

Scalable Video Streaming over Mobile WiMAX

Hung-Hui Juan, Hsiang-Chun Huang, ChingYao Huang, and Tihao Chiang

Department and Institute of Electronics Engineering

National Chiao Tung University (NCTU), Hsinchu, 30050, Taiwan

Tel: +886-35712171 ext 54175, Fax: +886-35724361

Abstract—In this paper, we will investigate the performance of scalable video streaming services in mobile WiMAX systems. We will show that for each user the implementation of multiple connections with feedback information of the available transmission bandwidth is critical for supporting H.264/AVC-based scalable video streaming in which the transmission packets can be further separated into multiple levels of importance. Based on a mobile WiMAX simulation platform including both the video encoder/decoder server and IEEE 802.16e MAC controls, the MAC performance evaluation and subjective test are used to quantify the visual performance between two-connection and one-connection scenarios.

Index Terms—Mobile WiMAX, IEEE 802.16e, SVC, scalable streaming service

I. INTRODUCTION

For video streaming services, to achieve flexible bitstream adaptation for multimedia transmissions, the Joint Video Team (JVT) formed by the ISO/IEC MPEG and ITU-T is developing the scalable extension of the H.264/AVC standard [1-3]. The scalable video coding (SVC) uses layered structure to provide spatial, temporal, and SNR (signal to noise ratio) scalability simultaneously. According to the network conditions and receiver capabilities, the pre-encoded SVC bitstream can be easily adapted by the streaming server to provide various spatial, temporal and quality (SNR) resolutions. Further, the SVC layered structure put the data of different importance into different layers. The unequal erasure protection (UEP) can be easily incorporated with SVC to provide more protection for the important data. With such features, the SVC bitstream is more suitable than the non-scalable bitstream when the video packets are transmitted over an error-prone channel with fluctuated bandwidth.

The IEEE 802.16 standard family [4][5] and the associated Worldwide Interoperability for Microwave Access (WiMAX) forum are developed and formed for supporting the broadband wireless access (BWA) in the wireless metropolitan area network (WMAN). Due to the flexibility and efficiency, the WiMAX is expected to provide many kinds of services, including voice, Internet, and multimedia services.

The rest of this paper is organized as follows: The scalable

extension of H.264/AVC is introduced in Section II. The basic medium access control (MAC) and physical (PHY) layers of IEEE 802.16e will be discussed in Section III. Section IV presents the simulation models and simulation results. Finally, the conclusions are drawn in Section V.

II. OVERVIEW OF SCALABLE VIDEO CODING

Scalable Video Coding (SVC) is a video coding technology that encodes the video at the highest resolution, and allows the bitstream to be adapted to provide various lower resolutions. There are three dimensions of scalability, including spatial, temporal, and quality (SNR) scalability. Spatial scalability means the bitstream can provide different spatial resolutions. Temporal scalability means various frame rates are available. And SNR scalability means the visual quality is scalable.

To achieve the scalability, the video data is encoded into several layers. The lower layers contain lower resolution data. This data is more important because it provided basic video quality with low bit-rate. The higher layers contain the refinement data. It refines the lower resolution data to provide higher resolution video. The refinement data is less important and can be removed when the bandwidth or decoding capability is not sufficient.

For wireless mobile communication, SVC has several advantages over the non-scalable video coding. For a single user, the transmission bandwidth is time-varying due to the mobility and the fluctuation of available resources. Besides, the users are located at different positions; the different signal quality leads to different transmission bandwidth. It is difficult to support all users with a single non-scalable bitstream. Moreover, there is no priority in the non-scalable bitstream. This leads to inefficient error protection because both the more important data and the less important data have the same quality of service in an error-prone channel. SVC provides simple solutions for these problems. The SVC bitstream can be easily adapted to the varying bandwidth and various receivers. And the layered data structure allows the more important data to get more protection easily. As a result, these features provide more efficient video transmission over the error-prone channel with fluctuated bandwidth.

