligned} & D_{k,m}^{display}\left(u_{k,m}, R_{k,m}\right) \\ &= u_{k,m} D_{k,m}^{act} + (1 - u_{k,m}) D_{k,m}^{drp} \\ &= u_{k,m}(D_{k,m}^{act} - D_{k,m}^{drp}) + D_{k,m}^{drp} \end{aligned} \tag{2}$$

Then, the optimization problem can be rewritten as

$$\begin{aligned} & \min_{\mathbf{R},\mathbf{u},\boldsymbol{\rho},\mathbf{p}} \sum_{k=1}^{K}\sum_{m=1}^{M_k} E\left\{u_{k,m}\left(D_{k,m}^{act} - D_{k,m}^{drp}\right) + D_{k,m}^{drp}\right\} \\ & subject\ to \\ & \text{(C1)}\ \sum_{m=1}^{M_k} u_{k,m} R_{k,m} \le \sum_{n=1}^{N} t_s \frac{B}{N} \rho_{k,n} \log_2\left(1 + \frac{p_{k,n}\left|h_{k,n}\right|^2}{N_0 B/N}\right) \\ & \text{(C2)}\ \sum_{n=1}^{N} \rho_{k,n} p_{k,n} \le P_k \\ & \text{(C3)}\ \sum_{n=1}^{N} \rho_{k,n} = 1, \quad \rho_{k,n} \in \{0, 1\} \\ & \text{(C4)}\ u_{k,m} \in \{0, 1\} \end{aligned} \tag{3}$$

C1 is the transmission capacity constraint which says that the number of transmitted bits must be less than the capacity. Here N is the total number of subcarriers, t_s is the time length of the OFDMA symbol, B is the bandwidth, N_0 is the power spectral density of additive white Gaussian noise, $h_{k,n}$ is the channel gain for user k in subcarrier n, $\rho_{k,n} = 1$ if subcarrier n is assigned to user k and $\rho_{k,n} = 0$ otherwise, and $p_{k,n}$ represents the power allocated on subcarrier n to user k.

C2 is the power constraint that ensures the consumed power of each user will not exceed their available power.

C3 represents the subcarrier constraints. Each subcarrier is assigned to one and only one user at each scheduling period.

C4 is the packet scheduling constraint. Each packet is either sent or not.

3 Cross-layer optimization algorithm

To solve the optimization problem in (3), a cross-layer algorithm is proposed. First, video bitrate adaptation is performed under the wireless resource constraints. Thus, the target number of encoding bits and the corresponding resource allocation is obtained. After video encoding at the target number of bits for every user, wireless resources are re-assigned due to the inaccuracy of the rate control algorithm and the rate-distortion function in the video encoder.

3.1 Joint rate adaptation and subcarrier/power assignment

Before encoding, the distortion $D_{k,m}^{act}$ cannot be accurately obtained. Thus, we use the estimated value $D_{k,m}^{est}$, which is predicted by the R-D function. The R-D function in the H.265/HEVC standard [12] is modeled as a hyperbolic function:

$$D_{k,m}^{est} = a_{k,m}^{est}\left(\frac{R_{k,m}}{I_{k,m}}\right)^{-b_{k,m}^{est}} \tag{4}$$

where $I_{k,m}$ is the number of pixels in packet m of user k. $a_{k,m}^{est}$ and $b_{k,m}^{est}$ are content-related parameters. To adaptively describe the relationship between rate and distortion, we periodically update the R-D parameters $a_{k,m}^{est}$ and $b_{k,m}^{est}$ during encoding by utilizing the least mean square (LMS) method, shown in Appendix, using the parameters in the same location in the previous frame.

At this point in the algorithm, it is assumed that all encoded packets will be scheduled for transmission, so $u_{k,m} = 1$. Also, $D_{k,m}^{act}$ is not known, so we use $D_{k,m}^{est}$ instead. We change the optimization problem in (3) to be:

$$\begin{aligned}
&\min_{\mathbf{R},\boldsymbol{\rho},\mathbf{p}} \sum_{k=1}^{K}\sum_{m=1}^{M_k} E\left\{D_{k,m}^{est}\right\} = \min_{\mathbf{R},\boldsymbol{\rho},\mathbf{p}} \sum_{k=1}^{K}\sum_{m=1}^{M_k} a_{k,m}^{est}\left(\frac{R_{k,m}}{I_{k,m}}\right)^{-b_{k,m}^{est}} \\
&subject\ to \\
&\text{(C1)}\quad \sum_{m=1}^{M_k} R_{k,m} \le \sum_{n=1}^{N} t_s \frac{B}{N}\rho_{k,n}\log_2\left(1+\frac{p_{k,n}\left|h_{k,n}\right|^2}{N_0 B/N}\right) \\
&\text{(C2)}\quad \sum_{n=1}^{N} \rho_{k,n} p_{k,n} \le P_k \\
&\text{(C3)}\quad \sum_{n=1}^{N} \rho_{k,n} = 1, \quad \rho_{k,n} \in \{0,1\}
\end{aligned} \tag{5}$$

To solve the optimization problem in (5), we propose an iterative algorithm by jointly considering the characteristics in the application layer, the media access control layer and the physical layer. In each iteration, one subcarrier is re-assigned to the user who has the potential to achieve the highest distortion reduction. The algorithm is iterative until no subcarrier can be re-assigned. By determining the subcarrier assignment, the target numbers of encoding bits for users can be obtained.

Step 1: Capacity-distortion (C-D) curve fitting

We consider the relationship between channel capacity and encoding distortion for each user. We examine S values of the channel capacity $C_k(s)$ evenly spaced over $[0, C_k^{\max}]$:

$$C_k(s) = \frac{(s-1)\cdot C_k^{\max}}{S-1} \tag{6}$$

where $C_k^{\max}$ is the channel capacity of user k when all subcarriers are assigned to user k, and $s = 1, 2, \cdots, S$. Then, we can derive the minimal estimated distortion $D_k(s)$ for each value of the channel capacity $C_k(s)$ by solving the problem

$$D_k(s) = \min\left\{ \sum_{m=1}^{M_k} a_{k,m}^{est} \left(\frac{R_{k,m}}{I_{k,m}}\right)^{-b_{k,m}^{est}} \middle| \sum_{m=1}^{M_k} R_{k,m} \le C_k(s) \right\} \tag{7}$$

The optimization in (7) is a convex problem, and can be solved by the Lagrange dual method [3]. For the S values of $(C_k(s), D_k(s))$, we find that the relationship between $C_k(s)$ and $D_k(s)$ can be depicted by an exponential relationship:

$$D_k = x_k \cdot {C_k}^{-y_k} \tag{8}$$

The parameters x_k and y_k are content-related, and are obtained by fitting. In Fig. 3, we show the curve fitting results for the 100th frame of three 720p videos (*old town cross, stockholm, in to tree*). For all three sequences, the R^2 values are greater than 0.999, and RMSE values are very small (3.4732, 1.7874 and 0.8955). The fit is performed per frame.

Step 2: Initial subcarrier assignment

Each subcarrier is initially assigned to the user with the highest channel gain $|h_{k,n}|^2$, as

$$k^* = \arg\max_k |h_{k,n}|^2 \tag{9}$$

for $1 \le n \le N$. The assignment is performed per frame. After this step, we have the initial subcarrier set Ω_k for each user $k \in \{1, 2, \cdots, K\}$.

Moreover, Ψ is defined as the set of users who have the potential to reduce the overall distortion given additional subcarriers. We initialize $\Psi = \{1, 2, \cdots K\}$.

