# Cross-Layer Optimization of Wireless Video Multicast Systems

Jounsup Park

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Jenq-Neng Hwang, Chair

Payman Arabshahi

James A. Ritcey

Program Authorized to Offer Degree:

Department of Electrical Engineering

University of Washington

## Abstract

Cross-Layer Optimization of Wireless Video Multicast Systems

Jounsup Park

Chair of the Supervisory Committee:

Professor Jenq-Neng Hwang

Department of Electrical Engineering

Multimedia data traffic occupies more than 70% of the internet traffic and half of the video consumers use their mobile devices to watch the videos. Moreover, Virtual Reality (VR) data traffic is increasing very fast. To provide the good quality of the multimedia service over the wireless channel, the large bandwidth is needed because users' quality of experience (QoE) usually proportional to the video rates they can receive. Moreover, the variation of the bandwidth in the wireless channel also affects to the users' experience. Therefore, it is necessary to apply the wireless transmission scheme that has high spectral efficiency and the video rate adaptation to provide the best quality to the users. Multicast system is one of the technologies that can improve the spectral efficiency, and Dynamic Adaptive Streaming over HTTP (DASH) is one of the most popular video rate adaptation platforms. DASH is a fast growing video streaming platform, which enables adaptive rate selection based on channel conditions. File Delivery over Unidirectional Transport (FLUTE) further enables multicasting of the DASH segments over LTE eMBMS

systems. We present the DASH multicast systems and its optimization in this paper to deliver high quality videos to the large number of mobile users. First, we show that the DASH can take advantages of the multicasting scheme in terms of utility by optimally allocating the resource. Second, we propose the VR video multicast system, which is compatible with DASH, and the optimization algorithms to improve users' experience for mobile VR applications.

An optimal DASH multicasting solution is proposed to allow more DASH clients in an LTE network to receive better videos by optimizing the resource allocation, application-layer Forward Error Correction (AL-FEC) code rate and modulation and coding scheme (MCS) selection of each multicasting group, which corresponds to a FLUTE session. Multiple FLUTE sessions are considered to deliver multiple videos, each with different video rates, for enhancing the overall utility. We have applied the convex optimization method to find the optimal resource allocation in terms of utility for multiple FLUTE sessions. We also find the optimal AL-FEC code rates to add redundancies to protect the video segments for each FLUTE session. Moreover, an efficient MCS selection is introduced to reduce the complexity of the algorithm. Simulation results, with realistic LTE parameters, are shown to prove the proposed scheme is optimal, with more DASH clients receiving better video representations within limited resources when compared to other existing algorithms.

VR video multicast systems and its optimization algorithms are presented. 360-degree videos for Virtual Reality (VR) applications are getting more popular because of their diverse applications. VR videos usually need more bandwidth than conventional videos to provide the same quality of experience (QoE). Tiled videos can help save bandwidth by selecting lower quality encoding for the tiles with lower probability of viewing. DASH enables the adaptive rate selection of tiles based on the channel conditions. Multicasting also can help save bandwidth, since many users share the spectrum when they request the same video contents. In this paper, we formulate the utility maximization problem to find which tiles should have which video representations to satisfy the most users in multicasting groups using limited resources. A cross-layer optimization framework, which includes user grouping, resource allocation, and the tile-based rate-selection algorithms, is proposed to maximize the total utility among all users. Saliency information extracted from the

videos is used to find the tiles which most users may be interested in. Simulation results show that the proposed cross-layer optimization framework can achieve a higher utility than the broadcasting solution or existing multicast solutions.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# DEDICATION

to my dear wife, Ahreum

# Chapter 1
# INTRODUCTION

## *1.1 Introduction*

The enhanced capabilities of mobile devices and the improved capacities of wireless networks have led to a massive growth in mobile video consumption. A recent report [1] shows that the video traffic occupies more than 70% of whole internet traffic in peak time, and moreover, half of the video consumers use mobile devices. Besides, VR/AR applications are getting more popular and users can enjoy diverse experience with them. VR/AR applications need more data than conventional video streaming services. As increasing multimedia data traffic being over wireless networks, efficient utilization of the wireless resources is getting more important to serve more users. In addition, the wireless channel condition varies frequently with channel environments and user behaviors. MPEG's Dynamic Adaptive Streaming over HTTP (MPEG-DASH) [2] is thus proposed as an effective video streaming platform, which enables the adaptive rate selection based on the channel conditions. DASH can provide superior video experience by giving clients a chance to receive the video quality based on their channel condition and buffer status, resulting in better quality of experience (QoE). Most of internet video service providers, such as Netflix and Youtube, support DASH in their video streaming platforms.

With the overloaded scenarios, for example, many people watch the popular live videos such as sports events, bandwidth can be easily used up and many people will suffer from delay or low video quality. To overcome the problem, video multicasting can be utilized. LTE allows using their spectrum for multicasting or broadcasting up to 60% of the spectrum and it is standardized as evolved Multimedia Broadcast Multicast Service (eMBMS) [3]. The Multicast Channel (MCH), that delivers eMBMS data, cannot get any advantage from either Hybrid Automatic Repeat-reQuest (HARQ) or retransmission since the MCH transfers the data in Radio Link Control (RLC) unacknowledged mode (UM) [3], due to the fact that a single user's channel condition cannot

represent all users' channel conditions. Besides, it is very inefficient to retransmit many lost packets to several user equipment (UEs) with poor channel conditions, resulting in further consumption of bandwidth. This situation makes Quality of Experience (QoE) worse because it cannot transmit the appropriate video representations to the subscribed users, resulting in very high packet loss rate or possibility that users cannot get a video with enough quality even when the channel condition is very good. To overcome the problem of DASH multicasting, the File delivery over Unidirectional Transport (FLUTE) [4] protocol is thus introduced. The IETF introduces FLUTE for unidirectional data transfer over the internet. To avoid packet loss, FLUTE adds redundant packets to help the recovery of the lost packet, which is done by Forward Error Correction (FEC) [5]. Moreover, if FEC is not enough to recover all lost packet, DASH clients can request packet recovery through reliable TCP unicast transmission [6].

Combining FLUTE and eMBMS of LTE makes DASH multicasting possible with capability of adaptive video quality control, however, it introduces more complexity to the systems. Since, there are multiple copies of the video with different rates, the system has to choose which video rates to be scheduled based on the channel information and user's requests. FLUTE sessions have to add redundant APP-layer FEC packets to protect the video data while not losing efficiency. Moreover, resources for each FLUTE session must be allocated in the OFDMA frames and PHY-layer modulation and coding schemes (MCS) for the chosen resource blocks also need to be selected for reliable communications. Its complexity exponentially increases as number of users and/or number of video increases to find the optimal solution. Moreover, channel condition always changes frequently, therefore, it is more difficult to optimize the whole system in real time.

In this paper, an optimal solution for video multicasting is proposed for DASH types of video sources. Through joint consideration of the video rates, the resource allocation, the FEC redundancy and the MCS for each subscribed multicast video, our proposed jointly optimized solution schedules the best set of videos to maximize the perceived QoE (i.e., utility values) without the need of additional packet repair requests, which call for additional resources and introduce more delays.

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 VR videos, most of the area is not visible to the users [7][8]. Therefore, more redundant data are transmitted than in conventional video if we directly use the DASH multicast for VR-video dissemination. To be more efficient, redundant data should be removed, and tiled video [9][10] allows the flexible removal and allocation of lower bits to the redundant parts of the video. For example, necessary parts are transmitted with multiple copies with different qualities to satisfy users with good channel quality, and the parts with lower probability of being viewed are transmitted just once with a single quality to save the spectrum. In this paper, a new approach to allocate the DASH-VR video on the LTE eMBMS systems. There are separate solutions for the rate adaptation of the tiled-video and the resource allocation of wireless video multicast using limited wireless resource, but the separate solutions cannot optimize the VR video multicast systems. Therefore, we propose a multicast system that can deliver the tiled-video with an adaptive rate selection and perform the optimal resource allocation for multiple FLUTE sessions.

### *1.2 Contributions*

The contributions of this dissertation are summarized as follows.

- The main contribution of this paper is that we formulate the video multicasting problem as a mathematical optimization task, which allows us to find a closed form solution for resource allocations for utility/QoE maximization.
- An efficient MCS selection algorithm is introduced to reduce the complexity of the selection procedure without degrading the performance, as evidenced by the simulation results.
- Comparisons with other state-of-the-art video multicasting algorithms shows that the proposed algorithm can achieve higher utility values and spectral efficiency than other methods, both for SVC and DASH video streaming systems.

- Multi-Session Multicast (MSM) is proposed to efficiently allocate tiled-videos on multicast systems, such as eMBMS.
- 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.
- We have derived the spectral efficiency as a function of user grouping to find the efficient and effective grouping algorithm. It reduces the number of parameters to be optimized.
- A convex optimization method is applied to allocate the optimal resources for each multicasting session.
- A modification of the existing state-of-the-art rate-allocation algorithm for tiled-videos is applied to a multicasting scenario.
- Extensive simulation results including comparisons with other state-of-the-art video multicasting algorithms shows that the proposed algorithms can achieve higher utility values and spectral efficiency than the other methods.

### *1.3 Dissertation Roadmap*

The rest of this dissertation is organized as follows.

**Chapter 2**: we review the related works on previously proposed video multicasting algorithms including user grouping, resource allocation, MCS selection, and AL-FEC code rate selection schemes. Tile-based video encoding and rate selection algorithms are also discussed for VR video streaming application.

**Chapter 3**: we present an optimal DASH multicasting method that works in OFDMA based systems such as LTE. The optimal resource allocation algorithm is proposed to improve the QoE, and an efficient MCS selection algorithm is applied to reduce the complexity of the algorithm without degrading the performance.

**Chapter 4**: we propose a VR video multicast system to deliver tiled video encoded with multiple representations. The proposed cross-layer optimization framework can allocate the best video representations of tiled video to efficiently utilize the limited resource. The proposed wireless VR video multicast system and the cross-layer optimization algorithm can provide superior quality to the users than existing algorithms.

**Chapter 5**: conclusions of this dissertation are given by summarizing the main contributions and discussing some extensions to this work that lead to potential research topics in the future.

# Chapter 2
# BACKGROUNDS AND RELATED WORKS

## *2.1. DASH*

A DASH service refers that clients who want to receive a video can request the video from the server based on the Media Presentation Description (MPD) [2] and channel information to dynamically adapt the video rate for continuous playing the video and guaranteeing the best video quality. An MPD file includes the information about the video, such as segment formats and available video representations, therefore, clients can request the video chunks to the server by using the information contained in the MPD file. DASH, as its name stands for, is standardized for unicast video delivery using HTTP [2]. However, the DASH defines formats which may be suitably delivered over non-HTTP/TCP transports. In case DASH is delivered over eMBMS, DASH clients can still operate on regular formats and interfaces, no eMBMS specific functionalities are required. FLUTE is introduced to permit DASH to deliver DASH video segments over eMBMS such that the client can receive them being delivered over HTTP/TCP. With this, clients can receive different video segments which are not requested by the users, but clients can still decode the video even if the received video segment is not exactly the same representation, and it is beneficial for utilizing the multicasting approach to save the spectrum and optimize the network at the eNB.

## *2.2. LTE eMBMS*

LTE supports multicasting of video streams by eMBMS [3] systems, and 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

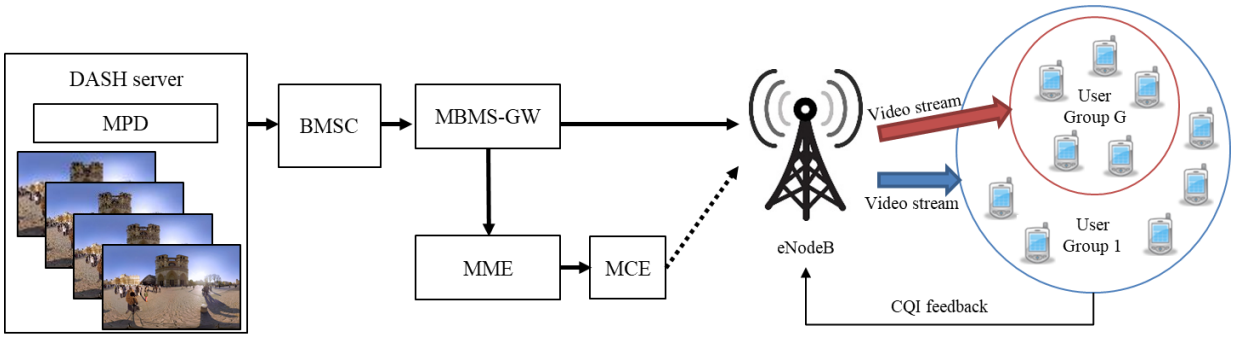



Figure 2.1. LTE eMBMS systems for DASH Multicast

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 for the multicast sessions and decides modulation and coding scheme (MCS). Multiple video multicasting sessions can thus be created and users can subscribe 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 [3]. 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, which is

$$c(\text{MCS}) = 12(\text{subcarriers}) \times 7(\text{symbols}) \times \text{efficiency}. \quad (2.1)$$

Table 2.1 shows the MCS along with efficiency [11] for various channel quality indication (CQI) indices. The more efficient modulation and coding scheme can be used when the channel condition is better, and higher CQI index indicate that the channel condition is better. eNBs should choose MCS that can be decoded on the receiver side. In this paper, we use the CQI index as an MCS index for notational convenience. Using the information in the table, we can find what is the expected data-rate when we know how many RBs are allocated to the FLUTE sessions.

Table 2.1. CQI-MCS MAPPING

| CQI index | Modulation | Code rate (x 1024) | efficiency |
|---|---|---|---|
| 0 | Out of range | | |
| 1 | QPSK | 78 | 0.1523 |
| 2 | QPSK | 120 | 0.2344 |
| 3 | QPSK | 193 | 0.3770 |
| 4 | QPSK | 308 | 0.6016 |
| 5 | QPSK | 449 | 0.8770 |
| 6 | QPSK | 602 | 1.1758 |
| 7 | 16QAM | 378 | 1.4766 |
| 8 | 16QAM | 490 | 1.9141 |
| 9 | 16QAM | 616 | 2.4063 |
| 10 | 64QAM | 466 | 2.7305 |
| 11 | 64QAM | 567 | 3.3223 |
| 12 | 64QAM | 666 | 3.9023 |
| 13 | 64QAM | 772 | 4.5234 |
| 14 | 64QAM | 873 | 5.1152 |
| 15 | 64QAM | 948 | 5.5547 |

### *2.3. File Delivery over Unidirectional Transport (FLUTE)*

The file delivery over unidirectional transport (FLUTE) protocol is proposed by Internet Engineering Task Force (IETF) [4] to multicast a file over the networks using User Datagram Protocol (UDP)-based protocol with application-layer Forward Error Correction (AL-FEC) being provided for protecting the file from the packet losses. Additional file repair procedures are allowed by the HTTP file repair request. The file repair response message consists of HTTP header and file repair payload. The file repair response message consists of HTTP header which informs that point-to-multipoint repair, instead of point-to-point repair, is used.

### *2.4. Application-layer Forward Error Correction (AL-FEC)*

The radio channel conditions vary among all the users receiving the multicast service. Therefore, the block error rate of the users that receive the video service delivered with a single MCS can

have a great variance. In order to increase the robustness and the reliability of the multicast transmissions, FEC redundancy packets are incorporated at the APP-layer [5].

The solution proposed by 3GPP to deliver video streaming over eMBMS uses the FLUTE protocol with UDP transport to send video segments with the corresponding APP-layer FEC over multicast. An FEC block consists of $N$ packets including $K$ source packets and $N$-$K$ redundancy packets, resulting in the encoding rate $K/N$. The FEC decoder can ideally recover the original $K$ source packets from any $K$ out of $N$ received packets with correction capability $t=N-K$. The Reed Solomon (RS) code [12] is a well-known FEC code which operates on non-binary symbols and has the ideal correction capability. However, the RS code has a high decoding complexity because of its non-binary operations, which is not suitable for high-definition (HD) video streaming applications. The Raptor code [13] is a more attractive solution for HD video streaming services due to the flexible parameters selection and linear decoding cost. The correction capability of a Raptor code is $t=N-(1+\epsilon)K$, where $\epsilon$ is the reception overhead efficiency. The correction capability of the Raptor code is sub-optimal, however, a standardized Raptor code can closely achieve the ideal correction capability with negligible $\epsilon$. Therefore, it is used in our scheme. In this paper, FEC block size is fixed as $N$ and the number of source blocks $K_m$ is determined to choose appropriate FEC code rates $K_m/N$ for the $m$-th FLUTE session.

### *2.5. Video Multicasting Algorithms*

Since the multicasting can take advantage of using common resource for many users, video streaming services including AR/VR video streaming [14][15] or interactive multi-view video streaming [16][17], can also be applicable. Many of the multicast researches are based on the SVC video source [18][19][20] because of its structural advantage when it is applied to the multicasting. However, SVC itself has encoding complexity which can possibly cause significant delay. Moreover, DASH platform has getting more popular and DASH can efficiently adapt the video rate and more suitable to internet platform. Therefore, in this paper, both SVC and DASH types of systems are all investigated.

