Wireless Netw (2010) 16:1143–1156
DOI 10.1007/s11276-009-0193-y

Cross-layer quality-driven adaptation for scheduling heterogeneous multimedia over 3G satellite networks

Hongfei Du · Xiaozheng Huang · Jie Liang · Jiangchuan Liu · Barry G. Evans · Imrich Chlamtac

Published online: 13 June 2009

Abstract Wireless networks are experiencing a paradigm shift from focusing on the traditional data transfer to accommodating the rapidly increasing multimedia traffic. Hence, their scheduling algorithms have to concern not only network-oriented quality-of-service (QoS) profiles, but also application-oriented QoS targets. This is particularly challenging for satellite multimedia networks that lack fast closed-loop power control and reliable feedbacks. In this paper, we present a cross-layer packet scheduling scheme, namely Hybrid Queuing and Reception Adaptation (HQRA), which performs joint adaptations by considering the traffic information and QoS targets from the applications, the queuing dynamics induced from the network, as well as the end-to-end performance and channel variations from respective users. By jointly optimizing multiple performance criteria at different layers, the scheme enjoys quality-driven, channel-dependant, and network-aware features. HQRA can well accommodate return link diversity and the imperfect feedbacks, whilst ensuring robustness in highly heterogeneous and dynamic satellite environments. We evaluate its performance over diverse network and media configurations in comparison with the state-of-the-art solutions. We observe noticeable performance gains on application-oriented QoS, bandwidth utilization, and objective video quality, together with favorable fairness and scalability measures.

Keywords Packet scheduling · Cross-layer · Quality-driven · Multimedia streaming · Radio resource management · SDMB

H. Du (✉) · X. Huang · J. Liang · J. Liu
Simon Fraser University, Burnaby, BC V5A 1S6, Canada
e-mail: hongfei.du@ieee.org

X. Huang
e-mail: xha13@sfu.ca

J. Liang
e-mail: jiel@sfu.ca

J. Liu
e-mail: jcliu@cs.sfu.ca

B. G. Evans
University of Surrey, Guildford, Surrey GU2 7XH, UK
e-mail: b.evans@surrey.ac.uk

I. Chlamtac
CREATE-NET, 38100 Trento, Italy
e-mail: chlamtac@create-net.org

1 Introduction

Recent advances in mobile multimedia broadcasting and satellite communications have generated many new challenges and barriers that require seamless internet working between satellite and terrestrial communication infrastructures, as well as efficient management of available radio resources. The satellite digital multimedia broadcasting (SDMB) [1] system implements a satellite-based broadcast layer over 2.5G and 3G terrestrial mobile cellular networks, aiming at the efficient delivery of multimedia broadcast multicast services (MBMS) [2]. Due to the unidirectional nature of the baseline SDMB system and the point-to-multipoint services it provides, the design of the packet scheduling algorithms in SDMB is quite challenging. An efficient packet scheduling scheme in SDMB is required not only to optimize the network-oriented QoS performance such as throughput, delay and jitter, but also to satisfy the application-specific QoS demands, whilst considering the transmission power constraints and the reception quality levels.

There has been an extensive research on efficient scheduling algorithm design for general wireless networks. One interesting approach in this context is delay differentiated scheduling, where waiting time and queuing delay are considered in the packet scheduling, as waiting time priority (WTP) and proportional delay differentiation (PDD) schemes as proposed in [3] for terrestrial differentiation networks. Authors in [4] propose a novel approach for Wireless CDMA networks via a QoS-oriented packet scheduling scheme for design optimization. Adaptive proportional fairness (APF) scheduling was proposed in high-speed downlink packet access (HSDPA) system [5], considering QoS demands for multimedia applications, where the CSI information for individual user can be tracked via return channel.

Given the absence of fast closed-loop power control (CLPC) and the difficulties in acquiring effective channel state information (CSI) and end-to-end (ETE) measures, existing packet scheduling algorithms [3, 4, 5] targeting at general wireless mobile networks cannot be directly applied in the SDMB system. Traditional SDMB applies a unidirectional transmission mode, in this paper, we consider more advanced scenario with return link available for gathering user/channel related performance. However, schemes in [3, 4, 5] applied in terrestrial wireless network still can not be applied directly to such a system, the main reasons are multi-folds: (1) the inherent features for a GEO-Sat based communication system is the long-trip delay, which is even negligible for terrestrial links. (2) There is a hard limit on GEO-Sat transmit power, which in turn limits the transmit power from the Sat-GW, the scheduler has to check this hard power constraints in each TTI for allocating the resource/bandwidth, however, the base station in a general wireless network is often considered with unlimited power supply. As such, adaptations as developed in this paper are needed for coping with the aforementioned constraints. (3) We consider multicast/broadcast service (MBMS), in another word, a single session at the Sat-GW will be transmitted to multiple receivers in different locations spreading over a very wide geographical area. Therefore, the effective collection and estimation of the overall performance from multiple receivers corresponding to a single session is an interesting but challenging task.

Previous studies [6] address the packet scheduling problems in SDMB via the adaptations of two well-known queuing models, namely multi-level priority queuing (MLPQ) and weighted fair queuing (WFQ). However, both of them suffer from major weaknesses in provisioning QoS-differentiated multimedia services with respect to efficiency and fairness. MLPQ always processes packets starting from the highest QoS class, with queues having the same priority served in a round robin fashion. This scheme favours the high QoS class service, assuring a delay bound for their packets, yet it provides no guarantees for lower QoS classes. Besides, there is no differentiation for queues with the same QoS class. However, rather than prioritizing queues in a strict manner, other QoS metrics (e.g., delay tolerance and guaranteed data rate) should also be considered in the scheduling discipline design. WFQ-based scheduling was motivated and developed in the SDMB system based on the well-known WFQ scheme. The weights are primarily set according to the data rates of the multiplexed service flows rather than its QoS class. Queues with the same QoS class will be served based on its timestamp on its head packet. The performance of WFQ is worse than that of MLPQ in terms of both delay and delay-variation.

On the other hand, the power allocation algorithm used in previous schemes is based on static rate matching (SRM), where the transmit power setting for a physical channel is based on the most demanding reception quality requirement (in terms of energy per bit to noise power spectral density ratio E_b/N_o) of all multiplexed service flows under the target block error rate (BLER). To improve this, the delay differentiation queuing (DDQ) [7] and the dynamic resource allocation (DRA) algorithm [8] were investigated in our previous works to optimize the resource allocation using the instantaneous data rate information. DDQ is developed based on the delay-sensitive schemes [3] proposed in differentiated services networks. DDQ-based scheduling performs slightly different as follows. For each TTI, DDQ derives the serving indices based on the average waiting delay for all packets currently in the queue, the average queuing delay for all the packets having left the queue, the packet arrival rate and QoS ratio. Compared with WFQ and MLPQ, DDQ offers improved performance in delay, delay-variation, and channel utilization. However, DDQ experiences unbalanced performance among multiple QoS attributes, namely the gain achieved in one performance attribute leads to the performance degradation in other attributes. Furthermore, multimedia services feature differentiated delay constraints and applies the delay constraints for differentiated services in an equal way may lead to poor QoS guarantee for high priority queues. Therefore the delay profile has to be considered against the respective delay constraints (i.e., maximum acceptable delay) specified by the class of service. Finally, rather than scheduling competing flows in a static manner, to provide more flexible QoS provisioning and maintain optimal resource utilization, it is highly desired that the scheduler is capable of choosing the best scheduling policy according to diverse QoS preferences of the services and instantaneous performance dynamics.

The aforementioned optimizations fail to take into account the CSI information at the receivers, which is difficult to be used in a unidirectional satellite broadcast/multicast (BC/MC) network. In this paper, we investigate the major problems encountered in the packet scheduling of conventional unidirectional mobile satellite networks, and propose a novel packet scheduling framework, namely the hybrid queuing and reception adaptation (HQRA), which effectively tracks the queuing dynamics induced from the network and accommodates users with diverse and fast-varying reception conditions.

The existing resource management strategies in the lower layers of SDMB are optimized without explicitly considering the application-specific QoS targets of multimedia applications. Therefore they lead to suboptimal multimedia performance on objective and/or perceptual quality. In this paper, to improve the final delivered reception quality, we employ a cross-layer methodology in HQRA that dynamically adapts the scheduling strategies across the protocol stack by jointly considering the application-layer traffic information and QoS targets, as well as the lower layer queuing and link dynamics. By jointly considering multiple performance criteria available at different layers, the HQRA features *quality-driven*, *channel-dependant*, *and network-aware* properties, thereby optimizing the network-oriented and application-oriented QoS. We discuss the respective issues on return link diversity, the delayed feedback as well as the scalability and flexibility. We evaluate the performance of HQRA over a set of simulation scenarios via discrete event simulations. Comparing with existing schemes in SDMB, our simulation results show that, HQRA not only improves network-oriented performance such as delay, jitter, and channel utilization, but also effectively optimizes the application-oriented QoS, fairness as well as objective video quality, with favorable flexibility and scalability features.

