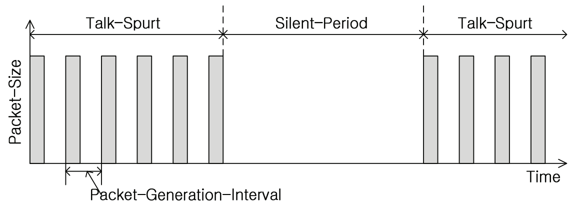
10

VoIP Features Oriented Uplink Scheduling Scheme in Wireless Networks

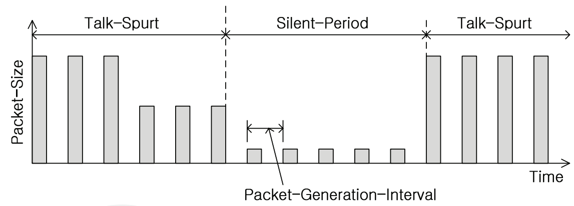
Sung-Min Oh and Jae-Hyun Kim
School of Electrical and Computer Engineering, Ajou University
Republic of Korea

1. Introduction

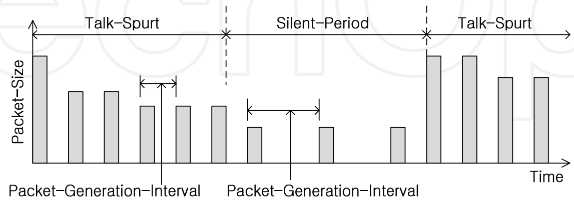
VoIP services have been considered as one of the most important services in the next generation wireless systems. VoIP service requires the same quality of service (QoS) requirement as constant bit rate services. For this reason, the IEEE 802.16 standard has defined an unsolicited grant service (UGS) to guarantee the QoS. However, the UGS is inadequate to support VoIP services with silence suppression because of the waste of radio bandwidth in the silent-periods. In the UGS, a base station (BS) periodically allocates a maximum-size radio bandwidth (grant) during the silent-periods even though a subscriber station (SS) does not have a packet to transmit in the silent-periods. To solve this problem, (Lee et al., 2005) proposed an extended real time polling service (ertPS) to support VoIP services with silence suppression. The ertPS can manage the grant-size according to the voice activity in order to save the radio bandwidth in silent-period. Unfortunately, the waste of radio bandwidth and the increase of access delay can still exist when the ertPS is applied to the system because the grant-size and grant-interval used by the ertPS cannot correspond with the packet-size and the packet-generation-interval of the VoIP services in the application layer.
Recently, the IEEE 802.16's Task Group m (TGm), which was approved by IEEE to develop an amendment to IEEE 802.16 standard in 2006, published the draft evaluation methodology document in which several kinds of VoIP speech codecs are considered such as G.711, G.723.1, G.729, enhanced variable rate codec (EVRC), and adaptive multi-rate (AMR) (Srinivasan, 2007). These VoIP speech codecs generate packets with different packet-size and packet-generation-interval as shown in Table 1. However, the IEEE 802.16 standard does not define the QoS parameter generation method, because they have focused on only medium access control (MAC) and physical (PHY) layer. For this reason, IEEE 802.16 based systems need the QoS parameter mapping algorithm to obtain the features of the VoIP services in the application layer. Hong and Kwon (Hong & Kwon, 2006) proposed the QoS parameter mapping algorithm to exploit the feature of the VoIP services in IEEE 802.16 systems which statistically measures the peak data rate of VoIP services and calculates the QoS parameters. However, the algorithm needs significant time to measure the VoIP traffic to perform the statistical analysis, and the QoS parameters cannot correspond to the features of the VoIP services when the number of samples of the VoIP traffic is not sufficient to analyze the features of the VoIP service. To overcome these problems, this chapter designs a cross-layer QoS parameter mapping scheme which exploits the information of the VoIP speech codec included in the session description protocol (SDP) to generate the QoS parameters for VoIP scheduling algorithms.