There have been researches on multicasting the videos over wireless networks. Chen et al. [21] consider the fair and optimal resource allocation for LTE multicast. They also consider the unicast for some users with lower SNR without considering FEC for packet protection. Belda et al. [22] introduce a Hybrid FLUTE/DASH video delivery system, which can multicast the video through FLUTE sessions and repair requests through the unicast channel to recover lost packets. They fix the FEC code rates of FLUTE sessions and provide the simulation results to show that the hybrid video delivery can improve the video quality compared to the video delivery systems using the unicast only. Nonetheless, in their approach, FEC code rates and resource allocation for multiple FLUTE sessions are not jointly optimized and it may create some repair requests through the unicast channel. Our research starts with the assumption that if we can jointly find the optimal resource allocation and FEC code rate selection, we can transmit the videos without repair requests which call for additional bandwidth. Our goal in this research is to jointly find the optimal resource allocations, the optimal MCS and FEC code rates for multiple FLUTE sessions so as to efficiently serve the DASH clients in the LTE networks without unicast channels for repairing the lost packets.

SVC-based video multicasting algorithms have been previously proposed, e.g., Conservative Multicasting Scheme (CMS) [23], Opportunistic Layered Multicasting (OLM) [24], Multicast Subgrouping for Multi-Layer video applications (MSML) [25], Median Quality Scheme (MQS) [26], and Median User Scheme (MUS) [27] and etc. Theses heuristic algorithms describe how to divide multiple users into several multicast groups (each group corresponds to one SVC layer) and select proper resource, MCS and FEC code rates, for the different groups based on the channel quality feedbacks from the users within the same group. More specifically, the CMS [23] first allocates each subchannel to a group of users in the multicast session based on their subchannel gains, then a greedy algorithm is adopted for resource allocation to achieve proportional fairness among sessions. OLM [24] can choose more aggressive MCS to achieve higher spectral efficiency and protect lost packets by using FEC for each group. On the other hand, MSML [25] utilizes the frequency diversity to achieve better throughput than other schemes. For example, a user with very low average SNR can possibly have some RBs that have high channel gains, and MSML utilizes these RBs to schedule lower video layers. Since MSML can choose the best RBs for each multicast group, it can select more efficient MCS than those chosen by other conservative schemes that are

constrained by the users with lowest channel quality, such as the less spectrally efficient CMS scheme. MUS and MQS choose the subgroups based on the number of users and the quality of the channel respectively. Their schemes can achieve better spectral efficiency than CMS, but less than those achieved by OLM and MSML.

### *2.6. Tiled VR Video-Streaming Systems*

In a tiled video scheme, the 360-degree video is divided into smaller tiles, which can be encoded independently [28]. There can be multiple copies of the same tile with different representation qualities. These tiles are transmitted through the wireless channel. DASH extends its standard to cover tiled 360-degree videos, i.e., DASH-VR [29]. DASH-VR has included virtual reality video descriptor (VRD) and spatial relationship descriptor (SRD), in addition to media presentation descriptor (MPD), to describe the projection types and spatial relationships among tiles. VRD contains the projection format and orientation information, SRD includes the region-wise quality of rectangular videos within the projected frame, and MPD includes the size of video chunks, locations of video files, and the codec information. As the clients join the multicast system, the MPD, VRD, and SRD are provided to the client, and the client can reconstruct the VR-video by using received tiles based on the received descriptor information.

A DASH multicast system [30] is introduced to efficiently utilize the limited resource and provide better videos to the users. The DASH multicast system allocates multiple copies of the same video with different quality to satisfy more users, but it inevitably generates redundant data that decreases the spectral efficiency. Especially in case of VR videos, most of area are not visible to users. Therefore, more redundant data than conventional video are transmitted if we directly use the DASH multicast for VR-video dissemination. To be more efficient, redundant data should be removed and the tiled video [28] allows to flexibly remove or allocate lower bits to the redundant part of the video. For example, necessary parts of video are transmitted with multiple copies with different quality to satisfy users with good channel quality and the parts with lower probability of view are transmitted just once with single quality to save spectrum.

The most popular and promising technology for controlling regional quality of the video is the use of tiled videos, which has been used for the panoramic interactive video [31], since the interactive video can change its view, and users cannot see the whole video at once. VR video is divided into smaller rectangular videos (tiles) and each video is encoded independently using legacy video encoders. Every tile has multiple copies with different encoding rates. Different representations of tiles are transmitted as users' viewport changes and network channel condition changes.

There are simple rate allocation algorithms for tiled-videos, which are Binary, Thumbnail, and Pyramid [32]. Binary allocates the higher representations on the visible tiles, and non-visible tiles have lowest representations to save the bandwidth. It is the most bandwidth efficient way to allocate the bits, but users can easily see the lowest quality when they move their viewport since the network has latency to respond with viewport changes. Thumbnail allocates the minimum bits of lowest representations for the whole video as the background video, and remaining bits are allocated for visible tiles with better representations. However, users still can see the lowest quality background video when they move the viewport faster than network latency. The Pyramid algorithm allocates the best representations on visible tiles and gradually lower the representations as 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 the users with variable network channel condition and viewport movement.

Alface et al. [33] propose a rate-selection algorithm to provide the best quality to users with a higher representation in the viewport and lower representations in 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 for tiles in order to maximize the total utility as long as there is available resource, the algorithm can achieve the best utility performance compared to other existing solutions.

However, none of the existing algorithm is directly applicable to the multicasting scenario. A new approach to perform the VR video multicasting is proposed in this paper.

# Chapter 3

# OPTIMAL DASH MULTICASTING OVER LTE

## *3.1 Overview*

This chapter describes the proposed optimal DASH multicasting over LTE method that achieves the best utility performance by optimally allocating the resource. Heuristic algorithms have been introduced to solve the video multicasting problem. These algorithms are differentiated based on types of the video sources. Scalable Video Coded (SVC) [18] videos are originally used for video multicasting systems because of its layer-dependent characteristics. More enhancement video layers can be combined with the base layer to create higher quality video for the users who have good channel conditions, while the users with poorer channel quality can only receive less enhancement video layers coded with more reliable but less efficient MCS. For the DASH systems, usually videos are encoded as multiple different video rates and stored at the server as small chunks, and they are transmitted to the clients who request the videos. Therefore, different video representations are independent each other and they can be scheduled independently for multicasting. DASH can also transmit SVC type video sources, but, in this paper, for notational convenience, DASH only denotes multiple video rates without dependencies among representations and SVC denotes the layered video with dependencies among layers.

## *3.2 System Model*

The system models are described as mathematical formats. Table 3.1. includes the notations used in this paper.

### A. Utility Model

There are several different ways to quantify the performance of wireless systems, e.g., delay/jitter, throughput, bit error rate (BER) or outage probabilities are the traditional quality of service (QoS)

metrics to describe the performance of a wireless system [34]. However, these parameters are not adequate to show how actual users are experiencing when media data are involved. Therefore, many of recent research uses utility as a metric to describe the performance of a wireless system [35][36].

Table 3.1. NOTATIONS USED IN THIS CHAPTER

| | |
|---|---|
| $c(MCS)$ | Capacity of an RB with selected MCS |
| $u_{v,m}$ | Utility of *m*-th representation of video *v* |
| $\tilde{u}_{v,m}$ | Effective utility |
| $a, b$ | Normalization coefficients of $u_{v,m}$ |
| $R_{v,m}$ | Video rate for *m*-th representation of video *v* |
| $R_{v,\max}$ | Maximum video rate of video *v* |
| $f_{mar}$ | Margin factor |
| $m$ | Video representation index |
| $v$ | Video index |
| $M$ | Number of video representations |
| $V$ | Number of videos |
| $\gamma$ | Signal-to-Noise Ratio (SNR) |
| $P_r$ | Received power at UE |
| $N_0$ | Thermal noise variance |
| $P_t$ | Transmit power at eNB |
| $d$ | Distance between eNB and UE |
| $d_0$ | Reference distance |
| $d_{th}$ | Threshold distance |
| $R$ | Cell radius |
| $\overline{\gamma}_{th}$ | Average threshold SNR |
| $\overline{\gamma}_{i,dB}$ | Average SNR in dB scale of UE *i* |
| $\gamma_{v,m}$ | Threshold SNR for FLUTE session *v, m* |
| $K_{v,m}$ | Number of data packets of FLUTE session *v, m* |
| $\mathbf{N}_{v,m}$ | Set of UEs subscribing FLUTE session *v, m* |
| $N_{RB}$ | Total number of RB |
| $U$ | Total utility |
| $U_{v,m}$ | Utility of FLUTE session *v, m* |

There are many different ways to define utility. In this paper, since we are working on video streaming, we have adopted the popular utility model based on a function of video transmission rates. If a user can receive higher video rates, then the utility will be higher. However, it is known that the user experience is not linearly proportional to the video rate, since it saturates with higher video rates.

In this paper, to quantify the quality of experience associated with a video, we adopt the most popular logarithmic law definition to quantify the video quality of experience because it considers human perception. Most of the mean opinion score (MOS) tests show that the human perception of video quality saturates with higher video rates [37][38]. Moreover, the proposed algorithm is not limited to use logarithmic law. It works with many utility functions to describe the quality of the video with some necessary modifications on the optimization procedures.

The utility function based on the logarithmic law [39] is,

$$u_{v,m} = \begin{cases} a \log b \dfrac{R_{v,m}}{R_{v,\max}}, & R_{v,m} > 0 \\ 0, & R_{v,m} = 0 \end{cases}, \tag{3.1}$$

where $u_{v,m}$ and $R_{v,m}$ denote the utility and rate for the $m$-th video representation of video $v$ respectively; $R_{v,\max}$ is the maximum video representation of video $v$ stored in the server; $a$ and $b$ are the coefficients to normalize the utility $u_{v,m}$ to stay in the range between 0 and 1, and they can be empirically determined for different applications. Note that $u_{v,m}$ equals 1 when $R_{v,m} = R_{v,\max}$, which implies that subscribers can watch the best video representation.

## B. User Distribution and Average SNR

We have considered the subscribing users in the LTE network to be uniformly distributed in a single cell area. Since the signal power attenuates as a function of distance from the eNB, we have

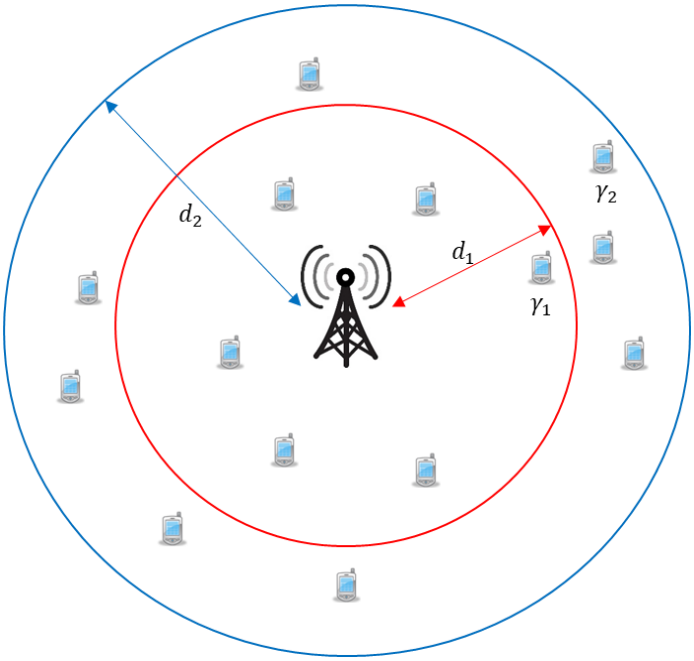



Figure 3.1. User distribution and grouping

derived the received power distribution of the users using the path loss model [40]. More specifically, the average SNR $\bar{\gamma}_i$ of the user $i$ is a function of distance from the eNB,

$$g(d) = \bar{\gamma}_i = \frac{P_r}{N_0} = \frac{P_t K}{N_0}\left(\frac{d_0}{d}\right)^{\lambda}, \tag{3.2}$$

where $P_t$ is the transmit power, $P_r$ is the receiving power at the UE, $d_0$ denotes reference distance, $\lambda$ is pathloss exponent, and $P_t K$ is the power measured at the reference distance. The inverse function of $g(d)$ is

$$g^{-1}(\bar{\gamma}_i) = d = d_0\left(\frac{P_t K}{N_0 \bar{\gamma}_i}\right)^{\frac{1}{\lambda}}. \tag{3.3}$$

The probability that the user is located within $d_{th}$ is

$$\Pr(d \le d_{th}) = \frac{d_{th}^2}{R^2}, \tag{3.4}$$

where $R$ is the cell radius. It is the same as the probability that the user's average Signal-to-Noise Ratio (SNR) is bigger than $\bar{\gamma}_{th}$, which is the average SNR at distance $d$. Therefore, the outage probability when the average threshold SNR is $\bar{\gamma}_{th}$ can be described as,

$$\Pr(\bar{\gamma}_i \le \bar{\gamma}_{th}) = 1 - \Pr(d \le d_{th}) = 1 - \frac{d_0^2}{R^2}\left(\frac{P_t K}{N_0 \bar{\gamma}_{th}}\right)^{2/\lambda}. \tag{3.5}$$

If we change the SNR into dB scale, we can get the function as shifted exponential form, and it is the cumulative distribution function (CDF) of user's SNR,

$$\Pr(\bar{\gamma}_{i,dB} \le \bar{\gamma}_{th,dB}) = F(\bar{\gamma}_{th,dB}) = 1 - \exp\left(-\frac{\bar{\gamma}_{th,dB} - \mu}{\beta}\right) \tag{3.6}$$

where $\mu = \beta \log\left(\frac{d_0}{R^2}\left(\frac{P_t K}{N_0}\right)^{1/\lambda}\right)$ and $\beta = \frac{5\lambda}{\log 10}$.

Fig 3.1. shows how users are distributed in a cell area and grouped into different multicast groups. As we decide the threshold SNR of group 1 as $\bar{\gamma}_1$, we can estimate the number of users who can join the group 1 (red area) by using (3.6), i.e., the circle with radius $d_1$. The remaining users outside radius $d_1$ will be assigned to group 2.

C. Exponential Effective SNR Mapping (EESM)

An RB is the smallest resource unit of an OFDMA system and each RB's channel quality is measured at the UE side and fed back to the eNB to decide which MCS has to be used. An RB consists of 12-subcarriers, therefore, it is difficult to describe the channel quality of the RB as a Signal-to-Noise Ratio (SNR) measured in one subcarrier. Exponential Effective SNR Mapping (EESM) [41] is used to describe effective SNR of the RB, and it works as an indicator to decide

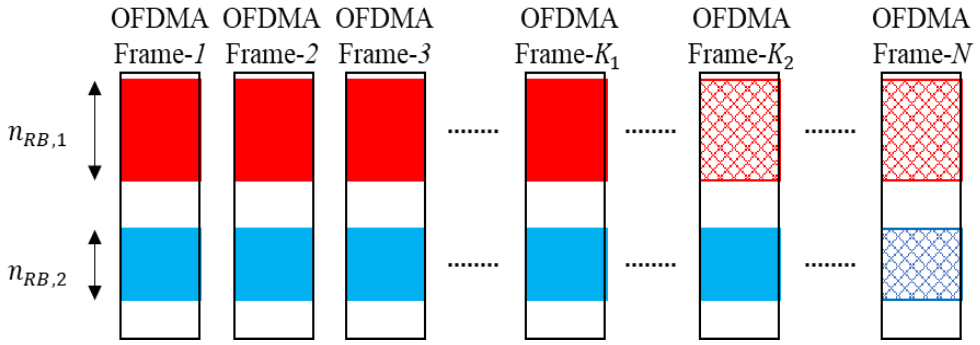



Figure 3.2. FEC block, grouping and resource allocation

which MCS has to be selected to ensure reliable communications. It has been proven that the probability distribution of EESM is Log-normal, while the wireless channel statistics follows the Rayleigh distribution [42]. Therefore, the outage probability of the user $i$ is shown as (3.7) when MCS for the *m*-th representation of video *v* is selected.

$$P_{out}(\gamma_{v,m}, \overline{\gamma}_{i,dB}) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\gamma_{v,m} - \overline{\gamma}_{i,dB}}{\sigma\sqrt{2}}\right)\right], \tag{3.7}$$

where $\gamma_{v,m}$ represents the minimum SNR ensuring a block error rate (BLER) value of the *m*-th FLUTE session of video *v* lower than 1%. Note that different MIMO or HARQ schemes can affect $\gamma_{v,m}$ [43], but since MCH cannot take advantage of HARQ, we do not consider HARQ. With different MIMO setting, we can adopt different $\gamma_{v,m}$ [44].

## D. Application-layer Forward Error Correction (FEC)

The radio channel conditions vary among all the users receiving the multicast service. Therefore, the block error rate of the users that receive the video service delivered with a single MCS can have a great variance. In order to increase the robustness and the reliability of the multicast transmissions, FEC redundancy packets are incorporated at the APP-layer [5].

The solution proposed by 3GPP to deliver video streaming over eMBMS uses the FLUTE protocol with UDP transport to send video segments with the corresponding APP-layer FEC over multicast [45][46]. An FEC block consists of $N$ packets including $K$ source packets and $N-K$ redundancy packets, resulting in the encoding rate $K/N$. The FEC decoder can ideally recover the original $K$ source packets from any $K$ out of $N$ received packets with correction capability $t = N - K$. The Reed Solomon (RS) code is a well-known FEC code which operates on non-binary symbols and has the ideal correction capability. However, the RS code has a high decoding complexity because of its non-binary operations, which is not suitable for high-definition (HD) video streaming applications. The Raptor code is a more attractive solution for HD video streaming services due to the flexible parameters selection and linear decoding cost. The correction capability of a Raptor code is $t = N - (1+\varepsilon)K$, where $\varepsilon$ is the reception overhead efficiency. The correction capability of the Raptor code is sub-optimal, however, a standardized Raptor code can closely achieve the ideal correction capability with negligible $\varepsilon$. Therefore, it is used in our scheme. In this paper, FEC block size is fixed as $N$ and the number of source blocks $K_{v,m}$ is determined to choose appropriate FEC code rates $K_{v,m}/N$ for the *m*-th FLUTE session of video *v*. Fig 3.2. shows the example of an FEC block with two multicasting groups sharing the resource with different FEC code rates. $n_{RB,1}$ and $n_{RB,2}$ are the number of resource blocks allocated in an OFDMA frame for groups 1 and 2 respectively. There are total $N$ OFDMA frames and $K_1$ and $K_2$ of them are video data and others are redundant data.