The rest of this paper is organized as follows. The system reference architecture and the radio resource management (RRM) concept are introduced in Sect. 2. The proposed HQRA strategy is then presented in Sect. 3. In Sect. 4, the evaluation methodology is described. We then proceed in Sect. 5 with the performance evaluation of the proposed scheme. We summarize our scheme and conclude this paper in Sect. 6.

2 Background

The reference architecture model for the bidirectional satellite multimedia broadcasting network, as shown in Fig. 1, typically consists of a satellite gateway (Sat-GW), a geostationary satellite (GEO-Sat), terrestrial gap-fillers, i.e., intermediate module repeaters (IMRs), and a wide variety of mobile satellite terminals (MSTs) with diverse capabilities and fast-varying channel conditions. Given the severe channel conditions associated with the satellite links, the system employs the Forward Link (FL) via either GEO-Sat (FL-GS) or IMRs (FL-IMR), whilst the interactive activities are maintained by the Return Link (RL) via either Terrestrial Network (RL-TN) or GEO-Sat (RL-GS). Due to the major discrepancies induced between satellite link and terrestrial link in terms of bandwidth, latency and loss, the preferred access link is defined as FL-GS with a RL-TN. Nevertheless, in the presence of blocking or fading of the preferred links, the interactive activities will be maintained via other available links in order to adapt to the fast-varying channel conditions and to maximize the geographical coverage. The functional components involved in this topology are described as follows:

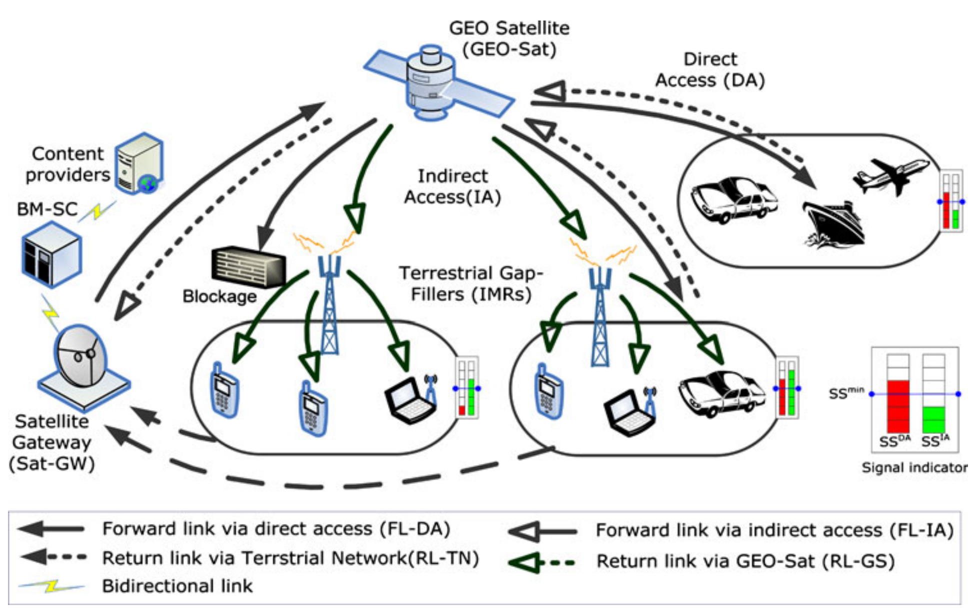


Fig. 1 Interactive satellite multimedia broadcasting system reference model

- Sat-GW: It is connected to an interactive broadcast network service provider via Broadcast/Multicast Service Center (BM-SC) and terrestrial core networks.
- GEO-Sat: It is capable of on-board processing, and is controlled by a remote Network Operation Center (NOC) through a dedicated high-bandwidth channel.
- IMR: It includes a full replacement with terrestrial core network, which can be used to complement the satellite unreachable/blocked coverage, e.g., serious multipath built-up areas, or deep fading/shadowing areas.
- MST: It applies the standard return channel satellite terminals, equipped with built-in channel measurements and evaluation model. The terminal is capable of collecting CSI information from the detectable signals coming from both direct access (DA) and indirect access (IA) links.

We focus on the scheduling area associated with a single spot beam from a Sat-GW, where the resource is allocated to the MSTs according to session QoS demands and user reception quality. Two types of reception signals can be identified at the MSTs: signal from GEO-Sat via direct access ($\mathrm{SS}^{\mathrm{DA}}$); signal from IMRs via indirect access ($\mathrm{SS}^{\mathrm{IA}}$), as illustrated in Fig. 1. Let $\mathrm{SS}^{\min}$ be the minimum acceptable signal strength for the specific MST in order to maintain the session communication, this threshold depends on the terminal capability. Therefore, three types of receiver reception conditions can be identified for the jth MST for the ith session at the nth TTI as:

- Type A ($\mathrm{SS}_{i,j}^{\mathrm{DA}}(n) > \mathrm{SS}_{i,j}^{\min}\,\&\,\mathrm{SS}_{i,j}^{\mathrm{IA}}(n) < \mathrm{SS}_{i,j}^{\min}$): An example of this scenario can be remote users in far-flung geographical locations without the access of terrestrial backhaul infrastructures. In this case, the RL-GS is the only available link.
- Type B ($\mathrm{SS}_{i,j}^{\mathrm{DA}}(n) > \mathrm{SS}_{i,j}^{\min}\,\&\,\mathrm{SS}_{i,j}^{\mathrm{IA}}(n) > \mathrm{SS}_{i,j}^{\min}$): This scenario applies to the users in unban/build-up area. The MSTs have excellent access to signals from both GEO-Sat and IMRs.
- Type C ($\mathrm{SS}_{i,j}^{\mathrm{DA}}(n) < \mathrm{SS}_{i,j}^{\min}\,\&\,\mathrm{SS}_{i,j}^{\mathrm{IA}}(n) > \mathrm{SS}_{i,j}^{\min}$): This scenario applies to the indoor/in-building users, where the satellite signal is currently blocked; the FL-IMR and RL-TN will be the only link available to maintain the interactive activities.

Based on the above discussions, the recipient member group associated with a typical broadcast scenario can be formed as a combination of MSTs with the reception conditions of Type A, B, and C, while the recipient member group associated with a multicast scenario is formed as any possible subsets of a combination of MSTs with the reception condition of Type A, B, and C.

The radio resource management (RRM) functionalities implemented at the SDMB access layer comprise three main separated but cooperated parts: packet scheduling, radio resource allocation (RRA), and admission control. Due to the long latency involved in satellite links and the point-to-multipoint services it provides, the design of an efficient RRM scheme is challenging.

The RRA is responsible for the radio bearer configuration at the admission for each session, which estimates the required number of logical/transport/physical channels and maps them from logical channels to transport/physical channels. Each service is mapped onto an MBMS point-to-multipoint Traffic CHannel (MTCH) logical channel [9], which is mapped onto the Forward Access Channel (FACH) transport channel. At the physical layer, the Secondary Common Control Physical CHannel (S-CCPCHs) can carry one or more FACH(s) via transport channel multiplexing.

The admittance decision of each incoming requested session is handled by the admission control function. An appropriate Transport Format Combination Set (TFCS) for each physical channel must be derived for the packet scheduler to perform the resource allocation.

The long-latency involved in satellite link renders the fast closed-loop power control [10] in the uplink unfeasible. Therefore, the packet scheduler, which is the single function performing short-term resource allocation, becomes the main mechanism for fast resource management and performs the priority handling and transport format (TF) selection tasks. On the other hand, the downlink power allocation should be effectively designed in conjunction with the resource management functions. More specifically, in SDMB, the packet scheduler is responsible for two important tasks that are executed in each transmission time interval (TTI) of the radio bearers:

- Time-multiplexing of service flows with different QoS requirements into physical channels with fixed spreading factor (SF), so as to satisfy these requirements.
- Adjusting the transmit power of the physical channels on the basis of the required reception service quality, i.e., block error rate (BLER), whilst being constrained by the limited transmission power within a satellite beam.

3 HQRA: hybrid queuing and reception adaptation

3.1 Overview of HQRA

The resource management function is physically implemented at the satellite gateway (Sat-GW). The MST is responsible for estimating the reception condition, which is periodically fed back to the Sat-GW via the terrestrial link with relatively low latency. Existing schemes adopted in the SDMB system are confined to the QoS traffic types, where the resource management entities operate without interacting with other layer functions. This leads to resource waste and an inferior performance on both network performance and application-oriented QoS guarantees.