The scalable extension of the H.264/AVC is the latest SVC standard. It is developed by the Joint Video Team (JVT) formed by the ISO/IEC MPEG and ITU-T and it is aimed to simultaneously provide three dimensional scalabilities with

good coding efficiency. To support the spatial scalability, the video is decomposed into several spatial layers. Each spatial layer can be encoded separately or get prediction from the lower spatial layers to remove the redundancy. In each spatial layer, the data can be separated into several SNR layers with two methods: Coarse Grain Scalability (CGS) can only allow the bitstream to be truncated at several pre-defined point, while the Fine Grain Scalability (FGS) allows the bitstream to be truncated at any point in the bitstream. Note that in the standard, the first SNR layer in a spatial layer is restricted to CGS. To support the temporal scalability, a hierarchical prediction structure is used. The pictures in a group of picture (GOP) are dyadic decomposed into several layers.

III. OVERVIEW OF IEEE 802.16E/MOBILE WiMAX

The IEEE 802.16 Working Group works on the technology of broadband wireless access and develops the standard of Wireless Metropolitan Area Networks (WMAN). The fixed broadband wireless access system IEEE 802.16d has been standardized in 2004, the standard is also called IEEE 802.16-2004. And the mobile broadband wireless access system IEEE 802.16e, or so-called IEEE 802.16e-2005 or mobile WiMAX, has been approved in December 2005. IEEE 802.16e provides mobility enhancements to IEEE 802.16-2004 for the purpose of supporting subscribers moving at the vehicular speed. The IEEE 802.16 Working Group mainly standardizes the specification of the medium access control (MAC) protocol and the physical layer (PHY).

■ 3.1 Overview of the PHY layer of WiMAX

The PHY layer of WiMAX comprises different configurations. In this study, the TDD (time division duplex) mode of WirelessMAN-OFDMA PHY layer will be considered. For more details of Wireless MAN-OFDMA frame structure, please refer to [4-5].

The IEEE 802.16 defines several burst profiles that are the combination of modulation and coding scheme in each PHY configuration. With link adaptation, the system can decide the proper modulation and coding scheme according to the current CINR (carrier to interference and noise ratio) value. In IEEE 802.16e, the CINR of each mobile station may change with time. At the beginning of the frame, the base station should decide the burst profile of each DL and UL data burst. For the DL data burst, the base station can decide the burst profile of the DL data burst according to the feedback DL channel condition in the UL fast feedback channel (or called channel quality indication channel, CQICH).

■ 3.2 Overview of the MAC layer of WiMAX

The MAC layer of WiMAX mainly supports a point-to-multipoint (PMP) architecture, and it can support mesh architecture optionally. The MAC layer is designed for supporting the different QoS (quality of service) requirements of different types of applications. The MAC layer of WiMAX consists of three sublayers including convergence sublayer, common part sublayer, and privacy sublayer. Because the MAC layer of WiMAX must support the various backhaul

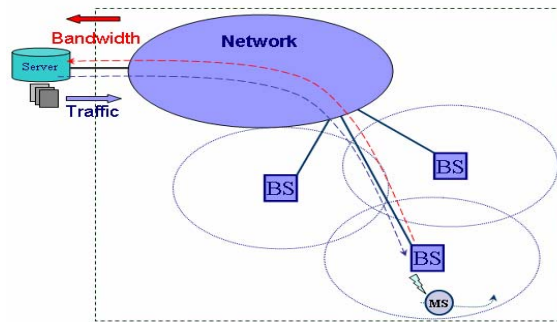


Figure 1. The end-to-end Mobile WiMAX system

network like asynchronous transfer mode (ATM) network and IP-based networks, **the convergence sublayer** is in charge of mapping different types of transport-layer traffic to a MAC connection. Each service is mapped to a connection in MAC layer. This sublayer classifies the service data units (SDUs) to a proper connection with specific QoS parameters. **The common part sublayer** controls most parts of MAC functionalities like fragmentation, packing, scheduling, and retransmission and so on. **The privacy sublayer** is for data encryption and it can provide the security on network transmission.