### *3.3 Problem Formulation*

The maximum utility problem under a limited number of resource blocks, MCS selection, and FEC code-rates can be defined as,

$$\mathrm{Q}: \underset{\mathbf{n}_{RB},\mathbf{MCS},\mathbf{K}}{\text{maximize}}\ U = \sum_{v=1}^{V}\sum_{m=1}^{M}\tilde{u}_{v,m}\,|\,\mathbf{N}_{v,m}\,|, \tag{3.8}$$

$$\text{subject to } \sum_{v=1}^{V}\sum_{m=1}^{M} n_{RB,v,m} \le N_{RB},\ n_{RB,v,m} \ge 0, \tag{3.9}$$

where $U$ denotes the total utility of the multicasting system; $n_{RB,v,m}$ is the allocated number of resource blocks to the *m*-th FLUTE session of video *v*; $MCS_{v,m}$ is the MCS index for the *m*-th representation for video *v* defined in the LTE standard, and $K_{v,m}$ is the number of OFDMA data frames out of *N*-OFDMA data frames, which composes one FEC-block.

The utility *U* is defined as a summation of number of users $|\mathbf{N}_{v,m}|$ who subscribe the *m*-th representation of video *v* times its effective utility value $\tilde{u}_{v,m}$, which reflects the user's received quality of experience (QoE) and is slightly different from the utility values defined in (3.1). The utilities of each FLUTE session are

$$\left\{\begin{array}{c} U_{v,M} = u_{v,M}\left|\mathbf{N}_{v,M}\right| \\ \vdots \\ U_{v,m} = u_{v,m}\left|\mathbf{N}_{v,m}\right| - u_{v,m}\left|\mathbf{N}_{v,m+1}\right| \\ \vdots \\ U_{v,1} = u_{v,1}\left|\mathbf{N}_{v,1}\right| - u_{v,1}\left|\mathbf{N}_{v,2}\right| \end{array}\right\}. \tag{3.10}$$

where $U_{v,m}$ denotes the utility of the *m*-th FLUTE session for video *v*. If we sum them together,

$$\sum_{m=1}^{M} U_{v,m} = \sum_{m=2}^{M} (u_{v,m} - u_{v,m-1})\left|\mathbf{N}_{v,m}\right| + u_{v,1}\left|\mathbf{N}_{v,1}\right|, \tag{3.11}$$

we can define the effective utility value as,

$$\tilde{u}_{v,m} = \begin{cases} u_{v,m} - u_{v,m-1}, & m > 1 \\ u_{v,1}, & m = 1 \end{cases}. \tag{3.12}$$

To maximize the utility of the system, the optimal $n_{RB,v,m}, MCS_{v,m}$, and, $K_{v,m}$ have to be determined. We assumed that the clients are desired to choose the best video representations they

can subscribe, and the eNB allocates the resource to maximize the total utility received by all the subscribing users under the limited total available resource blocks $N_{RB}$.

The condition for the user $i$ to successively decode the video is

$$f_{mar} \cdot (1 - P_{out}(\gamma_{v,m}, \bar{\gamma}_{i,dB})) \geq \frac{K_{v,m}}{N}, \tag{3.13}$$

where $f_{mar}$ is the error margin factor, which is introduced to ensure the Raptor code can closely achieve the ideal correction capability with negligible $\varepsilon$ so that the received video can be effectively decoded by the Raptor FEC decoder [13]. If we assume that the wireless channel follows the Rayleigh fading model and effective SNR of the resource block follows the log-normal distribution, the outage probability is (3.7). Applying (3.13) to the (3.7), we can find the condition that the video can be decoded without any error,

$$\bar{\gamma}_{i,dB} \geq \gamma_{v,m} - \sigma\sqrt{2}\,\mathrm{erf}^{-1}\left(1 - \frac{2}{f_{mar}}\frac{K_{v,m}}{N}\right). \tag{3.14}$$

Note that, (3.14) shows the condition for average SNR of user $i$ to successively decode the video. The number of users who can subscribe the $m$-*th* video representation can be found by using the statistical model of the average SNR in the LTE network. The distribution of the users average SNR follows the shifted exponential distribution. Since the CDF shows the cumulative probability that the user's average SNR is lower than the threshold SNR $\bar{\gamma}_{th,dB}$, (3.6) exactly denotes the failure probability of decoding the video. Therefore, to find the estimated number of users who can decode the video in the LTE network is

$$|\mathbf{N}_{v,m}| = \left\lfloor |\mathbf{N}_v| (1 - F(\bar{\gamma}_{th,dB})) \right\rfloor \tag{3.15}$$

$$= \left\lfloor |\mathbf{N}_v| \exp\left(-\frac{\bar{\gamma}_{th,dB} - \mu}{\beta}\right) \right\rfloor. \tag{3.16}$$

The allocated resource for the *m*-th FLUTE session of video *v* must be enough to carry whole video data, therefore, the following inequality must be hold.

$$R_{v,m} \le n_{RB,v,m} c(MCS_{v,m}) \frac{K_{v,m}}{N}, \tag{3.17}$$

$$\frac{K_{v,m}}{N} \ge \frac{R_{v,m}}{n_{RB,v,m} c(MCS_{v,m})}. \tag{3.18}$$

Applying (3.14) to (3.16) we can then make $|\mathbf{N}_{v,m}|$ as a function of the threshold SNR, error margin factor, and FEC rate,

$$|\mathbf{N}_{v,m}| = \left\lfloor |\mathbf{N}_v| \exp\left( -\left( \gamma_{v,m} - \sigma\sqrt{2}\,\mathrm{erf}^{-1}\left(1 - \frac{2}{f_{mar}} \frac{K_{v,m}}{N}\right) - \mu \right) \Big/ \beta \right) \right\rfloor \tag{3.19}$$

$$|\mathbf{N}_{v,m}| = \left\lfloor |\mathbf{N}_v| \exp\left( -\left( \gamma_{v,m} - \sigma\sqrt{2}\,\mathrm{erf}^{-1}\left(1 - \frac{2}{f_{mar}} \frac{R_{v,m}}{n_{RB,v,m} c(MCS_{v,m})}\right) - \mu \right) \Big/ \beta \right) \right\rfloor \tag{3.20}$$

If (3.18) holds, FEC coding rate does not have to be larger than the estimated condition on the right side of (3.18), which means that we can use equality to describe (3.18). More intuitive understanding is that there is no reason to choose larger FEC rate, since it means less protection for packet loss. We could then find the estimated number of users who can subscribe the *m-th* video representation is a function of $n_{RB,v,m}$, when we fix the MCS as shown in (3.18).

The goal of this research is to find the optimal resource allocation, specifically to find the best number of RBs allocated for each multicast group so as to maximize the total utility, subject to the limited number of total consumed RBs. Therefore, a problem solving strategy is proposed in the next section.

### *3.4 Problem-Solving Strategy*

There are three parameters to be determined for optimal video multicasting problem, more specifically, they are resource allocation of RBs, FEC code rate, and MCS. The number of DASH video representations to be scheduled is an optimization parameter too, but we did not include it as a part of optimization problem, because we can perform the proposed optimization algorithm for each combination set of video representations and choose the best set to maximize the total utility.

These parameters are discrete values, therefore it is difficult to find a closed-form solution. We divide the problem solving process into three steps. The first step is finding optimal resource allocation and the corresponding APP-layer FEC code rate, assuming the scheduled set of video representations and MCS are known. The second step is to find the best MCS combination by choosing the best MCS for each FLUTE session to maximize the utility using allocated resource and FEC rate found in first step. Finally, it finds the best set of video representations to be scheduled as FLUTE sessions to maximize the total utility. This step depends on applications especially type of the video sources. In case of SVC videos, dependencies among layers need to be considered when the algorithm schedules the video layers. However, with the general DASH scenarios, any different combination of video representations can possibly be scheduled, therefore, it has larger search space than SVC videos.

A. Optimal Resource Allocation and FEC Code Rate

Since $\left|\mathbf{N}_{v,m}\right|$ and $n_{RB,v,m}$ are integer values, to find the optimal solution effectively, we have relaxed the variables as continuous variables. After finding the optimal solution, we can round the variables into integer values again. Optimality of the solution could be shown by comparing with the results based on exhaustive searches.

Since $|\mathbf{N}_{v,m}|$ is a concave function of $n_{RB,v,m}$, the optimal solution for resource allocation can be found by a standard convex optimization process. A simple way to prove convexity is to use the property of the convex function [47].

As shown in (3.20), $|\mathbf{N}_{v,m}|$ is an exponential function, which is concave w.r.t. its arguments. The inverse error function is also concave function with negative arguments. Moreover, $-1/n_{RB,v,m}$ is a concave function w.r.t. $n_{RB,v,m}$. In conclusion, $|\mathbf{N}_{v,m}|$ is a concave function of $n_{RB,v,m}$. Therefore, we can rewrite the original problem (9) as a standard convex optimization problem as

$$\tilde{\mathrm{Q}}: \underset{\mathbf{n}_{RB},\mathbf{MCS},\mathbf{K}}{\text{minimize}} \; -U = -\sum_{v=1}^{V}\sum_{m=1}^{M}\tilde{u}_{v,m}\,|\mathbf{N}_{v,m}|, \tag{3.21}$$

$$\text{subject to} \sum_{v=1}^{V}\sum_{m=1}^{M} n_{RB,v,m} \le N_{RB}, \quad n_{RB,v,m} \ge 0. \tag{3.22}$$

By using the Lagrangian method, the objective function of the problem $\tilde{\mathrm{Q}}$ is shown in (3.23), where $\varphi$ and $\zeta_{v,m}$ are the Lagrangian multipliers of the equality and inequality constraints respectively.

$$L(\mathbf{n}_{RB},\varphi,\boldsymbol{\zeta}) = -\sum_{v=1}^{V}\sum_{m=1}^{M}\tilde{u}_{v,m}\,|\mathbf{N}_{v,m}| + \varphi\left(\sum_{v=1}^{V}\sum_{m=1}^{M} n_{RB,v,m} - N_{RB}\right) - \sum_{v=1}^{V}\sum_{m=1}^{M}\zeta_{v,m} n_{RB,v,m} \ge 0, \tag{3.23}$$

The gradient of Lagrangian (3.24)

$$\frac{\delta L(\mathbf{n}_{RB},\varphi,\boldsymbol{\zeta})}{\delta n_{RB,v,m}} = \tilde{u}_{v,m}\frac{\sqrt{\pi}\,|\mathbf{N}_v|\,A_{v,m}B_{v,m}}{2n_{RB,v,m}^2}\exp\left\{\left|\mathrm{erf}^{-1}\left(1-\frac{A_{v,m}}{n_{RB,v,m}}\right)\right|^2 + B_{v,m}\mathrm{erf}^{-1}\left(1-\frac{A_{v,m}}{n_{RB,v,m}}\right)\right\} + \varphi - \zeta_{v,m} \tag{3.24}$$

can now be derived from (3.23), where

$$A_{v,m} = \frac{2R_{v,m}}{f_{\text{mar}} c(MCS_{v,m})}, \quad B_{v,m} = \frac{\sigma\sqrt{2}}{\beta}.$$

The optimality condition is derived by using Karush-Kuhn-Tucker (KKT) conditions.

1. Primal feasibility : $\sum_{v=1}^{V}\sum_{m=1}^{M} n_{RB,v,m} = N_{RB}, \quad n_{RB,v,m} \geq 0$.

2. Dual feasibility : $\zeta_{v,m}^{\ *} \geq 0.$

3. Complementary slackness : $\zeta_{v,m}^{\ *} n_{RB,v,m}^{*} = 0.$

The gradient of Lagrangian vanishes when (3.25) holds.

$$\zeta_{v,m}^{*} = \varphi^{*} - \tilde{u}_{v,m} \frac{\sqrt{\pi} \mid \mathbf{N}_v \mid A_{v,m} B_{v,m}}{2n_{RB,v,m}^{2}} \exp\left\{ \left| \text{erf}^{-1}\left(1 - \frac{A_{v,m}}{n_{RB,v,m}}\right)\right|^{2} + B_{v,m}\text{erf}^{-1}\left(1 - \frac{A_{v,m}}{n_{RB,v,m}}\right)\right\} \geq 0 \qquad (3.25)$$

Thus, the optimal $n_{RB,v,m}$ can be found by solving these optimality conditions. The dual feasibility implies (3.26) and the complementary slackness implies (3.27).

$$\left( \varphi^{*} - \tilde{u}_{v,m} \frac{\sqrt{\pi} \mid \mathbf{N}_v \mid A_{v,m} B_{v,m}}{2n_{RB,v,m}^{2}} \exp\left\{ \left| \text{erf}^{-1}\left(1 - \frac{A_{v,m}}{n_{RB,v,m}}\right)\right|^{2} + B_{v,m}\text{erf}^{-1}\left(1 - \frac{A_{v,m}}{n_{RB,v,m}}\right)\right\} \right) n_{RB,v,m}^{*} = 0, \qquad (3.26)$$

$$g(n_{RB,v,m}) = \varphi^{*} = \tilde{u}_{v,m} \frac{\sqrt{\pi} \mid \mathbf{N}_v \mid A_{v,m} B_{v,m}}{2n_{RB,v,m}^{2}} \exp\left\{ \left| \text{erf}^{-1}\left(1 - \frac{A_{v,m}}{n_{RB,v,m}}\right)\right|^{2} + B_{v,m}\text{erf}^{-1}\left(1 - \frac{A_{v,m}}{n_{RB,v,m}}\right)\right\} \qquad (3.27)$$

Table 2.2. BISECTION SEARCH ALGORITHM

1. $upper = \min(g_{v,m}(1))$, for $m = 1,2,\ldots,M$, $v = 1,2,...,V$
2. $lower = \max(g_{v,m}(N_{RB}))$, for $m = 1,2,\ldots,M$, $v = 1,2,...,V$
3. $n_{RB,v,m} = 0$, for $m = 1,2,\ldots,M$, $v = 1,2,...,V$
4. while $\left( \left| \sum_{v=1}^{V} \sum_{m=1}^{M} n_{RB,v,m} - N_{RB} \right| > \epsilon \right)$
5. $\mu = (upper + lower)/2$;
6. $n_{RB,v,m} = g_{v,m}^{-1}(\mu)$, for $m = 1,2,\ldots,M$, $v = 1,2,...,V$
7. if $\left( \sum_{v=1}^{V} \sum_{m=1}^{M} n_{RB,v,m} - N_{RB} < 0 \right)$ $lower = \mu$;
8. else $upper = \mu$; end if
9. end while
10. $n^{*}_{RB,v,m} = \lfloor n^{*}_{RB,v,m} \rfloor$, for $m = 1,2,\ldots,M$, $v = 1,2,...,V$

The optimal $\{ n^{*}_{RB,v,m} \}$ for all FLUTE sessions satisfying above optimality conditions can now be found by using the bisection search algorithm shown in Table 3.2, since (3.27) is a decreasing function with respect to $n^{*}_{RB,v,m}$. Note that $\epsilon$ is a small positive number to stop the searching process when $n^{*}_{RB,v,m}$ is close to the optimal value. Therefore, we can find the optimal resource allocation matrix $\boldsymbol{n}_{RB}^{V \times M}$ and FEC rate matrix $\mathbf{K}^{V \times M} / N$ as determined by (3.18).

To show the optimality of the solution, we have compared the proposed algorithm with an exhaustive search algorithm when three different videos representations are scheduled. The details of the optimal resource allocation results of the proposed algorithm is shown in Table 4, where 100 resource blocks per OFDMA frame are allocated in total with an average SNR of 20dB. There are 100 users in total subscribing the same video and there are 3-different video representations (1.0, 2.8, 4.7 Mbps) associated with this video. It needs too many resource blocks if all the users

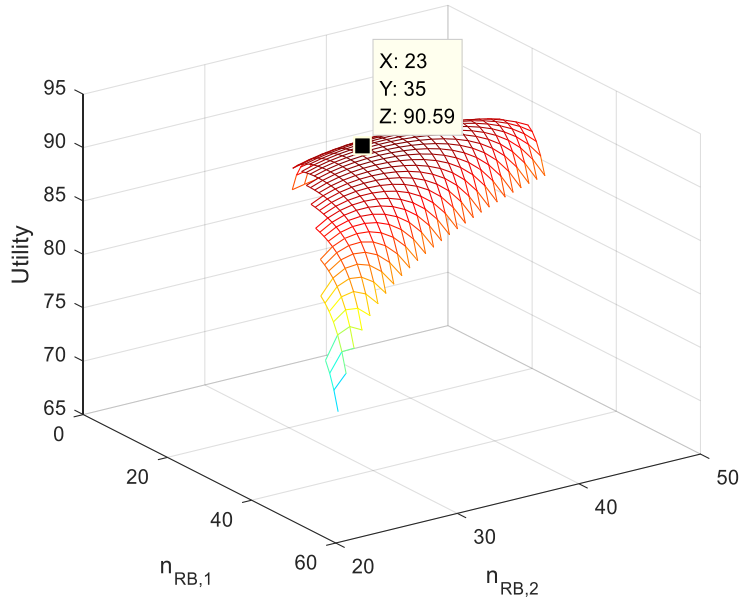



Figure 3.3. Optimal resource allocation

want to receive the best video rate, since it should either select very inefficient MCS or use many redundancy FEC packets. If all the users select lowest video rate, i.e., everyone might receive the video, but users may not be satisfied with the video quality. The proposed algorithm groups the subscribers into 3-groups and chooses a different video representation for each of them, resulting in different numbers of resource blocks usage for each group based on also different choice of MCS and FEC rates. Accordingly, some of users can receive the highest video rates while some of the users can receive the medium/lowest rates, even with worse channel conditions. The results of our proposed scheme are exactly matched with the results from the exhaustive search (see Fig. 4). For the exhaustive search, we try all possible combinations of resource allocations for 3-different video rates and all MCS combinations, as well as choose the maximum utility. Fig. 3.3 shows the utility with all possible combinations of the resource allocations. The *x-axis* denotes the number of RBs for the first representation R1, and *y-axis* denotes the number of RBs for the second representation R2. Remaining RBs are allocated to the third representation R3. The *z-axis* denotes the utility of the combinations. The point having maximum utility is exactly matched with the result from the proposed algorithm.