To address these problems, we propose HQRA (hybrid queuing and reception adaptation), a novel cross-layer quality-driven packet scheduling scheme for the bidirectional SDMB systems. HQRA jointly utilizes the multiple performance measures from across the protocol stacks. First, to cope with the highly vibrating wireless fading channels, the CSI information from physical layer is considered into the scheduling decision policy. Second, it effectively tracks the queuing dynamics in the RLC layer queuing buffer for each competing flows, as an important multi-dimensional performance metrics. Last, the ETE performance is considered via terrestrial feedback channels, aimed at guaranteeing their application layer QoS demands. We address the proposed HQRA cross-layer framework mathematically in subsequent section. Important notations used in this article are listed in Table 1.

Table 1 Important notation table

Symbol	Semantics
$\gamma_{i,j}^{\mathrm{DA}}(n)$	SNR associated with the $\mathrm{SS}_{i,j}^{\mathrm{DA}}(n)$ estimated from the jth MST for the ith session at the nth TTI
$\gamma_{i,j}^{\mathrm{IA}}(n)$	SNR associated with the $\mathrm{SS}_{i,j}^{\mathrm{IA}}(n)$ estimated from the jth MST for the ith session at the nth TTI
$\gamma_{i,j}^{\min}$	Minimum SNR associated with the jth MST in the ith session
$\phi_{i,j}^{\mathrm{DA}}(n)$	Reception index for the ith session at the jth MST from DA
$\phi_{i,j}^{\mathrm{IA}}(n)$	Reception index for the ith session at the jth MST from IA
$\phi_{i,j}(n)$	Overall reception index for the ith session at the jth MST
$\Delta_{i,j}(n)$	Delay index for the ith session at the jth MST at the nth TTI
$P_{i,j}(n)$	PLR index for the ith session at the jth MST at the nth TTI
$\Psi_{i,j}(n)$	Throughput index for the ith session at the jth MST at the nth TTI
$\mathbf{R}_{i,j}(n)$	Effective reception index for the ith session at the jth MST
$\mathbf{Q}_i(n)$	Queuing status function (QSF) for the ith transport channel
$\mathbf{R}_{i,x}^{\mathrm{TN}}(n)$	IMR reception condition for the ith session in the RL-TN
$\hat{\mathbf{R}}_i(n)$	Overall reception condition for the ith session
$\mathbf{R}_i^{\mathrm{GS}}(n)$	Reception condition for the ith session with RL-GS
$\mathbf{R}_i^{*}(n)$	Reception condition threshold for the ith transport channel
α_i	Application prescribed QoS rank for the ith session
$T_i(n)$	Queuing delay profile of ith FACH at the nth TTI
$\Lambda_i(n)$	Buffer occupancy profile of ith FACH at the nth TTI
$\Gamma_i(n)$	Data rate profile of ith FACH at the nth TTI
$\Xi_i(n)$	Buffer overflow probability profile of ith FACH at the nth TTI
$H_i(n)$	Buffer throughput profile of ith FACH at the nth TTI

3.2 Effective reception evaluation

Given the diverse and time-varying propagation channel conditions a receiver may experience, it is expected that the MSTs can effectively capture their instantaneous channel and link variations. As shown in Fig. 2, the MST performs the measurements and the evaluations on the received channel quality, and generates a "reception status table (RST)", which includes essential user-oriented performance metrics. We define the instantaneous SNR associated with the $\mathrm{SS}_{i,j}^{\mathrm{DA}}(n)$ and $\mathrm{SS}_{i,j}^{\mathrm{IA}}(n)$ estimated from the jth MST for the ith session at the nth TTI as $\gamma_{i,j}^{\mathrm{DA}}(n)$, $\gamma_{i,j}^{\mathrm{IA}}(n)$, respectively. The minimum acceptable SNR associated with the jth MST in the ith BC/MC group is denoted as $\gamma_{i,j}^{\min}$. To determine the most appropriate mode of return link from either RL-GS or RL-TN, we define the reception indices for the ith session at the jth MST from DA and IA as $\phi_{i,j}^{\mathrm{DA}}(n)$ and $\phi_{i,j}^{\mathrm{IA}}(n)$, respectively, which are given by:

$$\begin{cases} \phi_{i,j}^{\mathrm{DA}}(n) = \frac{\gamma_{i,j}^{\mathrm{DA}}(n)}{\gamma_{i,j}^{\min}} & \text{for } \mathrm{SS}^{\mathrm{DA}} \\ \phi_{i,j}^{\mathrm{IA}}(n) = \frac{\gamma_{i,j}^{\mathrm{IA}}(n)}{\gamma_{i,j}^{\min}} & \text{for } \mathrm{SS}^{\mathrm{IA}} \end{cases} \tag{1}$$

Based on the reception index derived in Eq. 1, the return link mode for the ith session at the jth MST is determined by:

$$\mathrm{RL}_{i,j}^{\mathrm{Mode}} = \begin{cases} \mathrm{RL\text{-}TN}, & \text{if } \phi_{i,j}^{\mathrm{IA}}(n) \geq 1 \\ \mathrm{RL\text{-}GS}, & \text{if } \phi_{i,j}^{\mathrm{IA}}(n) < 1 \,\&\, \phi_{i,j}^{\mathrm{DA}}(n) \geq 1 \\ \mathrm{none(without\,RL)}, & \text{else} \end{cases} \tag{2}$$

Therefore, the overall reception index $\phi_{i,j}(n)$for the ith session at the jth MST can be obtained by:

$$\phi_{i,j}(n) = \frac{\gamma_{i,j}^{\mathrm{DA}}(n) + \gamma_{i,j}^{\mathrm{IA}}(n)}{2 \cdot \gamma_{i,j}^{\min}} \tag{3}$$

It reflects the current SNR performance for the corresponding MST, and is included in the RST as one of the CSI metrics.

The transmission delay considered is the ETE performance metric measured at the respective MSTs. We define the mean delay $\overline{d_{i,j}}(n)$ for the ith session at the jth MST until the nth TTI as:

$$\overline{d_{i,j}}(n) = \frac{\sum_{k=1}^{N_{i,j}(n)} \left(t_{i,j}^{\mathrm{MST}}(k) - t_{i,j}^{\mathrm{Sat-GW}}(k)\right)}{N_{i,j}(n)} \tag{4}$$

where the $N_{i,j}(n)$ is the number of packets that have been received by the jth MST until the nth TTI, $t_{i,j}^{\mathrm{Sat-GW}}(k)$ is the time for the kth packet in the ith session leaving from Sat-GW, $t_{i,j}^{\mathrm{MST}}(k)$ is the time for the kth packet in the ith session arriving at the MST.

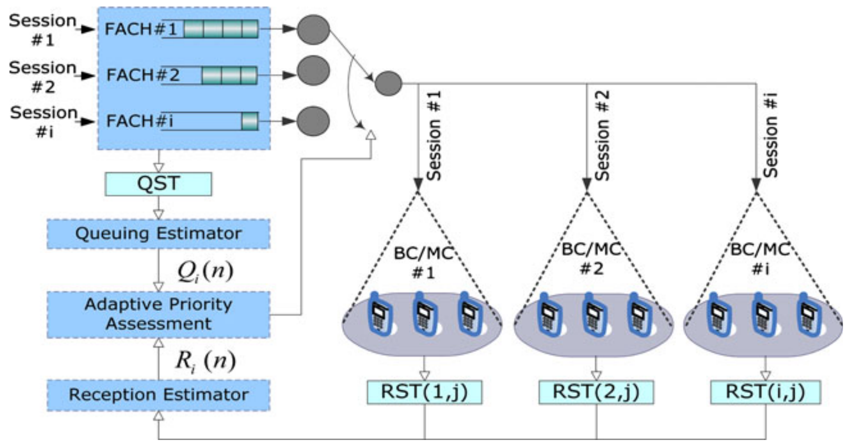


Fig. 2 The effective reception evaluation process

The ETE delay threshold is characterized by the application-specified maximum acceptable delay. Let d_i^* denote the maximum acceptable delay for the ith session. We associate with the jth MST an *delay index* $\Delta_{i,j}(n)$, which is given by:

$$\Delta_{i,j}(n) = \begin{cases} 1 & \text{if } \overline{d_{i,j}}(n) \leq d_j^* \\ \frac{\overline{d_{i,j}(n)}}{d_i^*} & \text{if } \overline{d_{i,j}}(n) > d_i^* \end{cases} \tag{5}$$

The packet loss in the propagation path is one of the most crucial factors impacting the QoS. Let $\rho_i^*(n)$ denote the target PLR for the ith session. The *PLR index* $P_{i,j}(n)$ is defined as:

$$P_{i,j}(n) = \begin{cases} 1 & \text{if } \overline{\rho_{i,j}}(n) \leq \rho_i^* \\ \frac{\overline{\rho_{i,j}(n)}}{\rho_i^*} & \text{if } \overline{\rho_{i,j}}(n) > \rho_i^* \end{cases} \tag{6}$$

where $\overline{\rho_{i,j}}(n)$ denotes the mean PLR at the jth MST in the ith session achieved until the nth TTI.