IV. SIMULATIONS

■ 4.1 Simulation Models

As depicted in Figure 1, an end-to-end transmission of a streaming video in the mobile WiMAX system is simulated, where the last mile transmission system is IEEE 802.16e/Mobile WiMAX which consists of base stations (BSs) and mobile stations (MSs). The streaming server and WiMAX subsystem are inter-connected by an IP-based backhaul network. The streaming service in this study is encoded by the scalable extension of H.264/AVC.

To support multiple connections between the BS and MS, the data sent to the BS from the streaming server is allocated into several connections according to the importance levels, as shown in Figure 2. The more important data is allocated to the connection that has more protection (i.e., higher transmission priority and MAC retransmission). To address the bandwidth fluctuation effect, in the proposed two-connection implementation, the server allocates the more important data

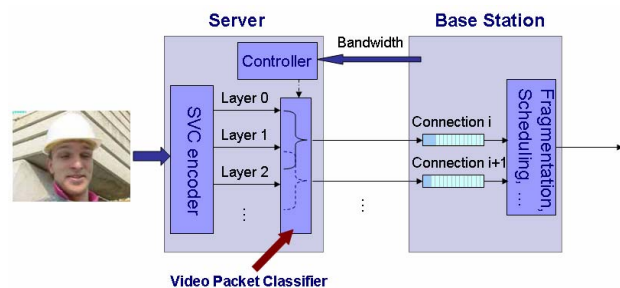


Figure 2. The interaction between the server and the base station

which occupies 80% of total data at the first connection, and put the remaining data at the second connection. This allows the BS need only re-transmit the more important data when the actual transmission bandwidth is smaller than the expected bandwidth.

Besides, the 802.16 MAC is flexible and it allows multiple connections for a mobile subscriber station. The multiple connections can belong to same or different service types and can be related to same or different QoS parameters. Because the streaming server adopts scalable video coding and it can decide which video packet is more important and which is not, the BS can establish two or more connections with the server. In this study, a two-connection scenario is adopted and compared with the performance of one-connection scenario. In the two-connection scenario, the first connection is for more important video packets and we give it higher priority. The second connection is for less important video packets (like some enhancement packets) and we give it lower priority as compared to the first connection.

In the simulation, WirelessMAN-OFDMA TDD mode and PUSC (Partial Usage of Subchannel) are adopted. The FFT size is 2048 and the bandwidth is 6 MHz. From the standard we can find the OFDMA channelization parameters [4-5]. The relative parameters are listed in Table 1.

Figure 3 shows the controls of the BS MAC which mainly comprise three parts: the first is the fragmentation, the second is the scheduling and the third is the MAC retransmission. For the fragmentation, we will set a target PDU error rate of 5% and decide the proper fragment PDU size. For the scheduling controls, the early deadline first algorithm is adopted. For the MAC retransmission, the ARQ mechanism of each MAC connection is enabled. Table 2 shows the parameters used in the simulations.

The SVC bitstream used in the simulation is encoded to provide two spatial layers including QCIF (Quarter Common Intermediate Format, 176x144) and CIF (Common Intermediate Format, 352x288). As shown in Table 3, the QCIF resolution has one SNR layer and the CIF resolution including four SNR layers.

In each of the spatial-SNR layers, the GOP size is limited to 8 pictures to reduce the buffer requirement at the mobile station but still provides four layers temporal scalability through the hierarchical prediction structure. The intra-pictures is inserted every 64 pictures (8 GOPs) to provide the error recovery point. With the five spatial-SNR layers and the four temporal layers in each spatial-SNR layer, up to 20 layers bitstream adaptation is allowed. Further, the FGS is used from the CIF-SNR1 to the CIF-SNR3 layers, such that the bitstream can be truncated at any point to achieve the requested bitrate.