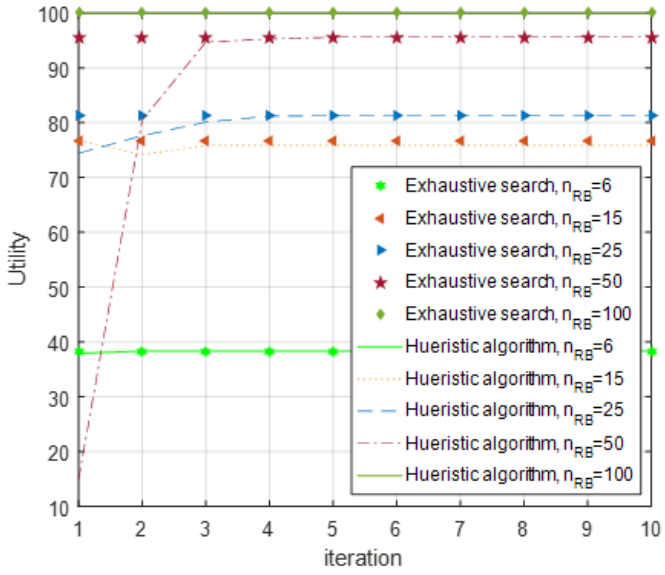



Figure 3.4. Heuristic MCS selection algorithm

Table 3.3. OPTIMAL SOLUTION for 3-REPRESENTATIONS

| Representations | R1 | R2 | R3 |
|---|---|---|---|
| Video rates | 1.0Mbps | 2.8Mbps | 4.7Mbps |
| $n_{RB,v,m}$ | 23 | 35 | 42 |
| $\lvert\mathbf{N}_{v,m}\rvert$ | 8 | 11 | 76 |
| $K_{v,m} / N$ | 0.45 | 0.56 | 0.58 |
| $MCS_{v,m}$ | 4 | 5 | 6 |

B. MCS Selection

There are $15^{MV}$ MCS combinations. All trials can eventually have their corresponding resource allocation matrix $\mathbf{n}_{RB}^{M\times V}$ and FEC rate matrix $\mathbf{K}^{M\times V}/N$ as determined by the proposed scheme. MCS selection is not a part of the convex optimization process, even though it has a significant effect on the system performance. Therefore, to find the optimal utility, we have to try all possible

combinations of MCS for $M$ different rates and for $V$ different videos. A heuristic-based MCS selection algorithm is shown in Table 3.4. If the number of resource blocks is fixed, the best MCS is obvious because we always need to maximize the number of subscribers who can receive the video within the limited number of resource blocks. The bisection algorithm in Table 3 can give us the optimal solution for the resource allocation and the FEC code rate when the MCS is fixed. The algorithm can be used to iteratively choose the best MCS with a fixed number of RBs. For the initial MCS set, all resource blocks are equally divided among all multicast groups. With these resource blocks, we can find the best MCS for each group. After performing the bisection search algorithm, we can get the optimal resource allocation for all groups. The algorithm is repeated iteratively until the solution converges. We find that the iteration converges within 10 iterations most of time. Therefore, finding the optimal MCS using the algorithm shown in Table 3.4 is significantly faster than performing an exhaustive search. Fig. 3.4 shows the utility with different total number of RBs when the heuristic MCS selection algorithm and exhaustive search are used. The exhaustive search tries all possible MCS combinations and take the best result. As iteration increases, the utility of the heuristic MCS selection algorithm get close to the utility result of exhaustive search. Therefore, the heuristic MCS selection algorithm is used for further simulation results to reduce the complexity.

Table 3.4. MCS SELECTION ALGORITHM

1. Initialize $n_{RB,v,m} = \left\lfloor \frac{N_{RB}}{MV} \right\rfloor$, for $m = 1, 2, \ldots, M$, $v = 1, 2, ..., V$
2. Repeat
3. $MCS'_{v,m} = \arg\max_{MCS_{v,m}} \left| \mathbf{N}_{v,m}(n_{RB,v,m}, \mathbf{MCS}) \right|$
4. Do the bisection search algorithm in Table 3 with $MCS'_{v,m}$ to find the best $\boldsymbol{n}_{RB}$
5. until $MCS'_{v,m}$ are converged for $m = 1, 2, \ldots, M$, $v = 1, 2, ..., V$

C. Rate or Layer Selection

The last part of the optimization is selecting the video representations to schedule. There can be maximum *M*-representations scheduled for one video, and if *m*-representations are chosen, there are $\binom{M}{m}$ possible combinations. If there are *V* distinct videos, the number of combinations is $\left(\sum_{m=1}^{M}\binom{M}{m}\right)^{V}$. To find the optimal solution, all possible combinations are tested. After testing all possible combination, we choose the rate combination that has the best utility.

In case of SVC, there are only $M$ different choices for layer selection of each video because of the layer dependency. Therefore, total searching space to find optimal solution is $M^{V}$.

### *3.5 Performance Analysis*

A. Simulation Setup

The performance analysis is conducted using standard LTE parameters [48], i.e., the exponential effective SNR mapping [41] is first used to map the channel state into effective SNR, which is then mapped onto the CQI level, ensuring a block error rate (BLER) value lower than 10%. The systems parameters adopted in our simulations are suggested by the LTE standard as summarized in Table 6. Only 1 Tx and 1 Rx antenna setting is used for simulations, but other MIMO configurations can be applied by choosing different SNR and CQI mapping table [11]. Massive MIMO system [49] is one of the important methods to improve received SNR cooperating with the proposed algorithm. For propagation loss, the COST 231 suburban model [50] with the standard deviation of lognormal shadowing 8dB is used. Besides, ITU Ped-B power delay profile is also used in our simulations.

The subscribing UEs of each multicast group are randomly distributed in a single cell area, with the eNodeB being located in the center. We have generated 100 uniformly distributed users

(2-Dimentional Poisson Point Process) for each of the 4 distinct videos (Crew, Football, City, and Harbour), therefore, there are in total 400 users subscribing videos. eNB can allocate the spectrum resource for multicasting from 15% to 55% of whole spectrum resource, which corresponds to 15, 25, 35, 45, and 55 RBs in one OFDMA frame.

Table 3.5. LTE PARAMETERS

| Parameter | Value |
|---|---|
| Cell Radius | 500m |
| 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 | 2GHz |
| Scheduling frame | 10ms |
| RB size | 12 sub-carrier, 0.5ms |
| Sub-carrier spacing | 15kHz |
| TTI | 1ms |
| EUTRA UE | Antenna gain 0dBi, NF 9dB |
| EUTRA eNB | Antenna gain 14dBi, NF 5dB |
| eNB transmit power | 43dBm |
| MIMO configuration | 1 Tx, 1 Rx |
| Thermal Noise | -174dBm/Hz |

B. SVC

Our proposed algorithm is compared with existing video multicasting algorithms, where OLM and MSML are the most efficient and advanced multicasting algorithm for OFDMA systems as reported in [25]. Since they are developed for layered video multicasting, in this paper, we compare the proposed algorithm both with SVC video sources and DASH video sources. The difference between these two video sources is layer dependency. Since the SVC videos always need lower layers to decode upper layers, lower layers must be scheduled first. Therefore, the scheduling

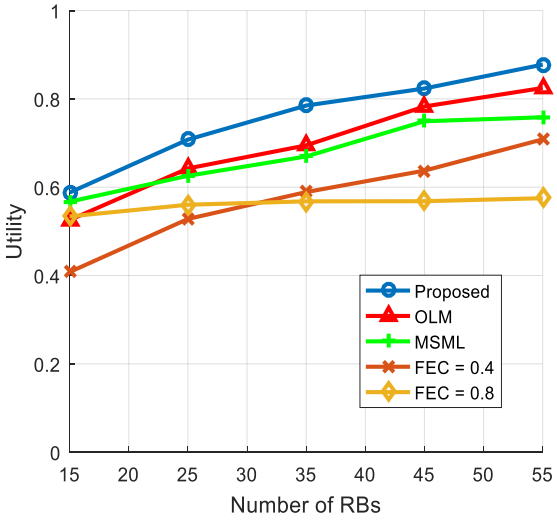



Figure 3.5. Utilities of SVC dataset

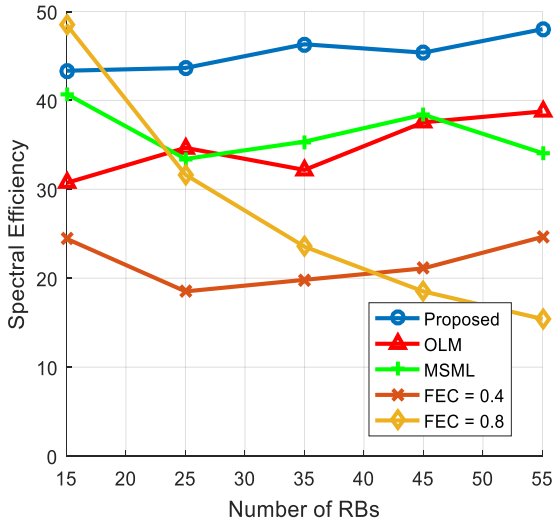



Figure 3.6. Spectral Efficiencies of SVC dataset

algorithm is restricted by the dependency. However, in DASH case, video rates are independent and every user needs only one video rate to playback the video. If some users can subscribe multiple video representations at the same time, it will choose better representation and discard other representations.

Table 3.6. VIDEO RATES (Kbps) for SVC SIMULATION

| | Video | BL | L1 | L2 | L3 |
|---|---|---|---|---|---|
| 1 | CREW | 306 | 578 | 814 | 1184 |
| 2 | FOOTBALL | 442 | 827 | 1114 | 1621 |
| 3 | CITY | 448 | 923 | 1288 | 1943 |
| 4 | HARDBOUR | 577 | 1025 | 1379 | 1929 |

Table 3.6 shows the accumulating video rates of 4 distinct SVC video sources stored in the video server to be adaptively multicasted through the eMBMS channels. These video layers are scheduled based on the proposed algorithm, OLM, and MSML separately. Static resource allocation schemes with FEC rates as 0.4 and 0.8 are also compared to see the effect of FEC. Utility,

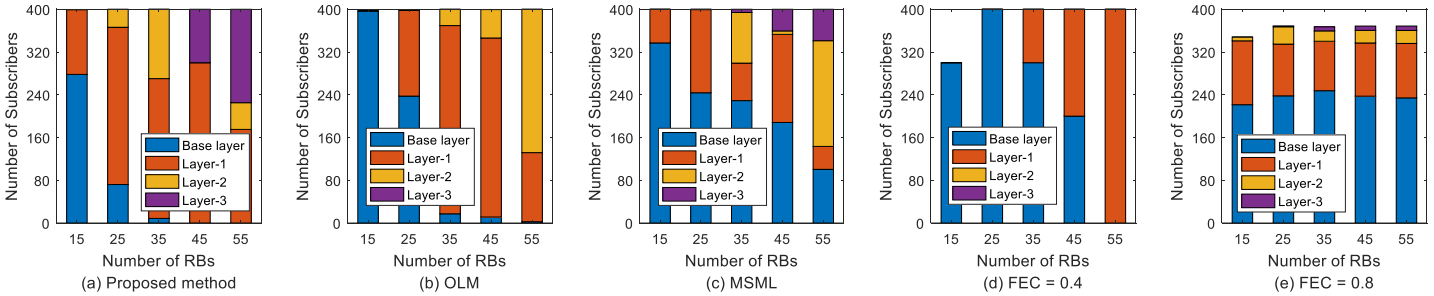



Figure 3.7. Number of subscribers

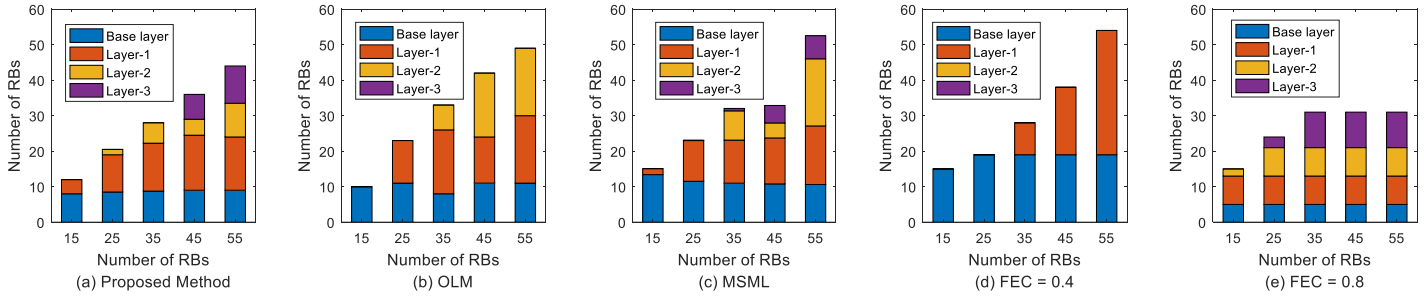



Figure 3.8. Resource allocation results

spectral efficiency, actual number of subscribers of each video and layer, and resource consumptions are measured during the simulations for performance evaluation.

Fig. 3.5. shows the obtained utility of the SVC streaming over eMBMS channels. The proposed algorithm can achieve the highest utility when compared to OLM and MSML, which are the state-of-the-art algorithms for SVC video multicasting. The proposed algorithm, OLM, and MSML all can achieve better utility than static schemes. Utility gradually increases as the number of available resource blocks increase. Moreover, we can also see that more subscribers can receive more video layers as the number of resource blocks increases, and the proposed algorithm can most efficiently allocate the resources so as to get the highest utility. Fig. 3.6. shows the spectral efficiencies of the comparing algorithms. Note that, MSML can achieve better spectral efficiency than OLM when more resources are provided, because MSML can take advantage of frequency diversity with more resource blocks. However, better spectral efficiency not exactly means better utility. As we can see in Fig. 3.6, OLM has slightly better utility than MSML. We could find the reason from Fig. 3.7, which shows the actual number of subscribers who can receive the video layers. More subscribers can receive layer-3 (L3) with MSML than OLM, but less subscribers can

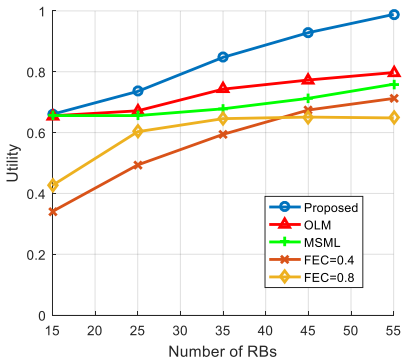



Figure 3.9. Utilities of DASH dataset

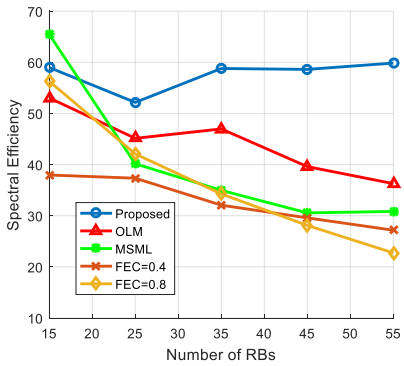



Figure 3.10. Spectral Efficiencies of DASH dataset

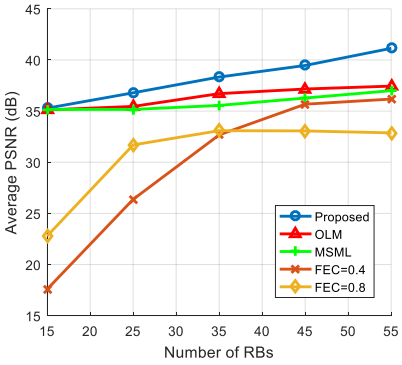



Figure 3.11. PSNR of DASH dataset

receive base-layer (BL) or L1 with OLM than MSML. Most of the subscribers using the proposed algorithm can receive L2 or L3 layers. Therefore, these subscribers can combine all received video layers to make higher quality video. Therefore, the proposed scheme can both achieve better utility and spectral efficiency than other multicasting algorithms. Fig. 3.8. shows the actual number of resource blocks allocated to each layer. The proposed method finds the optimal FEC rates and MCS which makes it possible to allocate optimal resource. Fixed FEC methods cannot utilize the resource efficiently. OLM and MSML can utilize the resource better than the fixed FEC methods, but less efficient than the proposed method.

