

## Research Article

# Multi-Session Multicasting for 360-Degree Video Multicast over OFDMA Systems

Jounsup Park 

Department of Electrical Engineering, University of Texas at Tyler, Tyler TX 75799, USA

Correspondence should be addressed to Jounsup Park; [jpark@uttyler.edu](mailto:jpark@uttyler.edu)

Received 8 January 2021; Revised 22 April 2021; Accepted 26 May 2021; Published 16 June 2021

Academic Editor: Marco Rocchetti

Copyright © 2021 Jounsup Park. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

360-degree video content provides a rich and immersive multimedia experience to viewers by allowing viewers to view the video from any angle. However, 360-degree videos require much higher bandwidth to be delivered over mobile networks compared to conventional videos. Multicasting of the videos is one of the solutions to efficiently utilize the limited bandwidth since many viewers share the wireless spectrum resource for popular videos, such as sports events or musical concerts. LTE eMBMS assigns the videos to the video sessions, and multiple viewers can subscribe to the same video allocated to the video sessions. Moreover, the tiling of the 360-degree video makes it possible to control the regional quality of the video. The tiles that are likely to be seen by many viewers should have higher quality than other tiles to satisfy more viewers. In this paper, we proposed the Multi-Session Multicast (MSM) system to optimally allocate the wireless resources to tiles with different qualities to maximize the expected user experience. The experimental results show that the proposed MSM system provides higher quality videos to viewers using limited wireless resources.

## 1. Introduction

Virtual Reality (VR) is growing more popular these days, and people can enjoy more realistic experiences with VR systems [1]. VR allows people to look around a virtual world and feel like they are in the environment. 360-degree video streaming is a key technology in the implementation of VR applications. However, streaming 360-degree videos to mobile devices is a challenging task. 360-degree videos need much higher resolution than conventional videos; therefore, providing viewers with satisfactory 360-degree videos is much more difficult. Viewers cannot watch the whole video frame at the same time; instead, they can only focus on the area that they want to watch, which is called *a viewport or a view*, and the viewport is usually only 20% of whole video [2]. In fact, 4-6 times more resolution is required for 360-degree videos to provide the same experience as conventional videos. On the other hand, there is potential for saving bandwidth, because 80% of the video is unseen by the user at a given time. In an ideal case, we could save 80% of the bandwidth, but in practice, we still need to transmit redundant areas of the video because it is difficult to predict how a user's viewport will change [3].

A multicast system helps to meet the bandwidth requirement of 360-degree video streaming by sharing the same spectrum with many viewers requesting the same video [4]. LTE-evolved Multimedia Broadcast Multicast Services (eMBMS) allow the utilization of up to 60% of their resources for the multicast [5]. An eNodeB (eNB) allocates the resources for multicast video sessions, and users subscribe to whatever videos they want with their user equipment (UE). In the case of a unicast, eNB must allocate the resource for every UE. The resource usage of multicasting is not a function of the number of users but of the number of video sessions. It helps save the spectrum on the LTE network when many users request the same video.

The dynamic adaptive streaming over HTTP (DASH) [6] multicast system [7–9] is applied to efficiently utilize the limited resources and provide better videos to users. The DASH multicast system allocates multiple copies of the same video with different qualities to satisfy more users, but it inevitably generates redundant data that decreases the spectral efficiency. Especially in the case of 360-degree videos, most of the area is not visible to the

users. Therefore, more redundant data are transmitted than in conventional videos if we directly use the DASH multicast for 360-degree video dissemination. To be more efficient, redundant data should be reduced, and tiled video [10] should allow the flexible allocation of fewer bits to the redundant parts of the video.

In a tiled video scheme, the 360-degree video is divided into tiles that can be encoded independently. There can also be multiple copies of the same tile with different representation qualities. These tiles are transmitted through a wireless channel. DASH can deliver the tiles along with a spatial relationship descriptor (SRD) [11], in addition to the media presentation descriptor (MPD), for describing the projection types and spatial relationships among the tiles. SRD includes the region-wise quality of rectangular videos within the projected frame, and an MPD includes the size of video chunks, location of the files, and the codec information.

In this paper, we propose Multi-Session Multicasting (MSM) for tiled-video allocation on a multicast system and the cross-layer optimization framework for QoE optimization of the system. We formulate the tiled-video multicast problem as a mathematical optimization task, which allows us to find a solution for the wireless resource allocation, user grouping, and tiled-video rate selections for utility/QoE maximization. The existing resource allocation algorithms and rate-selection algorithms can only support conventional video multicasting and tiled-video rate selection independently. However, separate solutions are not efficient for 360-degree video multicasting systems. In this research, we aim to design a video multicast system that is suitable for delivering 360-degree tiled videos. The proposed cross-layer optimization framework and algorithms jointly optimize operational components in VR video multicast systems with reasonable complexity. The framework includes algorithms to do the user grouping, wireless-resource allocation, and tiled-video rate selections. The algorithms use the tiled video encoded by the conventional video encoder as their source for the 360-degree video, and LTE eMBMS is used as the wireless video multicast system. Simulation results show that the proposed algorithms can achieve a better utility, which quantitatively denotes the Quality of Experience (QoE), than other existing algorithms. Here is the summary of the contributions of the paper.

- (1) MSM is proposed to efficiently allocate tiled videos on multicast systems, such as eMBMS
- (2) A cross-layer optimization framework is proposed to jointly optimize the algorithms for grouping users, allocating wireless resources, and selecting the tiled-video rates to achieve the best utility value and visual quality
- (3) The spectral efficiency as a function of user grouping was derived to find an efficient and effective grouping algorithm that reduces the number of parameters to be optimized

- (4) A convex optimization method is applied to allocate optimal resources for each multicasting session

A preliminary version of this work appeared in [12]. In addition to giving a more detailed description of our proposed method, the main addition of the content in this paper includes 1) performance comparisons in different user distributions resulting in different numbers of groups 2) the performance gap by opening the multiple multicast sessions, 3) the curve fitting results of the resource-utility curve that allow us to apply the convex optimization method 4) more in-depth discussions and analyses on the proposed method.

This paper is organized as follows. Section 2 summarizes the related work. Section 3 presents our Multi-Session Multicasting (MSM) system for 360-degree video multicast. Section 4 shows the problem formulation to optimize the MSM system. Section 5 introduces the cross-layer optimization algorithms for 360-degree video multicasting over LTE system. Section 6 presents a performance analysis, and Section 7 concludes the paper.

## 2. Related Work

This section summarizes the related work. Since we aim to design a multicasting system for tiled media, we first review the 360-degree video tiling and rate adaptation schemes. Second, we reviewed existing video multicasting systems and discuss drawbacks of existing systems. Finally, the user experience is discussed to reveal the design criteria of the 360-degree video streaming systems.

*2.1. Tiled 360-Degree Video-Streaming Systems.* The most popular and promising technology for controlling regional quality of 360-degree videos is the use of tiles [13]. The tiling scheme has been used for panoramic interactive videos [14], since the interactive video allows users to change their view, and users cannot watch the whole video at once. A 360-degree video is divided into smaller rectangular videos (tiles), and each video is encoded independently using conventional video encoders. OpTile [15] is introduced to optimally divide the video, but it is not practical for real-time streaming systems because of its longer processing time. In many practical systems, the 360-degree videos are divided into same-size tiles to make the smaller videos. Every tile has multiple copies with different encoding rates. Different representations of the tiles are transmitted as users' viewport changes and network channel conditions changes.

There are simple rate allocation algorithms for tiled videos: Binary, Thumbnail, and Pyramid [16]. Binary allocates higher representations on the visible tiles, and non-visible tiles have the lowest representations to save the bandwidth. It is the most bandwidth efficient way to allocate the bits, but users can easily watch the lowest quality when they move their viewport since the network has latency to respond with viewport changes. Thumbnail allocates the minimum bits for lowest representations to the whole video as the background video, and the remaining bits are allocated to visible tiles for better representations. However, users still can watch the lowest quality background video when they move the

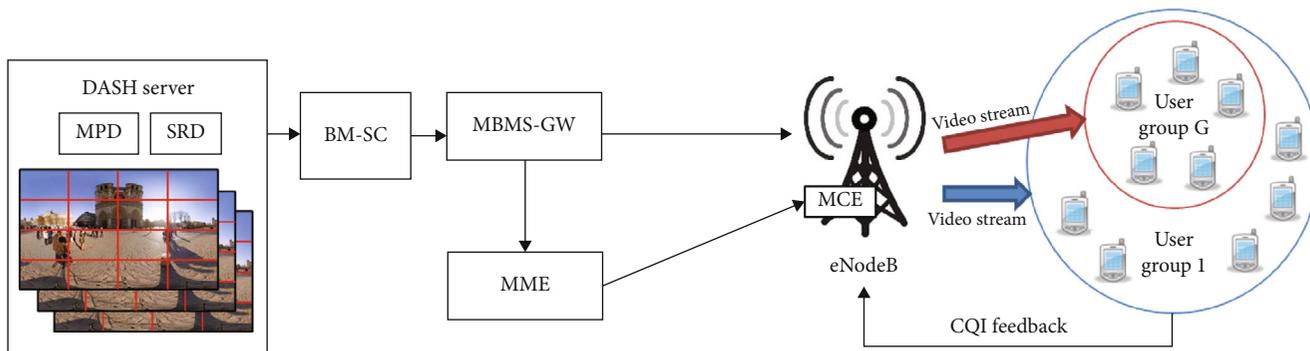


FIGURE 1: 360-degree Video Multicasting over LTE eMBMS.

viewport faster than network latency. The Pyramid algorithm allocates the best representations to visible tiles and gradually lowers the representations for the tiles located far from the viewport. However, these rate allocation algorithms are not network-aware and not flexible enough to provide best quality to users with variable network channel conditions and viewport movement.

Alfaced et al. [17] propose a rate-selection algorithm to provide the best quality to users with a higher representation for the viewport and lower representations for the other tiles. The algorithm allocates the video rates on the tiles based on utility-over-cost ratios. The utility includes the video bitrates and a probability of view. Since it allocates the best representations to tiles to maximize the total utility if there is available network bandwidth, the algorithm can achieve better utility performance than other existing heuristic algorithms.

To improve the performance of 360-degree video streaming systems, a deep learning-based rate adaptation algorithm [18] and a layered video coding scheme [19] were introduced. However, these algorithms are not applicable for multicasting scenarios.

The rate allocation algorithms introduced in [17, 18, 19] for the tiled-video streaming cannot be directly applied to multicast scenarios. Multicast systems divide the users into smaller groups, and these groups have different channel qualities. It makes the total available bitrate, which is used as resource constraint to solve the optimization problem in [17–19], change with its grouping strategy and the wireless resource allocation algorithm. Therefore, the efficient user grouping and the wireless resource allocation algorithms that can work together with the tiled video rate selection algorithm are needed to optimize the 360-degree video multicast systems.

**2.2. Multicasting over LTE eMBMS Systems.** LTE supports multicasting of video streams by eMBMS [5] systems. Figure 1 describes 360-degree video multicast systems using LTE eMBMS systems. Broadcast Multicast Service Center (BM-SC) is responsible for managing multicast sessions. It provides membership, session and transmission, proxy and transport, service announcement, security, and content synchronization. An MBMS gateway (MBMS-GW) distributes the video data to the eNBs. It performs session control signaling towards the mobile management entity (MME). Multi-

cell/multicast coordination entities (MCEs) are part of eNBs, and they provide admission control. They allocate the radio resource to the multicast sessions and decide modulation and coding scheme (MCS). Multiple video multicasting sessions can thus be created, and users can subscribe to those sessions at the same time.