The ETE throughput is measured as the ratio between the total bits arrived in the jth MST up to current time to the total bits that have been transmitted from the Sat-GW. Let θ_i^* denote the target throughput for the ith session, the *throughput index* $\Psi_{i,j}(n)$ associated with the jth MST is given by:

$$\Psi_{i,j}(n) = \begin{cases} 1 & \text{if } \overline{\theta_{i,j}}(n) \geq \theta_i^* \\ \frac{\theta_i^*}{\overline{\theta_{i,j}(n)}} & \text{if } \overline{\theta_{i,j}}(n) < \theta_i^* \end{cases} \tag{7}$$

where $\overline{\theta_{i,j}}(n)$ denotes the mean throughput of the jth MST in the ith session that has been achieved so far.

At the Sat-GW, upon receiving the RSTs from respective MSTs, an *effective reception index* $\mathbf{R}_{i,j}(n)$ is obtained:

$$\mathbf{R}_{i,j}(n) = \phi_{i,j}(n) \cdot \Delta_{i,j}(n) \cdot P_{i,j}(n) \cdot \Psi_{i,j}(n) \tag{8}$$

3.3 Return link adaptation

To effectively manage the radio resources and maximize the channel capacity, the HQRA adopts a novel return link adaptation (RLA) scheme to adapt the scheduling policies based on whether a return path from terrestrial network or satellite is used. A unified reception estimation (URE) is defined in HQRA to effectively retrieve the reception conditions for a MST, based on the measurements on both $\mathrm{SS}^{\mathrm{DA}}$ and $\mathrm{SS}^{\mathrm{IA}}$.

To increase the reliability and scalability of the overall scheduling performance, we perform an intermediate evaluation at the IMRs. It conducts the measurements and assessments on all the BC/MC members in its IMR cell, and then reports the overall status of the IMR cell to the Sat-GW. The URE applies differentiated treatments on the MSTs accordingly, based on whether a RL-TN or RL-GS is used for the channel feedback.

Return Link (RL) via either Terrestrial Network scenario: for the MSTs with an accessible terrestrial return link, each IMR performs the initial gathering of the channel status for these MSTs in the ith BC/MC group and reports it altogether to the Sat-GW. The overall reception condition for all the BC/MC members in the xth IMR cell is defined as a metric:

$$\mathbf{R}_{i,x}^{\mathrm{TN}}(n) = \left\{\mathbf{R}_{i,1}(n), \mathbf{R}_{i,2}(n), \ldots, \mathbf{R}_{i,q}(n)\right\} \tag{9}$$

where q is the total number of active BC/MC clients for the session i with RL-TN in the IMR cell. This initial reception condition gathering scheme is capable of presenting the performance for entire IMR cell in the form of a metrics including the respective performance of its MST members whilst distributing the involved computational complexity from the Sat-GW to respective IMRs.

Return Link (RL) via GEO-Sat scenario: for the MSTs with only DA signal, the RL-GS will be the only way for the Sat-GW interworking with the remote MSTs. The overall reception condition for all the BC/MC members with RL-GS is gathered at the GEO-Sat, and is defined as:

$$\mathbf{R}_{i}^{\mathrm{GS}}(n) = \left\{\mathbf{R}_{i,1}(n), \mathbf{R}_{i,2}(n), \ldots, \mathbf{R}_{i,p}(n)\right\} \tag{10}$$

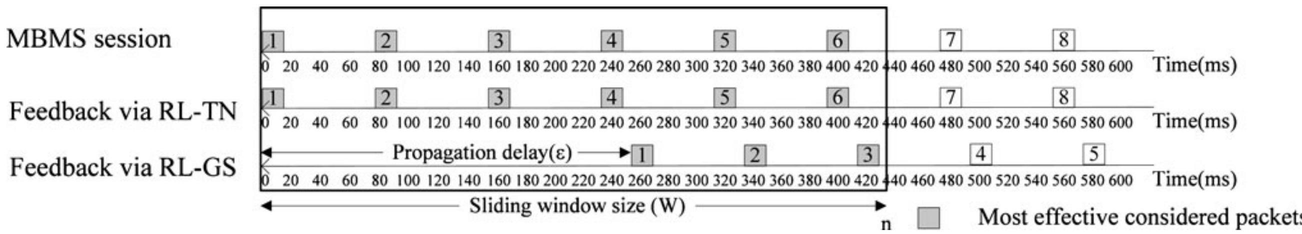


Fig. 3 Analysis on delay-differentiation reception condition capturing

where p is the total number of active BC/MC clients for session i with RL-GS. The delay difference between RL-TN and RL-GS is estimated as around 240–250 ms, indicated by ε, depending on the locations of the MSTs. Therefore, differentiation and adaptation on delayed feedback from MSTs are of significant importance on the overall reliability and scalability of the hybrid network. We illustrate the delay-differentiation issues in Fig. 3.

Upon receiving the RSTs from all the members associated BC/MC group, from both IMRs and GEO-Sat. To consider both the aforementioned scenarios, an overall reception condition for the ith session is defined as:

$$\hat{\mathbf{R}}_i(n) = \left\{ \bigcup_{x=1}^{X} \bigcup_{m=n-W}^{n} \mathbf{R}_{i,x}^{\mathrm{TN}}(m) \right\} \cup \left\{ \bigcup_{m=n-\lfloor W-\varepsilon/T_{tti} \rfloor}^{n} \mathbf{R}_i^{\mathrm{GS}}(m) \right\} \tag{11}$$

where W is the sliding window size, X is the total number of IMRs involved in the BC/MC group, $\lfloor x \rfloor$ computes the biggest integer that is smaller than x. Equation 11 considers two metrics:

- Diverse return link related metrics, i.e., from terrestrial network or from GEO-Sat. This delay adaptation is of prime importance in long-latency GEO-Sat scenario.
- Historical CSI of the most recent effective packets that best reflect the current reception condition. It can greatly enhance the robustness against highly vibrating wireless channels.

3.4 Queuing status estimation at Sat-GW

As another essential criterion in HQRA, the queuing dynamics induced in the queuing buffer at the Sat-GW is fed into the scheduler via "queuing status table" (QST), which includes the essential metrics indicating the queuing and buffer status. To perform the queuing status estimation, we define the queuing status function (QSF) $\boldsymbol{Q_i(n)}$ for the ith transport channel at the nth TTI as:

$$Q_i(n) = \alpha_i \cdot T_i(n) \cdot \Lambda_i(n) \cdot \Gamma_i(n) \cdot \Xi_i(n) \cdot H_i(n), \quad *20ci = 1,\ldots,I;\ n = 1,\ldots,N \tag{12}$$

where α_i is the application-specific traffic rank for the ith session, I is the total number of competing FACHs, N is the total number of TTI during the session transmission. For each TTI, the instantaneous queuing behaviors in the ith FACH can be characterized by a multi-dimensional vector. Each profile denotes instantaneous performance coefficient based on its targeted performance.

The first involved profile α_i, namely the *QoS profile*, is a time-independent parameter designated for each queue, reflecting the comparative traffic priority level of the service carried by the ith FACH. A higher α_i indicates a higher priority of the session. It is worth noting that the QoS profile is the premier criterion in the QSF, which means that for majority of the time, the high QoS session will be served ahead of their low QoS counterparts.