The test sequence is making up from 13 commonly used MPEG test sequences to form a 3600 pictures video, which including bus, football, foreman, mobile, city, crew, harbour, soccer, coastguard, container, mother daughter, stefan, and table tennis. The average bitrate of the SVC bitstream at various spatial-SNR and temporal resolutions are shown in Table 3.

4.2 Simulation Results

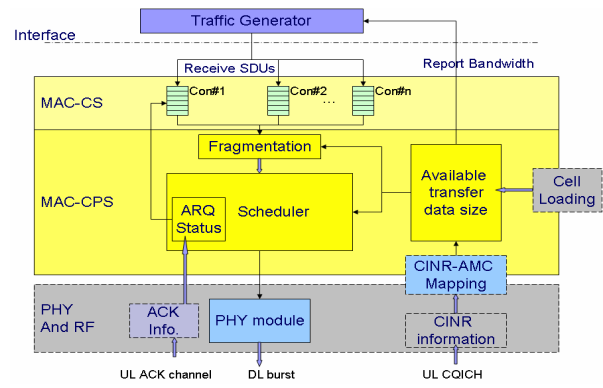


Figure 3. Base Station MAC Control modules

Table 1. The WirelessMAN-OFDMA PHY parameters

Parameters	Values
Channel Bandwidth	6MHz
FFT size (N_{FFT})	2048
Number of Sub-channels	60
Useful Symbol Time (T_b)	298.667 μ s
Guard Time ($T_g = 1/8 T_b$)	37.333 μ s
OFDMA symbol Time ($T_s = T_g + T_b$)	336 μ s
Frame duration	5ms

Table 2. The parameters used in the simulation

Parameters	Value
Maximum Latency	550 ms
Max. MAC retransmission Times	6
Average BER of the channel model	$2 \cdot 10^{-4}$
Bad State BER of the channel model	10^{-3}
Good State BER of the channel model	10^{-4}
B_w , (defined in [6])	2

TABLE 3. THE AVERAGE BITRATE OF THE SVC BITSTREAM AT VARIOUS SPATIAL-SNR AND TEMPORAL RESOLUTIONS

SPATIAL-SNR LAYERS	BITRATE@ 3.75FPS	BITRATE@ 7.5FPS	BITRATE@ 15FPS	BITRATE@ 30FPS
QCIF-SNR0	10.9801	14.8971	19.3172	23.9405
CIF-SNR0	56.6238	67.7029	81.5013	103.4000
CIF-SNR1	152.9394	172.2780	205.6717	247.6827
CIF-SNR2	335.3745	387.9433	456.7107	539.1969
CIF-SNR3	616.8722	747.8123	911.1053	1095.1940

Figure 4 shows the SDU failure rate of important video packet in two-connection scenario and one-connection scenario. In the one-connection scenario, by defining the Type0 SDU as important SDU for the scalable video streaming, the Type0 SDU failure rate will then be the failure rate of the important video packets. In one-connection scenario, the SVC-encoded video packets of different layers are sent within

one connection. In the two-connection scenario, the first connection (CON#1) carries more important video packets and the second connection (CON#2) carries less important video packets. By considering the failure rate of the important video packets only, from Figure 4, it is obvious that the two-connection scenario has the better control of the failure rate of the important video packets even when the cell loading increases.

4.3 Subjective test

In the subjective test, there are 18 viewers and we compare the visual performance between two-connection scenario and one-connection scenario. There are four cases, each case has different average data rate. The average data rate of each case is: case 1 > case 2 > case 3 > case 4. For each video sequence, there are five scores, ranged from 1 to 5.