The physical layer of an LTE downlink is based on the OFDMA technology, and the basic unit of the resource in the LTE system is a physical Resource Block (RB), which has 180 KHz bandwidth with 12 subcarriers and 7 symbols [20]. Within an RB, the same Modulation and Coding Scheme (MCS) is applied for all subcarriers. Therefore, if we define the MCS of an RB, there is corresponding number of bits that one RB can carry.

DASH or scalable video coding (SVC)-based multicast algorithms have been introduced to efficiently utilize the limited resources and give more users better video quality [21, 22]. Park et al. [8] show that the total utility can be improved, and more users can watch better video by using DASH multicast over LTE. This algorithm allocates one video representation to one multicast video session; therefore, there is corresponding video quality when the resource is allocated to the video sessions. However, in case of tiled 360-degree videos, multiple tiles share the resource, and many combinations of tiles with different representations may be allocated in a single multicast video session. Therefore, the video quality not only depends on the allocated resource but also on the tile-based rate-selection algorithm. In this paper, resource allocation and user-grouping algorithms are proposed together with the tile-based rate-selection algorithm to optimize the system.

There are two possible ways to do 360-degree video multicasting. The multicasting featured by grouping the users to share the same resource. First, users with the same view can be grouped into a multicast group. The number of multicast groups is the same as the number of views [23, 24]. Some resources can be saved by sharing the same view with many users, but we cannot take advantage of using a multicast scheme when users have different channel quality. All the multicasting groups will suffer with the user who has very bad channel quality. Moreover, all the users eventually need to receive all the tiles because there is latency between the server and the client which is difficult to overcome. Second, users can be grouped by their channel quality [8, 21, 22]. This

grouping strategy helps to select more efficient MCS and application layer forward error correction (AL-FEC) code rate to allocate better videos [25]. As the number of users joining the group with better video increases, total utility is also improved. Therefore, we have designed the multicast systems based on the second scheme that groups the users by their channel quality.

*2.3. User Experience.* User experience of streamed multimedia is determined by the resolution of video, loading delay, and stall events. These factors contribute user experience in different ways and there are many efforts to quantify the user experience [2, 21]. However, recent literature shows that user experience of multimedia differs by individual [26]. Content itself can also affect multimedia experience [27]. Moreover, it is much more difficult to define user experience in mathematical form for 360-degree videos since they are interactive media. Every individual can have different experience while watching the 360-degree videos depending on their behaviour during watching the video. Machine learning-based Quality of Experience (QoE) metric was introduced, but it requires a human assistance to quantify QoE [28].

The 360-degree video multicast system should aim to maximize average utility using limited resources. It is not a practical goal trying to satisfy every single user. We know the user experience is not linearly proportional to the video rate and is gradually saturated with higher video rates. The logarithmic law [29] is applied to quantify the quality of user experience based on the video data rate. Therefore, in this paper, we define the QoE of the 360-degree video multicast service as an average utility.

### 3. 360-Degree Video Multicast Systems

The clients in a 360-degree video multicast system request video chunks from the server based on MPD and SRD information. The DASH server starts to deliver the tiled-video data. BM-SC creates the multiple video sessions that will deliver the tiled-videos with multiple video representations. BM-SC is also responsible for adding AL-FEC redundant blocks for lost packet recovery. Multiple video multicast sessions are created to deliver multiple 360-degree videos and multiple video representations to different user groups. A video multicast session can contain a single tile or multiple tiles. MBMS-GW passes the video data to the eNBs, and MCE allocates the resource for video sessions and assigns the proper MCS for the resource. Users participate in video sessions, and the users who can participate on the multiple video sessions have chances to choose better representations. eNB receives the CQI feedback information from the UEs to help allocate resource blocks (RB) and choose AL-FEC code rate and MCS for the multicasting sessions.

We can consider two different ways to create video multicast sessions. One is the per-tile multicasting (PTM), that considers the tiles as independent videos, in which each tile has its own resource, and every UE subscribes to all necessary sessions to regenerate the 360-degree video. It needs to create as many multiple multicasting sessions as the number of tiles times the number of representations for a single 360-degree

video. All the possible video representations of all the tiles are available to the users based on their channel quality, and the users regenerate the 360-degree video with the best quality tiles that they can decode. For example, if there are  $T$  tiles and  $M$  representations for each tile total  $T \times M$  multicast sessions can be created. MCS, AL-FEC, and resource for all multicast sessions must be determined to maximize the total utility. The search space to find optimal solution is  $M^T$ . Each user selects one representation for one tile and subscribes to  $T$  multicast sessions to regenerate the 360-degree video. It generates too many control signals, and the complexity of the solution increases with the number of multicast sessions.

The other is the multi-session multicasting (MSM), which creates the same number of multicast sessions as the number of user groups. Each multicast session includes multiple tiles with different qualities. Figure 2 shows an example of MSM system with 3 groups and 3 multicast sessions. It first creates a multicast session and allocates most of the tiles with lower representations to all the users. Since the first video multicast session chooses lower MCS index and more redundant AL-FEC packets with lower efficiency, all users can subscribe to the first multicast session. After assigning the first multicast session, better representations for more important tiles can be allocated on additional multicast sessions using the remaining resources. This multicast session uses higher MCS index and less redundant AL-FEC packets than the first multicast session for groups of users with better channel quality. The users in group 2 can subscribe to both the first and the second multicast sessions. Therefore, a user in group 2 can decode multiple representations for every single tile and can choose better representations to play. UEs in group 3 have very good channel quality; therefore, higher video representations can be assigned on multicast session 3 using a smaller number of resources.

The difference between a multicast session and a multicast group is that a multicast session denotes a video session that uses the radio resource controlled by the MCE, while a multicast group denotes a set of users grouped by their channel conditions and subscribing to the same video. Note that users can subscribe to multiple multicast sessions at the same time; therefore, the number of multicast sessions and the number of multicast groups are not necessarily the same. The multicast groups are arranged based on the channel condition, and the user groups with high channel quality can take advantages of subscribing to multiple multicast sessions. In MSM scheme, we only consider opening the same number of multicast sessions as the number of multicast groups.

There are six parameters to optimize in the MSM system for 360-degree video multicast, including the number of multicast sessions for the video, user groups, resource allocations, AL-FEC code rate, MCS index, and video data rate of tiles. These parameters are mathematically formulated to maximize the total utility.

### 4. Problem Formulation

The mathematical expressions of Quality of Experience (QoE), Spectral Efficiency (SE), and utility maximization

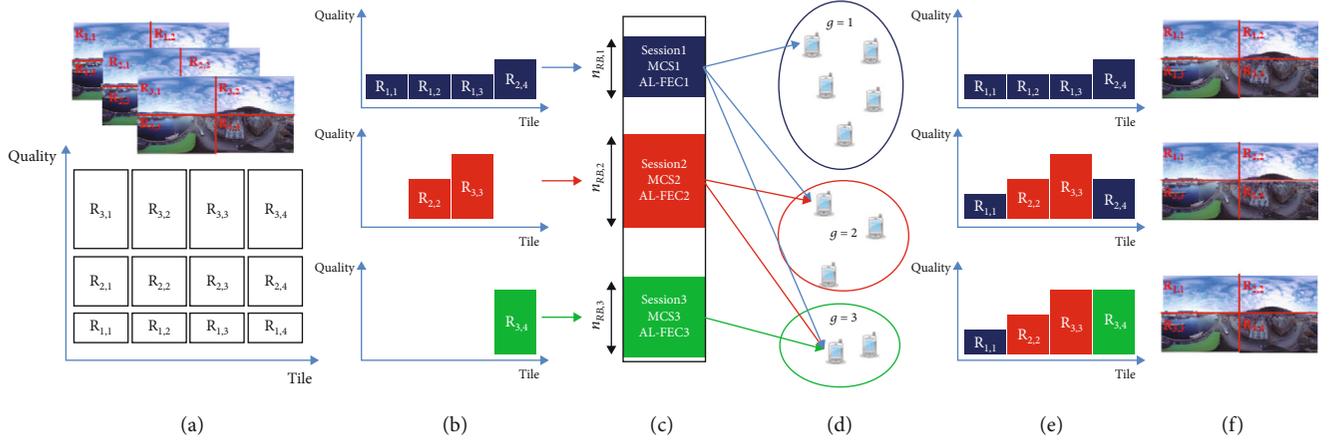


FIGURE 2: Multisession multicasting (MSM): (a) tiled videos with multiple representations; (b) rate selection; (c) resource allocation; (d) user grouping; (e) selected tiles; (f) users' scenes.

problem are described in this section. The QoE to be maximized is modeled as utility that describes the expected QoE when the video is delivered to the users. The SE is derived based on grouping results. Since MCS and AL-FEC directly affect the SE, we can reduce the number of control parameters by formulating the relationship between MCS, AL-FEC, and SE. Utility maximization problem is formulated using the widely accepted QoE model and SE with wireless resource constraints. Table 1. includes the notations used in this paper.

**4.1. QoE Model.** It is known that the user experience is not linearly proportional to the video rate and is gradually becoming saturated with higher video rates. In this paper, the well accepted logarithmic law [29] is considered to quantify the quality of user experience based on the video data rate:

$$u(r) = \begin{cases} \alpha \log \beta \frac{r}{R_M}, & r > 0 \\ 0, & r = 0 \end{cases}, \quad (1)$$

where  $r \in \{0, R_1, R_2, \dots, R_M\}$ ,  $u(r)$  denotes the utility, which is the function of the allocated video rate;  $R_m$  is the video rate of the  $m$ -th video representation;  $R_M$  is the maximum video rate in the server; and  $\alpha$  and  $\beta$  are the normalization coefficients to ensure utility  $u(r)$  staying in the range between 0 and 1. They can be empirically determined for different applications.

Since the tiles are encoded independently with many different qualities, the quality of the received video is represented by a combination of multiple tiles with different qualities. We estimate the utility of a video to quantify the user's experience, which is a weighted combination of multiple tiles. We give the weighting of each tile based on the view probability. Therefore, the total utility of the video that users really receive is

$$\sum_{t=1}^T u(r_t) S_t, \quad (2)$$

TABLE 1: Notations used in this paper.

$c(\text{MCS})$	Capacity [bits/s] of an RB with selected MCS
$f_{mar}$	Margin factor of AL-FEC
$g$	Group index
$n_{RB,g}$	Number of RBs assigned to group $g$
$r$	Video rate
$t$	Tile index
$G$	Number of groups
$K$	Number of data packets in an AL-FEC block
$N$	Number of packets in an AL-FEC block
$\mathbb{N}_g$	Set of UEs in group $g$
$N_{RB}$	Total number of available RB
$Q_{t,g}$	Video representation of tile $t$ for multicast session $g$
$R_M$	Maximum video rate
$S_t$	View probability of tile $t$
$SE_i$	Spectral efficiency [bits/s/RB] of user $i$
$T$	Number of tiles
$U_g(n_{RB})$	Approximated utility of group $g$
$\alpha, \beta$	Normalization coefficients of $u(r)$
$\gamma_{MCS}$	Threshold SNR with selected MCS index
$\gamma_i$	Average SNR in dB scale of UE $i$

where  $S_t$  is the view probability of the tile  $t$ . The utility model implies that we can achieve higher utility value when we allocate more bits to the tiles with higher view probability. One of the ways estimating the view probability is measuring the saliency score of the tiles [30–33]. The higher saliency scores of the tiles imply that people may be more interested in these tiles. For example, the tiles might have moving objects or higher contrast so that the saliency score is higher than other tiles.