Step 3: Selection of the user with steepest C-D curve slope

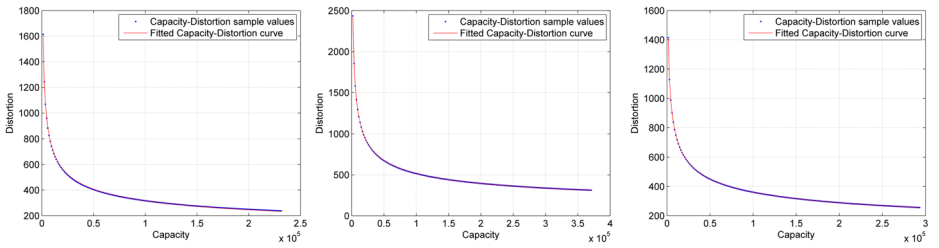


Fig. 3 Capacity-distortion sample points curve fitting result of **a** *old town cross*, **b** *stockholm*, **c** *in to tree*.

Water filling is utilized to allocate the power P_k to Ω_k. Then, the current transmission capacity of each user $C_k(\Omega_k)$ can be obtained by

$$C_k(\Omega_k) = \sum_{n\in\Omega_k} t_s \frac{B}{N}\log_2\left(1+\frac{p_{k,n}\left|h_{k,n}\right|^2}{N_0 B/N}\right) \tag{10}$$

Based on the fitted relationship in (8), the current distortion of each user can be calculated by

$$D_k(\Omega_k) = x_k \cdot C_k(\Omega_k)^{-y_k} \tag{11}$$

The user with the steepest slope of the *C-D* curve achieves the highest distortion reduction when receiving an increment of additional transmission capacity. Therefore, we select user k^* satisfying

$$k^* = \arg\min_{k\in\Psi} S_k \tag{12}$$

where S_k is the slope of the *C-D* curve for user k, which is calculated by

$$S_k = \frac{dD_k(\Omega_k)}{dC_k(\Omega_k)} = -x_k y_k C_k(\Omega_k)^{-(1+y_k)} \tag{13}$$

Step 4: Subcarrier re-assignment

Given user k^* in **Step 3**, we consider all the subcarriers which are not assigned to user k^*. We aim to re-assign to user k^* one of the subcarriers, which can bring the maximal performance improvement. Here, we use the method proposed in [25].

For subcarrier n_0 which is assigned to user k_0 at this time, the performance improvement for subcarrier n_0 transferring from user k_0 to k^* can be defined as

$$\Delta(n_0) = \Delta^g_{k^*}(n_0) - \Delta^l_{k_0}(n_0) \tag{14}$$

where $\Delta^g_{k^*}(n_0)$ is the distortion reduction when user k^* gains subcarrier n_0, and $\Delta^l_{k_0}(n_0)$ is the distortion increment when user k_0 loses subcarrier n_0, which are calculated by

$$\Delta^g_{k^*}(n_0) = D_{k^*}(\Omega_{k^*}) - D_{k^*}(\Omega_{k^*}\cup\{n_0\}) \tag{15}$$

$$\Delta^l_{k_0}(n_0) = D_{k_0}\left(\Omega_{k_0}-\{n_0\}\right) - D_{k_0}\left(\Omega_{k_0}\right) \tag{16}$$

Figure 4 illustrates the process of subcarrier re-assignment. In Fig. 4a, subcarrier n_0 belongs to user $k_0 \in \{1,2,\cdots,K\}-\{k^*\}$, and the current frame distortions of users k_0 and k^* are marked by squares in the left two R-D diagrams. After re-assignment, subcarrier n_0 is given to user k^* in Fig. 4b, whose distortion goes down to the circular mark. Thus, for each subcarrier $n \in \Omega - \Omega_{k^*}$, we calculate the performance improvement if it were assigned to user k^* by (14). Then we find the subcarrier n^* which achieves the maximal performance improvement by

$$n^* = \arg\max_{n\in\Omega-\Omega_{k^*}} \Delta(n) \tag{17}$$

If $\Delta(n^*) > 0$, we take subcarrier n^* from user k_0 and assign it to user k^*. So Ω_{k^*} is updated to be $\Omega_{k^*}\cup\{n^*\}$, and Ω_{k_0} is updated to be $\Omega_{k_0}-\{n^*\}$.

If $\Delta(n^*) < 0$, there is no subcarrier which can be re-assigned to user k^* to produce an overall distortion improvement. So we remove user k^* from the set Ψ. So Ψ is updated to be $\Psi-\{k^*\}$.

Then we go back to **Step 3** for the next iteration. The algorithm stops when there is only one user in the set Ψ.

After iterations, the current allocated subcarrier set is denoted Ω^*_k. Moreover, we obtain the subcarrier assignment decision $\rho^*_{k,n}$. Then, water filling is utilized to allocate user k's

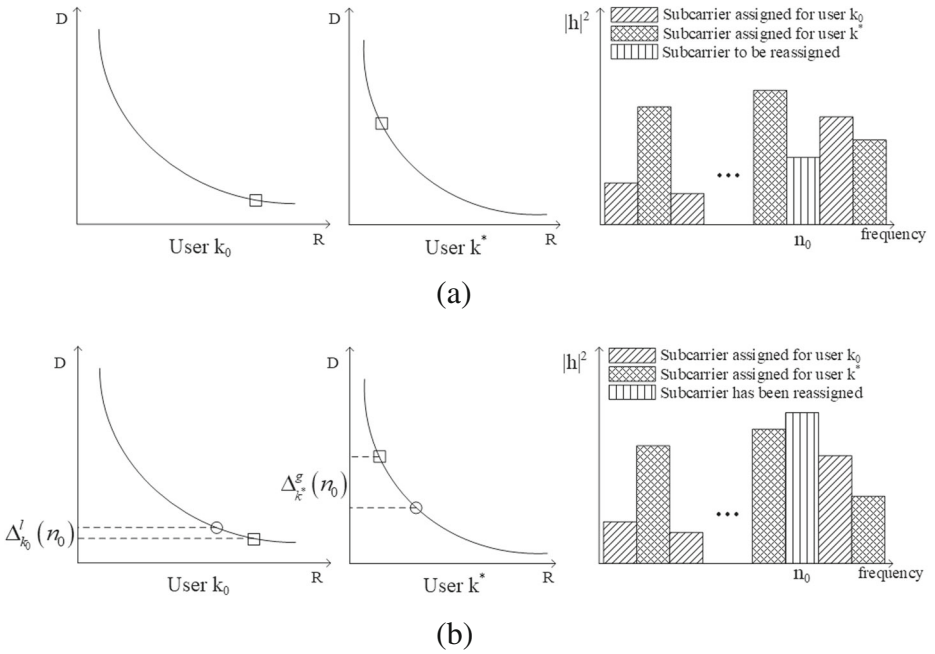


Fig. 4 Illustration of subcarrier re-assignment. **a** Before re-assignment **b** After re-assignment

total power P_k on the assigned subcarrier set, yielding the power distribution $p^*_{k,n}$. Getting the transmission capacity $C_k\left(\Omega^*_k\right)$ of each user by (10), we obtain the target number of encoding bits $R^*_{k,m}$ of each user by solving the problem

$$R^*_{k,m} = \arg\min_{R_{k,m}} \left\{ \sum_{m=1}^{M_k} a^{est}_{k,m} \left(\frac{R_{k,m}}{I_{k,m}} \right)^{-b^{est}_{k,m}} \middle| \sum_{m=1}^{M_k} R_{k,m} \le C_k(\Omega^*_k) \right\} \tag{18}$$

Thus, subcarrier assignment $\rho^*_{k,n}$, power distribution $p^*_{k,n}$, and the target number of encoding bits $R^*_{k,m}$ are obtained.