(a) G.7xx with silence suppression



(b) EVRC



(c) AMR

Fig. 1. Traffic models for various VoIP speech codecs

Moreover, this chapter proposes a new cross-layer VoIP scheduling algorithm which exploits the QoS parameters generated by the proposed QoS parameter mapping scheme. The conventional VoIP scheduling algorithms have been designed considering a specific VoIP speech codec. The UGS has been designed to guarantee a QoS for G.7xx (i.e. G.711, G.723.1, and G.729) without silence suppression, and the ertPS has been developed to support EVRC. In particular, the ertPS is designed to compensate for the resource inefficiency of the UGS in the silent-periods. Unfortunately, the ertPS is not an optimal VoIP scheduling algorithm for the whole VoIP speech codecs. In the ertPS, a BS periodically allocates a minimum-size grant to a SS every 20 msec regardless of the voice activity in the silent-period. However, the AMR speech codec generates a packet every 160 msec in the silent-period. Thus, the ertPS can cause the waste of radio bandwidth in the silent-period when it supports the AMR speech codec. To overcome this inefficiency of the ertPS, Oh et al (Oh et al., 2008) proposed a new VoIP scheduling algorithm, which is called as a hybrid VoIP (HV) algorithm in this chapter. The HV algorithm adapts a random access scheme in the silent-period to save radio bandwidth. However, it can suffer from an overhead occurred in the silent-period when the EVRC is applied in the application layer. The problems of VoIP scheduling algorithms according to the VoIP speech codecs are detailed in section 3. Consequently, this chapter proposes the cross-layer VoIP scheduling algorithm to support all available VoIP speech codecs. The main feature of the cross-layer VoIP scheduling algorithm is that it can dynamically adjust the grant-interval in the silent-period according to the VoIP speech codec applied in the application layer. By this feature, the proposed scheduling algorithm can save radio bandwidth guaranteeing a QoS for all VoIP speech codecs in the silent-period. The description of the proposed scheduling algorithm is presented in section 4.

2. Traffic models for various VoIP speech codecs

This section describes traffic models for various VoIP speech codecs which are presented in Fig. 1 where each VoIP speech codec has individual features in their packet generation policy (ITU-T-G711, 2000; ITU-T-G7231, 1996; ITU-T-G729, 2007; 3GPP2-EVRC, 2004; 3GPP-TS-26201, 2001; 3GPP-TS-26092, 2002; 3GPP-TS-26071, 1999). Fig. 1 (a) represents a traffic model for

VoIP Speech Codec	PS (bytes)	PGI (msec)
G.711	160	20
G.723.1	19.88	30
G.729	10	10
EVRC	21.375, 10, 2	20
AMR	Voice frame: 11.875, 12.875, 14.75, 16.75, 18.5, 19.875, 25.5, 30.5 SID frame: 5	Talk-spurt: 20 Silent-period: 160

Table 1. Features of VoIP Speech Codecs (PS: Packet-Size, PGI: Packet-Generation-Interval)

G.7xx with silence suppression. In the silent-period, this model does not generate any packets so as to save radio bandwidth but a receiver side can suffer from deterioration in the QoS performance in these situations when the background noise at the transmitter side is high. The reason for this is that the source controlled rate (SCR) switching in a VoIP speech codec of the receiver side can take place rapidly so that the EVRC and AMR speech codecs periodically send packets which include the information of the background noise at the transmitter side every grant-interval in the silent-period. However, these speech codecs have different grant-interval; namely the EVRC generates the packets every 20 msec, whereas the AMR speech codec generates silence indicator (SID) frames every 160 msec in the silent-periods, as depicted in Fig. 1 (b) and (c).