The utility model is not restricted with saliency weighting.  $S_t$  can be estimated by other methods such as number

of users who watch the tile. We used saliency score because it does not require feedback information from users, which helps save the spectrum for the feedback information. Moreover, we do not need to consider the latency between the server and the users, which causes the view probability estimation error.

**4.2. Spectral Efficiency (SE).** Spectral efficiency (SE) indicates the actual information bits an RB can carry over the LTE channel. Therefore, redundant bits or packets are not counted as information. The SE is a function of MCS and AL-FEC code rate:

$$SE = c(MCS) \frac{K}{N}, \quad (3)$$

where  $c(MCS)$  is the efficiency [34] of an RB with a specific MCS index, and  $K/N$  is the AL-FEC code rate. The redundant bits or packets also help improve SE, with a certain level of packet-loss rate, and signal-to-noise ratio (SNR) because they help to recover information from lost packets. Using the fountain code, the AL-FEC code rate for successful packet recovery is given as [8, 35]:

$$(1 - P_{\text{out}}(\gamma_{MCS}, \bar{\gamma}_i)) f_{\text{mar}} \geq \frac{K}{N}, \quad (4)$$

where  $P_{\text{out}}(\gamma_{MCS}, \gamma_i)$  is the outage probability of user  $i$  given average SNR  $\gamma_i$ , threshold SNR  $\gamma_{MCS}$ , and selected MCS [36];  $f_{\text{mar}}$  denotes the margin factor for describing the non-ideal AL-FEC decoding capability. The received SNR is modelled as a Log-normal random variable because of exponential effective SNR mapping (EESM) [37] based link error prediction used for the orthogonal frequency domain multiple access (OFDMA) systems. Therefore, the outage probability  $P_{\text{out}}(\gamma_{MCS}, \gamma_i)$  is CDF of a Log-normal distribution. The packet loss rates can be used rather than the outage probability. However, in this paper, we assume that the eNB is unaware of users' application-level information.

We can select an AL-FEC code rate directly from (4) since the most efficient AL-FEC code rate is the largest code rate that satisfies (3), and the SE of user  $i$  becomes

$$SE_i = c(MCS)(1 - P_{\text{out}}(\gamma_{MCS}, \bar{\gamma}_i)) f_{\text{mar}}. \quad (5)$$

Spectral efficiency is a function of MCS and the average SNR of user  $i$ . Every user has his/her own MCS that maximizes the SE, because the efficiency increases with a higher MCS index, but the outage probability also increases. Therefore, the best MCS for user  $i$  is given as

$$MCS^*_i = \arg \max_{MCS} c(MCS)(1 - P_{\text{out}}(\gamma_{MCS}, \bar{\gamma}_i)) f_{\text{mar}}. \quad (6)$$

The spectral efficiencies of the groups depend on how we group them. If multiple users with different SEs are grouped together, the SE of the group is determined by the user with the lowest SE of the group,  $SE_{\text{min},g}$ , because users in the same group share the same resources and watch the same video; thus, the video rate should be low enough to be successfully

delivered to all the users in the group, especially the user with the lowest SE.

**4.3. Utility Maximization Problem.** The purpose of this work is to find the optimal solution to maximize the total utility of the users in the LTE network. Therefore, we sum together all users' expected utility to formulate the problem as

$$Q_1 : \underset{\mathbf{r}}{\text{maximize}} \sum_{i=1}^N \sum_{t=1}^T u(r_{i,t}) S_t, \quad (7)$$

where  $N$  is the number of users,  $T$  is the number of tiles,  $r_{i,t}$  is the video rate that is allocated to the tile  $t$  for user  $i$ , and  $S_t$  is the saliency score of tile  $t$ . Since the saliency score is measured from the original video, it is the same for all users.

To efficiently utilize the spectrum, we can apply the multicasting scenario on the original problem  $Q_1$  by grouping the users based on their channel quality instead of grouping them based on tiles. A multicast session includes multiple tiles, with each tile being encoded with the same MCS and AL-FEC code rate. Multiple multicast sessions can be created with different MCS and AL-FEC code rates. Since a single multicast session can contain multiple tiles, some users can reconstruct a 360-degree video by subscribing only one multicast session. If users can decode the video from multiple multicast sessions, users can choose which video to play. It is a much more efficient way to utilize limited resources when the users need multiple tiles at the same time.

The utility maximization problem of the 360-degree video multicast system can be formulated as follows. Four parameters must be determined to solve the problem: the number of multicasting groups  $G$ , a user grouping assignment  $\mathbb{N}_g$ , the resource allocation  $n_{RB,g}$ , and the rate selection on the tiles:

$$Q_2 : \underset{G, \mathbb{N}_g, n_{RB}, \mathbf{Q}}{\text{maximize}} \sum_{g=1}^G |\mathbb{N}_g| \sum_{t=1}^T u(r_{Q_{t,g}}) S_t, \quad (8)$$

subject to

$$\begin{cases} \sum_{t \in \{\tau \in \{1, \dots, T\} | Q_{\tau,g} > Q_{\tau,1 \dots g-1}\}} r_{Q_{t,g}} \leq n_{RB,g} SE_{\text{min},g}, & g > 1 \\ \sum_{t \in \{1, \dots, T\}} r_{Q_{t,g}} \leq n_{RB,g} SE_{\text{min},g}, & g = 1 \end{cases} \quad (9)$$

$$r_{Q_{t,g}} \in \{0, R_{t,1}, R_{t,2}, \dots, R_{t,M}\} \quad (10)$$

$$\sum_{g=1}^G n_{RB,g} \leq N_{RB} \quad (11)$$

$$n_{RB,g} \geq 0 \quad (12)$$

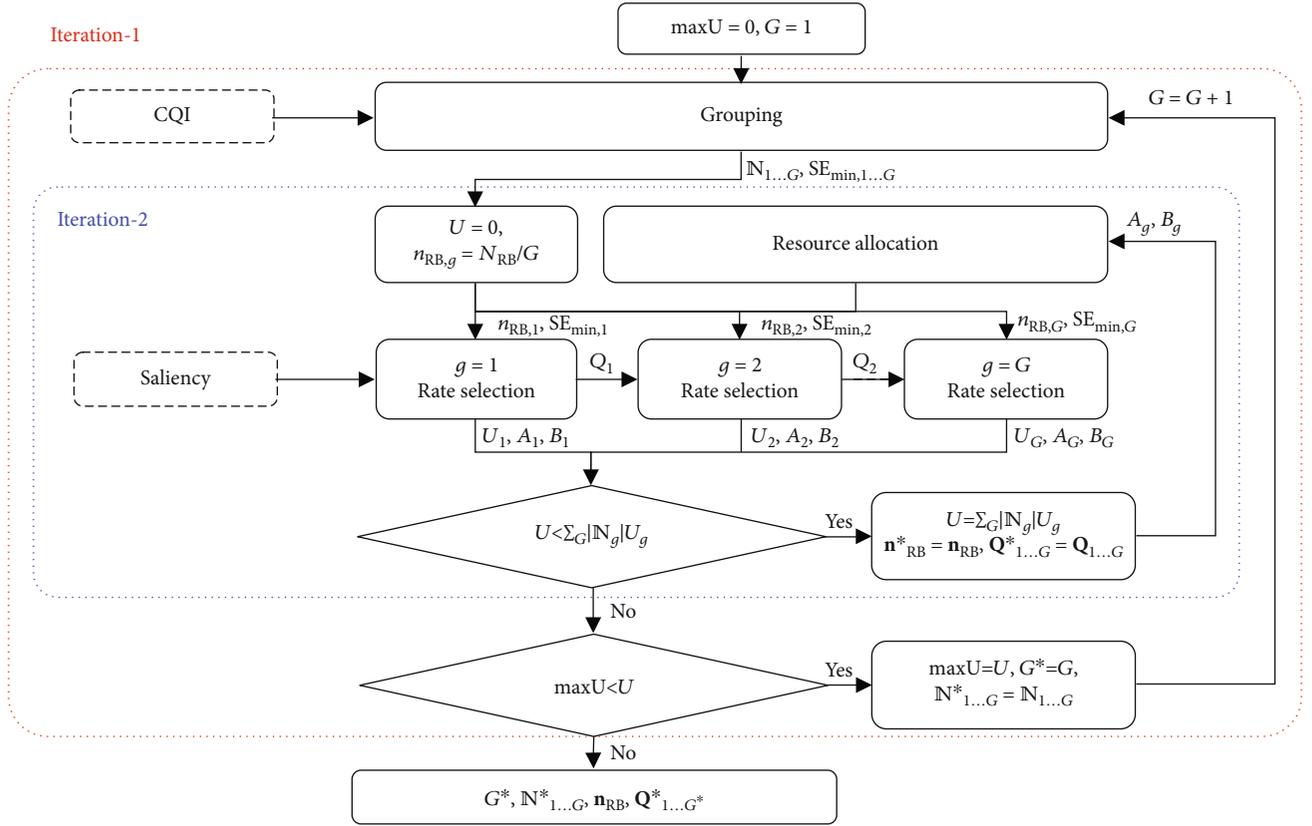


FIGURE 3: Cross-layer optimization framework.

where  $\mathbb{N}_g$  is the set of users joining a group  $g$ ,

$$\mathbb{N}_g = \begin{cases} \{i \in \mathbb{N} \mid SE_{\min,g} \leq SE_i < SE_{\min,g+1}\}, & g < G \\ \{i \in \mathbb{N} \mid SE_{\min,g} \leq SE_i\}, & g = G \end{cases}. \quad (13)$$

$|\mathbb{N}_g|$  is the number of users in the group  $g$ ;  $u(r_{Q_{t,g}})$  is the utility determined by the allocated rate  $r_{Q_{t,g}}$ , and  $Q_{t,g} \in \{1, 2, \dots, M\}$  is the index of the selected video representation for the user group  $g$  and tile  $t$ . The total utility is the summation of the utilities of each group, which is a combination of the tiles' utilities weighted by  $S_t$ , times the number of users in a group. If more people can view better representations on their salient parts, the total utility can achieve a higher value. There are resource constraints (11), (12), e.g., groups share the total resources, and the tiles in a group share the resources allocated to each group.

Four parameters  $G^*$ ,  $\mathbb{N}_{1...G}^*$ ,  $n_{RB}^*$ , and  $\mathbf{Q}_{1...G}^*$  jointly contribute to the total utility. A cross-layer optimization framework is proposed in the next section to maximize the utility by configuring these parameters.