Queuing delay experienced in the RLC buffer is employed in defining the delay-related metric in the Sat-GW. We define the mean queuing delay for the ith FACH until the nth TTI as:

$$\overline{\tau_i}(n) = \frac{\sum_{k \in \Delta_i(n)} \tau_{i,k}^q(n) + \sum_{k \in \Theta_i(n)} \tau_{i,k}^q(n)}{N_i^l(n) + N_i^q(n)}, \quad \begin{array}{l} i = 1,\ldots,I \\ n = 1,\ldots,N \end{array} \tag{13}$$

where $N_i^l(n)$ is the number of packets that have left the ith queue before the nth TTI, $N_i^q(n)$ is the number of packets that are queuing in the ith FACH buffer at the nth TTI, $\Delta_i(n){:} = \{1, 2, \ldots, N_i^l(n)\}$, $\Theta_i(n){:} = \{N_i^l(n)+1, N_i^l(n)+2,\ldots,N_i^l(n)+N_i^q(n)\}$, $\tau_{i,k}^q(n)$ is the current queuing delay for the kth packet arrived in the ith FACH, which is defined as:

$$\tau_{i,k}^q(n) = \begin{cases} T_i^{\mathrm{lev}}(k) - T_i^{\mathrm{avl}}(k) & \text{if } k < N_i^l(n) \\ n \cdot T_{tti} - T_i^{\mathrm{avl}}(k) & \text{if } k > N_i^l(n) \end{cases}, \quad \begin{array}{l} i = 1,\ldots,I \\ n = 1,\ldots,N \end{array} \tag{14}$$

where $n \cdot T_{tti}$ represents current timing, $T_i^{\mathrm{avl}}(k)$and $T_i^{\mathrm{lev}}(k)$ denote the arrival time and leaving time for the kth packet in the ith queue.

A queuing delay threshold is assigned to each session based on the application-specified quality demands on queuing delay. Let τ_i^* denote the maximum acceptable queuing delay for the ith FACH specified by session's QoS requirements. We associate with the ith FACH a *queuing delay profile* $T_i(n)$, which is given by:

$$T_i(n) = \begin{cases} 1 & \text{if } \overline{\tau_i}(n) \le \tau_i^* \\ \frac{\overline{\tau_i}(n)}{\tau_i^*} & \text{if } \overline{\tau_i}(n) > \tau_i^* \end{cases}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{15}$$

Once the finite length buffer at the Sat-GW is employed, it is vital, especially for loss-sensitive service, to maintain the queue length to prevent excessive buffer overflow. Let λ_i^* denote the maximum buffer length for the ith FACH. A *buffer occupancy profile* $\Lambda_i(n)$ is defined for the ith FACH, which is given by:

$$\Lambda_i(n) = \begin{cases} 1 & \text{if } \lambda_i(n) \le \lambda_i^* \cdot \sigma_i \\ \frac{\lambda_i(n)}{\lambda_i^* \cdot \sigma_i} & \text{if } \lambda_i(n) > \lambda_i^* \cdot \sigma_i \end{cases}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{16}$$

where σ_i is the buffer occupancy threshold, providing a safe bound for the buffer length, $\lambda_i(n)$ denotes the instantaneous queue length of the ith FACH at current TTI.

The *date rate profile* is calculated as the ratio of the service required/guaranteed data rate against the mean data rate at current time. Let γ_i^* denote the guaranteed data rate for the ith FACH, the data rate profile of the ith FACH is defined as:

$$\Gamma_i(n) = \begin{cases} 1 & \text{if } \overline{\gamma_i}(n) > \gamma_i^* \\ \frac{\gamma_i^*}{\overline{\gamma_i}(n)} & \text{if } \overline{\gamma_i}(n) \le \gamma_i^* \end{cases}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{17}$$

where $\overline{\gamma_i}(n)$ denotes the mean data rate of ith FACH achieved until the nth TTI, which is determined by:

$$\overline{\gamma_i}(n) = \sum_{k=1}^{N_i^l(n)} S_{i,k} \bigg/ n \cdot T_{tti}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{18}$$

where $S_{i,k}$ represents packet size for the kth packet in the ith FACH.

The packet loss available at the Sat-GW is also confined to the buffer overflow probability (BOP). Let ξ_i^* denote the acceptable BOP for the ith FACH, δ_i is the BOP threshold for the ith FACH. The *BOP profile* $\Xi_i(n)$ is defined as:

$$\Xi_i(n) = \begin{cases} 1 & \text{if } \overline{\xi_i}(n) \le \xi_i^* \cdot \delta_i \\ \frac{\overline{\xi_i}(n)}{\xi_i^* \cdot \delta_i} & \text{if } \overline{\xi_i}(n) > \xi_i^* \cdot \delta_i \end{cases}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{19}$$

where $\overline{\xi_i}(n)$ denotes the mean BOP of the ith FACH achieved until the nth TTI, which is defined as:

$$\overline{\xi_i}(n) = \frac{N_i^d(n)}{N_i^q(n) + N_i^l(n)}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{20}$$

where $N_i^d(n)$ represents the total number of packets that are dropped due to buffer overflow for the ith FACH until the nth TTI.

We define the buffer throughput as the ratio between the total bits arrived in a queue to the total bits that have been successfully transmitted to the physical channel for radio frame transmission. Let η_i^* denote the target buffer throughput for the ith FACH, the *buffer throughput profile* $H_i(n)$ for the ith FACH is given by:

$$H_i(n) = \begin{cases} 1 & \text{if } \overline{\eta_i}(n) > \eta_i^* \cdot \varphi_i \\ \frac{\eta_i^* \cdot \varphi_i}{\overline{\eta_i}(n)} & \text{if } \overline{\eta_i}(n) \le \eta_i^* \cdot \varphi_i \end{cases}, \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{21}$$

where φ_i is the buffer throughput ratio threshold of the ith FACH, $\overline{\eta_i}(n)$ denotes the mean buffer throughput of the ith FACH that has been achieved so far, which is defined as:

$$\overline{\eta_i}(n) = B_i^s(n) / B_i^a(n), \quad \begin{array}{l} i = 1, \ldots, I \\ n = 1, \ldots, N \end{array} \tag{22}$$

where $B_i^s(n)$ is the total number of bits that are successfully transmitted for the ith FACH until current TTI, $B_i^a(n)$ represents the total number of bits that are arrived in the ith FACH so far.

3.5 Adaptive priority assessment at Sat-GW

Based on the $\mathbf{Q}_i(n)$ for each session at the Sat-GW, in conjunction with the $\hat{\mathbf{R}}_i(n)$ evaluated from all the MSTs in a BC/MC group, HQRA at the Sat-GW performs the adaptive priority assessment (APA) to derive an overall priority for each FACH.

Throughout this paper, let us assume the feedback report is perfectly generated at the MSTs and is reliably fed back to the Sat-GW, reporting the current CSI and E2E conditions. Upon receiving the feedback information from each MST member j of the intended BC/MC group in a dynamic and periodical manner, the MST subsequently derives the overall reception level $\mathbf{L}_i(n)$ for each session i associated with the entire BC/MC group as:

$$\begin{aligned} \mathbf{L}_i(n) = Prob\,\{\mathbf{R}_{i,j}(n) | \mathbf{R}_{i,j}(n) \ge \mathbf{R}_i^*(n)\}, \\ \mathbf{R}_{i,j}(n) \in \hat{\mathbf{R}}_i(n), \quad j = 1, 2, \ldots, J \end{aligned} \tag{23}$$

Equation 23 is a measure of the conditional probability of the number of members given a good reception condition over the total number of members; where J is the total number of members in the ith BC/MC group, $\mathbf{L}_i(n)$ represents the instantaneous percentage of MSTs whose reception level are above the predefined thresholds $\mathbf{R}_i^*(n)$ within the specific BC/MC group. The scheduler then estimates the overall priority of each session as:

$$\mathbf{P}_i(n) = \mathbf{L}_i(n) \cdot \mathbf{Q}_i(n) \tag{24}$$

The queues with a higher $\mathbf{P}_i(n)$ will be scheduled ahead of theirs lower priority counterparts. In this way, HQRA performs distributed and hybrid considerations on retrieving the CSI and ETE metrics from users, whilst keeps tracking the queuing dynamics involved in the Sat-GW.

3.6 Flexibility and scalability analysis

By employing the RLA and APA schemes, HQRA provides essential adaptations to both satellite link and terrestrial link, tolerating a wide range of network and link variations. In the above context, we assume all the contributing profiles behave and influence the priority assessment in an equal way during the session transmission. However, fixed settings upon all performance criteria may not work well in provisioning multimedia data with different QoS demands and fast-varying traffic dynamics. The performance gain achieved in one profile may sacrifice the performance on other profiles, which may be even more critical for the specific service. To offer more flexibility and enhance the bandwidth utilization, HQRA provides a tuning mechanism over different performance dimensions to further optimize the scheduling efficiency. By observing the QoS preferences specified by the service and the behaviors of queuing dynamics, HQRA dynamically adjusts the following "tuning knobs" on a TTI-scale: (1) reception condition thresholds ($SS_{i,j}^{\min}$, $\gamma_{i,j}^{\min}$, $\mathbf{R}_i^*(n)$), (2) ETE QoS targets (d_i^*, ρ_i^*, θ_i^*), and (3) queuing state thresholds (τ_i^*, λ_i^*, γ_i^*, ξ_i^*, η_i^*). By selecting an appropriate combination of the above threshold parameters, the serving orders of competing flows can be effectively managed. According to the sensitivity preferences from the service-oriented QoS targets, through giving flexible importance on delay, loss and throughput, it is therefore possible to adaptively select the best possible scheduling policy to allow for different treatments of diverse quality demands and maintain optimal resource utilization.