Figure 5 shows the average scores from 18 viewers. It indicates that the visual performance of two-connection scenario is better than one-connection scenario in all cases especially when the data rate is lower. Figure 6 shows the distribution of the score comparisons. For each case, if the ranking of two-connection scenario is higher than the ranking of one-connection scenario, we classify it as the “Better” case, and vice versa. We can see that most viewers consider that the visual performance of two-connection scenario is better than

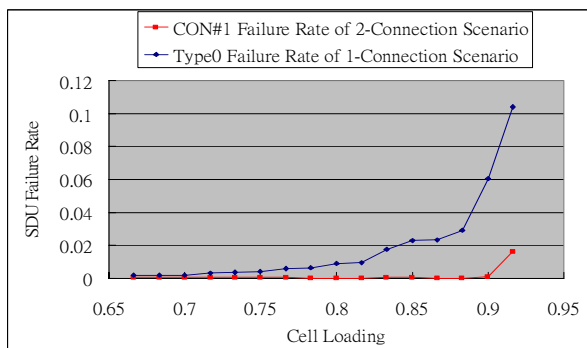


Figure 4. The important SDU failure rate of the two-connection and one-connection scenario under the different cell loading.

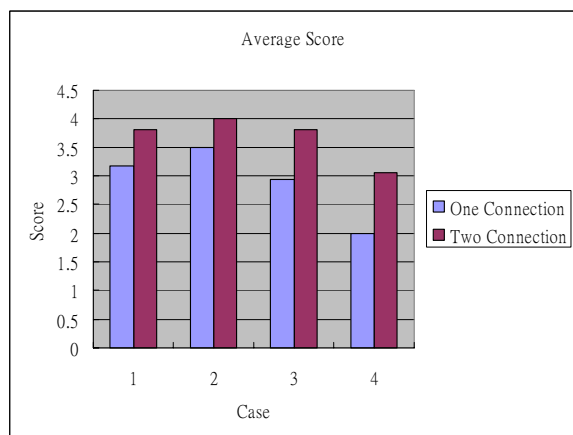


Figure 5. The average scores of the subjective test

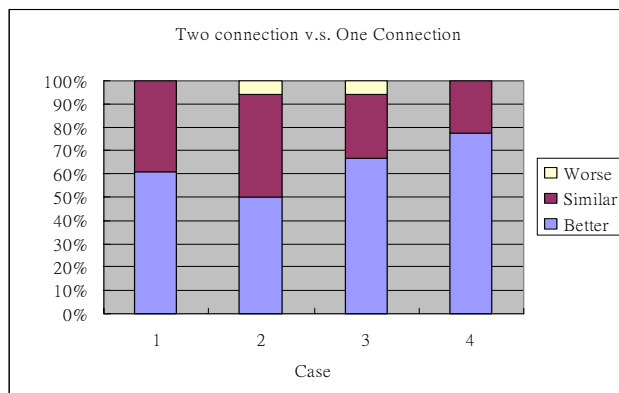


Figure 6. The score comparison of the subjective test

that of one-connection scenario.

V. CONCLUSIONS

In this study, the transmission performance of the scalable video streaming services over the mobile WiMAX system is investigated. We show that the implementation of multiple connections for the scalable video streaming with feedback information of the available transmission bandwidth is critical in which the transmission packets can be further separated into multiple levels of importance. Finally, simulation results and subjective tests are done to verify above conclusions by performing user perceived video quality tests.

REFERENCES

- [1] [1] “Advanced Video Coding for Generic Audiovisual Services,” ITU-T and ISO/IEC JTC1, ITU-T Recommendation H.264 – ISO/IEC 14496-10 AVC, 2003
- [2] [2] “Joint Draft 5: Scalable Video Coding,” ITU-T and ISO/IEC JTC1, JVT-R201, Jan. 2006
- [3] [3] “Joint Scalable Video Model JSVM-5,” ITU-T and ISO/IEC JTC1, JVT-R202, Jan. 2006
- [4] [4] IEEE 802.16-2004, “IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems,” June, 2004
- [5] [5] IEEE 802.16e-2005, “IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1,” February, 2005
- [6] K. Sakakibara, “Performance Analysis of the Error-Forecasting Decoding for Interleaved Block Codes on Gilbert-Elliott channels,” IEEE Transactions on Communications, March 2000