## 5. Cross-Layer Optimization Framework

The goal of a cross-layer resource allocation framework, as shown in Figure 3, is to find the number of the multicast groups ( $G^*$ ), user groups  $\mathbb{N}_g$ , resource allocation  $n_{RB,g}^*$ ,

and rate selection vector  $\mathbf{Q}_g^*$  for all  $g$ . These parameters are found by performing three functional blocks/algorithms, which are (A) a grouping algorithm that decides the number of the multicast group ( $G^*$ ), user groups  $\mathbb{N}_g$ , and the groups' spectral efficiency  $SE_{\min,g}$ , (B) a resource allocation algorithm that decides  $n_{RB,g}^*$  for all  $g$ , and (C) a rate selection algorithm that decides the utility  $U_g$ , the curve fitting results  $A_g, B_g$ , and the tiled-video rate selection result  $\mathbf{Q}_g^*$  for all  $g$ . These functional blocks, which are mutually dependent, jointly work to find the optimal solution based on two-nested loop iterations. The first loop iteration (Iteration-1) finds the best number of groups and their spectral efficiency by measuring utility result  $maxU$ , which is initialized as 0. Iteration-1 starts from  $G=1$  and increases  $G$  until the resulting utility of Iteration-2,  $U$ , is larger than the utility achieved in the previous iteration,  $maxU$ . We can achieve better utility and spectral efficiency by dividing all users into smaller groups, since the groups with better spectral efficiency can achieve better utility using the same amount of resources; but total utility will decrease when there are too many groups because each group can only have a small number of resource blocks. The second iteration (Iteration-2) makes the (B) resource allocation and (C) rate selection algorithm work together to maximize the  $U$  with given grouping results  $G$  and  $SE_{\min,g}$ . Within the first round of Iteration-2, there is no information about the utility as a function of resource allocation; therefore, we cannot perform the optimal resource allocation algorithm directly. The rate selection algorithms (C) perform with resources evenly

allocated to all groups and returns the  $A_g, B_g$  of approximated utility functions.

$$U_g(n_{RB}) = A_g \log(B_g n_{RB}) \quad (14)$$

for all  $g$ . The resource allocation algorithm (B) uses the information  $A_g, B_g$  to find the optimal solution. The resource allocation results are passed into the rate selection algorithms (C) again, and they return the utilities. The summation of utility values is compared to the initial utility value  $U$ , and if it is larger than  $U$ ,  $U$  will be updated with the new utility result. The  $A_g$  and  $B_g$  of the utility function (14) found from the rate allocation algorithm are approximated values; therefore, Iteration-2 continues to find better solutions.

The proposed cross-layer optimization framework optimizes the multicast system every 1 second, which is corresponds to 100-OFDMA frames, and every OFDMA frame contains one AL-FEC block. Therefore,  $K$ -OFDMA frames delivers the data, and  $(N-K)$ -OFDMA frames deliver the redundant data, where  $N=100$ . User groups are also updated every 1 second, therefore, users can change their video quality every one second, and 1 second is consistent with the segment duration defined in DASH. The optimization period is adjustable by controlling the segment duration and  $N$  values. The detailed operations of the functional blocks (A), (B), and (C) are in the following section.

**5.1. Grouping Algorithm.** The goal of the grouping algorithm is to determine the number of multicast groups, the group sizes of all multicast groups, and their spectral efficiencies. The grouping results  $G$ ,  $|\mathbb{N}_g|$ , and  $SE_{\min,g}$  are given to the resource allocation algorithm to find the optimal resource allocation solution.  $G$  is determined by Iteration-1.  $|\mathbb{N}_g|$  is determined by the minimum spectral efficiency of the group  $g$ ,  $SE_{\min,g}$ . An exhaustive search (i.e., trying all possible groupings) can certainly achieve the optimal solution when combined with the optimal resource-allocation algorithm. However, its complexity increases exponentially with the number of groups. Therefore, a simple heuristic search, which maximizes the overall spectral efficiency (MaxSE), is proposed for optimization. Initially, all the users are in group 1; therefore, the spectral efficiency of group 1 is set to the spectral efficiency of the user with the poorest channel quality. Then the users are divided into two groups to maximize the total spectral efficiency. The users in group 2 can be further divided into two groups to generate the third group. This procedure is repeated to generate more user groups. The spectral efficiency of the group  $g$  determined by the MaxSE method is

$$SE_{\min,g} = SE_{k_g} \quad (15)$$

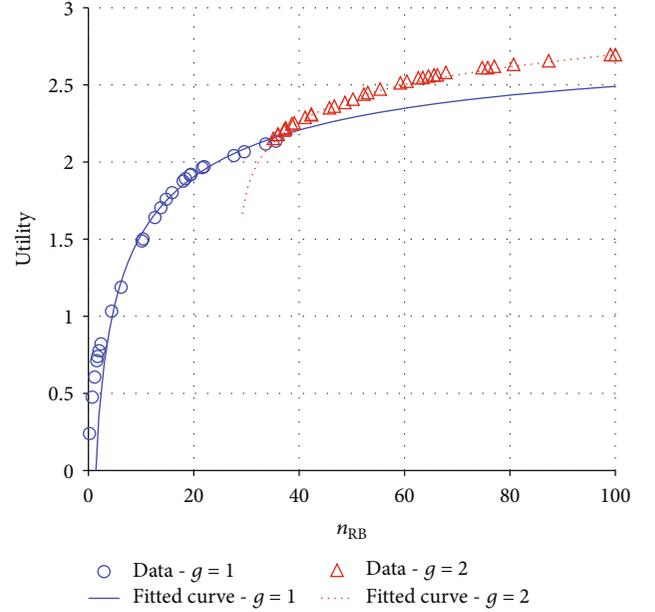


FIGURE 4: Curve fitting results of video utility functions.

where the boundary user index  $k_g$  of group  $g$  is

$$k_g = \begin{cases} \underset{i \in \mathbb{N}}{\operatorname{argmin}}(SE_i), & g = 1 \\ \underset{i \in \{n \in \mathbb{N} | SE_n > SE_{\min,g-1}\}}{\operatorname{argmax}} \{ |\mathbb{N}_{g-1}| SE_{\min,g-1} + |\mathbb{N}_g| SE_i \}, & g > 1 \end{cases} \quad (16)$$

$|\mathbb{N}_g|$  for all  $g$  are determined by (13).

**5.2. Resource Allocation.** The goal of the resource allocation algorithm is to allocate radio resources to each multicast group. The rate-selection algorithm shown in the next section gives the estimated utilities of all the groups, each of which is a function of the allocated number of RBs. Moreover, for every possible grouping solution, an optimal resource-allocation solution exists. Note that the original problem  $Q_2$  is formulated as a convex optimization problem, since the utility functions (14) have convex forms and the constraint is convex:

$$Q_3 : \underset{n_{RB}}{\operatorname{maximize}} \sum_{g=1}^G |\mathbb{N}_g| U_g(n_{RB,g}), \quad (17)$$

subject to (11), (12).

Using the Lagrangian method,

$$L(n_{RB}, \lambda, \zeta) = \sum_{g=1}^G |\mathbb{N}_g| U_g(n_{RB,g}) - \lambda \left( \sum_{g=1}^G n_{RB,g} - N_{RB} \right) + \sum_{g=1}^G \zeta_g n_{RB,g}, \quad (18)$$

INPUTS:  $n_{RB,g}, SE_{\min,g}$   
 OUTPUTS:  $A_g, B_g, Q_{t,g}$   
 1. Make a list of  $\tilde{u}_{t,m}S_t/C_{t,m}$  for all  $(t,m)$  pairs except the pair with  $C_{t,m} = 0$   
 2. Set  $C_{\text{current}} = 0, U_{\text{current}} = 0$   
 3. While  $C_{\text{current}} \leq n_{RB,g}SE_{\min,g}$   
 4. Allocate  $(t, m)$  pair that has maximum  $\tilde{u}_{t,m}S_t/C_{t,m}$ , and exclude the pair from the list  
 5. Mark  $Q_{t,g} = m$   
 6. Update  $C_{\text{current}} = C_{\text{current}} + C_{t,m}$   
 7. Update  $U_{\text{current}} = U_{\text{current}} + \tilde{u}_{t,m}S_t$   
 8. Collect the pairs  
      $(x, y) = ((C_{\text{current}}/SE_{\min,g}), U_{\text{current}})$   
 9. End while  
 10. Do the curve fitting with  $(x, y)$  to find  $A_g$  and  $B_g$  of  $y = A_g \log(B_g x)$

ALGORITHM 1: Rate-selection algorithm.

Gradient of (18) is

$$\frac{\delta L(\mathbf{n}_{RB}, \lambda, \zeta)}{\delta \mathbf{n}_{RB}} = \frac{|\mathbb{N}_g| A_g}{n_{RB,g}} - \lambda + \zeta_g, \quad (19)$$

The optimality condition can thus be derived by using Karush-Kuhn-Tucker (KKT) conditions [38].

- (1) Primal feasibility:  $\sum_{g=1}^G n_{RB,g} = N_{RB}, n_{RB,g} \geq 0$
- (2) Dual feasibility:  $\zeta_g^* \geq 0$ .
- (3) Complementary slackness:  $\zeta_g^* n_{RB,g}^* = 0$ .

The gradient of Lagrangian vanishes when

$$\lambda^* = \frac{|\mathbb{N}_g| A_g}{n_{RB,g}} = f_g(n_{RB,g}) \quad (20)$$

holds for all  $g$ . Therefore, the optimal resource allocation solution is  $n_{RB,g}^*$ , which satisfies (20) for all  $g$ . Actual  $n_{RB,g}^*$  for all  $g$  could be found using bisection search algorithm.

**5.3. Rate Selection.** The goal of the rate selection algorithm is to determine the representation of the tiles and to allocate radio resources for each tile within a multicast group. Each multicast session has its own resource that is assigned by the resource allocation algorithm. In this way, each group should achieve the maximum utility using the assigned resource. We can write the problem as

$$Q_4 : \underset{Q_{g,1 \dots T}}{\text{maximize}} U_g = \sum_{t=1}^T u(r_{Q_{g,t}}) S_t, \quad (21)$$

subject to (9), (10).

where  $n_{RB,g}SE_{\min,g}$  is an achievable rate for group  $g$ . A greedy algorithm is introduced to solve the problem. It makes a list of the total utility per cost (rate) with candidates  $R_{t,m}$  (10). The  $Q_{g,t}$  that can increase the total utility per cost (rate)

TABLE 2: LTE Parameters.

Parameter	Value
Distance attenuation	128.1 + 37.6*log(d), d[km]
Shadow fading	Log-normal, 0 mean, sigma=8
Fast fading	ITU-R PedB
Carrier frequency	2.6GHz
Cellular layout	3 cell sites
System bandwidth	20 MHz
Scheduling frame	10 ms
RB size	12 sub-carrier, 0.5 ms
Sub-carrier spacing	15 kHz
Extended CP duration	16.7 $\mu$ s
TTI	1 ms
EUTRA UE	Antenna gain 0dBi, NF 9 dB
EUTRA eNB	Antenna gain 14dBi, NF 5 dB
eNB transmit power	43 dBm
MIMO configuration	1 Tx, 1 Rx
Thermal noise	-174 dBm/Hz

the most is chosen first and excluded from the list. The algorithm continuously chooses the  $Q_{g,t}$  until all the rate resource is used up. The algorithmic difference between our proposed algorithm and the one introduced in [17] is that ours considers the condition of the lower group (i.e. the user group with the worse channel quality). Since the upper groups (i.e. the user groups with better channel quality) can subscribe to the lower groups' tiles, some of the tiles already have available videos, even though the upper groups do not allocate the video to the tiles. Therefore, an upper group does not need to pay the cost to subscribe to the same representations as a lower group, but it needs to pay its own cost to subscribe to better representations than the lower group. The cost matrix  $[C_{t,m}]$  shows the "marginal" cost to pay for subscribing the  $m$ -th representation on tile  $t$ . After the first representation is selected, the algorithm can improve the representation, and it only needs to pay the difference

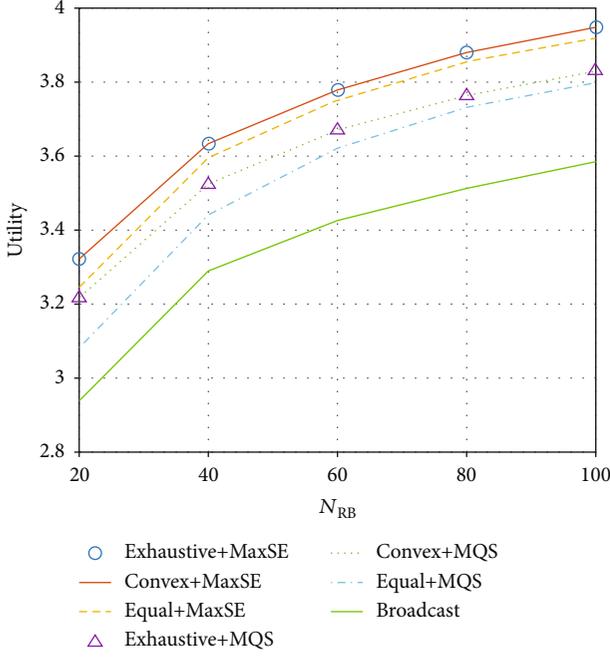


FIGURE 5: Utility.