3.2 Resource re-allocation based on actual number of encoding bits

Each video packet is encoded with the target number of encoding bits $R^*_{k,m}$ using the H.265/HEVC encoder. The rate control algorithm in [12] is utilized in the video encoder. After encoding, we can get the actual number of encoding bits $R^{act}_{k,m}$ and the encoding distortion $D^{act}_{k,m}$.

Since rate control in the encoder may be inaccurate, $R^{act}_{k,m}$ would be unequal to $R^*_{k,m}$. Then, the transmission capacity allocated can be deficient or excess for transmitting the video packets. Besides, the R-D model's inaccuracy also leads to some difference between $D^{est}_{k,m}$ and $D^{act}_{k,m}$. Therefore, wireless resources are re-allocated to perform packet scheduling, subcarrier assignment and power distribution based on actual number of encoding bits for

each video user. According to the actual encoding distortion $D_{k,m}^{act}$, the re-allocation problem can be formulated as

$$\min_{\mathbf{u},\boldsymbol{\rho},\mathbf{p}} \sum_{k=1}^{K}\sum_{m=1}^{M_k} E\left\{u_{k,m}\left(D_{k,m}^{act}-D_{k,m}^{drp}\right)+D_{k,m}^{drp}\right\}$$

$$\Rightarrow \min_{\mathbf{u},\boldsymbol{\rho},\mathbf{p}} \sum_{k=1}^{K}\sum_{m=1}^{M_k} E\left\{u_{k,m}\left(D_{k,m}^{act}-D_{k,m}^{drp}\right)\right\}$$

$$subject\ to$$

$$(C1)\ \sum_{m=1}^{M_k} u_{k,m}R_{k,m}^{act} \le \sum_{n=1}^{N} t_s\frac{B}{N}\rho_{k,n}\log_2\left(1+\frac{p_{k,n}\left|h_{k,n}\right|^2}{N_0 B/N}\right)$$

$$(C2)\sum_{n=1}^{N}\rho_{k,n}p_{k,n} \le P_k$$

$$(C3)\sum_{n=1}^{N}\rho_{k,n}=1,\quad \rho_{k,n}\in\{0,1\}$$

$$(C4)\ u_{k,m}\in\{0,1\} \tag{19}$$

We propose an iterative subcarrier re-assignment algorithm to solve the problem in (19). Similar to the basic idea in Section 3.1, we iteratively re-assign one subcarrier to the user who achieves the highest quality contribution improvement until meeting the iteration stop condition.

Step 1: Packet sorting

Define the quality contribution $Q_{k,m}$ as $(D_{k,m}^{drp}-D_{k,m}^{act})$, where $D_{k,m}^{drp}$ is estimated by the distortion caused by the TR error concealment strategy when the packet is lost. Among all the packets of user k, the one with highest $Q_{k,m}/R_{k,m}$ will be scheduled first. We give priority to packets which can achieve a higher quality enhancement at the cost of less wireless resources. So we sort the video packets of each user $k\in\{1,2,\cdots,K\}$ by $Q_{k,m}/R_{k,m}$ in descending order as $\left\{\frac{Q_{k,(1)}}{R_{k,(1)}},\frac{Q_{k,(2)}}{R_{k,(2)}},\cdots,\frac{Q_{k,(M_k)}}{R_{k,(M_k)}}\right\}$. It is noted that the packet sorting is performed per frame.

Here, we start the subcarrier re-assignment from the result Ω_k^* obtained in Section 3.1. And we initialize $\Theta=\{1,2,\cdots,K\}$ as the set of users who have the potential to increase the overall quality contribution when given additional subcarriers.

Step 2: Selection of the user with steepest C-Q curve slope

The user with the steepest slope of the $C-Q$ (Transmission Capacity-Quality Contribution) curve can achieve the highest increase in quality contribution for a given increment of transmission capacity. Unlike the algorithm of Section 3.1 which is executed prior to video encoding, the algorithm here is executed after encoding. So packets can either be discarded or transmitted. Because partially received video packets can not be decoded, the relationship is stepwise between the accumulated video quality contribution Q_k^{ac} and the

transmission capacity C_k based on the scheduling order in **Step 1**, as shown in Fig. 5. Here Q_k^{ac} is calculated by

$$Q_k^{ac}(\Omega_k) = \begin{cases} 0, & if \ C_k(\Omega_k) < R_{k,(1)} \\ \sum_{i=1}^{m} Q_{k,(i)}, & if \ \sum_{i=1}^{m} R_{k,(i)} \le C_k(\Omega_k) < \sum_{i=1}^{m+1} R_{k,(i)} \\ \sum_{i=1}^{M_k} Q_{k,(i)}, & if \ C_k(\Omega_k) \ge \sum_{i=1}^{M_k} R_{k,(i)} \end{cases} \tag{20}$$

where the transmission capacity $C_k(\Omega_k)$ is obtained by using water filling.

Since $Q_k^{ac}(\Omega_k)$ changes discretely with different values of $C_k(\Omega_k)$, the increase and decrease of $Q_k^{ac}(\Omega_k)$ may not be reflected by one-by-one subcarrier re-assignment. We use the slope-like relationship instead of the step-like $C_k - Q_k^{ac}$ relationship, as the red line in Fig. 5. The slope of user k's $C_k - Q_k^{ac}$ curve is calculated by

$$S_k = \begin{cases} \frac{Q_{k,(1)}}{R_{k,(1)}}, & if \ C_k(\Omega_k) \le R_{k,(1)} \\ \frac{Q_{k,(i)}}{R_{k,(i)}}, & if \ \sum_{i-1} R_{k,(i)} \le C_k(\Omega_k) \le \sum_{i} R_{k,(i)} \\ 0, & otherwise \end{cases} \tag{21}$$

The user who has the maximal S_k has the steepest $C_k - Q_k^{ac}$ curve slope under the current C_k, which is determined by

$$k^* = \arg\min_{k \in \Theta} S_k \tag{22}$$

where k^* is the user who can achieve the highest increase in quality contribution when receiving an increment of additional rate.

Step 3: Subcarrier re-assignment

Based on the slope-like $C_k - Q_k^{ac}$ relationship, given the subcarrier set Ω_k of user k, we can calculate the corresponding accumulated packet quality contribution $Q_k^{ac}(\Omega_k)$.

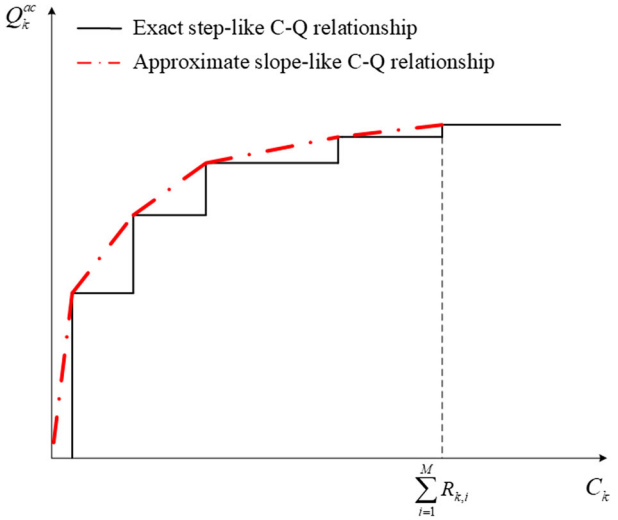


Fig. 5 Step-like and slope-like $C_k - Q_k^{ac}$ relationship

We define the quality contribution increment when user k^* gains subcarrier n_0 as

$$\Delta_{k^*}^{g}(n_0) = Q_{k^*}^{ac}(\Omega_{k^*} \cup \{n_0\}) - Q_{k^*}^{ac}(\Omega_{k^*}) \tag{23}$$

and the quality contribution reduction when user k_0 loses subcarrier n_0 as

$$\Delta_{k_0}^{l}(n_0) = Q_{k_0}^{ac}(\Omega_{k_0}) - Q_{k_0}^{ac}(\Omega_{k_0} - \{n_0\}) \tag{24}$$

Then, the performance improvement for subcarrier n_0 transferring from k_0 to k^* can be calculated by (14).