In talk-spurts, the G.7xx generates fixed-size packets, whereas the EVRC and AMR speech codecs generate variable-size packets according to the wireless channel or the network condition. The packet-size is as specified in Table 1. The EVRC generates packets according to three data rate which are full rate (21.375 bytes), half rate (10 bytes), and eighth rate (2 bytes), where the eighth rate is for the background noise. The AMR speech codec generates variable-size packets every 20 msec in the talk-spurts and Table 1 represents the variable packet-sizes for the AMC speech codec.

IEEE 802.16e/m systems can suffer from several problems in supporting these various features of the VoIP speech codecs. These problems are detailed in the following section.

3. Challenges for VoIP services in IEEE 802.16e/m systems

IEEE 802.16 defined UGS and ertPS to support VoIP services with a QoS guarantee. However, the conventional VoIP scheduling algorithms can suffer from the waste of radio bandwidth and the increase of access delay. These problems can be caused by two challenges in the IEEE 802.16e/m systems such as the absence of a QoS parameter mapping scheme and the resource inefficiency of the conventional VoIP scheduling algorithms.

3.1 Absence of the QoS parameter mapping scheme

The convergence sublayer (CS) defined in (Handley & Jacobson, 1998) connects the MAC layer with the IP layer. When a session is generated in the application layer, a connection identifier (CID) is created in the CS. At this time, QoS parameters are needed to guarantee the QoS of the session. However, the IEEE 802.16 standard does not define a QoS parameter generation method and hence mismatches between QoS parameters in the MAC layer and the features of a session in the application layer can occur. Such mismatch problems can cause the waste of