From the viewpoint of implementation, HQRA introduces extra computation overhead due to its nonlinear nature. However, as the reception condition collections are implemented at different parts of the network, compared with a centralized scheduler that operates in the Sat-GW, the complexity is distributed and the overall reliability of the scheduling is increased. The *Big O notation* [11] is employed for determining the involved computational complexity for the HQRA algorithm. It is assumed that there are n sessions to be transmitted to MSTs in a number of multicast groups, located within multiple sectors of a satellite beam. Derived from the worst case scenario, where the processing time is the most expensive among all possible scenarios, with the input size of n (i.e., total number of FACHs), the involved computational time complexity required for MLPQ and DDQ are derived as $O(n)$ and $O(n^2)$, respectively, whilst HQRA requires an overall computational complexity of $O(n^2)$, featuring typical quadratic statistics.

4 Evaluation methodology

To evaluate the performance of HQRA, we have built a system-level SDMB simulation model using the *ns2* network simulator [12] and the H.264/MPEG-4 AVC JM reference software [13], where different traffic scenarios and physical channel capacities are involved. Three types of service, namely streaming, hot download and cold download, as specified in SDMB, are considered in the simulation. These services correspond to UMTS QoS classes streaming and background, respectively [14]. The performance of the proposed scheme is examined over different scenarios: RL-TN, RL-GS, and without RL. The radio bearer mapping configuration deployed for the transport/physical channel multiplexing is shown in Table 2. Interested reader may refer to [10] for more channel multiplexing solutions in SDMB.

We built the discrete event simulator, where the video streaming trace is generated as inter-packet arrival time from the JM reference model. In our simulation, we apply *Pareto* distribution for modelling traffic arrival pattern of Internet download service, as Internet download is found to be heavy-tailed and is approximately Pareto-distributed. Hot download and cold download traffic follow the classical *Pareto* distribution (shape factor = 1.5), with assigned different QoS ranks. Packets are buffered at the Sat-GW, the head-of-line packet in each queue is scheduled for transmission over air according to scheduling policies. For simplicity, we consider a single IMR scenario, where the sliding window size W is set to 10, ε equals to 250 ms. The satellite channel are characterized by a classical Finite-State Markov channel (FSMC) [15], which is proven to be a good approximate for the received signal envelop in a typical multipath propagation environment. We assume the receivers have perfect knowledge of the channel status and feedbacks to the transmitter without errors. Thus, the transmission errors are confined to the forward link only. In our simulation, we consider a TTI of 80 ms, different TTI can be set with various performance features. Simulation run period is set to 20,000 s for each scenario, where 250,000 TTIs are elapsed. Our link budget simulation results [10] provide the E_b/N_o *v.s BLER* look-up curves of

Table 2 Radio bearer mapping configuration (KBPS)

S-CCPCH	1	2	3
Bit rate	384	384	384
Streaming FACH	256 × 1; 64 × 1	256 × 1; 128 × 1	–
Hot download FACH	64 × 1	–	–
Cold download FACH	–	–	384 × 1

Table 3 System simulation parameter

Simulation parameter	Value
Frequency of operation (GHz)	2.5
Chip rate (Mchip/s)	3.84
Spreading factor	8
TTI (s)	0.08
Modulation	QPSK
Coding	Turbo code
Code rate	1/3
Maximum bit rate (kbit/s)	384
Packet size (bytes)	1,280
Terrestrial channel model	Rayleigh
Satellite channel model	Ricean

each FACH. The physical layer settings for our test bed are given in Table 3.

5 Simulation results

5.1 Round trip delay

The mean round trip delay (RTD) statistics for all the MSTs in a BC/MC group are measured at the Sat-GW. It is considered as the sum of the propagation delay experienced by the GEO satellite link, the processing and queuing delay, and the transmission delay over the mobile return link. As illustrated in Fig. 4, the cumulative distribution function (CDF) of RTD demonstrates that, compared with MLPQ and DDQ, the service under HQRA scheme experiences much lower RTD and enjoys better fairness than existing schemes in that, it follows the steepest convergence with smoothest slope amongst all the schemes. Therefore, it can be inferred that remarkable gain on the end-to-end delay and delay variation (jitter) is achievable. It is worth noticing that the streaming service in SDMB is quite sensitive to delay variation, thereby the jitter of the traffic flow should be limited in order to preserve the time variation between packets in the stream [16]. Numerically, hot and cold download classes have a reduction of 43.1 and 56.6% on their mean RTD, respectively. The average jitter reduction for download services is as much as 65.8.0%, while the average jitter reduction for the streaming service is 32.9%. Therefore, the proposed HQRA packet scheduling scheme achieves better queuing jitter than existing schemes. It is noted that, by using HQRA, significant reduction on delay of the lower class service (i.e., download service) has been achieved.

5.2 Fairness and channel utilization

The parameter of interest representing the fairness of a packet scheduling algorithm is the throughput ratio, which is obtained by dividing the total bits successfully received by the MSTs with the total bits released from the service provider; lower variance indicates a fairer scheduling scheme. As shown in Fig. 5, different schemes are investigated during a sample simulation period (i.e., >250,000 TTIs), where the HQRA achieves the lowest variance value with the fastest convergence curve and lowest max-min variations, which mean that it can provide MSTs with better throughput equality in a short period of time.

The mean physical channel utilization for different scheduling schemes is shown in Fig. 6. The performance indicates that HQRA is capable of increasing the overall spectrum efficiency on the S-CCPCH channel. Numerically, it reaches 98.5, 97.4, and 85.8% of the total capacities, respectively. Its nearest candidate, DDQ, reaches 93.7, 95.6, and 85.7% channel utilization, respectively. It is noted that the performance gain achievable on S-CCPCH 3