## C. DASH

Table 3.7. shows the video rates of DASH dataset. Each video is encoded independently with 4 different video rates, with R4 being the highest rate for all the 4 distinct videos. These DASH videos are scheduled based on the proposed algorithm, OLM, MSML and a static resource allocation scheme with fixed APP-layer FEC rates as 0.4 and 0.8 to serve multiple users requesting multiple videos with different channel conditions. These videos have higher video rates than SVC video sources, because SVC creates one video by combining multiple layers but in DASH video dataset, every video rate itself is one complete video and they are High Definition (HD) videos.

Table 3.7. VIDEO SOURCES [51] for DASH SIMULATION

| | Video | R1 | R2 | R3 | R4 |
|---|---|---|---|---|---|
| 1 | Big Buck Bunny | 1662 | 2617 | 3842 | 4726 |
| 2 | Elephant Dream | 1568 | 2543 | 3603 | 4516 |
| 3 | Valkaama | 1516 | 2366 | 3676 | 4909 |
| 4 | The Swiss Acount | 1371 | 2182 | 3679 | 4712 |

We again assume that there are 100 LTE users are requesting each video, and the local cache has 4 different video (rate) representations of each video ranging from maximum 4.909Mbps video rate to minimum 1.371Mbps vide rate, to be multicasted from the eNB using eMBMS channels. The eNB has to allocate the RBs to 4 multicasting groups for each video and also find the optimal MCS and FEC coderates. Note that RB, MCS, FEC are updated every FEC block, which corresponds to 1second of time duration, i.e., 100 OFDMA frames. Fig. 3.9. shows the utility of the videos for our proposed, OLM, MSML algorithms and static resource allocation scheme with fixed FEC code rates as 0.4 and 0.8. All the algorithms can achieve better utility as the total number of available resources increases while the proposed algorithm can achieve the highest utility. The static resource allocation algorithm with FEC code rate 0.4 can achieve better utility than the algorithm with FEC code rate 0.8, because it can recover more of lost packets. Figure 8 also shows that the proposed algorithm achieves the best spectral efficiency. Unlike SVC video's spectral efficiency, Fig. 3.10. shows that the spectral efficiency of OLM, MSML and the static resource allocation scheme decrease with more resource. OLM and MSML first schedule the lowest video representation (R1) to guarantee all the users to receive at least one video rate, and schedule higher video rates with remaining resources choosing subsets of the users. The resources allocated for lower video representations could be wasted, because subscribers who can receive multiple video representations will discard lower video representations.

Fig. 3.11. shows the average PSNRs, and, by comparing with Fig. 3.9., it shows that video subscribers can see the video with higher PSNR with the scheme that can achieve higher utility. Since the higher PSNR means that the decoded video is more similar with original video than the

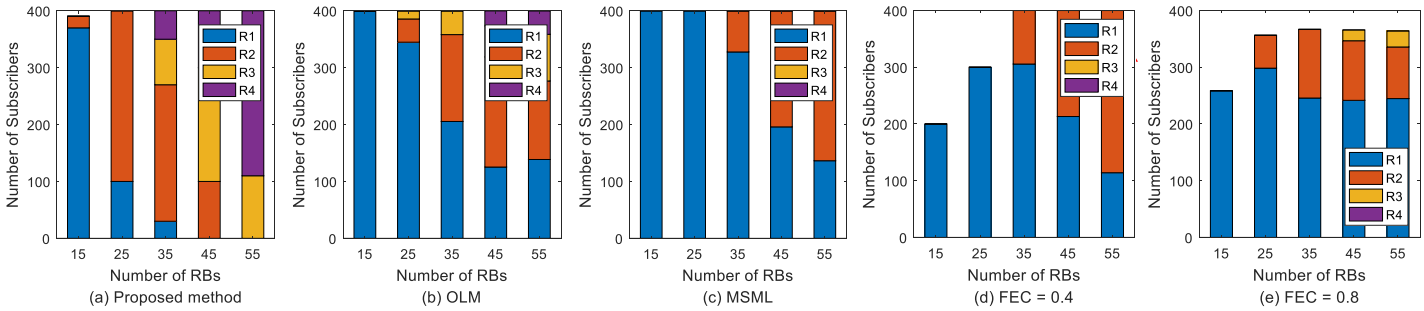



Figure 3.12. Number of subscribers of DASH dataset

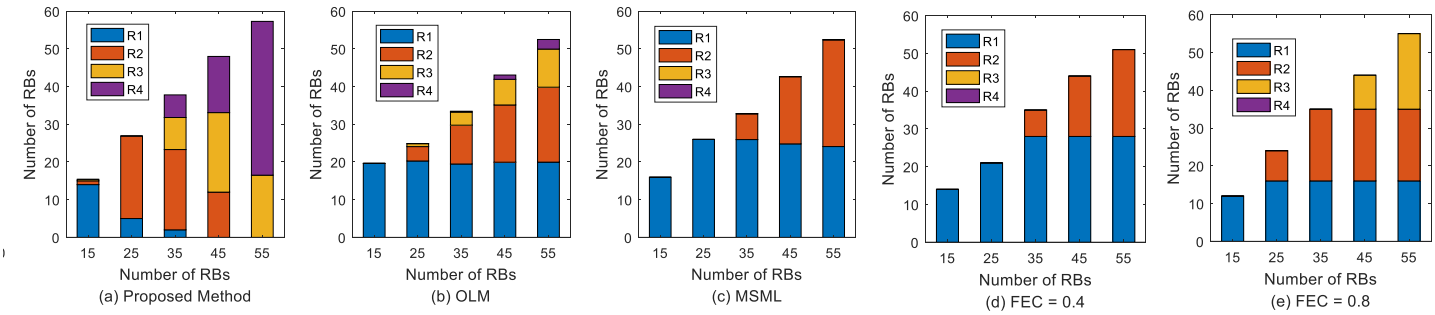



Figure 3.13. Resource allocation results of DASH dataset

decoded video with lower PSNR, video subscribers will see better video with higher utility. On the other hand, the proposed algorithm can choose the best video set that jointly optimizes RB, MCS, and FEC. Therefore, the proposed algorithm can achieve much higher utility and spectral efficiency than OLM or MSML. Moreover, the proposed algorithm also makes all the users receive at least one video rate. Fig. 3.12. shows the actual number of subscribers who receive each video rate. Resource blocks consumed by each video rate are also shown in Fig. 3.13. The proposed algorithm allocates resources to higher video representations and many users can receive higher video representations, which mean more users can experience better video quality.

## D. Complexity Analysis

Complexities of the proposed algorithm, OLM, and MSML are also analyzed. Complexity of OLM is in the order of number of subscribers as described in [24]. Since OLM performs the MCS selection by extracting one subscriber from a group until it finds the best MCS and grouping that can maximize the spectral efficiency. This has to be done for all multicasting groups, but the

number of groups is small when compared to the number of subscribers. Therefore, complexity order is linearly proportional to the number of subscribers. Complexity of MSML [25] is in the order of square of number of subscribers times number of available resource blocks since MSML needs to search all possible resource blocks of all subscribers, this makes the searching space as $N_{UE}$ by $N_{RB}$-sized square grids. Moreover, MSML also needs to perform grouping in similar manner with OLM.

Proposed algorithm's complexity is order of logarithm of number of resource blocks and possible combination of video representations. The heaviest part in terms of computational complexity is a bisection search algorithm to find optimal resource allocation, and it must be performed for all possible combinations of video representations. Therefore, the computational complexity of the proposed algorithm depends the number of videos and representations, while OLM mainly depends on the number of subscribers. Both parameters affect MSML algorithm's computational complexity. Table 3.8. shows the comparisons of complexities. CPU time is measured using Matlab when there are 4-different videos with 4-different representations. This experiment is repeated with two different number of subscribers, which are 400 and 800. Proposed scheme takes same amount of time to complete its computation for different number of subscribers, while OLM and MSML takes twice longer to complete their computation with more UEs.

Table 3.8. COMPUTATIONAL COMPLEXITIES

| Algorithms | | Proposed | OLM | MSML |
|---|---|---|---|---|
| Complexity | | $O(M^V \log(N_{RB}))$ | $O(N_{UE})$ | $O(M^V N_{RB}^2 N_{UE})$ |
| CPU time (sec) | $N_{UE}$=400, $n_{RB}$ =15 | 0.28 sec | 0.08 sec | 3.53 sec |
| | $N_{UE}$=800, $n_{RB}$ =15 | 0.30 sec | 0.10 sec | 8.15 sec |
| | $N_{UE}$=400, $n_{RB}$ =55 | 0.72 sec | 0.09 sec | 11.07 sec |
| | $N_{UE}$=800, $n_{RB}$ =55 | 0.75 sec | 0.11 sec | 22.69 sec |

The proposed algorithm can complete its whole computation much faster than MSML with better utility and spectral efficiency. This experiment shows that the proposed algorithm need more time to finish its computation than OLM, but it can achieve better performance than OLM and its complexity does not change with the number of subscribers. Moreover, OLM assumes that the server receives the feedback information from all the users to measure the packet loss rate, which implies that it uses RTP/RTCP which is not always available in DASH platform. The proposed algorithm is compatible with DASH because the server does not need the RTP/RTCP information. The processing time of OLM and the proposed algorithm are less than a second because the number of RBs are smaller than the number of subscribers. Moreover, the performance gain we can get from using the proposed algorithm is high especially when DASH type of video streaming platform is used.

# Chapter 4
# QOE-DRIVEN VR VIDEO MULTICAST OVER LTE

## *4.1 Overview*

Virtual Reality (VR) is getting more popular these days, and more people can enjoy more realistic experiences with VR systems [7]. Moreover, it allows people to look around the virtual world and feel like they are actually in the environment. VR gaming also can provide a more exciting experience to gamers. However, it is a more challenging task to make users satisfied with the quality of the VR videos, because VR videos need much higher resolution than conventional videos. Users cannot see the whole video at the same time, they can only focus on the area that they want to see and the area is usually only 20% of whole video [8]. Therefore, 4~6 times more resolution is required for VR videos to provide the same experience as conventional videos. On the other hand, this fact allows the saving of 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.

The original DASH system allows the clients to do the video rate adaptation, but it is difficult to do the individual rate adaptation in multicast systems because the users grouped in to the same group share the same spectrum resource and they receive the same video rate even though they have different channel quality. Rather than doing the individual rate adaptation, the server can adaptively choose the video rates of the tiles to maximize the expected total utility of the users. Feedback information of users' viewport can be used to decide which tiles should have better video rates to satisfy more users. Another way to decide which tiles are more important than others is analyzing the video at the server side. The server can analyze the video contents first and then decide which part may have higher interest from users. Saliency [52][53] of the video is one of the useful indicator to find the important area of the video. Therefore, we can give more bits to the area that has higher saliency to satisfy more users. Saliency detection algorithms usually find the area that has high contrast or active movement in the video [54][55], because those area usually

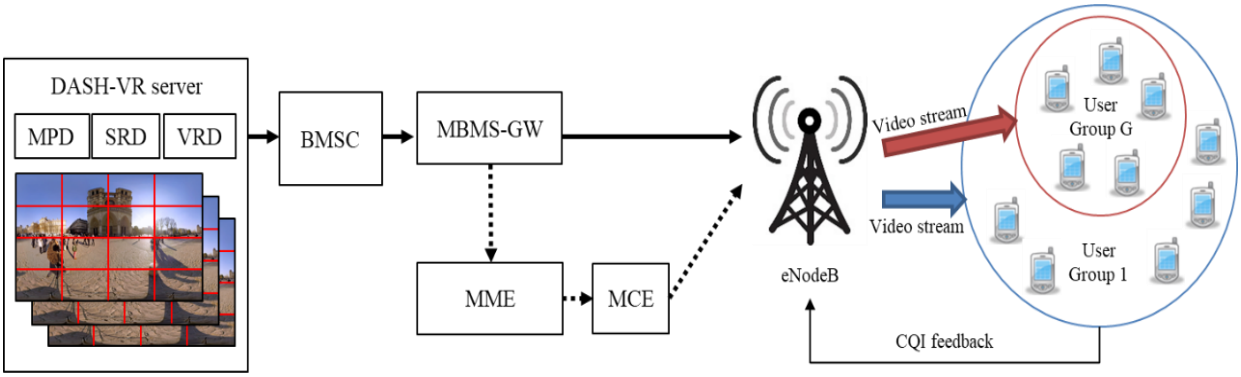



Figure 4.1. LTE eMBMS systems for VR video Multicast

have richer or more appealing information such as important texture or moving objects. By using the saliency information, the server can allocate more bits to the areas that have higher saliency scores to make them clearer. There are many video saliency detection algorithms to find which parts are more important and interesting to users.

There are two possible ways to do the VR video multicasting. Since the multicasting featured by grouping the users to share the same resource. First, the users with the same view can be grouped into a multicast group. The number of multicast group is the same as the number of views [56]. It can save some resource 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 with 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 with their channel quality. This grouping strategy helps to select more efficient MCS and the application layer forward error correction (AL-FEC) code rate to allocate better video. As the number of users could join 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 with their channel quality.

The clients in a DASH-VR multicast system request video chunks to the server based on MPD, SRD, and VRD 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 the lost packet recovery. Multiple video multicast sessions are created to deliver multiple VR videos and multiple video representations to different user groups. A video 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 on 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 allocating resource blocks (RB) and choosing AL-FEC code rate and MCS for the multicasting sessions.

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 multicasting group denotes a set of users grouped by their channel conditions and subscribing the same video. Note that users can subscribe 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 multiple multicast sessions.

We can consider two different ways to create the video multicast sessions. One is the per-tile multicasting (PTM), that considers the tiles as independent videos, where each tile has its own resource and every UE subscribes all necessary sessions to regenerate the VR video. It needs to create multiple multicasting sessions as many as the number of tiles times the number of representations for a single VR-video content. All possible video representations of all tiles are available for the users based on their channel quality, and the users regenerate the VR-video with the tiles that has the best quality 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 have to be determined to maximize the total utility. Its search space to find optimal solution is $M^T$. Each user selects one representation for one tile and subscribe $T$ multicast sessions to regenerate the VR-video. It generates too much control signal and the complexity of the solution increases with its number of multicast sessions.

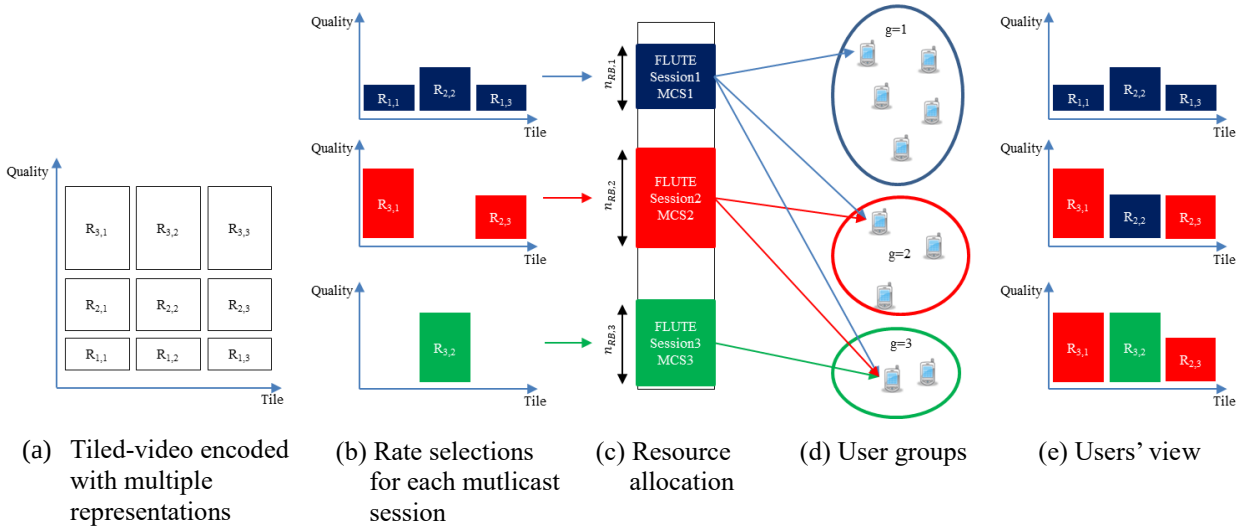



Figure 4.2. Multi-Session Multicast mechanism for VR video multicast

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 quality. Fig. 4.2. shows an example of MSM system with 3 groups and 3 multicast sessions. Fig. 4.2.(a) shows the tiled-video encoded with multiple representations. Every tile has multiple copies with different representations (qualities) and they are generated by legacy video encoder with different quantization parameters (QP). Higher representations indicate better qualities, and they need more bandwidth to be transmitted. Fig. 4.2.(b) shows the rate selection results for multiple multicast sessions. The first multicast session has all the tiles with lower representations to guarantee all the users requesting the VR video to receive at least lower quality video. The second and third multicast sessions do not need to have all the tiles. They allocate higher representations to improve the quality of the tiles for the users with better channel quality. Therefore, they are allocated on the wireless resource with more efficient MCS and AL-FEC code rates (Fig. 4.2. (c)). The users can subscribe multiple multicast sessions at the same time, but their channel quality should be good enough to decode the data packets assigned with certain MCS and AL-FEC. In Fig. 4.2.(e), the user group 1 only can receive the data in the multicast session 1, while user group 2 can receive multicast session 1 and 2. The user group 3 can receive all three multicast sessions. Therefore, the user group 2 and 3 have chance to choose better representations from multiple representations they can receive.

Since MSM's multicast session includes the multiple tiles, it creates less multicast sessions than PTM. Another advantage of MSM is that it can use existing rate selection algorithms introduced in Chapter 2. The rate selection algorithm such as [33] can work to allocate the tiles of different representations with the bit rate constraint of each multicast session. The bit rate constraints of each multicast sessions are determined by the resource allocated on the multicast sessions, MCS, and AL-FEC code rates.