Like **Step 4** in Section 3.1, for each subcarrier $n \in \Omega - \Omega_{k^*}$, we calculate the performance improvement if it were assigned to user k^* by (14). Then we find the subcarrier n^* which achieves the maximal performance improvement by (17).

If $\Delta(n^*) > 0$, we take the subcarrier n^* from user k_0 and assign it to user k^*. So Ω_{k^*} is updated to be $\Omega_{k^*} \cup \{n^*\}$, and Ω_{k_0} is updated to be $\Omega_{k_0} - \{n^*\}$.

If $\Delta(n^*) < 0$, we remove user k^* from the set Θ. So Θ is updated to be $\Theta - \{k^*\}$.

Then, we go back to **Step 2** for the next iteration. The algorithm stops when there is only one user in the set Θ.

After iterations, the subcarrier set Ω_k^{**} of each user is determined. Hence, we obtain the subcarrier assignment decision $\rho_{k,n}^{**}$. Then, water filling is utilized to allocate user k's total power P_k on the assigned subcarrier set. Thus, we obtain the power distribution decision $p_{k,n}^{**}$. Getting the transmission capacity $C_k\left(\Omega_k^{**}\right)$ of each user by (10), we determine the packet scheduling $u_{k,m}^{**}$ as:

- If $C_k\left(\Omega_k^{**}\right) < R_{k,(1)}$, then $u_{k,(i)}^{**} = 0$ for $1 \le i \le M_k$
- If $\sum_{i=1}^{m} R_{k,(i)} \le C_k\left(\Omega_k^{**}\right) < \sum_{i=1}^{m+1} R_{k,(i)}$, then $u_{k,(i)}^{**} = 1$ for $1 \le i \le m$ and $u_{k,(i)}^{**} = 0$ otherwise.
- If $C_k\left(\Omega_k^{**}\right) \ge \sum_{i=1}^{M_k} R_{k,(i)}$, then $u_{k,(i)}^{**} = 1$ for $1 \le i \le M_k$.

Then, we have the final subcarrier assignment $\rho_{k,n}^{**}$, power distribution $p_{k,n}^{**}$ and packet scheduling decision $u_{k,m}^{**}$.

4 Baseline algorithms

In this section, we introduce four baseline algorithms which are used as comparisons. The first baseline algorithm is the rate adaptation algorithm which is based on the one proposed in Section 3.1. The second one is based on the algorithm proposed in [24].The other two algorithms are designed based on the baselines in [25], one of which uses only application layer RD information and the other of which uses only physical layer CSI. For each baseline algorithm, the packet scheduling method is applied after encoding to determine which packets will be dropped.

4.1 Rate Adaptation (RA) Algorithm

For the RA Algorithm, video rate allocation, power distribution and subcarrier assignment are performed as proposed in Section 3.1. With the estimated parameters $a_{k,m}^{est}$ and $b_{k,m}^{est}$ of the application layer R-D model and CSI of the physical layer, we can acquire the target number of encoding bits $R_{k,m}$ for each video packet by the RA algorithm. The power distribution $p_{k,n}$ and subcarrier set Ω_k of each user can also be determined.

After encoding, the actual number of bits for some video packets may exceed the target. Thus, for the video packets in the transmission queue of user k, the packet scheduling method is performed as follows.

The video packet size $R_{k,m}$ and the quality contribution $Q_{k,m}$ are obtained after encoding. The $Q_{k,m}$ values will be sorted in descending order as $\left\{Q_{k,(1)}, Q_{k,(2)}, ..., Q_{k,(M_k)}\right\}$.

For $1 \leq m \leq M_k$, calculate $\sum_{i=1}^{m} R_{k,(i)}$ for each m.

- If $\sum_{i=1}^{M_k} R_{k,(i)} \leq C_k\,(\Omega_k)$, every packet in the queue can be successfully transmitted.
- However, when $R_{k,(1)} \geq C_k\,(\Omega_k)$, it means all packets have to be dropped.
- When $\sum_{i=1}^{L} R_{k,(i)} \leq C_k\,(\Omega_k) \leq \sum_{i=1}^{L+1} R_{k,(i)}$, the first L video packets can be transmitted while the remaining packets have to be dropped.

4.2 Average PSNR optimization (APO) algorithm

Based on [24], a cross-layer average PSNR optimization algorithm is designed as a comparison of our method. The objective function of this optimization is maximizign the average PSNR of all users, which is equal to minimizing the product of all uses' distortion, as

$$\min_{\underline{\underline{P}}} \prod_{k=1}^{K} MSE_k \tag{25}$$

where $\underline{\underline{P}}$ is the power assignment matrix whose (k, m) entry $P_{k,m}$ is the allocated power for user k on subcarrier m. The optimization period of APO is one GOP, which is set as 4 frames based on the *encoder lowdelay P main.cfg* configuration in HEVC. For each GOP, a rate-distortion model is fitted by four RD samples, which are coded with different rates. In this way, with the power and bandwidth constraint, APO can optimize the target encoding bits and resource allocation for each user. After encoding, if the size of video packets exceeds the transmission capacity, the packets will be sorted and dropped based on the method in Section 4.1.

It should be mentioned that in [24], the author also apply the MU-MIMO uplink transmission system to improve the information rate and video quality. But we remove this part when comparing our algorithm with APO, because APO and our proposed algorithm can get the same promotion if apply MIMO transmission system.

4.3 Application layer optimization (APP) algorithm

We design the baseline application layer algorithm according to "APP Baseline Algorithm" in Section V-A of [25].

The subcarriers are assigned based on the R-D function. Then every subcarrier is regarded as having the same channel gain. Given the same target PSNR value for each video, a corresponding rate R_k can be acquired by their R-D relationships. Therefore the number of subcarriers assigned to user k can be calculated based on the proportion of rate among users, which can be expressed as

$$N_k \sim N \cdot \frac{R_k}{\sum_{i=1}^{K} R_i} \tag{26}$$

where N is the total number of subcarriers and K is the number of users. The subcarriers will be randomly assigned to users according to their target subcarrier number N_k. Thus, the subcarrier set Ω_k is determined.

It should be noted that "APP Baseline Algorithm" in Section V-A of [25] bases the R-D information on the GOP level and is obtained off-line before transmission. Here, we get the R-D function online at the packet level.

Based on CSI, water filling is applied to distribute the power for the subcarriers in Ω_k to get the power allocation $p_{k,n}$ and transmission capacity $C(\Omega_k)$ of each user. Then the target number of encoding bits for each video packet can be acquired by solving the problem in (18). Here CSI is only used to determine the transmission but is not applied when allocating resources. After encoding, the same packet scheduling method as Section 4.1 is applied here.

4.4 Physical layer optimization (PHY) algorithm

The physical layer baseline algorithm is the same as the one in Section V-B of [25]. The subcarriers are assigned to users based on their channel gains $h_{k,n}$ without considering the video contents. Afterwards, we allocate the power for each subcarrier by water filling and we can get the transmission capacity $C(\Omega_k)$ of each user. Then we distribute $C(\Omega_k)$ equally to each video packet m, $1 \leq m \leq M$ of user k as their target number of encoding bits. After encoding, the same packet scheduling method as Section 4.1 is applied here.