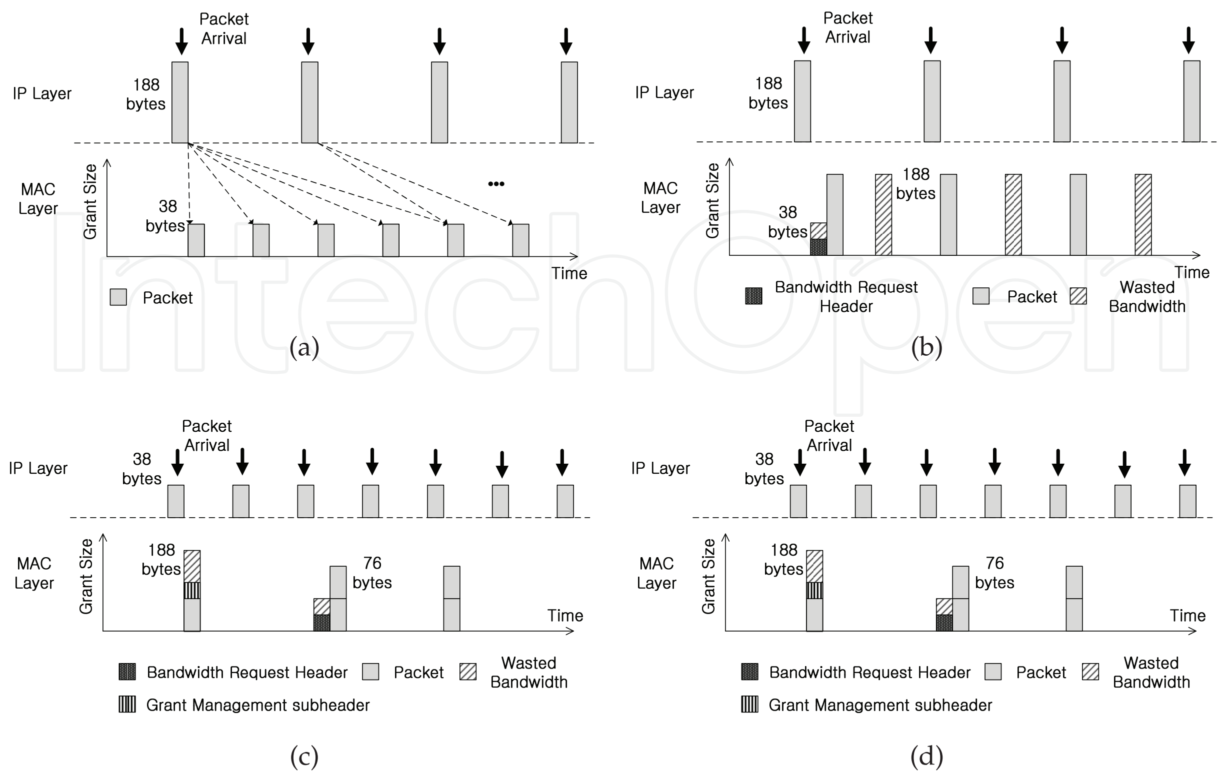


Fig. 2. Examples of the mismatch problem between the QoS parameters in the MAC layer and the features of VoIP services in the application layer; default QoS parameters (grant-size: 38 bytes and grant-interval: 10 msec), VoIP speech codec (G. 711 without silence suppression), and VoIP scheduling algorithm ((a) UGS and (b) ertPS), {default QoS parameters (grant-size: 188 bytes and grant-interval: 20 msec), VoIP speech codec (G. 729 without silence suppression), and VoIP scheduling algorithm ((c) UGS and (d) ertPS)}

radio bandwidth or the increase of access delay. To describe the mismatch problems in detail, this chapter gives examples as shown in Fig. 2.

Figs. 2 (a) and (b) represent the mismatch problems. In this case, the default values of the QoS parameters are set by considering the G.729 . In addition, VoIP scheduling algorithms, as shown in Fig. 2 (a) and (b), are UGS and ertPS, respectively. As depicted in Fig. 2 (a), the access delay increases by 40 msec to transmit a packet due to the mismatch problem. A BS periodically allocates a fixed-size grant (38 bytes) every 10 msec even though a SS needs additional bandwidth to transmit a packet, because the UGS cannot request any additional bandwidth. Due to this problem, the access delay can increase linearly when the system continuously receives data packets from the upper layer. This anomalistic phenomenon can cause serious deterioration of the QoS performance for VoIP services. Unlike UGS, the ertPS can prevent the increase of access delay, as shown in Fig. 2 (b). The reason is that the ertPS can request additional bandwidth by the bandwidth-request-header. However, the radio bandwidth (188 bytes) can be wasted every 20 msec; because a BS periodically allocates a grant (188 bytes) every 10 msec even though data packets are generated every 20 msec.

Figs. 2 (c) and (d) also represent the mismatch problem. In this case, the default values of the QoS parameters are set by considering the G.711. As depicted in Fig. 2 (c), a packet can experience an access delay of 10 msec every 20 msec. In addition, 112 bytes of bandwidth is wasted every 20 msec when UGS is applied to the system. The ertPS can save the waste of radio bandwidth as shown in Fig. 2 (d). However, the access delay still exists because of the mismatch between the grant-interval and the packet-generation-interval.

As mentioned above, the mismatch problem can cause the waste of radio bandwidth or the increase of access delay. To solve the mismatch problem, this chapter proposes a new cross-layer QoS mapping scheme, which will be described in section 4.