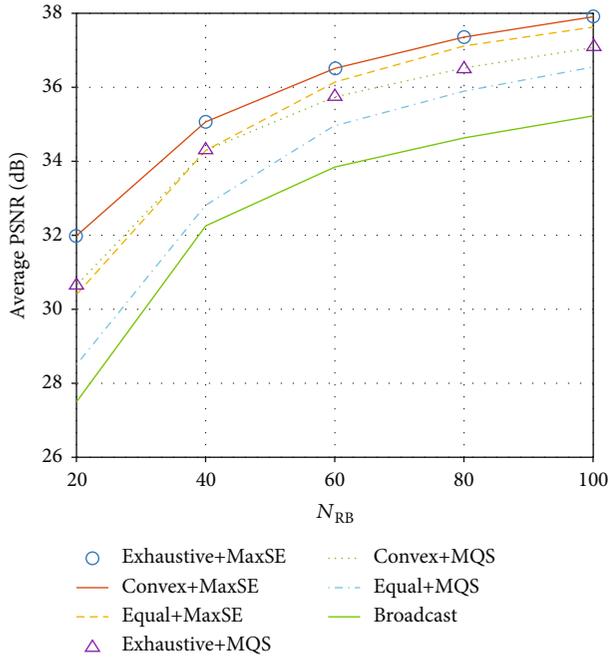


FIGURE 6: Average PSNR.

between the allocated representation and the new representation. Therefore, the cost matrix is

$$C_{t,m} = \begin{cases} R_{t,m} - R_{t,m-1}, & m \geq Q_{t,g-1} + 2 \\ R_{t,m}, & m = Q_{t,g-1} + 1. \\ 0, & m \leq Q_{t,g-1} \end{cases} \quad (22)$$

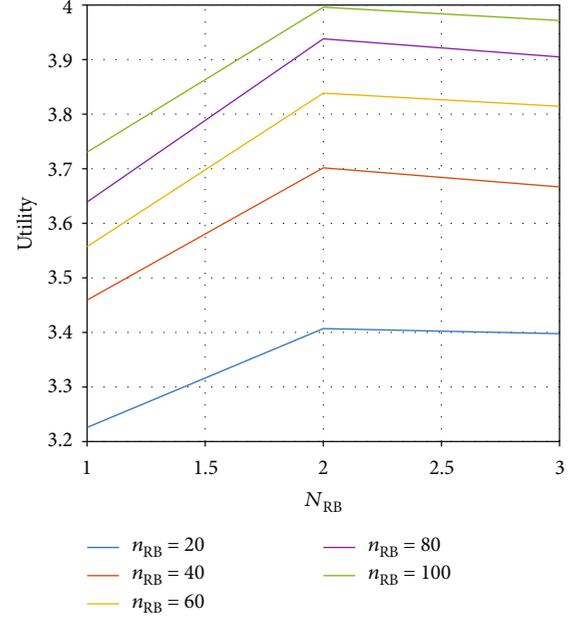


FIGURE 7: Utility with the Number of Groups.

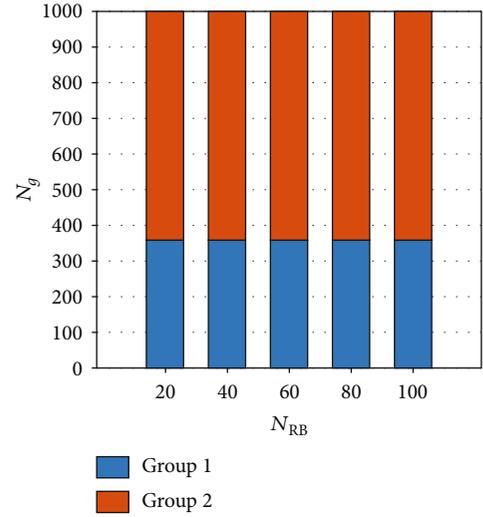


FIGURE 8: Number of users.

where  $Q_{t,g}$  indicates the selected representation index of tile  $t$  for group  $g$ . The utility matrix is given as

$$\tilde{u}_{t,m} = \begin{cases} u(R_{t,m}) - u(R_{t,m-1}), & m \geq 2 \\ u(R_{t,m}), & m = 1 \end{cases} \quad (23)$$

$\tilde{u}_{t,m}$  is the effective utility, which shows how much we can improve the total utility, when the  $m$ -th representation is selected for tile  $t$ .

As the algorithm performs the iteration, it collects the increased utility per consumed resource. We can draw a curve with the data to find the relationship between the resource usage and the utility. The curves differ with the spectral efficiency curves of multicast sessions because every

multicast session has different spectral efficiency, and the utility depends on the condition of other multicast sessions. We have used logarithmic function (14) to model the curves and performed curve fitting to find the parameters  $A_g$  and  $B_g$ . Figure 4. shows an example of the curve fitting results of video utility function under different radio resource allocations when there are two video multicast sessions. The first multicast session ( $g=1$ ) makes the utility increase fast by allocating the lower video representations, but the slope of the curve decreases because it needs more resources to allocate the higher video representations with its lower spectral efficiency. The second multicast session can allocate the tiles with higher video representations more efficiently than the first multicast session; therefore, its utility increases quickly again by allocating the videos to the second multicast session. The gap between the blue curve ( $g=1$ ) and the red curve ( $g=2$ ) shows the utility gain achieved by allocating the tiles to the second multicast session. The slopes of the curves indicate the efficiencies of the multicast sessions; therefore, this information is used for optimal resource allocation. The detailed algorithm is described in Algorithm 1.

## 6. Performance Analysis

**6.1. Simulation Setup.** Performance analysis is conducted using the standard LTE parameters [39, 40] described in Table 2. Exponential effective SNR mapping [41] is used to map the channel state onto an effective SNR. Finally, the effective SNR is mapped onto the MCS table [42], ensuring a block error rate (BLER) value lower than 10%. For propagation loss, the COST 231 suburban model [43] with a standard deviation of lognormal shadowing 8 dB is used. Also, the ITU Ped-B power delay profile is used in our simulations. 1000 UEs are generated and randomly distributed in the cell area. 20 to 100 RBs are used to transmit 360-degree tiled 360-degree video over the LTE eMBMS channel. The videos of resolution  $3840 \times 1920$  with 16 tiles are used for the simulation. Each tile has four different representations encoded with different QP values (25, 30, 35, and 40) [44], which result in tiled video rates of 13Kbps to 7.7Mbps. Each tile is encoded by a conventional video encoder (ffmpeg) [45], and the total bitrate to deliver the best representations for all tiles is 40Mbps.

The proposed resource algorithm, a convex-optimization method (Convex), is compared with some of the existing solutions for resource-allocation: 1) the exhaustive search (Exhaustive), 2) the equal-resource allocation (Equal), and 3) the broadcast algorithm (Broadcast). Exhaustive search tries all possible resource allocations, and the equal-resource allocation allocates the same amount of the resource to each multicast session. The broadcasting algorithm allocates all available resource to group 1. The median-quality scheme (MQS) [46] algorithm is a grouping algorithm that selects the user with the median SNR in a group as the boundary for dividing the group in half. The MaxSE algorithm, which maximizes the SE, is our proposed grouping algorithm. Six combinations of these algorithms and broadcast are tested. The proposed tiled video rate-allocation algorithm is applied in all simulations.

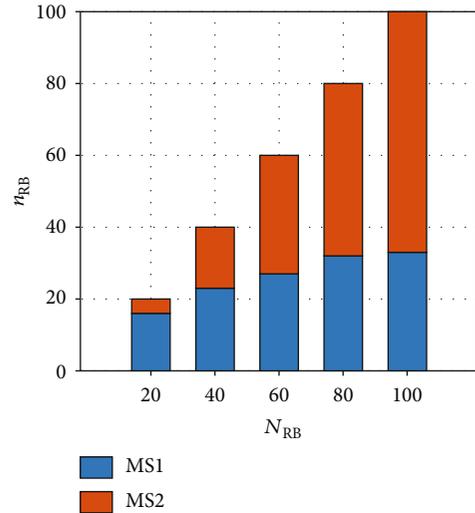


FIGURE 9: Resource allocation.

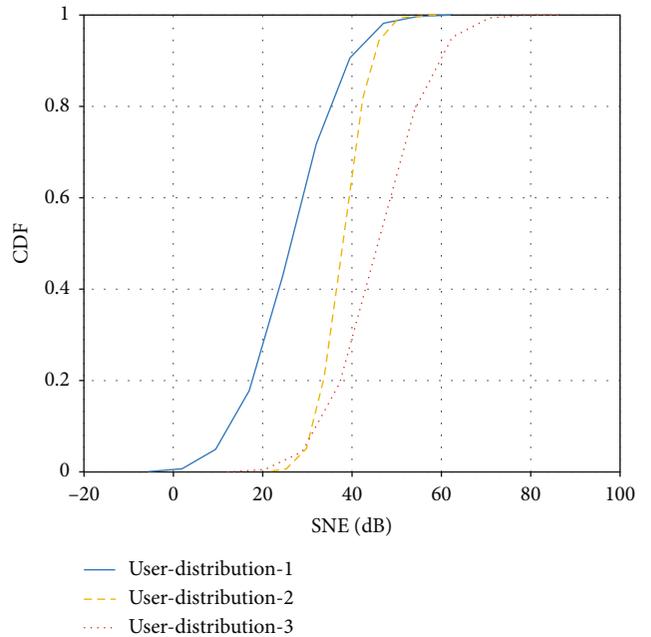


FIGURE 10: Users' average SNR distribution.

**6.2. Utility and PSNR.** Figure 5 shows the total utilities achieved by the competing methods. The proposed resource-allocation algorithm (Convex) combined with any grouping algorithm achieves the best utility performance, having the same results as the exhaustive search. Figure 6 shows the average peak signal-to-noise-ratio (PSNR) of the videos that are delivered using the introduced algorithms. It shows that we can deliver more similar videos to the users using the proposed algorithm. It also shows that the utility value is a good way to describe the users' experience in mathematical form. Figure 7 shows the utility performance of Convex+MaxSE with different total numbers ( $G$ ) of groups. We can observe that the utility performance is the best with  $G=2$ ; therefore, the proposed algorithm stops after testing the utility with  $G=3$  and decides the best solution is  $G=2$ .