Table 4.1. NOTATIONS USED IN THIS CHAPTER

| | |
|---|---|
| $c(MCS)$ | Capacity of an RB with selected MCS |
| $u(r)$ | Utility model, a function of video rate $r$ |
| $\alpha, \beta$ | Normalization coefficients of $u(r)$ |
| $r$ | Video rate |
| $Q_{t,g}$ | Video representation of tile $t$ for multicast session $g$ |
| $R_M$ | Maximum video rate |
| $f_{mar}$ | Margin factor of AL-FEC |
| $g$ | Group index |
| $t$ | Tile index |
| $G$ | Number of groups |
| $T$ | Number of tiles |
| $\gamma_{MCS}$ | Threshold SNR with selected MCS index |
| $\overline{\gamma}_i$ | Average SNR in dB scale of UE $i$ |
| $S_t$ | Saliency score of tile $t$ |
| $K$ | 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 $g$ |
| $SE_i$ | Spectral Efficiency of user $i$ |
| $SE_{\min,g}$ | Spectral Efficiency of user group $g$ |
| $N_{RB}$ | Total number of available RB |
| $n_{RB,g}$ | Number of RBs assigned to group $g$ |
| $U_g(n_{RB})$ | Approximated utility of group $g$ |

There are six parameters to be optimize in the MSM system for VR-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.

### *4.2 Problem Formulations*

The mathematical expressions of Quality of Experience (QoE), Spectral Efficiency (SE), and utility maximization 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 result. Since MCS and AL-FEC directly affect the SE, we can reduce the number of control parameters by formulating the relationship among MCS, AL-FEC and SE. Utility maximization problem is formulated using the widely accepted QoE model and SE with wireless resource constraints. Table 4.1. includes the notations used in this paper.

A. QoE Model

It is known that the user experience is not linearly proportional to the video rate and is gradually getting saturated with higher video rates. In this paper, the well accepted logarithmic law [39] is considered to quantify the quality of user experience based on the video data rate:

$$u(r)=\begin{cases}\alpha \log \beta \dfrac{r}{R_M}, & r>0\\ 0, & r=0\end{cases}, \tag{4.1}$$

where $r \in \{0, R_1, R_2, ..., R_M\}$, $u(r)$ 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, every tile has its own utility value with assigned rate. 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 on each tile based on the saliency information extracted from the video contents. Therefore, the total utility of the video that users really receive is

$$\sum_{t=1}^{T} u(r_t) S_t , \tag{4.2}$$

where $S_t$ is the saliency score of the tile *t*. Higher saliency scores of the tiles mean that people may be interested in the tiles more than other tiles with lower saliency scores. The utility model implies that we can achieve higher utility value when we allocate more bits to the tiles with higher saliency scores.

The utility model is not restricted with saliency weighting. $S_t$ can be other weighting coefficient such as number of users who see the tile. We used saliency score because it does not require the feedback information from users. However, if feedback information is applicable at the server side, users' viewport information could be used to define the weighting coefficients.

B. 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:

$$SE = c(MCS)\frac{K}{N}, \tag{4.3}$$

where $c(MCS)$ is the efficiency [11] 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 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 [13]:

$$(1-P_{out}(\gamma_{MCS},\overline{\gamma}_i))\,f_{\text{mar}} \geq \frac{K}{N}, \tag{4.4}$$

where $P_{out}(\gamma_{MCS},\overline{\gamma}_i)$ is the outage probability of user *i* given average SNR $\overline{\gamma}_i$, threshold SNR $\gamma_{MCS}$, and selected MCS [43]; $f_{\text{mar}}$ denotes the margin factor for describing the non-ideal AL-FEC decoding capability. The received SNR is modeled as Log-normal random variable because of exponential effective SNR mapping (EESM) [41] based link error prediction used for OFDMA systems. Therefore, the outage probability $P_{out}(\gamma_{MCS},\overline{\gamma}_i)$ 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

$$SE_i = c(MCS)(1-P_{out}(\gamma_{MCS},\overline{\gamma}_i))\,f_{\text{mar}}. \tag{4.5}$$

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 will also increase. Therefore, the best MCS for user *i* is given as

$$MCS^*_i = \arg\max_{MCS} c(MCS)(1-P_{out}(\gamma_{MCS},\overline{\gamma}_i))\,f_{\text{mar}}. \tag{4.6}$$

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 smallest SE of the group, $SE_{\min,g}$, because users in the same group share the same resource and see the same video; and the video rate should be small enough to be successfully delivered to all the users in the group, especially the user with the smallest SE.

C. 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

$$\mathrm{Q}_1 : \underset{\mathbf{r}}{\text{maximize}} \sum_{i=1}^{N}\sum_{t=1}^{T} u(r_{i,t})S_t , \tag{4.7}$$

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 save the spectrum, we can apply the multicasting scenario on the original problem $\mathrm{Q}_1$ by grouping the users based on their channel quality instead of grouping 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 VR-video by subscribing only one multicast session. If users are able to decode the video from the multiple multicast sessions, users can choose which one to play. It is a much more efficient way to utilize the limited resource when the users need multiple tiles at the same time.

The utility maximization problem of the VR-video multicast system can be formulated as follow. Four parameters to 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:

$$\mathrm{Q}_2 : \underset{G,|\mathbb{N}_g|,\mathbf{n}_{RB},\mathbf{Q}}{\text{maximize}} \sum_{g=1}^{G} |\mathbb{N}_g| \sum_{t=1}^{T} u(r_{Q_{t,g}})S_t , \tag{4.8}$$

subject to

$$\begin{cases} \sum_{t=\{\tau\in\{1,...,T\}|Q_{\tau,g}>Q_{\tau,1...g-1}\}} r_{Q_{t,g}} \le n_{RB,g} SE_{\min,g}, & g>1 \\ \sum_{t=\{1,...,T\}} r_{Q_{t,g}} \le n_{RB,g} SE_{\min,g}, & g=1 \end{cases} \tag{4.9}$$

$$r_{Q_{t,g}} \in \{0, R_{t,1}, R_{t,2}, ..., R_{t,M}\} \tag{4.10}$$

$$\sum_{g=1}^{G} n_{RB,g} \le N_{RB} \tag{4.11}$$

$$n_{RB,g} \ge 0 \tag{4.12}$$

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

$$\mathbb{N}_g = \begin{cases} \left\{ i \in \mathbb{N} \mid SE_{\min,g} \le SE_i < SE_{\min,g+1} \right\}, & g < G \\ \left\{ i \in \mathbb{N} \mid SE_{\min,g} \le SE_i \right\}, & g = G \end{cases}. \tag{4.13}$$

$|\mathbb{N}_g|$ 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 $Q_{t,g} \in \{1,2,...,M\}$ 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 see better representations on their salient part, the total utility can achieve a higher value. There are resource constraints (4.11),(4.12), e.g., groups share the total resource, and the tiles in a group share the resource allocated to each group.

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

### *4.3 Cross-Layer Optimization Framework*

The goal of the cross-layer resource allocation framework, as shown in Fig. 4.3., is to find the number of the multicast group (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) 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) resource allocation algorithm that decides $n^*_{RB,g}$ for all *g*, and (C) the 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 the resource; 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 evenly allocated resource to all groups return the $A_g, B_g$ of approximated utility functions

$$U_g(n_{RB}) = A_g \log(B_g n_{RB}) \tag{4.14}$$

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

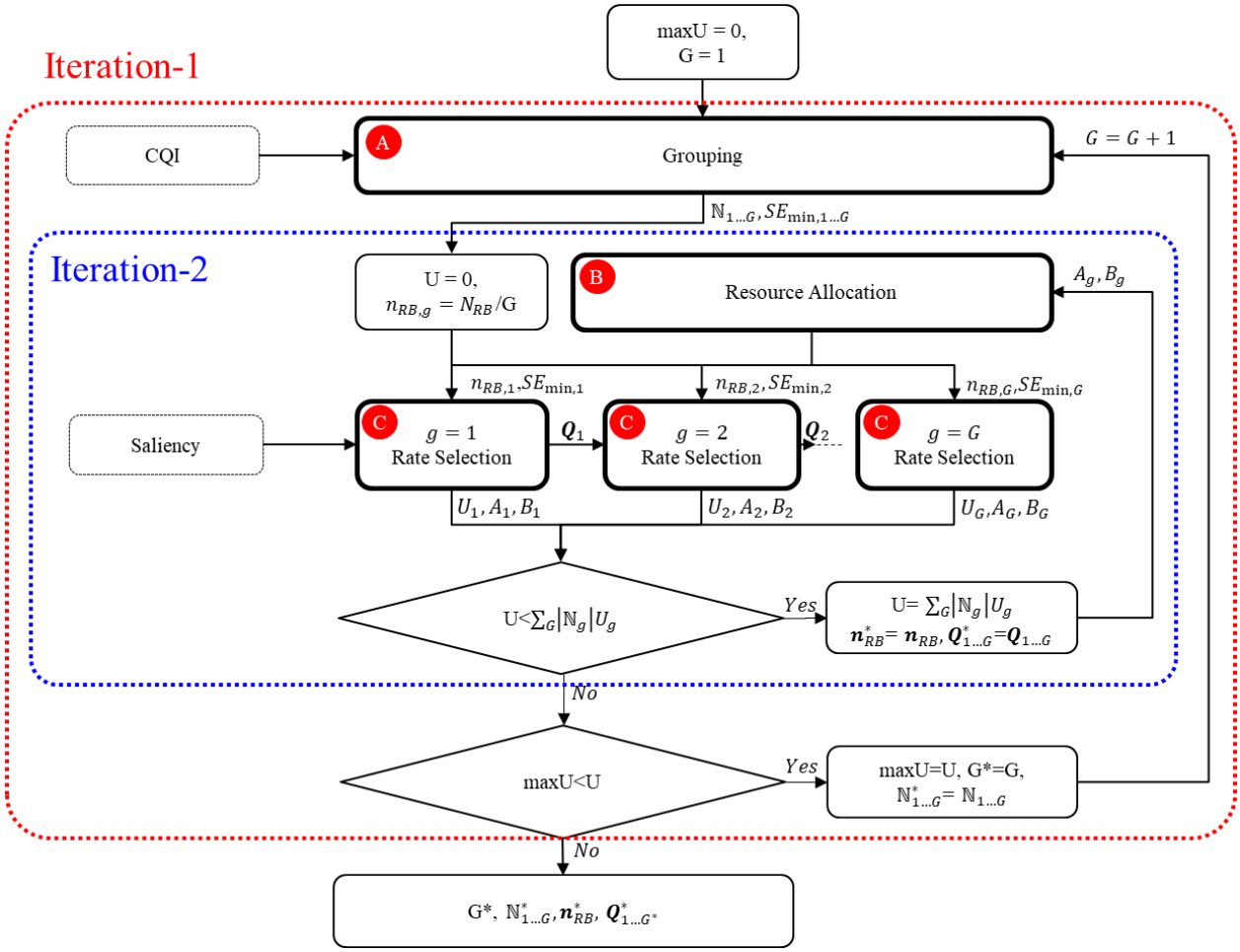



Figure 4.3. Cross-layer optimization framework

they return the utilities. The summation of utility values is compared to initial utility value U and if it is larger than U, which will be updated with the new utility result. The $A_g$ and $B_g$ of the utility function (4.14) found from the rate allocation algorithm are approximated values, therefore, Iteration-2 continues to find better solution.

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 following.

## A. Grouping Algorithm

The goal of 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*, $SE_{\min,g}$. The 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 the optimization. Initially, all the users are in group 1, therefore, the spectral efficiency of the 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 the 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

$$SE_{\min,g} = SE_{k_g} \tag{4.15}$$

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

$$k_g = \begin{cases} \arg\min_{i\in\mathbb{N}}(SE_i), & g = 1 \\ \arg\max_{i\in\{n\in\mathbb{N}|SE_n > SE_{\min,g-1}\}} \left\{|\mathbb{N}_{g-1}|\,SE_{\min,g-1} + |\mathbb{N}_g|\,SE_i\right\}, & g > 1 \end{cases} \tag{4.16}$$

$|\mathbb{N}_g|$ for all *g* are determined by (4.13).

B. Resource Allocation

The goal of resource allocation algorithm is to allocate radio resource to each multicast group. The rate-selection algorithm shown in the next section gives the estimated utilities of all 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 (4.14) have convex forms and the constraint is convex:

$$Q_3 : \underset{\mathbf{n}_{RB}}{\text{maximize}} \sum_{g=1}^{G} |\mathbb{N}_g| U_g(n_{RB,g}), \tag{4.17}$$

subject to (4.11), (4.12).

Using the Lagrangian method,

$$\begin{aligned} L(\mathbf{n}_{RB}, \lambda, \boldsymbol{\zeta}) = & \sum_{g=1}^{G} |\mathbb{N}_g| U_g(n_{RB,g}) \\ & - \lambda \left( \sum_{g=1}^{G} n_{RB,g} - N_{RB} \right) + \sum_{g=1}^{G} \zeta_g n_{RB,g} \end{aligned}, \tag{4.18}$$

Gradient of (4.18) is

$$\frac{\delta L(\mathbf{n}_{RB}, \lambda, \boldsymbol{\zeta})}{\delta \mathbf{n}_{RB}} = \frac{|\mathbb{N}_g| A_g}{n_{RB,g}} - \lambda + \zeta_g, \tag{4.19}$$

The optimality condition can thus be derived by using Karush-Kuhn-Tucker (KKT) conditions [47].

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

The gradient of Lagrangian vanishes when

$$\lambda^* = \frac{|\mathbb{N}_g| A_g}{n_{RB,g}} \tag{4.20}$$

holds for all $g$. Therefore, the optimal resource allocation solution is $n^*_{RB,g}$, which satisfies (4.20) for all $g$.

C. Rate Selection

The goal of the rate selection algorithm is to determine the DASH-VR representation of the tiles and to allocate radio resource 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

$$\mathrm{Q}_4 : \underset{Q_{g,1\ldots T}}{\text{maximize}} \; U_g = \sum_{t=1}^{T} u(r_{Q_{g,t}}) S_t , \tag{4.21}$$

$$\text{subject to (4.9), (4.10)}$$

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 $R_{t,m}$ (4.10). The $Q_{g,t}$ that can increase the total utility per cost (rate) 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 [33] is that ours considers the condition of the lower group (i.e. the user group with worse channel quality). Since the upper groups (i.e. the user group with better channel quality) can subscribe the lower groups' tiles, some of the tiles already have available videos, even though the upper groups do not allocate the video for the tiles. Therefore, an upper group does not need to pay the cost to subscribe the same representations as a lower group, but it needs to pay its own cost to subscribe better representations than the lower group. The cost matrix [$C_{t,m}$] shows the "marginal" cost to pay for subscribing the

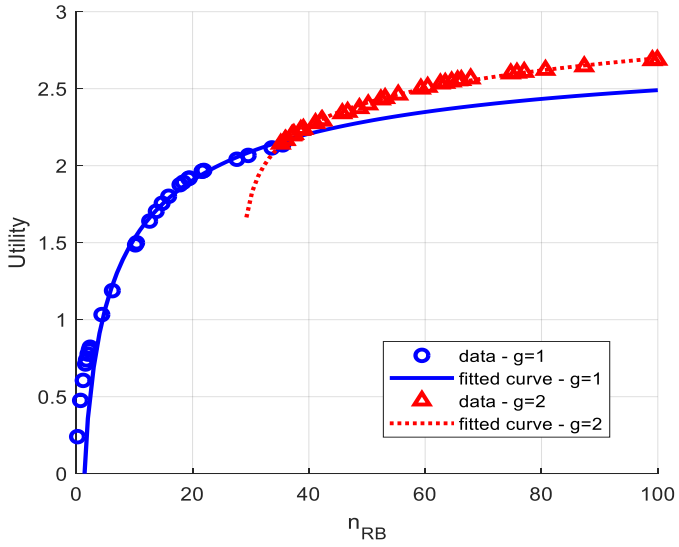



Figure 4.4. Curve fitting results of video utility function

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

$$C_{t,m} = \begin{cases} R_{t,m} - R_{t,m-1}, & m \geq Q_{t,g-1} + 2 \\ R_{t,m}, & m = Q_{t,g-1} + 1 \\ 0, & m \leq Q_{t,g-1} \end{cases}. \tag{4.22}$$

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

$$\tilde{u}_{t,m} = \begin{cases} u(R_{t,m}) - u(R_{t,m-1}), & m \geq 2 \\ u(R_{t,m}), & m = 1 \end{cases}. \tag{4.23}$$

$\tilde{u}_{t,m}$ 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 the 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 (4.14) to model the curves, and performed the curve fitting to find the parameters $A_g$ and $B_g$. Fig. 4.4. shows the 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 increases fast by allocating the lower video representations but the slope of the curve decreases because it needs more resource to allocate the higher video representations with its lower spectral efficiently. The second multicast session can allocate the tiles with higher video representations more efficiently than the first multicast session, therefore, its utility increases fast again by allocating the videos on 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 on the second multicast session. The slopes of the curves indicate the efficiencies of the multicast sessions, therefore, this information is used to do the optimal resource allocation. The detailed algorithm is described in Table 4.2.