5 Numerical results

In this section, we provide numerical simulation results to show the performance of our proposed cross-layer optimization (CLO) algorithm and the baseline algorithms introduced in Section 4. In the simulations, three different videos (*Old town cross, Stockholm, In to tree*) with various contents are used. Each of them has 300 frames in 720p (1280*720) format with frame rate of 25 fps. The video encoder is H.265/HEVC (JCTVC reference software, HM-15.0 [11]). Thus, each frame has 240 CTUs (Coding Tree Units). All frames, except for the first frame, are encoded as P frames. To increase the error resilience, 22 random forced intra CTUs are inserted into each frame. Each frame can be partitioned into one or more slices in the video sequences. With the slice header serving for resynchronization, each slice can be decoded independently at the receiver. In this paper, a slice consists of 10 CTUs, which can be a good balance between error robustness and compression efficiency [18]. After encoding, each slice will be packetized into a transport packet. Therefore, there are 24 packets in one frame. The transport packet would be divided into several data packets before transmission in an OFDMA wireless system based on the restriction of the maximal transmission unit.

In the OFDMA system, the total transmission bandwidth is 5 MHz. The bandwidth is divided into 256 subcarriers. The link between the users and the BS is modeled as a frequency-selective channel that consists of six independent Rayleigh multipaths. The component of each path is calculated by the Clarke's flat-fading model [7]. It is assumed that the power delay profile is exponentially decaying with e^{-2l}, where l is the multipath index. The relative power values of the six multipath components are [0, −8.69, −17.37, −26.06, −34.74, −43.43] (dB) and the power spectral density of AGWN is -70 (dBW/Hz). The maximum Doppler shift is 50 Hz. The maximum available transmission power of the users is 1.0 W.

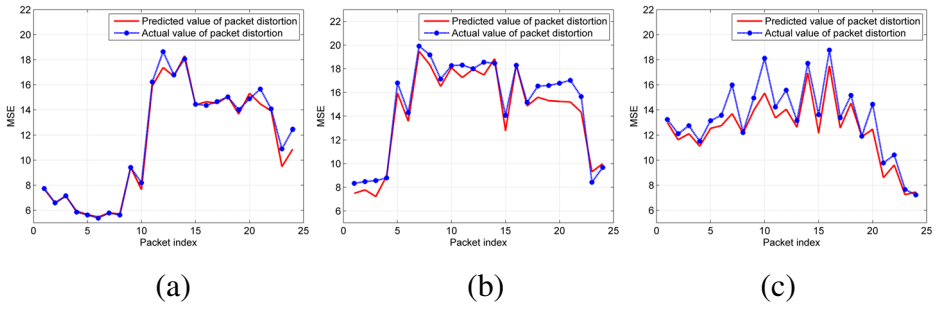


Fig. 6 Comparison of predicted value of packet distortion and actual value of packet distortion. **a** *old town cross*, **b** *stockholm*, **c** *in to tree*

5.1 Performance of rate control and distortion estimation in video coding

In Fig. 6, we show the predicted and actual values of packet distortion for the 100th frame of the sequences "*old town cross*", "*stockholm*", and "*in to tree*". The predicted values are calculated by the R-D function using the actual number of packet bits. The deviations of the predicted distortion from the actual distortion will cause suboptimal wireless resource allocation.

Figure 7 shows the comparison between the target number and the actual number of encoding bits for the 100th frame of the same three sequences. In Fig. 7a and b, the target numbers of encoding bits for most packets are larger than the actual ones. While in Fig. 7c, the actual number of encoding bits exceeds the target ones for some packets. The allocated wireless resource is deficient to transmit these packets. Therefore, it may be useful to assign more subcarriers to transmit these frames. The deviations shown in Figs. 6 and 7 motivate the resource re-allocation process of Section 3.2.

5.2 Performance comparison among different algorithms

Figure 8 shows the average PSNRs for video sequences by using different algorithms. The average PSNR of our proposed CLO is 36.01dB, whereas the average PSNRs of RA, APO, APP, PHY are 35.14 dB, 34.00 dB, 29.57 dB and 33.55 dB, respectively. As expected, the cross-layer algorithms (CLO, RA and APO) outperform other baseline algorithms (APP

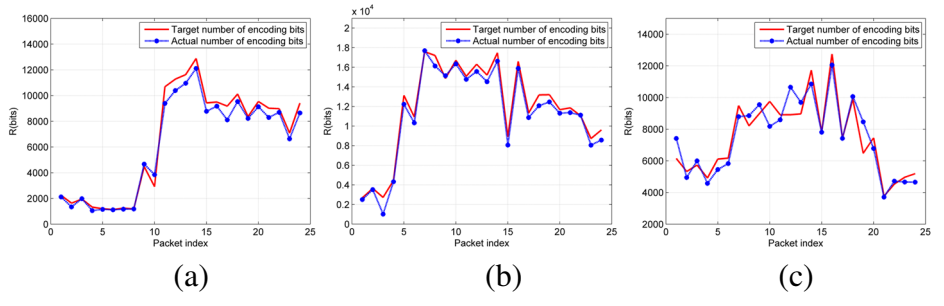


Fig. 7 Comparison of target number of packet encoding bits and actual number of packet encoding bits. **a** *old town cross*, **b** *stockholm*, **c** *in to tree*

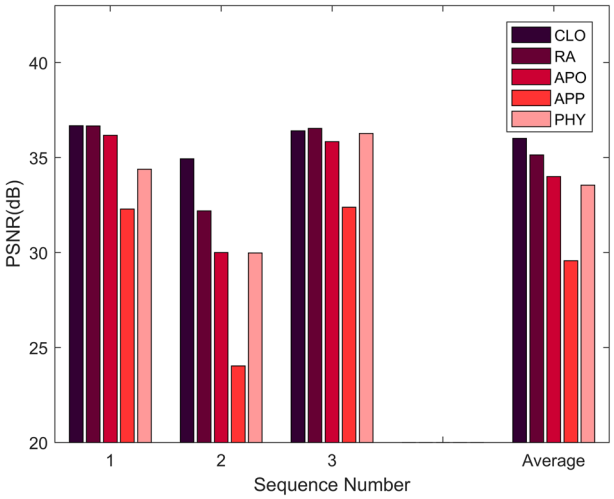


Fig. 8 Average received PSNR for each user for different algorithms. Sequence numbers are represented as follows: 1-*old town cross*; 2-*stockholm*; 3-*in to tree*

and PHY), because the rate adaptation and the resource allocation are determined by jointly taking video contents and channel states into consideration. Meanwhile, the resource re-allocation performs a better subcarrier allocation, which makes CLO perform better than RA and APO. APP tends to assign more subcarriers to users who have complex video contents even if they experience poor channel states. However, this kind of user might only get a small quality improvement at the cost of a large number of subcarriers. For PHY, the subcarriers are assigned to users who have better channel states, which wastes wireless resource for videos with simple contents.

Based on our simulations, the performance gains of CLO are mainly obtained by videos with more complex contents such as *stockholm* (user 2). The C-Q slope of the user with complex video content is the steepest among all three users. In this way, during subcarrier re-assignment, user 2 is more likely to get subcarriers from the other two users. Meanwhile, for users 1 and 3, the C-Q curve slopes are shallow. Therefore, the video qualities of these users will not suffer great degradation when some of their subcarriers are removed. With the resource re-allocation, CLO can obtain a higher quality improvement compared with RA and APO. Among three cross-layer algorithms, APO performs the worst because the optimization period of this algorithm is GOP, the channel condition can be changed during transmission of one GOP. This might cause the waste of transmission resources or packet loss.