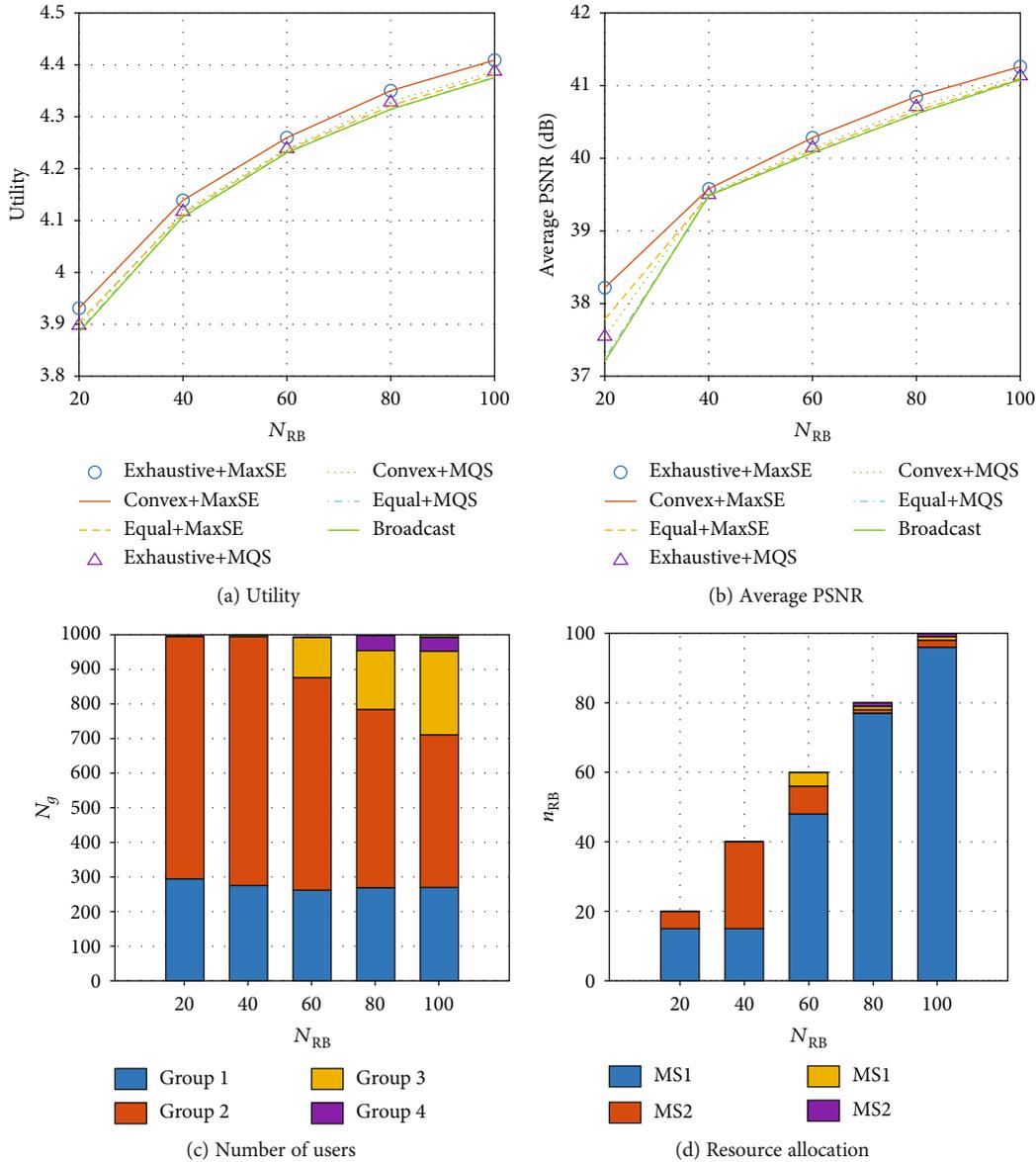


FIGURE 11: MSM in User-distribution-2.

This fact implies that creating two multicast sessions to allocate better representations improves the total utility when compared to broadcasting the same representation of tiles to all users. However, creating too many multicast sessions cannot improve the utility because multiple copies of the same tile are allocated using the wireless resource. Figure 8 shows the number of users in each group when Convex+MaxSE scheme is used. There are two groups, and the number of users who can join group 2 is larger than the number of users who join group 1. Figure 9 shows the number of resource blocks allocated for each multicast session (MS) when we use the Convex+MaxSE scheme.

The proposed algorithm is tested in different user SNR distributions to see how it works in different environments. The resulting users' average SNR distribution generated by using Table 2 is shown by user-distribution-1 in Figure 10, with an average SNR range from 0 dB to 50 dB. User-distribution-2, and user-distribution-3 are also generated to show the performance of the proposed algorithms in the case that most of the users in the LTE system have higher average SNRs. User-distribution-2 assumes that the users are located densely; therefore, variance of the average SNR is very small. User-distribution-3 assumes that the users have good channel conditions, but the variance of the average SNR is the same as user-distribution-1. Figure 11(a) shows the utilities that can be achieved by applying the algorithms. The performance gap between the proposed algorithm and the broadcast algorithm is smaller than in user-distribution-1. The reason is found in Figure 11(d), which shows the resources allocated for each multicast session. Most of the resources are allocated in multicast session 1 (MS1) because it contributes the most to improving the total utility. It is different from the resource allocation results for user-distribution-1 shown in Figure 10. Since there are more users with lower

SNRs, the proposed algorithm allocates more resources to MS1 to improve the utility of those users. In user-distribution-2, since the average SNR is high and the variance is small, the performance gap between the proposed algorithm and the broadcast algorithm is smaller than in user-distribution-1. The reason is found in Figure 11(d), which shows the resources allocated for each multicast session. Most of the resources are allocated in multicast session 1 (MS1) because it contributes the most to improving the total utility. It is different from the resource allocation results for user-distribution-1 shown in Figure 10. Since there are more users with lower

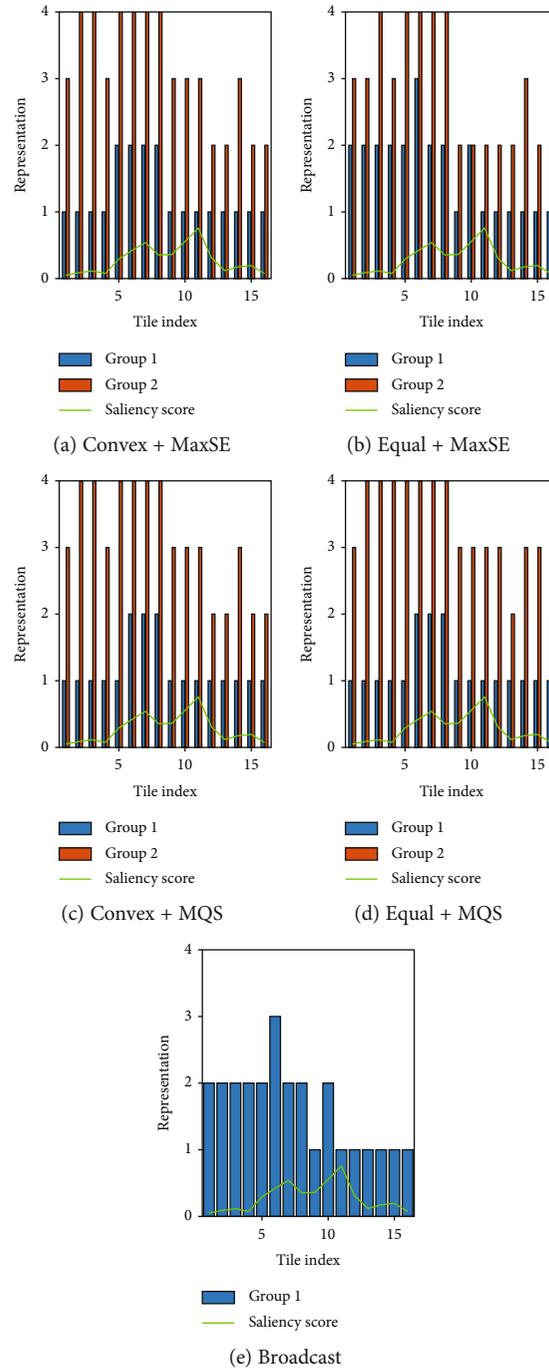


FIGURE 12: Rate selection results with User distribution-1.

average SNR in user-distribution-1, the proposed algorithm gave minimal resources to multicast session 1 (MS1) to satisfy the users with lower SNR and more to multicast session 2 (MS2) to improve the total utility more efficiently than in multicast session 1 (MS1). Moreover, user-distribution-2 opens more multicast sessions than user-distribution-1 because using a smaller number of RBs by allocating only a few tiles with better representations improves the total utility. Most of the tiles already had the best representations in multicast session 1.

Figure 12 shows the selected representations for all groups of users for all 16 tiles in user-distribution-1. Figure 12(a) shows the representations that user groups 1 and 2 could watch in their views. Those in group 2 have much better quality since most of the tiles have better representations than those of user group 1. Figure 12(b) shows the simulation results with the Equal resource allocation combined with the MaxSE grouping algorithm. The difference between Figures 12(a) and 12(b) is the representation of tile 5 for user group 1. It shows that the proposed resource

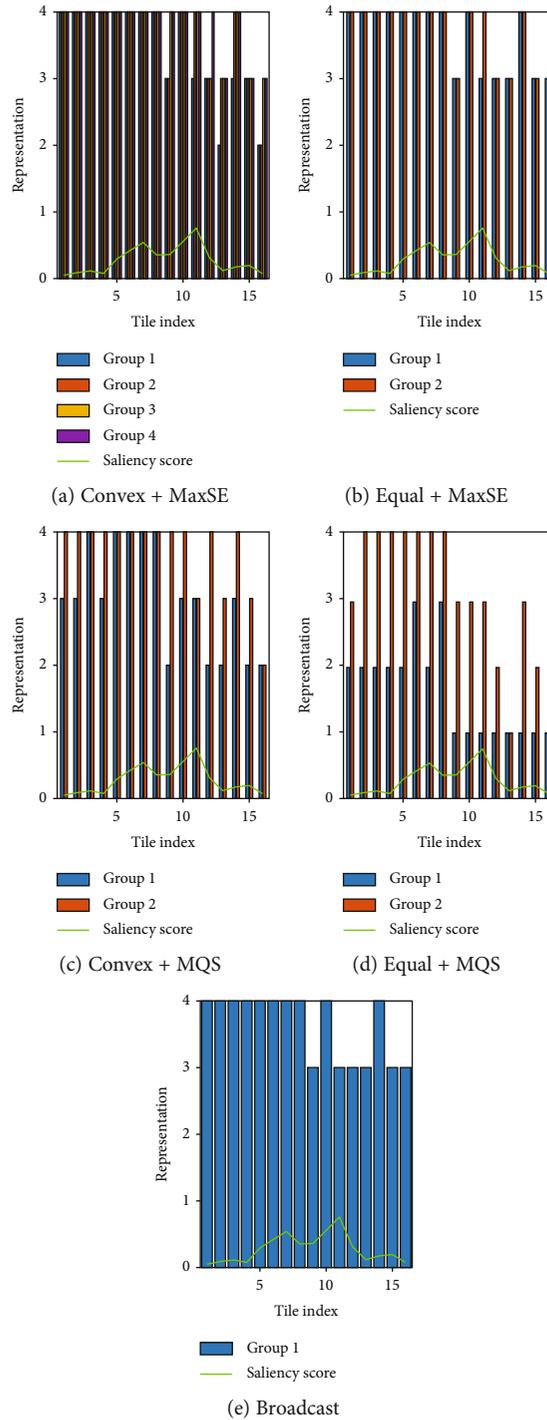


FIGURE 13: Rate selection results with User distribution-2.

allocation method could allocate the resource more efficiently than the Equal resource allocation method. Figure 12(c) shows the simulation results using Convex+MQS. Users in group 1 had better representations than the Convex+MaxSE method, and users in group 2 had worse representations because the MQS algorithm led more users to join group 1 and the Convex resource allocation algorithm gave more resources to group 1 to maximize the utility.