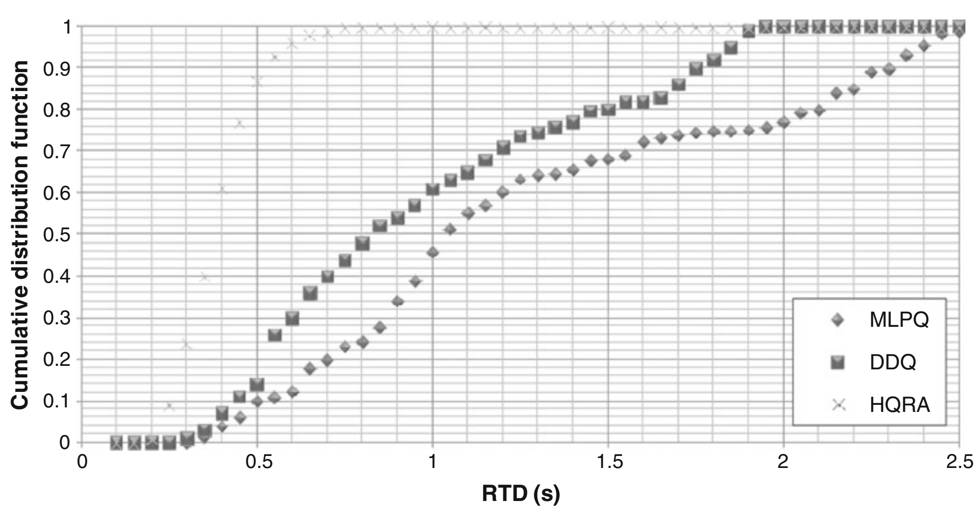


Fig. 4 CDF of RTD for streaming under different scheme

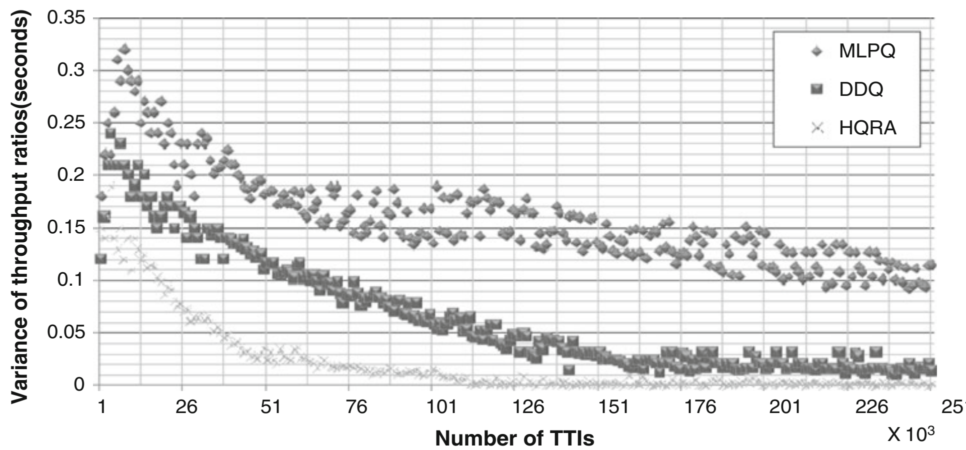


Fig. 5 PDF of S-CCPCH throughput ratios

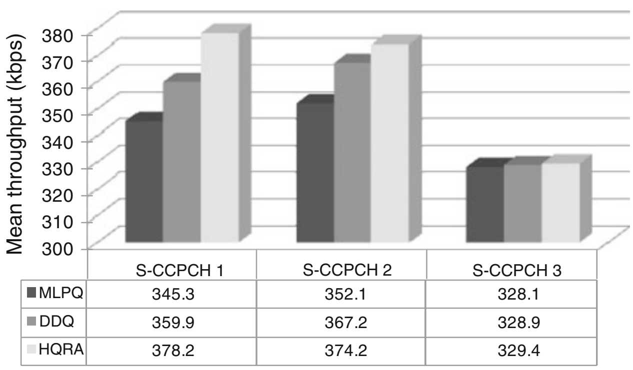


Fig. 6 Mean physical channel utilization

are largely limited since there is a maximum power limit on each satellite spot beam, and the power is allocated from S-CCPCH 1 to S-CCPCH 3. From the results, it can be inferred that the proposed HQRA scheme has essential impact on increasing the throughput and physical channel utilization.

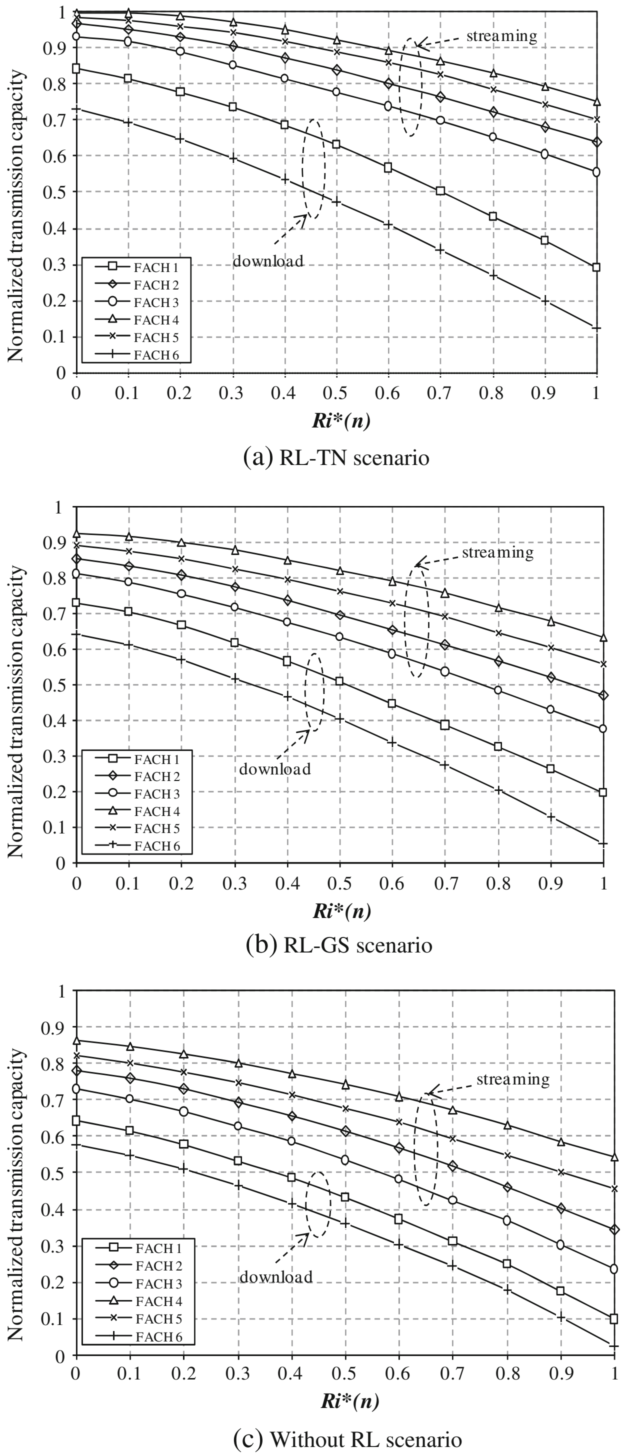
5.3 Analysis of return-link diversity

In this section, we study the performance variations of HQRA under diverse return links. From our discussions, we found that perfect feedback on CSI metrics can be presumed if RL-TN is in presence, while the RL-GS is characterized by imperfect feedbacks with large delay-loss products. We define the normalized transmission capacity as the ratio between instantaneous data rate and the maximum supportable data rate of the corresponding session. Figure 7 analyzes the average normalized transmission capacity achievable with different reception condition threshold $\mathbf{R}_i^*(n)$. It is shown that the transmission capacity declines with more stringent reception condition threshold, while it achieves the best performance for all the FACHs when the MSTs are accessible to RL-TN. It is therefore proves that, effective utilization on important feedback information, even if it is imperfect in RL-GS, can improve the transmission efficiency and thereby optimizing the overall system performance.

5.4 PSNR performance

Table 4 summarizes the overall performances of MLPQ, DDQ, and HQRA for streaming the video sequence *Coastguard* over four FACHs in terms of the number of lost packets, the number of affected video frames, and the average peak-signal-to-noise-ratio (PSNR) of all FACHs. The BLER of 1 and 2% are tested, respectively. The video sequences are encoded with group of picture (GOP) of 30 frames, including one I-frame and nine P-frames. There are two B-frames between neighboring P frames. No error control coding is used in the video codec. The encoded bitstream is concatenated, packetized and converted to a trace file, which serves as a traffic generator during the simulation. The trace file repeats itself to simulate the transmission of 10,000 frames for each sequence in each FACH.

At the video decoder, a frame is affected whenever a part of its data is in a lost packet. A lost B-frame will affect only one frame, whilst a lost P-frame or I-frame will affect all subsequent frames in the GOP. In our simulation, a simple frame repetition scheme is used to reconstruct all affected frames, where the last perfectly decoded frame repeats itself until the next perfectly decoded frame is available. Table 4 shows that, compared with MLPQ and DDQ, the HQRA reduces the overall number of lost packets and affected frames by 17–35%. It is also proved that the PSNR improvement is greater as the BLER increases to 2%, where more than 0.2 dB gain is achievable, compared with <0.1 dB in BLER = 1%. Note that although the average PSNR improvement across all FACHs is limited due to the large number of frames, the significant reduction on the affected number of frames will greatly enhance the visual quality of the received video.



(a) RL-TN scenario

(b) RL-GS scenario

(c) Without RL scenario

Fig. 7 Normalized transmission capacity under different return-link scenarios

6 Conclusions

In this paper, we proposed a novel cross-layer quality-driven packet scheduling scheme, namely HQRA, for bidirectional satellite multimedia networks. By jointly optimizing the cross-layer problem from MAC layer scheduling, application layer quality targets and physical layer channel and data rate information, the proposed scheme is capable of improving the performance on service QoS guarantees and transmission efficiency. Performance evaluation of the HQRA scheme has been carried out via simulation studies. The results show that, compared with the existing packet scheduling schemes adopted in unidirectional satellite broadcasting system, by considering various aspects of multimedia QoS provisioning, our scheme not only improves the network performance on delay, jitter, channel utilization, but also optimizes the application-oriented QoS, fairness and video quality.

Table 4 Performances comparison for streaming the video sequence coastguard under different scheduling schemes

BLER	Scheme	No. of lost packets	No. of affected frames	Y-PSNR (dB)
1%	MLPQ	38 (0.51%)	757 (1.87%)	33.35
	DDQ	34 (0.46%)	754 (1.86%)	33.36
	HQRA	26 (0.35%)	567 (1.40%)	33.43
2%	MLPQ	92 (1.24%)	1,750 (4.32%)	33.02
	DDQ	73 (0.99%)	1,742 (4.30%)	33.00
	HQRA	61 (0.81%)	1,142 (2.82%)	33.24

Acknowledgments The authors would like to thank Dr. Linghang Fan and Dr. Atta Quddus from University of Surrey for their valuable discussions and providing physical layer link budget data for the simulation part of this paper. J. Liu's work was supported by a Canada NSERC Discovery Grant and an NSERC Strategic Project Grant.