Figure 9 shows average PSNRs per frame across the users for different algorithms. For clarity of exposition, we only give the results from the 70th to the 140th frame. We can see that our proposed CLO performs better than other counterparts. APP performs worst among all the algorithms. Moreover, PHY experiences sharp drops in quality several times, especially around the 100th, 112th and 120th frames. It is because PHY assigns subcarriers based on channel states regardless of video contents, which will be strongly influenced by the time-varying characteristic of channels. The average PSNRs per frame of CLO and RA have similar trends with frame index. Due to resource re-allocation, the average PSNR of each frame in CLO is 1.0dB higher than that in RA. APO is more than 1.0dB lower than RA because of the channel variation.

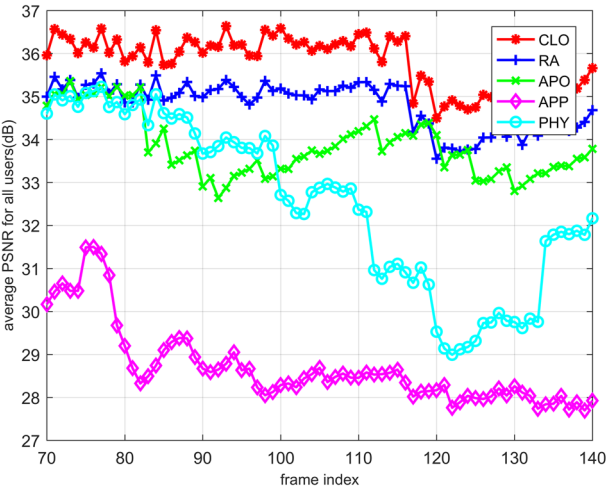


Fig. 9 Frame-by-frame quality over all users for different algorithms

5.3 Comparison of packet loss rates and corresponding average PSNRs

Table 2 demonstrates the results of the packet loss rates and the corresponding average PSNRs. Two interesting phenomena emerge from the results.

First, for CLO and RA, the packet loss rates are 2.91% and 2.52% on average, but the average PSNR of CLO is 0.87dB higher than RA. As the video qualities are the same for CLO and RA after encoding, the quality improvement is entirely brought by the resource re-allocation algorithm. User 2 achieves a 0.51% decrement in packet loss rate and obtains a 2.74dB increment in average PSNR. Meanwhile, the average PSNRs of users 1 and 3 in CLO are approximately the same as those in RA. However, the packet loss rates of users 1 and 3 in CLO are obviously higher than those in RA. This illustrates that the packets of users 1 and 3 make less contribution for quality improvement. In CLO, more wireless resources are re-allocated to the packets of user 2, who can make a large contribution to overall video quality. From Table 2, the performance of APO is worse than CLO and RA. It is caused by

Table 2 Average Packet Loss Rates and PSNRs of three users for different algorithms

		CLO	RA	APO	APP	PHY
User 1	PLR	4.49%	3.08%	4.96%	9.62%	2.95%
	PSNR(dB)	36.67	36.67	36.17	32.29	34.38
User 2	PLR	1.70%	2.21%	3.51%	8.30%	3.91%
	PSNR(dB)	34.94	32.20	30.00	24.03	29.98
User 3	PLR	2.53%	2.28%	2.88%	9.97%	1.92%
	PSNR(dB)	36.41	36.54	35.84	32.39	36.27
Average	PLR	2.91%	2.52%	3.78%	9.29%	2.93%
	PSNR(dB)	36.01	35.14	34.00	29.57	33.55

the high packet loss rate since the optimization period is longer than CLO and RA, thus the change of channel condition cannot be acquired by APO.

For user 1, the packet loss rates of RA and PHY are 3.08% and 2.95%, respectively. Despite the lower packet loss rate, PHY has a 2.29dB decrement in PSNR. Based on the channel states, user 1 in PHY is allocated fewer subcarriers and hence is set a low target number of encoding bits. The encoding distortion of user 1 is high under the low target number of encoding bits. However, for RA, user 1 is allocated more encoding bits by jointly considering the video contents and the channel state, and then the video packets will have lower encoding distortion. In this way, although the packet loss rate is higher, RA also performs better than PHY for user 1.

5.4 Performance comparison under different average SNRs

To verify the robustness of our proposed algorithms, two different scenarios are set for simulation. In the first scenario, the average SNRs for users 1 and 3 are set to be 25 dB, and the average SNR for user 2 is 15 dB. In the second scenario, the average SNRs for users 1 and 3 are 15 dB, and the average SNR for user 2 is 25 dB. The two scenarios correspond to the situations that users have different channel states caused by large-scale fading. As shown in Fig. 10, in both scenarios, CLO performs best in terms of the average PSNRs.

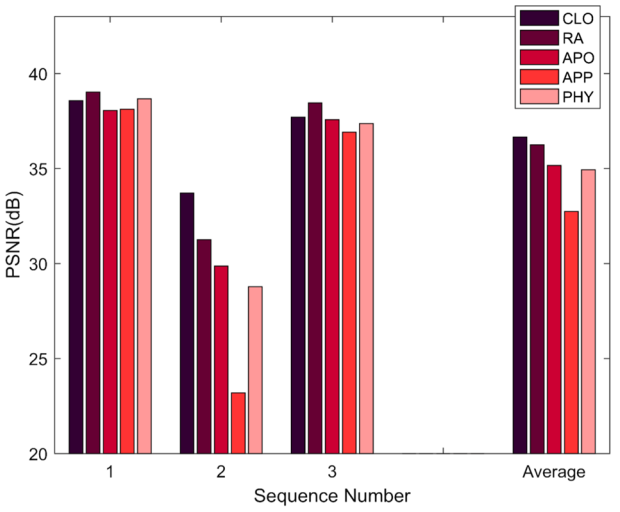
In the first scenario, user 2, who has complex video contents, experiences the worst channel state. For baseline algorithms, the average PSNR obviously decreases, especially for APP. Our proposed CLO produces high PSNR for user 2. Most packets of user 2 have fewer target bits under the low average SNR. Thus the C-Q curve slope is steep. It is opposite for the packets of user 1 and 3. Therefore, many subcarriers will be adjusted to user 2 in resource re-allocation. Thus, CLO achieves higher overall video quality across all users. In the second scenario, CLO gets higher performance for the users with poor channel states, and achieves the highest average PSNR. The simulation results of the two scenarios verify the robustness of CLO.

5.5 Performance Comparison under Different Power Constraints

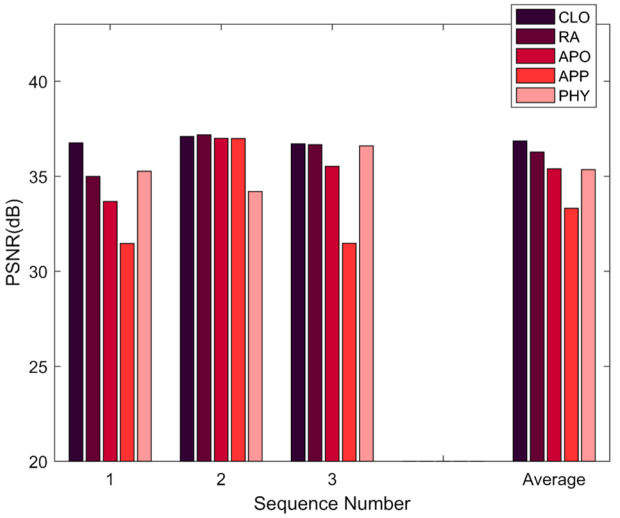
In Fig. 11, the performance of all algorithms with different user power constraints are shown. For CLO, RA, APP and PHY, we find approximately 0.44 dB, 0.65 dB, 1.27 dB and 1.05 dB performance gains respectively when the total powers are 1.0 W, 1.1 W, 1.2 W and 1.3 W. In Fig. 11, we see that CLO and RA achieve a PSNR of 35 dB where the total power constraint is 1W. While APO can achieve a PSNR of 35dB only at the power of 1.3 W, but APP and PHY do not attain that level even when the power is increased to 1.3 W. In uplink OFDMA systems, energy consumption is an important issue for user equipment. The simulation results show that CLO, in which the video contents, channel states and power constraints are exploited sufficiently, requires lower transmission power for a user to achieve the same video quality compared to the baseline algorithms.