Figure 12(d) shows the results with Equal+MQS. Users in group 1 received worse representations than Convex+MaxSE, but users in group 2 received better ones. However, the number of users who could join group 2 was smaller when we used MQS; therefore, the total utility and average PSNR with Equal+MQS were worse than with Convex+MaxSE. Figure 12(e) shows the result with the Broadcast scheme, which allocated all resources to group 1 and

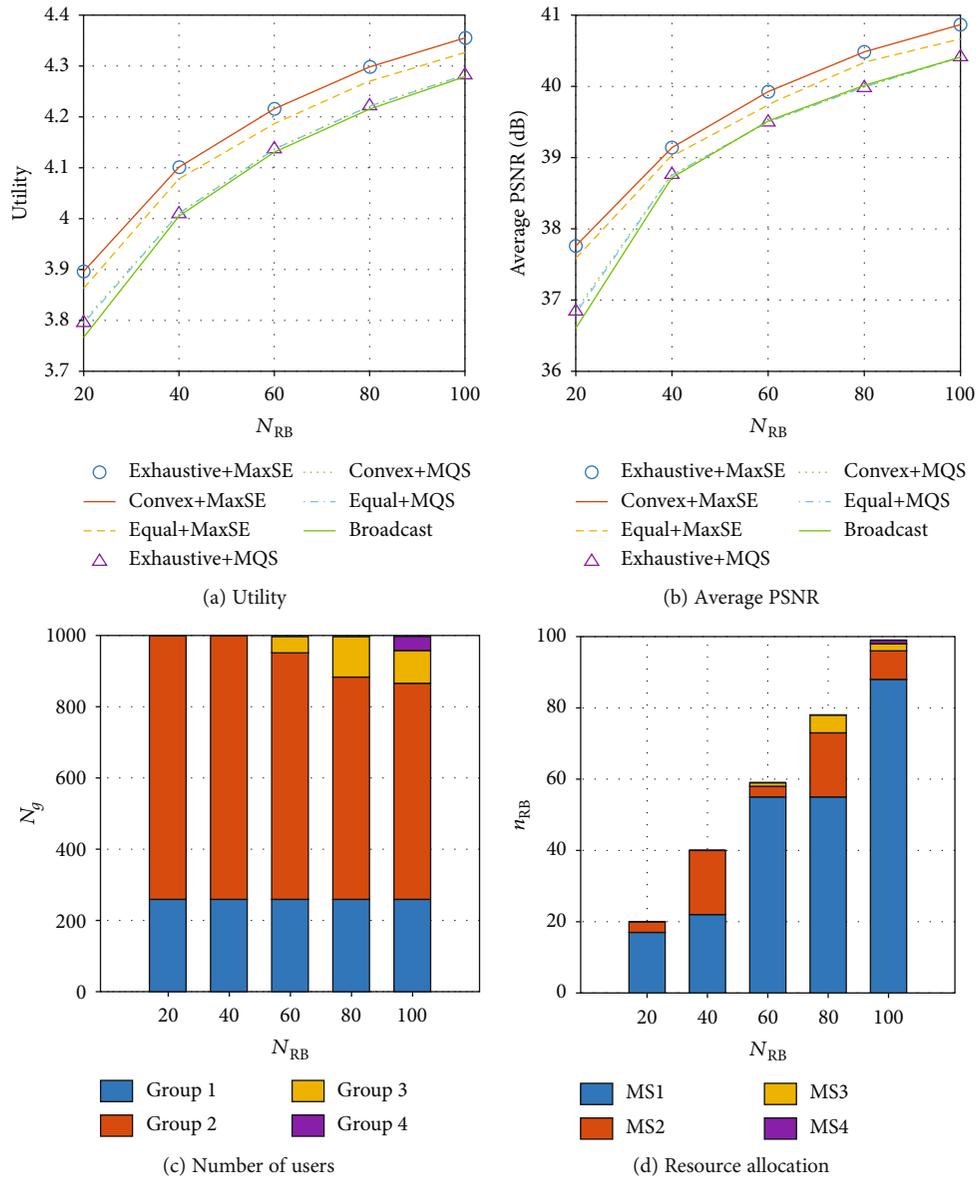


FIGURE 14: MSM in User-distribution-3.



FIGURE 15: 360-degree video example.

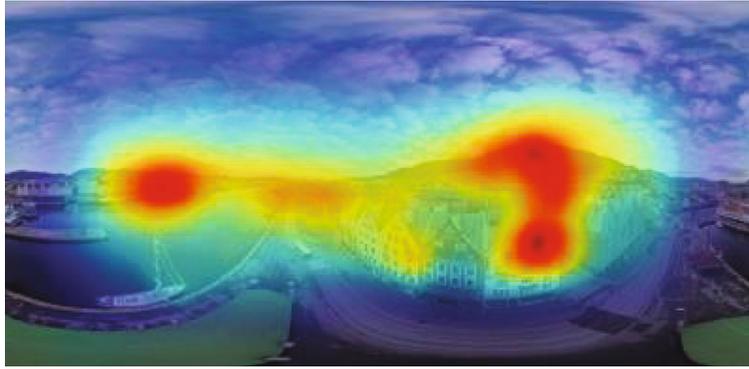


FIGURE 16: Saliency map.

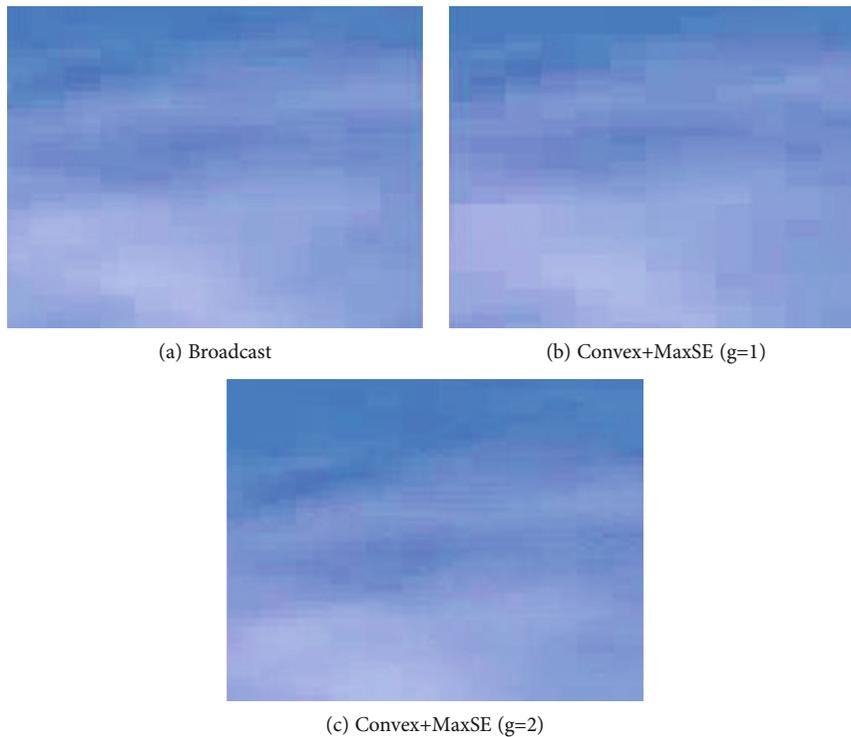


FIGURE 17: Visual Qualities – Scene 1.

performed the same rate-selection algorithm. Since it could not utilize the resources because of the users with low SNR, most of the tiles could not have higher quality representations.

Figure 13 shows selected representations for user-distribution-2. Since all the users in user-distribution-2 had very good channel quality, most of the tiles have high representations through the Broadcast scheme. However, there is still some space to improve the visual quality for users with better channel quality. Figure 13(a) shows that the Convex+MaxSE scheme creates four groups. The users in group 1 received worse representations on tiles 13 and 16, but the users in groups 3 and 4 received better ones on tiles 9, 11, and 12 than with the Broadcast scheme. Since tiles 9, 11, and 12 have higher saliency scores than tiles 13 and 16, we can expect users to have better quality views.

Figures 14(a), 14(b). shows the utilities and the average PSNRs of the algorithms. Figures 14(c), 14(d) shows the number of users in each group, and the wireless resource allocation results with the user-distribution-3 using the Convex+MaxSE scheme. Since the variance of the users' average SNR in user-distribution-3 is greater than that of user-distribution-2, the utility and the average PSNR performance gap between Convex+MaxSE and Broadcast scheme is larger than the performance gap in user-distribution-2. Since most of the users have very good channel quality, multicast session 1 (MS1) occupies most of the resources, and MS2, MS3, and MS4 only have small number of resources to improve the total utility.

6.3. *Visual Quality Comparison.* Figure 15 shows examples of 360-degree video footage. We took two different viewpoints of

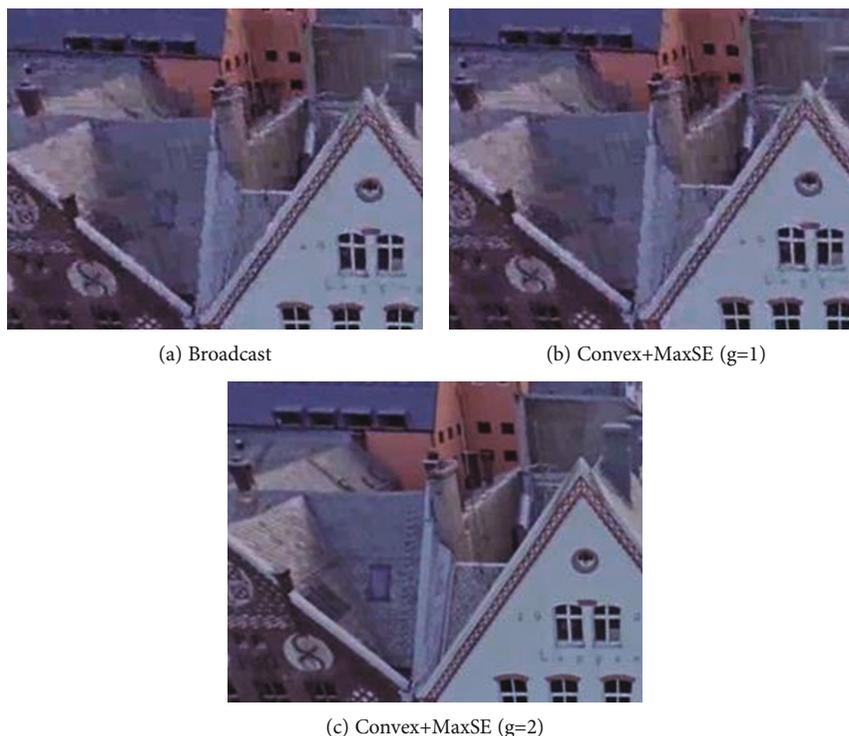


FIGURE 18: Visual Qualities – Scene 2.

a 360-degree video to qualitatively compare the performances. The scene 1 includes sky and cloud, where users may not give much attention. Scene 2 includes buildings, where users can see the texture. Scene 2 has a higher saliency score than the scene 1. Figure 16 shows the saliency score of the video. Figures 17 and 18 show the visual quality of scene 1 and scene 2, respectively, where Broadcast scheme and Convex+MaxSE scheme are used with user-distribution-1. Figures 17(a) and 18(a) are two different scenes shown to users using the Broadcast. All the users have the same quality. There is recognizable pixelation in the scene with lower video quality, which caused the users to recognize the video quality as poor. The Convex+MaxSE method divided the users into two groups based on their channel quality. Group  $g=2$  has better quality than group  $g=1$ . Figures 17(c) and 18(c) are the scenes shown to user group  $g=2$ , and Figures 17(b) and 18(b) were shown to user group  $g=1$ . The multicasting group  $g=2$  received better video quality than the users with the broadcasting method. Multicasting group  $g=1$ 's video quality is similar to that of the broadcasting method. Note that only 35% of users joined group  $g=1$ , while 65% joined group  $g=2$ , which had much better quality.