Table 4.2. RATE-SELECTION ALGORITHM

| | |
|---|---|
| INPUTS: $n_{RB,g}$, $SE_{\min,g}$ | |
| OUTPUTS: $A_g$, $B_g$, $Q_{t,g}$ | |
| 1. | Make a list of $\frac{\tilde{u}_{t,m}S_t}{C_{t,m}}$ 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** $C_{\text{current}} \le n_{RB,g}SE_{\min,g}$ |
| 4. | Allocate $(t,m)$ pair that has maximum $\frac{\tilde{u}_{t,m}S_t}{C_{t,m}}$, 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 $y=A_g\log(B_g x)$ |

### *4.4 Performance Analysis*

The performance analysis is conducted using the standard LTE parameters [46][57] described in Table 4.3. 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 [11], ensuring a block error rate (BLER) value lower than 10%. 1000 UEs are generated and randomly distributed in the cell area. 20 to 100 RBs are used to transmit 360-degree tiled VR video over the LTE eMBMS channel. The videos of resolution 3840×1920 with 16 tiles are used for the simulation. Each tile has four different representations encoded with different QP values (25, 30, 35, and 40) [58], which resulted in tiled video rates 13Kbps to 7.7Mbps. Each tile is encoded by a legacy video encoder (ffmpeg) [59] and the total bitrate to deliver the best representations for all tiles is 40Mbps.

Table 4.3. 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 | 20MHz |
| Scheduling frame | 10ms |
| RB size | 12 sub-carrier, 0.5ms |
| Sub-carrier spacing | 15kHz |
| Extended CP Duration | 16.7$\mu s$ |
| TTI | 1ms |
| EUTRA UE | Antenna gain 0dBi, NF 9dB |
| EUTRA eNB | Antenna gain 14dBi, NF 5dB |
| eNB transmit power | 43dBm |
| MIMO configuration | 1 Tx, 1 Rx |
| Thermal Noise | -174dBm/Hz |

The proposed resource algorithm, a convex-optimization method (Convex), is compared with some existing solutions for resource-allocation: 1) the exhaustive search (Exhaustive), 2) the

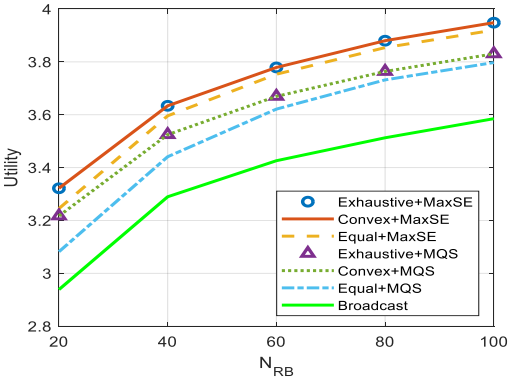



Figure 4.5. Utilty

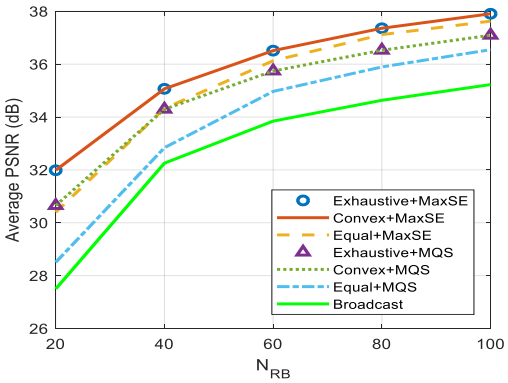



Figure 4.6. Average PSNR

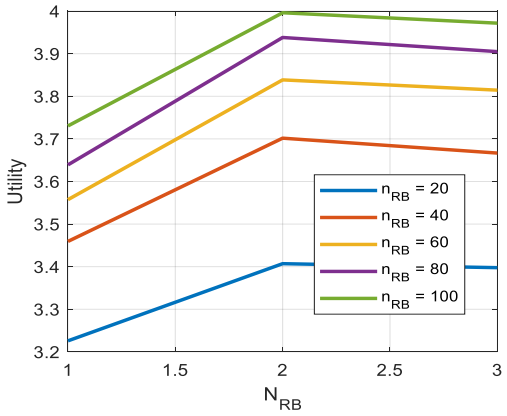



Figure 4.7. Utility with the number of groups

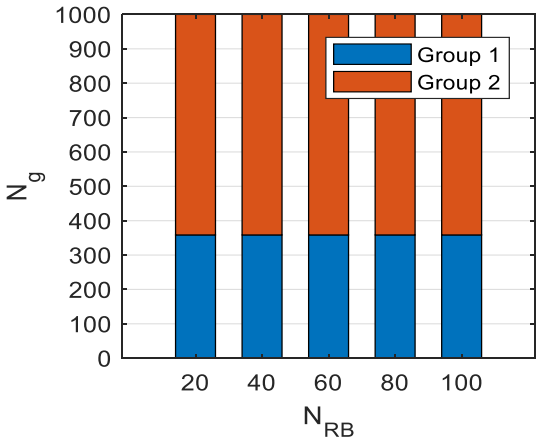



Figure 4.8. Number of users

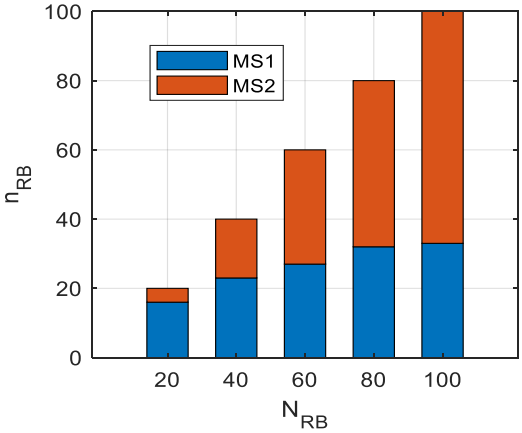



Figure 4.9. Resource allocation

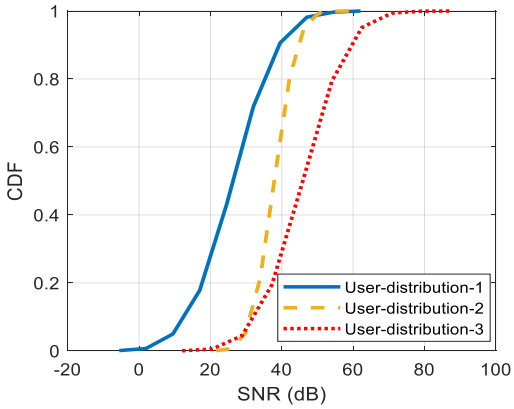
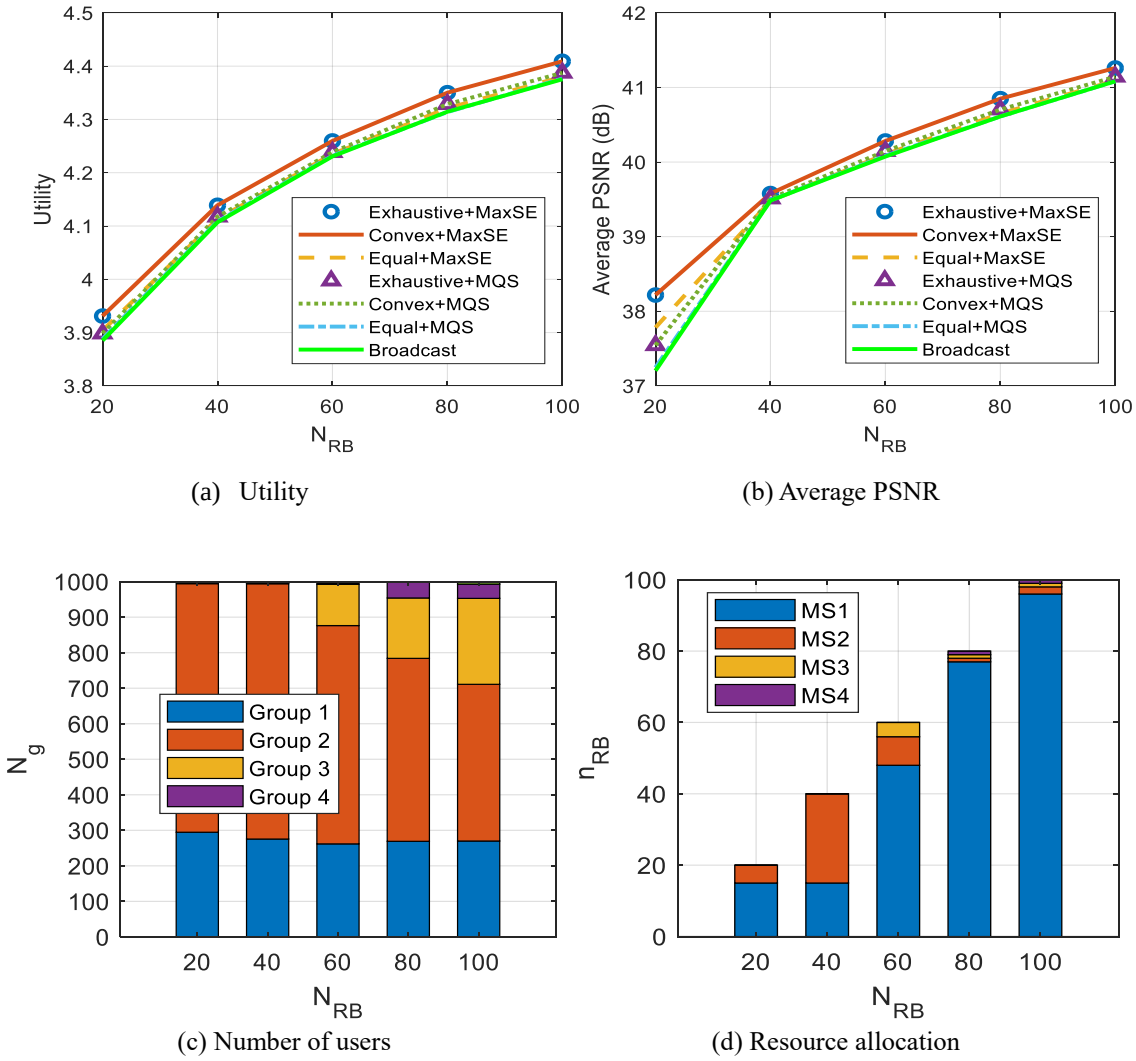



Figure 4.10. CDF of Users' average SNR

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) [26] algorithm is a grouping algorithm




(a) Utility (b) Average PSNR

(c) Number of users (d) Resource allocation

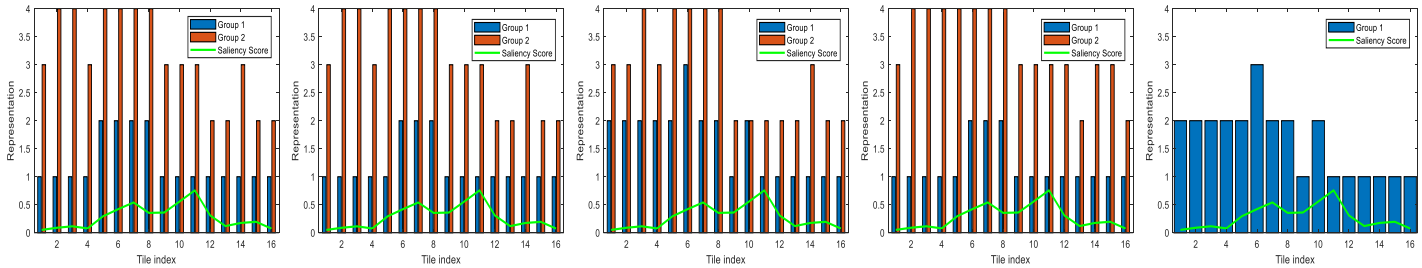
Figure 4.11. User distribution 2

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.

Fig. 4.5. 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. Fig. 4.6. 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

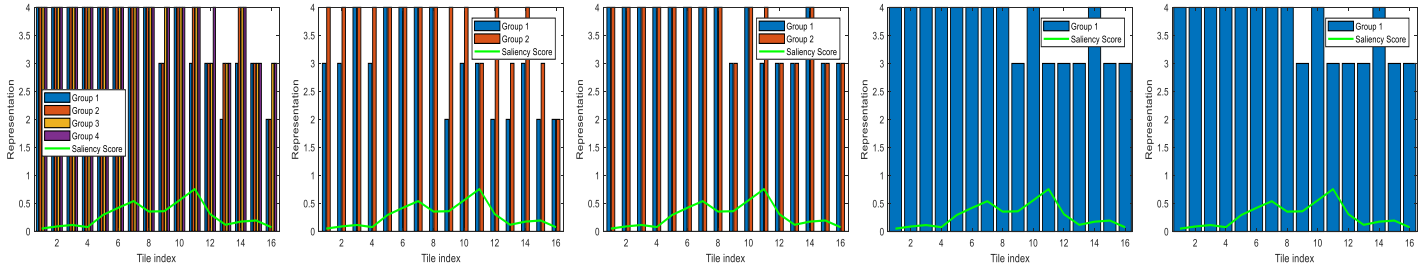
shows that the utility value is a good indicator to describe the users' experience in mathematical form. Fig. 4.7. shows the utility performance of Convex+MaxSE with different total numbers ($G$) of groups. We can see that the utility performance is the best with $G$=2; therefore, the proposed algorithm stops after testing the utility with $G$=3, and decides the best solution is $G$=2. 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. Fig. 4.8. 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 the group 2 is larger than the number of users who join the group 1. Fig. 4.9. 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 12 is shown by user-distribution-1 in Fig. 4.10, with an average SNR range from 0dB to 50dB. 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. Fig. 4.11(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. We can find the reason in Fig. 4.11(d), which shows the resources allocated for each multicast session. Most of resources are allocated in multicast session 1 (MS1) because it contributes to improving the total utility the most. It is different from the resource allocation results for user-distribution-1 shown in Fig. 4.10. Since there are more users with lower average SNR in user-distribution-1, the proposed algorithm gave minimum 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




(a) Convex + MaxSE (b) Equal + MaxSE (c) Convex + MQS (d) Equal + MQS (e) Broadcast

Figure 4.12. Rate selection results with User distribution 1

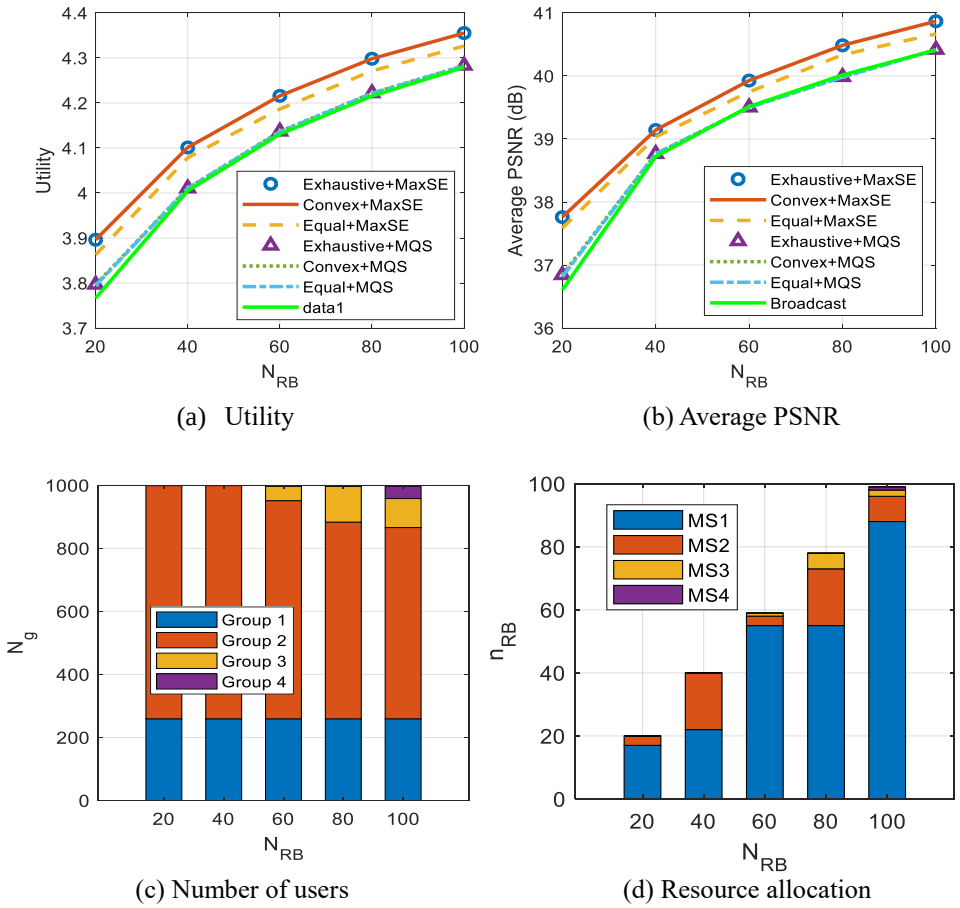



(a) Convex + MaxSE (b) Equal + MaxSE (c) Convex + MQS (d) Equal + MQS (e) Broadcast

Figure 4.13. Rate selection results with User distribution 2

user-distribution-1 because using a smaller number of RBs by allocating only a few tiles with better representations improved the total utility. Most of the tiles already had the best representations in multicast session 1.

Fig. 4.12. shows the selected representations for all groups of users for all 16 tiles in user-distribution-1. Fig. 4.12(a) shows the representations that user groups 1 and 2 could see in their views. Those in group 2 could see much better quality since most of the tiles have better representations than those of user group 1. Fig. 4.12(b) shows the simulation results with the Equal resource allocation combined with the MaxSE grouping algorithm. The difference with Fig. 4.12(a) is the representation of tile 5 for user group 1. It shows that the proposed resource allocation method could allocate the resource more efficiently than the Equal resource allocation method. Fig. 4.12(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 ones 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. Fig. 4.12(d) shows the results with




(a) Utility (b) Average PSNR

(c) Number of users (d) Resource allocation

Figure 4.14. User distribution 3

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. Fig. 26(e) shows the result with the Broadcast scheme, which allocated all resources to group 1 and 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 representations.