5.6 Performance Comparison for Different Video Resolution

To evaluate the robustness of our proposed algorithm, we set the resolution of each user to be different. The video sequences tested here are *Basketball* (1080p), *Parkjoy* (720p) and *BQMall* (480p). Other experiment settings are the same as in Section 5.2. The average PSNR comparison is shown in Fig. 12. We can find that cross-layer algorithms outperforms



(a)



(b)

Fig. 10 Average received PSNR under different average SNRs. **a** Average SNRs for users 1 and 3 are set as 25dB, and the average SNR for user 2 is set as 15 dB. **b** Average SNRs for users 1 and 3 are set as 15 dB, and the average SNR for user 2 is set as 25 dB

single-layer algorithms. Among 3 cross-layer algorithms, our proposed CLO can obtain the highest average PSNR of all three users. Compare with RA, CLO outperform about 2dB for user 1, and only slightly worse for user 2 and 3. Because the video sequence of user 1 is with the highest resolution and complex motion content, thus when the channel condition is bad, the allocated coding rate for user 1 can be insufficient, and the real packet size might exceed

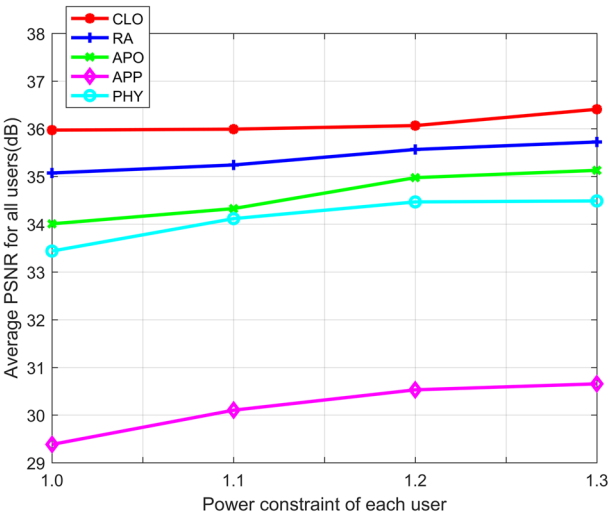


Fig. 11 Average received PSNR versus transmit power

the transmission capacity that allocated in the optimization. So with the re-allocation process in CLO, the packets which have very high utility can acquire more resources from the unimportant packets of user 2 and 3, and transmitted successfully. APO has the worst performance of all three cross-layer algorithms because of the long optimization period, which cause the waste of channel resources or high packet loss rate when the channel fluctuation happens.

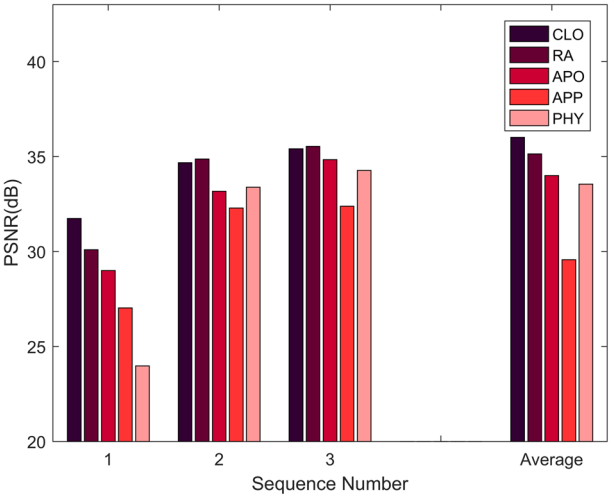


Fig. 12 Average received PSNR for each user with different video resolution for different algorithms. Sequence numbers are represented as follows: 1-*BasketballDrive (1080p)*; 2-*Parkjoy (720p)*; 3-*BQMall (480p)*

6 Conclusion

In this paper, we propose a cross-layer algorithm of joint encoding rate adaptation, packet scheduling and wireless resource allocation for multiuser video communication over uplink OFDMA systems. Based on online content-based rate-distortion optimization, rate adaptation is performed for users under wireless resource constraints. Considering the inaccuracy of the rate control in the encoder, wireless resources are re-allocated based on the actual number of encoded bits after videos are encoded by H.265/HEVC. Iterative algorithms are presented both for video rate adaptation and wireless resource re-allocation. Simulation results verify the performance of our proposed algorithm. Compared to the baseline algorithms, our proposed CLO outperforms by 0.87–6.44 dB in terms of average received PSNR, and is robust in scenarios with various channel states and different total power constraints.

Acknowledgements This research work was supported in part by the National Science Foundation of China Project No.61671365, and Joint Foundation of Ministry of Education of China No.6141A02022344.

Appendix: R-D parameter updating

We use the LMS method to derive the update formulation of the R-D parameters (4). For simplicity, we use D^{est}, a, b and r to represent $D^{est}_{k,m}$, $a^{est}_{k,m}$, $b^{est}_{k,m}$ and $R_{k,m}/I_{k,m}$, respectively. We take the logarithm for both side of (4).

$$\begin{aligned} & D^{est} = a \cdot r^{-b} \\ & \Rightarrow \ln D^{est} = \ln a - b \cdot \ln r \triangleq a' - b \cdot \ln r \end{aligned} \tag{27}$$

where $a' = \ln a$. The squared error between the predicted distortion D^{est} and actual distortion D^{act} can be expressed as

$$e^2 = \left(\ln D^{est} - \ln D^{act}\right)^2 \tag{28}$$

Taking the derivatives with respect to a' and b, we have

$$\begin{aligned} \frac{\partial e^2}{\partial a'} &= 2\left(\ln D^{est} - \ln D^{act}\right) \\ \frac{\partial e^2}{\partial b} &= 2\left(\ln D^{est} - \ln D^{act}\right)(-\ln r) \end{aligned} \tag{29}$$

According to the adaptive Least Mean Square (LMS) method

$$a'_{new} = a'_{old} - \delta_a \left(\ln D^{est} - \ln D^{act}\right) \tag{30}$$

where δ_a is the update step of a. Therefore,

$$\begin{aligned} \ln a_{new} &= \ln a_{old} - \delta_a \left(\ln D^{est} - \ln D^{act}\right) \\ a_{new} &= a_{old} e^{-\delta_a \left(\ln D^{est} - \ln D^{act}\right)} \end{aligned} \tag{31}$$

After Taylor's expansion and ignoring high-order terms,

$$\begin{aligned} a_{new} &= a_{old}\left(1 - \delta_a \left(\ln D^{est} - \ln D^{act}\right)\right) \\ &= a_{old} - \delta_a \left(\ln D^{est} - \ln D^{act}\right) a_{old} \end{aligned} \tag{32}$$

For b,

$$b_{new} = b_{old} + \delta_b \left(\ln D^{est} - \ln D^{act}\right) \ln r \tag{33}$$

where δ_b is the update step of b.