## 7. Conclusion

In this paper, the Multi-Session Multicast system for the 360-degree video multicasting is proposed to achieve a higher utility with a limited spectrum in overloaded situations. We have theoretically formulated a relationship among user groupings, MCS and the AL-FEC code rate to simplify the problem solution. A grouping algorithm for achieving better spectral efficiency and an optimal resource-allocation solu-

tion that uses a convex optimization technique to maximize utility are introduced. The algorithm can allocate the optimal resource and find the best user grouping to maximize utility. The simulation results show that the multicasting can take advantage of sharing the resource among many users requesting the same 360-degree video. Saliency information is used for our simulation, but the proposed algorithm is not limited to saliency information. The users' viewport information or object detection results can also be used to provide the same formulation and solution.

## Data Availability

The source code data used to support the findings of this study have been deposited in <https://github.com/jsup517/DASH2>.

## Conflicts of Interest

The author(s) declare(s) that they have no conflicts of interest.

## References

- [1] A. R. Costa, *Storytelling for Cinematic Virtual Reality: Audiences Needs And Expectations*, 2018, <http://handle.net/10216/114279>.
- [2] X. Corbillon, A. Devlic, G. Simon, and J. Chakareski, "Viewport-Adaptive Navigable 360-Degree Video Delivery," in *Proceedings of IEEE International Conference on Communications (ICC)*, Paris, France, 2017.
- [3] Y. Bao, H. Wu, T. Zhang, A. A. Ramli, and X. Liu, "Shooting a moving target: Motion-prediction-based transmission for 360-

- degree videos,” in *2016 IEEE International Conference on Big Data (Big Data)*, Washington, DC, USA, 2016.
- [4] J. Liu, W. Chen, Z. Cao, and K. B. Letaief, “Dynamic Power and Sub-Carrier Allocation for OFDMA-Based Wireless Multicast Systems,” in *2008 IEEE International Conference on Communications*, Beijing, China, 2008.
  - [5] D. Lecompte and F. Gabin, “Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: overview and Rel-11 enhancements,” *IEEE Communications Magazine*, vol. 50, no. 11, pp. 68–74, 2012.
  - [6] ISO/IEC IS 23009-1, *Information technology - Dynamic adaptive streaming over HTTP (DASH) - Part 1: Media presentation description and segment formats*, 2012.
  - [7] T. Stockhammer and M. G. Luby, “Dash in mobile networks and services,” in *2012 Visual Communications and Image Processing*, San Diego, CA, USA, 2012.
  - [8] J. Park, J.-N. Hwang, Q. Li, Y. Xu, and W. Huang, “Optimal DASH-multicasting over LTE,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4487–4500, 2018.
  - [9] R. Belda, I. de Fez, F. Fraile, P. Arce, and J. C. Guerri, “Hybrid FLUTE/DASH video delivery over mobile wireless networks,” *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 11, pp. 1070–1082, 2014.
  - [10] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth, and M. Zhou, “An overview of tiles in HEVC,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 969–977, 2013.
  - [11] L. Dacunto, J. V. D. Berg, E. Thomas, and O. Niamut, “Using MPEG DASH SRD for zoomable and navigable video,” in *Proceedings of the 7th International Conference on Multimedia Systems-MMSys 16*, Klagenfurt, Austria, 2016.
  - [12] J. Park, J. N. Hwang, and H. Y. Wei, “Cross-Layer Optimization for VR Video Multicast Systems,” in *IEEE Globecom 2018*, Abu Dhabi, UAE, 2018.
  - [13] J. L. Feuvre and C. Concolato, “Tiled-based adaptive streaming using MPEG-DASH,” in *Proceedings of the 7th International Conference on Multimedia Systems-MMSys 16*, Klagenfurt, Austria, 2016.
  - [14] H. Kimata, S. Shimizu, Y. Kunita, M. Isogai, and Y. Ohtani, “Panorama video coding for user-driven interactive video application,” in *2009 IEEE 13th International Symposium on Consumer Electronics*, Kyoto, Japan, 2009.
  - [15] M. Xiao, C. Zhou, Y. Liu, and S. Chen, “OpTile: Toward Optimal Tiling in 360-degree Video Streaming,” in *Proceedings of the 25th ACM international conference on Multimedia (MM '17)*, pp. 708–716, New York, NY, USA, 2017.
  - [16] V. R. Gaddam, M. Riegler, R. Eg, C. Griwodz, and P. Halvorsen, “Tiling in interactive panoramic video: approaches and evaluation,” *IEEE Transactions on Multimedia*, vol. 18, no. 9, pp. 1819–1831, 2016.
  - [17] P. R. Alfance, J.-F. Macq, and N. Verzijp, “Interactive omnidirectional video delivery: a bandwidth-effective approach,” *Bell Labs Technical Journal*, vol. 16, no. 4, pp. 135–147, 2012.
  - [18] Y. Zhang, P. Zhao, K. Bian, Y. Liu, L. Song, and X. Li, “DRL360: 360-Degree Video Streaming with Deep Reinforcement Learning,” in *IEEE INFOCOM 2019- IEEE Conference on Computer Communications*, pp. 1252–1260, Paris, France, 2019.
  - [19] J. Son, D. Jang, and E.-S. Ryu, “Implementing 360 video tiled streaming system,” in *Proceedings of the 9th ACM multimedia systems conference (MMSys '18)*, pp. 521–524, New York, NY, USA, 2018.
  - [20] J. Zyren and W. McCoy, *Overview of the 3GPP long term evolution physical layer*, Freescale Semiconductor, Inc., white paper 22, 2007.
  - [21] M. Condoluci, G. Araniti, A. Molinaro, and A. Iera, “Multicast resource allocation enhanced by channel state feedbacks for multiple scalable video coding streams in LTE networks,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 2907–2921, 2016.
  - [22] C.-W. Huang, S.-M. Huang, P.-H. Wu, S.-J. Lin, and J.-N. Hwang, “OLM: opportunistic layered multicasting for scalable IPTV over Mobile WiMAX,” *IEEE Transactions on Mobile Computing*, vol. 11, no. 3, pp. 453–463, 2012.
  - [23] H. Ahmadi, E. Omar, and H. Mohamed, “Adaptive multicast streaming of virtual reality content to mobile users,” in *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pp. 170–178, Mountain view, CA, USA, 2017.
  - [24] Y.-C. Chen, D.-N. Yang, and W. Liao, “Efficient multi-view 3D video multicast with depth image-based rendering in LTE networks,” in *2013 IEEE Global Communications Conference (GLOBECOM)*, Atlanta, GA, USA, 2013.
  - [25] M. Luby, A. Shokrollahi, M. Watson, and T. Stockhammer, *Raptor Forward Error Correction Scheme for Object Delivery*, RFC 5053 (Proposed Standard), IETF, 2007.
  - [26] M. S. Anwar, J. Wang, W. Khan, A. Ullah, S. Ahmad, and Z. Fei, “Subjective QoE of 360-degree virtual reality videos and machine learning predictions,” *IEEE Access*, vol. 8, pp. 148084–148099, 2020.
  - [27] T. Keskinen, V. Mäkelä, P. Kallioniemi et al., “The Effect of Camera Height, Actor Behavior, and Viewer Position on the User Experience of 360° Videos,” in *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 423–430, Osaka, Japan, 2019.
  - [28] S. Van Damme, M. T. Vega, and F. De Turck, “Human-centric Quality Management of Immersive Multimedia Applications,” in *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, pp. 57–64, Ghent, Belgium, 2020.
  - [29] W. W. Zhang, Y. G. Wen, Z. Z. Chen, and A. Khisti, “QoE-Driven cache management for HTTP adaptive bit rate streaming over wireless networks,” *IEEE Transactions on Multimedia*, vol. 15, no. 6, pp. 1431–1445, 2013.
  - [30] X. Hou, J. Harel, and C. Koch, “mage signature: highlighting sparse salient regions,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, 2012.
  - [31] W. Kim and J.-J. Han, “Video saliency detection using contrast of spatiotemporal directional coherence,” *IEEE Signal Processing Letters*, vol. 21, no. 10, pp. 1250–1254, 2014.
  - [32] J. Xu, X. Guo, Q. Tu, C. Li, and A. Men, “A novel video saliency map detection model in compressed domain,” in *MILCOM 2015-2015 IEEE Military Communications Conference*, Tampa, FL, USA, 2015.
  - [33] W. Wang, J. Shen, and L. Shao, “Consistent video saliency using local gradient flow optimization and global refinement,” *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4185–4196, 2015.
  - [34] D. Lopez-Perez, A. Ladanyi, A. Juttner, H. Rivano, and J. Zhang, “Optimization method for the joint allocation of modulation schemes, coding rates, resource blocks and power in self-organizing LTE networks,” in *2011 Proceedings IEEE INFOCOM*, Shanghai, China, 2011.

- [35] A. D. L. Fuente, C. M. Lentisco, L. Bellido, R. P. Leal, E. Pastor, and A. G. Armada, "Analysis of the impact of FEC techniques on a multicast video streaming service over LTE," in *2015 European Conference on Networks and Communications (EuCNC)*, Paris, France, 2015.
- [36] M. Kawser, "Downlink SNR to CQI mapping for different multiple antenna techniques in LTE," *International Journal of Information and Electronics Engineering*, vol. 2, no. 5, pp. 757–760, 2012.
- [37] S. N. Donthi and N. B. Mehta, "An accurate model for EESM and its application to analysis of CQI feedback schemes and scheduling in LTE," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3436–3448, 2011.
- [38] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2013.
- [39] A. Awada, E. Lang, O. Renner et al., "Field trial of LTE eMBMS network for TV distribution: experimental results and analysis," *IEEE Transactions on Broadcasting*, vol. 63, no. 2, pp. 321–337, 2017.
- [40] C. Yahiaoui, M. Bouhali, and C. Gontrand, "Simulating the Long Term Evolution (LTE) Downlink Physical Layer," in *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, Cambridge, UK, 2014.
- [41] R. Giuliano and F. Mazzenga, "Exponential effective SINR approximations for OFDM/OFDMA-based cellular system planning," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4434–4439, 2009.
- [42] X. Li, Q. Fang, and L. Shi, "A effective SINR link to system mapping method for CQI feedback in TD-LTE system," in *2011 IEEE 2nd International Conference on Computing, Control and Industrial Engineering*, Wuhan, China, 2011.
- [43] G. F. Pedersen, *COST 231-Digital mobile radio towards future generation systems*, EU, 1999.
- [44] H. Wang, V.-T. Nguyen, W. T. Ooi, and M. C. Chan, "Mixing tile resolutions in tiled video," *Proceedings of Network and Operating System Support on Digital Audio and Video Workshop-NOSSDAV 14*, 2013.
- [45] FFmpeg, 2018, <https://github.com/FFmpeg/FFmpeg>.
- [46] K. Bakanoglu, W. Mingquan, L. Hang, and M. Saurabh, "Adaptive Resource Allocation in Multicast OFDMA Systems," in *IEEE Wireless Communication and Networking Conference*, Sydney, NSW, Australia, 2010.