Fig. 4.13. shows the selected representations for user-distribution-2. Since all the users in user-distribution-2 had very good channel quality, most of the tiles could have high representations through the Broadcast scheme. However, there is still some space to improve the visual quality for users with better channel quality. Fig. 4.13(a) shows that the Convex+MaxSE scheme creates four


Figure 4.15. VR video example

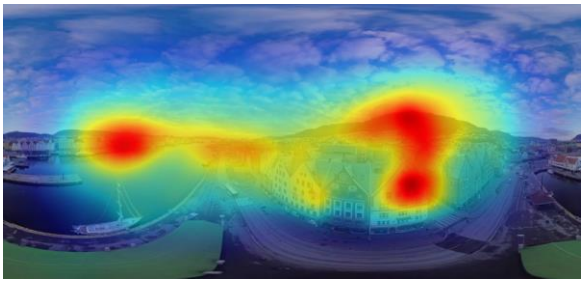


Figure 4.16. Saliency map

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 the tiles 9, 11, and 12 have higher saliency scores than the tiles 13 and 16, we can expect users to see better quality in their views.

Fig. 4.14(a)(b). shows the utilities and the average PSNRs of the algorithms. Fig. 4.14(c)(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 the 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 amount of resources to improve the total utility.

Fig. 4.15. shows the VR-video examples. We took two different viewports of a VR-video to qualitatively compare the performances. The scene 1 includes sky and cloud, where users may not give much attention. The scene 2 includes buildings, where users can see texture. Scene 2 has higher saliency score than the scene 1. Fig. 4.16. shows the saliency score of the video. Fig. 4.17. and Fig. 4.18. show the visual quality of the scene 1 and the scene 2 respectively, where Broadcast scheme and Convex+MaxSE scheme are used with user-distribution-1. Fig. 4.17(a) and 4.18(a) are two different scenes shown to users using the Broadcast. All of the users see the same quality. There are recognizable blocks in the scene with the lower video quality that made the users feel the poor quality. The Convex+MaxSE method divided the users into two groups based on their

(a) Broadcast (b) Convex+MaxSE (g=1) (c) Convex+MaxSE (g=2)

Figure 4.17. Visual Qualities – Scene 1

(a) Broadcast (b) Convex+MaxSE (g=1) (c) Convex+MaxSE (g=2)

Figure 4.18. Visual Qualities – Scene 2

channel quality. Group g=2 see better quality than group g=1. Fig. 4.17(c) and 4.18(c) are the scenes shown to user group g=2, and Fig. 4.17(b) and 4.18(b) are 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 has similar video quality with the broadcasting method. Note that only 35% of users joined group g=1, while 65% joined group g=2, which could see much better quality.

# Chapter 5
# CONCLUSION AND FUTURE WORKS

## *5.1. Conclusion*

In this dissertation, we present the video multicast systems and its optimization algorithms. The system includes the optimal resource allocation algorithms, video rate selection algorithm, MCS selection algorithms, and AL-FEC rate selection algorithms. The proposed algorithms jointly optimize the video multicast systems that deliver the SVC, DASH, and VR videos.

An optimal DASH multicasting algorithm is proposed to achieve the higher utility with the limited spectrum in overloaded situations. An optimal solution for maximizing utility is provided by using a convex optimization technique. The proposed scheme can allocate the optimal resource and choose the best MCS and FEC code rates for multiple FLUTE sessions to maximize utility. Efficient MCS selection algorithm is also introduced to reduce the complexity of the algorithm. The simulation results show that the proposed algorithm is optimal. By using the proposed algorithm, more users in the network can get the video with better quality when the eNB multicast the multiple representations. The proposed algorithm is applicable to both SVC and DASH type of video sources.

We propose MSM system to allocate the multiple tiles on a single multicast session and generates the multiple multicast session to provide set of tiles with different representations. The proposed tiled video multicasting scheme uses the limited wireless resource more efficiently than other VR multicasting schemes. A cross-layer optimization framework for the VR-video multicast system 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 solution 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 VR-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.

### *5.2. Future works*

In this dissertation, the cross-layer optimization of wireless video multicast systems is presented. The proposed video multicast systems are based on DASH, which is designed for video down streams. We can deliver SVC, DASH, and VR videos efficiently over the wireless channel using the proposed systems and algorithms. However, most recent and future multimedia systems, such as Augmented Reality systems, require faster user interactions than DASH or VR video systems.

The future work is extending the proposed method for the point cloud data streaming. The point cloud data can provide users real 3D experience by mixing the volumetric media on the users' view. User can see the objects through Hololens [60], head mounted display (HMD), or conventional 2D displays. However, it is more challenging task to deliver the point cloud data over the network, because it requires more bandwidth than VR or other multimedia data streaming. Moreover, predicting users' viewport is more important in AR applications. In case of VR video streaming, we have applied the saliency detection algorithm to predict users' viewport, but the saliency detection is not helpful for AR applications, because users' viewport can be changed by both the users' movements and the real world environments that the point cloud data is mixed with. Therefore, the server should receive the users' feedback information and react immediately to provide the high quality multimedia services over the internet. We can consider two methods to solve this problem, which are the 3D tiling of the point could data [61] and the client buffer management [62] for fast interactions.

By grouping points together into independently coded tiles, it is possible to extend the basic idea of tile-based streaming from spherical video to volumetric media. 3D tiling of the point clouds can help to adaptively selecting the regional quality and reducing the required bandwidth. However, unlike the tiles in VR video, the set of occupied tiles is sparse, changing, content-

dependent, and possibly large. Furthermore, the set of visible tiles is not only content-dependent but also viewpoint-dependent due to occlusions. Therefore, the selection of qualities of 3D tiles based on users' feedback information is the future research topic.

In addition to this, the client side buffer management algorithm is important to deal with point cloud data with fast response. There are many buffer management algorithms introduced for DASH systems. To provide good quality of video using variable network, most of the algorithms fetch several seconds of future frames [63][64][65]. However, AR applications need sub-seconds latency to provide the services. Therefore, we can apply a new buffer management algorithm that both receives future frames to prevent buffer depletion and updates existing frames with better quality to improve the quality of 3D tiles.

# BIBLIOGRAPHY

[1] Index, Cisco Visual Networking. "Forecast and methodology, 2014-2019 white paper." Retrieved 23rd September (2015).

[2] ISO/IEC 23009-1, "Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats", April 2012.

[3] Boni A, Launay E, Mienville T, Stuckmann P.,"Multimedia broadcast multicast service-technology overview and service aspects.", *Fifth IEE International Conference on. IET*, 2004, pp. 634-638

[4] IETF RFC 3926, "FLUTE - File Delivery over Unidirectional Transport", October 2004.

[5] D. Gozálvez, D. Gómez-Barquero, T. Stockhammer and M. Luby, "AL-FEC for Improved Mobile Reception of MPEG-2 DVB-T Transport Streams", *International Journal of Digital Multimedia Broadcasting*, vol. 2009, pp. 1-10, 2009.

[6] B. Wang, J. Kurose, P. Shenoy and D. Towsley, "Multimedia streaming via TCP", *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 4, no. 2, pp. 1-22, 2008.

[7] S. E. Chen, "QuickTime VR," *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques - SIGGRAPH 95*, 1995.

[8] X. Corbillon, G. Simon, A. Devlic, and J. Chakareski, "Viewport-adaptive navigable 360-degree video delivery," *2017 IEEE International Conference on Communications (ICC)*, 2017.

[9] Y. Bao, H. Wu, T. Zhang, A. A. Ramli, and X. Liu, "Shooting a moving target: Motion-prediction-based transmission for 360-degree videos," *2016 IEEE International Conference on Big Data (Big Data)*, 2016.

[10] L. Dacunto, J. V. D. Berg, E. Thomas, and O. Niamut, "Using MPEG DASH SRD for zoomable and navigable video," *Proceedings of the 7th International Conference on Multimedia Systems - MMSys 16*, 2016.

[11] M. Kawser, "Downlink SNR to CQI Mapping for Different MultipleAntenna Techniques in LTE", *International Journal of Information and Electronics Engineering*, 2012.

[12] Agrawal, D., Tarokh, V., Naguib, A. and Seshadri, N., "Space-time coded OFDM for high data-rate wireless communication over wideband channels." in *Proc. Vehicular Technology Conference, 1998. VTC 98. 48th IEEE* (Vol. 3, pp. 2232-2236).

[13] Luby, Michael, et al. Raptor forward error correction scheme for object delivery. No. RFC 5053. 2007.

[14] X. Ge, L. Pan, Q. Li, G. Mao and S. Tu, "Multipath Cooperative Communications Networks for Augmented and Virtual Reality Transmission", *IEEE Transactions on Multimedia*, vol. 19, no. 10, pp. 2345-2358, 2017.

[15] J. Lee, D. Yang, Y. Chen and W. Liao, "Efficient Multi-View 3D Video Multicast with Depth-Image-Based Rendering in LTE-Advanced Networks with Carrier Aggregation", *IEEE Transactions on Mobile Computing*, vol. 17, no. 1, pp. 85-98, 2018.

[16] Z. Liu, G. Cheung, J. Chakareski and Y. Ji, "Multiple Description Coding and Recovery of Free Viewpoint Video for Wireless Multi-Path Streaming", *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 1, pp. 151-164, 2015.

[17] Z. Liu, G. Cheung and Y. Ji, "Optimizing Distributed Source Coding for Interactive Multiview Video Streaming Over Lossy Networks", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 10, pp. 1781-1794, 2013.

[18] H. Schwarz, D. Marpe and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103-1120, 2007.

[19] J. Li, Z. Bao, C. Zhang, Q. Li and Z. Liu, "Joint MCS and power allocation for SVC video multicast over heterogeneous cellular networks", *Computer Communications*, vol. 83, pp. 16-26, 2016.

[20] C. Guo, Ying Cui, D. Ng and Z. Liu, "Power-efficient multi-quality multicast beamforming based on SVC and superposition coding," *IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017.

[21] Chen, Jiasi, Mung Chiang, Jeffrey Erman, Guangzhi Li, K. K. Ramakrishnan, and Rakesh K. Sinha. "Fair and optimal resource allocation for LTE multicast (eMBMS): Group partitioning and dynamics." In *Computer Communications (INFOCOM), 2015 IEEE Conference on*, pp. 1266-1274. IEEE, 2015.

[22] Belda, I. de Fez, F. Fraile, P. Arce and J. Guerri, "Hybrid FLUTE/DASH video delivery over mobile wireless networks", *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 11, pp. 1070-1082, 2014.

[23] K. Bakanoglu, W. Mingquan, L. Hang, and M. Saurabh, "Adaptive resource allocation in multicast OFDMA systems," in *Proc. IEEE WCNC*, Apr. 2010, pp. 1–6.

[24] C. Huang, S. Huang, P. Wu, S. Lin and J. Hwang, "OLM: Opportunistic Layered Multicasting for Scalable IPTV over Mobile WiMAX", *IEEE Transactions on Mobile Computing*, vol. 11, no. 3, pp. 453-463, 2012.

[25] 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.

[26] Bakanoglu, Kagan, et al. "Adaptive resource allocation in multicast OFDMA systems." in *Proc. IEEE Wireless Communication and Networking Conference.* Apr 2010, pp. 1-6.

[27] T. Low, M. Pun, Y. Hong and C. Kuo, "Optimized opportunistic multicast scheduling (OMS) over wireless cellular networks", *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 791-801, 2010.

[28] 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.

[29] J. L. Feuvre and C. Concolato, “Tiled-based adaptive streaming using MPEG-DASH,” *Proceedings of the 7th International Conference on Multimedia Systems - MMSys 16*, 2016.

[30] J. Park, J.-N. Hwang, Q. Li, Y. Xu, and W. Huang. "Optimal DASH-multicasting over LTE." *IEEE Transactions on Vehicular Technology* (2018).

[31] H. Kimata, S. Shimizu, Y. Kunita, M. Isogai, and Y. Ohtani, “Panorama video coding for user-driven interactive video application,” *2009 IEEE 13th International Symposium on Consumer Electronics*, 2009.

[32] 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.

[33] P. R. Alface, 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.

[34] J. Proakis, Digital communications, McGraw-Hill, New York, 1995

[35] X. Chen, J. Hwang, C. Lee, S. Chen, "A Near Optimal QoE-Driven Power Allocation Scheme for Scalable Video Transmissions Over MIMO Systems", *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 1, pp. 76-88, 2015.

[36] J. Park, X. Chen, J. Hwang. "Optimal power allocation and rate adaptation for scalable video over multi-user MIMO." in *Proc. Global Communications Conference (GLOBECOM), Dec. 2015,* pp. 1-6.

[37] Reichl, Peter, Sebastian Egger, Raimund Schatz, and Alessandro D'Alconzo. "The logarithmic nature of QoE and the role of the Weber-Fechner law in QoE assessment." in *Proc. IEEE International Conference on Communications (ICC)*, 2010, pp. 1-5.

[38] P. Reichl, B. Tuffin and R. Schatz, "Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience", *Telecommunication Systems*, 2011.

[39] W. Zhang, Y. Wen, 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.

[40] D. Tse, P, Viswanath, Fundamentals of Wireless Communications, Cambridge University Press, 2005

[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] S. Ohno and G. Giannakis, "Capacity Maximizing MMSE-Optimal Pilots for Wireless OFDM Over Frequency-Selective Block Rayleigh-Fading Channels", *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2138-2145, 2004.

[43] S. Donthi and N. 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.

[44] M. Kawser, "Downlink SNR to CQI Mapping for Different MultipleAntenna Techniques in LTE", *International Journal of Information and Electronics Engineering*, 2012.

[45] de la Fuente, Alejandro, et al. "Analysis of the impact of FEC techniques on a multicast video streaming service over LTE." in *Proc. European Conference on Networks and Communications (EuCNC)*, Jun. 2015, pp.219-223.

[46] Lentisco, Carlos M., et al. "A model to evaluate MBSFN and AL-FEC techniques in a multicast video streaming service." in *Proc. IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob),* Oct. 2014, pp. 691-696.

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

[48] Mehlführer, Christian, et al. "Simulating the long term evolution physical layer." in *Proc. 17th European. IEEE Signal Processing Conference*, Aug. 2009, pp. 1471-1478.

[49] X. Ge, H. Wang, R. Zi, Q. Li and Q. Ni, "5G multimedia massive MIMO communications systems", *Wireless Communications and Mobile Computing*, vol. 16, no. 11, pp. 1377-1388, 2016.

[50] Pedersen, Gert Frølund. COST 231-Digital mobile radio towards future generation systems. EU, 1999.

[51] Lederer, Stefan, Christopher Müller, and Christian Timmerer. "Dynamic adaptive streaming over HTTP dataset." in *Proc. of the 3rd Multimedia Systems Conference*. ACM, Feb. 2012, pp. 89-94.

[52] X. Hou, J. Harel, and C. Koch, "Image Signature: Highlighting Sparse Salient Regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, 2012.

[53] 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.

[54] J. Xu, X. Guo, Q. Tu, C. Li, and A. Men, "A novel video saliency map detection model in compressed domain," *MILCOM 2015 - 2015 IEEE Military Communications Conference*, 2015.

[55] 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.

[56] Y.-C. Chen, D.-N. Yang, and W. Liao, "Efficient multi-view 3D video multicast with depth image-based rendering in LTE networks," *2013 IEEE Global Communications Conference (GLOBECOM)*, 2013.

[57] A. Awada, E. Lang, O. Renner, K.-J. Friederichs, S. Petersen, K. Pfaffinger, B. Lembke, and R. Brugger, “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.

[58] 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.

[59] FFmpeg (http://www.ffmpeg.org)

[60] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou, V. Tankovich, C. Loop, Q. Cai, P. A. Chou, S. Mennicken, J. Valentin, V. Pradeep, S. Wang, S. B. Kang, P. Kohli, Y. Lutchyn, C. Keskin, and S. Izadi, “Holoportation: Virtual 3D Teleportation in Real-time,” ACM UIST, Oct. 2016.

[61] J. Park, P.A. Chou, J-N. Hwang, “Rate-Utility Optimized Streaming of Volumetric Media for Augmented Reality”. *arXiv preprint arXiv:1804.09864*.

[62] L. De Cicco, et al. "Elastic: a client-side controller for dynamic adaptive streaming over http (dash)." Packet Video Workshop (PV), 2013 20th International. IEEE, 2013.

[63] Li, Zhi, et al. "Probe and adapt: Rate adaptation for HTTP video streaming at scale." IEEE Journal on Selected Areas in Communications 32.4 (2014): 719-733.

[64] T.-Y. Huang, R. Johari, and N. McKeown, "Downton abbey without the hiccups: Buffer-based rate adaptation for http video streaming." Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking. ACM, 2013.

[65] K. Spiteri, R. Urgaonkar, and R. K. Sitaraman. "BOLA: near-optimal bitrate adaptation for online videos." Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on. IEEE, 2016.

# VITA

Jounsup Park received the B.S., and M.S. degree in electronical engineering from Korea University, Seoul, Korea. in 2006, and 2008, respectively. He is currently pursuing the Ph.D. degree in electrical engineering in University of Washington, Seattle. From 2008 to 2012, he was with Cooperate Research Institute of Samsung Electro-Mechanics, Inc., Suwon, Korea. His research interest includes multimedia networking, wireless communication, and AR/VR quality of experience.