References

1. Awad MK, Mahinthan V, Mehrjoo M, Shen X, Mark JW (2010) A dual-decomposition-based resource allocation for OFDMA networks with imperfect CSI. IEEE Trans Vehi Technol 59(5):2394–2403
2. Biagioni A, Fantacci R, Marabissi D, Tarchi H (2009) Adaptive subcarrier allocation schemes for wireless OFDMA systems in WiMAX networks. IEEE J Sel Areas Commun 27(2):217–225
3. Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge
4. Cicalo S, Tralli V (2014) Distortion-fair cross-layer resource allocation for scalable video transmission in OFDMA wireless networks. IEEE Trans Multimed 16(3):848–863
5. Dani MN, AI-Abbasi ZQ, So DKC (2017) Power allocation for layered multicast video streaming in non-orthogonal multiple access system. In: Proceedings of the IEEE GLOBECOM workshops, pp 1–6
6. de la Fuente A, Escudero-Garzás JJ, García-Armada A (2017) Radio resource allocation for multicast services based on multiple video layers. IEEE Trans Broadcast. https://doi.org/10.1109/TBC.2017.2781121
7. Feng T, Field TR (2008) Statistical analysis of mobile radio reception: an extension of Clarke's model. IEEE Trans Commun 56(12):2007–2012
8. Gao L, Cui S (2008) Efficient subcarrier, power, and rate allocation with fairness consideration for OFDMA uplink. IEEE Trans Wireless Commun 7(5):1507–1511
9. He L, Liu G (2014) Quality-driven cross-layer design for H.264/AVC video transmission over OFDMA system. IEEE Trans Wireless Commun 13(12):6768–6782
10. He Z, Cai J, Chen CW (2002) Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding. IEEE Trans Circ Syst Video Technol 12(6):511–523
11. HEVC Test Model 15.0 (2014) [Online]. Available: http://hevc.kw.bbc.co.uk/git/w/jctvc-tmuc.git
12. JCTVC-K0103 RatecontrolbyR-lambdamodelforHEVC Shanghai (2012)
13. Ji X, Huang J, Chiang M, Lafruit G, Catthoor F (2009) Scheduling and resource allocation for SVC streaming over OFDM downlink systems. IEEE Trans Circ Syst Video Technol 19(10):1549–1555
14. Jung YH, Song Q, Kim KH, Cosman P, Milstein L (2017) Cross-layer resource allocation using video slice header information for wireless transmission over LTE. IEEE Trans Circ Syst Video Technol. https://doi.org/10.1109/TCSVT.2017.2701503
15. Kim K, Han Y, Kim S-L (2005) Joint subcarrier and power allocation in uplink OFDMA systems. IEEE Commun Lett 9(6):526–528
16. Le H, Behboodi A, Wolisz A (2015) Quality driven resource allocation for adaptive video streaming in OFDMA uplink. In: Proceedings of the 26th annual international symposium on personal indoor, and mobile radio communications (PIMRC), pp 1277–1282
17. Li F, Liu G, He L (2009) Application-driven cross-layer approaches to video transmission over downlink OFDMA networks. In: Proceedings of the IEEE GLOBECOM workshops, pp 1–6
18. Li F, Ren P, Du Q (2012) Joint packet scheduling and subcarrier assignment for video communication over downlink OFDMA systems. IEEE Trans Vehi Technol 61(6):2753–2767
19. Li F, Fu S, Liu Z, Qian X (2018) A Cost-constrained video quality satisfaction study on mobile devices. IEEE Trans Multimed 20(5):1154–1168
20. Ng CY, Sung CW (2008) Low complexity subcarrier and power allocation for utility maximization in uplink OFDMA systems. IEEE Trans Wireless Commun 7(5):1667–1675
21. Qian L, Cheng Z, Fang Z, Ding L, Yang F, Huang W (2017) A QoE-driven encoder adaptation scheme for multi-user video streaming in wireless networks. IEEE Trans Broadcast 63(1):20–31
22. Rohling M, May T, Bruninghaus K, Grunheid R (1999) Broadband OFDM radio transmission for multimedia application. Proc IEEE 87(10):1778–1789
23. Sabir MF, Heath RW, Bovik AC (2009) Joint source-channel distortion modeling for MPEG-4 video. IEEE Trans Image Process 18(1):90–105
24. Tseng S-M, Chen Y-F (2018) Average PSNR optimized cross layer user grouping and resource allocation for uplink MU-MIMO OFDMA video communications. IEEE Access 6:50559–50571
25. Wang D, Toni L, Cosman PC, Milstein LB (2013) Uplink resource management for multiuser OFDM video transmission systems: analysis and algorithm design. IEEE Trans Commun 61(5):2060–2073
26. Wu D, Yu D, Cai Y (2008) Subcarrier and power allocation in uplink OFDMA systems based on game theory. In: Proceedings of the IEEE international conference on neural networks and signal processing, pp 522–526
27. Wu P-H, Huang C-W, Hwang J-N, Pyun J-Y, Zhang J (2015) Video-quality-driven resource allocation for real-time surveillance video uplinking over OFDMA-based wireless networks. IEEE Trans Veh Technol 64(7):3233–3246

28. Zhang R, Regunathan SL, Rose K (2000) Video encoding with optimal inter/intra-mode switching for packet loss resilience. IEEE J Sel Areas Commun 18(6):966–976
29. Zhang H, Ma Y, Yuan D, Chen HH (2011) Quality-of-service driven power and subcarrier allocation policy for vehicular communication networks. IEEE J Sel Areas Commun 29(1):197–206
30. Zhang Z, Liu D, Wang X (2018) Joint carrier matching and power allocation for wireless video with general distortion measure. IEEE Trans Mobile Comput 17(3):577–589
31. Zhou N, Zhu X, Huang Y, Lin H (2010) Low-complexity cross-layer design with packet-dependent scheduling for heterogeneous traffic in multiuser OFDM systems. IEEE Trans Wireless Commun 9(6):1912–1923

Fan Li obtained his B.S. and Ph.D. degrees in information engineering from Xi'an Jiaotong University, Xi'an, China, in 2003 and 2010, respectively. From 2017 to 2018, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of California, San Diego. He is currently a Professor with the School of Electronic and Information Engineering, Xi'an Jiaotong University. He has published more than 30 technical papers. His research interests include multimedia communication and video quality assessment. He served as the Local Chair for ICST Wicon 2011, and was a member of the Organizing Committee for IET VIE 2008.

Taiyu Wang obtained his B.S degrees in Information Engineering from Xi'an Jiaotong University, Xi'an, China, in 2015. He is currently a Ph.D. candidate at the School of Electronic and Information Engineering, Xi'an Jiaotong University. His research interests mainly focus on video compression and communication.

Pamela C. Cosman obtained her B.S. with Honor in Electrical Engineering from the California Institute of Technology in 1987, and her Ph.D. in Electrical Engineering from Stanford University in 1993. Following a postdoc at the University of Minnesota, in 1995, she joined the department of Electrical and Computer Engineering at the University of California, San Diego, where she is currently a Professor. Her past administrative positions include Director of the Center for Wireless Communications (2006–2008), ECE Department Vice Chair (2011–2014), and Associate Dean for Students (2013–2016). Her research interests are in image and video compression and processing, and wireless communications. Dr. Cosman's awards include the ECE Departmental Graduate Teaching Award, NSF Career Award, Powell Faculty Fellowship, Globecom 2008 Best Paper Award, HISB 2012 Best Poster Award, 2016 UC San Diego Affirmative Action and Diversity Award, and 2017 Athena Pinnacle Award. She was an associate editor of the IEEE Communications Letters, and of the IEEE Signal Processing Letters. She served as the Editor-in-Chief (2006–2009) and a Senior Editor of the IEEE Journal on Selected Areas in Communications, and as Technical Program Co-Chair of ICME 2018. She is a Fellow of the IEEE.