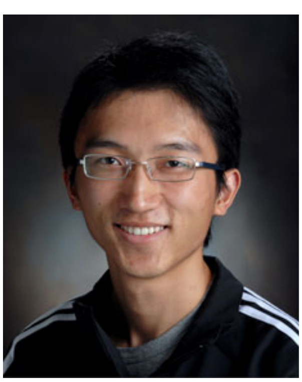
References

1. Chuberre, N., et al. (2004). Satellite digital multimedia broadcasting for 3G and beyond 3G systems, *13th IST mobile & wireless communication summit 2004*, Lyon, France.
2. 3GPP TS 22.146 V6.6.0, Multimedia broadcast/multicast service: Stage 1, October 2004.
3. Dovrolis, C., et al. (2002). Proportional differentiated services: Delay differentiation and packet scheduling. *IEEE Transactions Networking, 10*(1), 12–26.
4. Huang, V., & Zhuang, W. (2004). QoS-oriented packet scheduling for wireless CDMA network. *IEEE Transactions Mobile Computing, 3*(1), 73–85.
5. Aniba, G., & Aissa, S. (2004). Adaptive proportional fairness for packet scheduling in HSDPA. *IEEE Globecom, 6*(29), 4033–4037.
6. Karaliopoulos, M., et al. (2004). Packet scheduling for the delivery of multicast and broadcast services over S-UMTS. *International Journal of Satellite Communications and Networking, 22*, 503–532.
7. Fan, L., Du, H., Mudugamuwa, U., Evans, B. G. (2008). A cross-layer delay differentiation packet scheduling scheme for multimedia content delivery in SDMB system. *IEEE Transaction on Broadcasting, 54*(4), 806–815.
8. Du, H., et al. (2007). A cross-layer packet scheduling scheme for multimedia broadcasting via satellite digital multimedia broadcasting system. *IEEE Communications Magazine, 45*(8), 94–101.
9. 3GPP, TS 25.301 v8.2.0, Radio interface protocol architecture, June 2008.

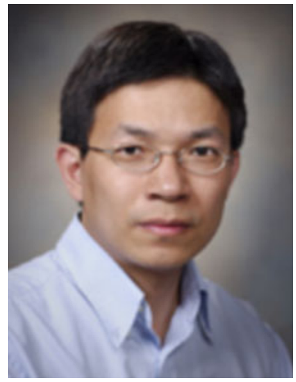
10. 3GPP, TS 25.214 v8.2.0, Physical layer procedures (FDD), May 2008.
11. Homer, S., & Selman, A. (2000). *Computability and complexity theory*. Berlin: Springer.
12. http://www.isi.edu/nsnam.
13. http://iphome.hhi.de/suehring/tml/.
14. Holma, H., & Toskala, A. (2002). *WCDMA for UMTS: Radio access for third generation mobile communications* (2nd ed.). New York: Willey.
15. Zhang, Q., & Kassam, S. A. (1999). Finite-state markov model for reyleigh fading channels. *IEEE Transactions Communications, 47*(11), 1688–1692.
16. 3GPP TS 23.107 V7.1.0, Quality of service (QoS) concept and architecture, September 2007.

Author Biographies

Hongfei Du (S'05-M'07) received the B.E. degree in Electronic Information Engineering from the Department of Electronic Engineering, Beijing University of Aeronautics & Astronautics, Beijing, China, in 2003. He received the M.Sc., M.Phil, and Ph.D. degrees in Wireless Communications from University of Surrey, United Kingdom, in 2004, 2005, and 2007, respectively. From 2007 to 2008, he was with *CREATE-NET International Research Institute*, Italy, as a member of research staff then project leader, coordinating and conducting EU research projects on middleware/software implementation, system architecture and protocol design for the convergence between heterogeneous broadcast and mobile networks. From 2008, he is with School of Computing Science & School of Engineering Science, Simon Fraser University, as a postdoctoral researcher and Ebco-Epic Fellow, working on adaptive video transmission over mobile WiMAX networks. Hongfei has been involved in extensive research projects in the area of mobile broadcasting convergence, mobile communications and satellite communications systems and has also served as a TPC and reviewer for many leading journals and international conferences/workshops including IEEE Wireless Communication, IEEE Transaction on Vehicular Technology, ICC, Globecom, etc. His research interests lie in the area of mobile and satellite multimedia broadcasting, focusing on radio resource management, packet scheduling, quality-of-service support, scalable video coding and cross-layer design.



Xiaozheng Huang received his B.E. degree in Telecommunication Engineering from Department of Information Science and Electronic Engineering, Zhejiang University, China in 2007. He is currently working towards the M.A.Sc. degree at the school of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, pursuing research in the field of image and video coding and joint source channel coding.



Jie Liang (S'99-M'04) received the B.E. and M.E. degrees from Xi'an Jiaotong University, China, in 1992 and 1995, the M.E. degree from National University of Singapore (NUS), in 1998, and the Ph.D. degree from the Johns Hopkins University, Baltimore, MD, in 2003, respectively. Since May 2004, he has been an Assistant Professor at the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. From 2003 to 2004, he was with the Video Codec Group of Microsoft Digital Media Division, Redmond, WA. Dr. Liang's research interests include image and video coding, multirate signal processing, and joint source channel-coding.

Jiangchuan Liu (S'01-M'03-SM'08) received the B.E. degree (cum laude) from Tsinghua University, Beijing, China, in 1999, and the Ph.D. degree from The Hong Kong University of Science and Technology in 2003, both in computer science. He was a recipient of Microsoft Research Fellowship (2000), a recipient of Hong Kong Young Scientist Award (2003), and a co-inventor of one European patent and two US patents. He co-authored the Best Student Paper of IWQoS'08 and the Best Paper (2009) of IEEE Multimedia Communications Technical Committee (MMTC). He is currently an Assistant Professor in the School of Computing Science, Simon Fraser University, British Columbia, Canada, and was an Assistant Professor in the Department of Computer Science and Engineering at The Chinese University of Hong Kong from 2003 to 2004. His research interests include multimedia systems and networks, wireless ad hoc and sensor networks, and peer-to-peer and overlay networks. He is an Associate Editor of IEEE Transactions on Multimedia, and an editor of IEEE Communications Surveys and Tutorials. He is a Senior Member of IEEE and a member of Sigma Xi.

Barry G. Evans received the B.Sc. and Ph.D. degrees in Electrical Engineering and microwave systems from the University of Leeds in 1965 and 1968, respectively. He is Director of the Centre for Communications Systems Research at the University of Surrey in the United Kingdom where he is a Professor and also Pro-Vice Chancellor for Research and Enterprise. He is editor of the International

Journal of Satellite Communications and Networking and a well known International consultant having researched for over thirty years in the field. He is the author of over 400 publications in the technical literature and of several books including 'Satellite Communication systems' IEE press. Outside of the University of Surrey, Professor Evans has been involved in several National and International committees including the UK Foresight programmes in Communications and ITEC; EPSRC Strategic Advisory Committees; MoD-DSAC Committees; Advisor to DG of OFTEL; Board Member of BNSC-TNAB as well as ITU, ETSI and EU Advisory Committees. He is currently a member of the Ofcom Spectrum Advisory Board and is Director of a small spin off company, Mulsys Ltd. He has recently been appointed to the Steering Council of the European Technology Platform—Integral Satcom Initiative and to the HEFCE Strategic Advisory Committee for Business and Community. He was elected a Fellow of the Royal Academy of Engineering in 1991.

Imrich Chlamtac is President of CreateNet, Distinguished Chair Professor at the University of Texas at Dallas, Bruno Kessler Honorary Professor at the University of Trento, the Sackler Professorship from Tel Aviv University, and the University Professor at the Budapest University of Technology and Economics. Dr. Chlamtac is known as the inventor of the fundamental lighpath and multihop ad-hoc networking concepts. He is a Fellow of the IEEE and of the ACM, recipient of the 2001 ACM Award for Outstanding Contributions to Research on Mobility and the 2002 IEEE Award for Outstanding Technical Contributions to Wireless Personal Communications. He published over three hundred and fifty refereed articles and four books, including the first textbook on LAN-s entitled "Local Networks" (1980) and "Wireless and Mobile Network Architectures", John Wiley & Sons (2000), an IEEE Network Editor's choice. He is the founder and past Chair of ACM Sigmobile, and of various leading conferences including Mobicom, Mobiquitous, Broadnets, and Wiopt. Dr. Chlamtac is the founding Editor in Chief of the ACM/URSI/Kluwer Wireless Networks (WINET) and of the ACM/Kluwer Journal on Special Topics in Mobile Networks and Applications (MONET).