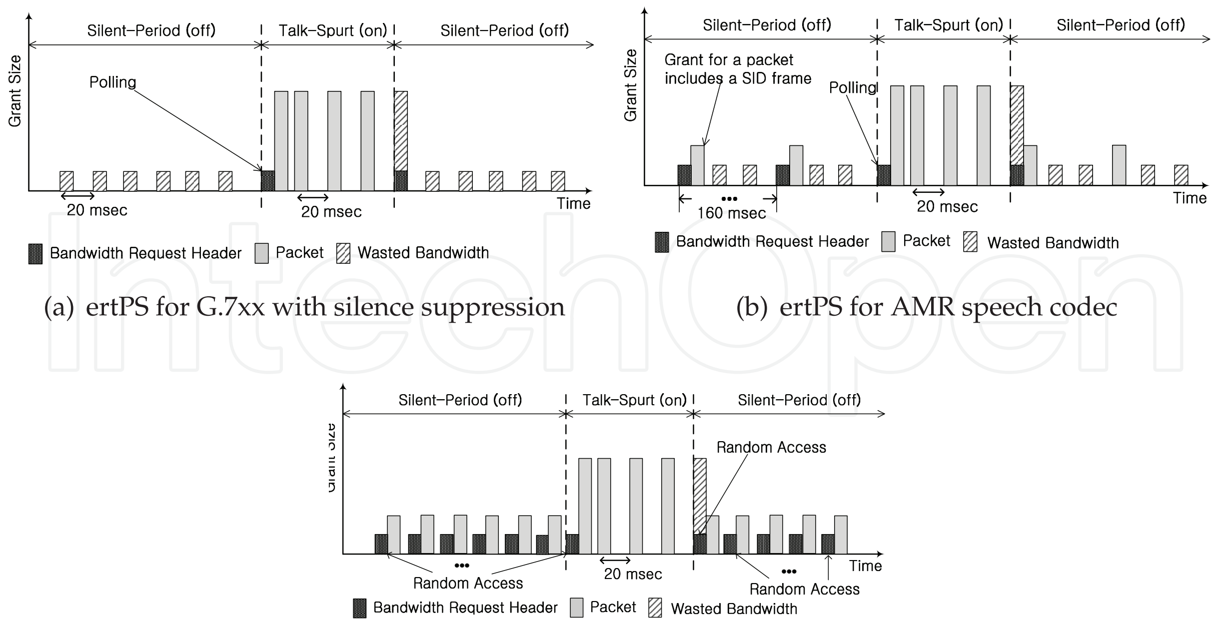
3.2 Resource inefficiency of the conventional VoIP scheduling algorithms

The UGS and ertPS methods are inefficient in their use of the wireless resource. In UGS, a BS periodically allocates the maximum-size grant to a SS regardless of the voice activity even though the data rate of the VoIP services with silence suppression decreases in the silent-periods. Because of this resource inefficiency of the UGS, the ertPS has been designed to support VoIP services with silence suppression. The ertPS can manage the grant-size according to the voice activity. In order to change this, the ertPS has two main features. Firstly, it exploits a generic-MAC-header to inform a BS of the SS's voice activity. Lee et al (Lee et al., 2005) defined a Grant-Me (GM) bit using a reserved bit in the generic-MAC-header. When in a silent-period the voice activity indicated by the GM bit is '0' whereas in a talk-spurt, the GM bit is '1'. Secondly, a BS periodically allocates a grant to transmit a generic-MAC-header in the silent-period. By using this feature, a SS can transmit a generic-MAC-header even though there is no packet to transmit in the silent-period.

On the other hand, the grant for a generic-MAC-header is wasted during the silent-period from considering the wireless resource aspects. As shown in Fig. 3 (a), a grant is wasted every 20 msec when the G.7xx situation with silence suppression is applied to the system. When the AMR speech codec is applied to the system, seven grants are wasted every 160 msec during the silent-period, as shown in Fig. 3 (b).

To overcome this inefficiency of the ertPS, (Oh et al., 2008) proposed a HV algorithm with three main features. Firstly, a BS does not periodically allocate a grant to a SS in the silent-period in order to save the uplink bandwidth. Secondly, the HV adopts the random access scheme to transmit a packet in the silent-period. Thirdly, it also uses the random access scheme when the voice activity changes from a silent-period to a talk-spurt, because the transition time from one to the other is unpredictable. The HV exploits a bandwidth-request-and-uplink-sleep-control (BRUSC) header in order to inform a BS of the SS's voice activity and request the required bandwidth. The BRUSC header has a reserved bit which is defined as a silence talkspurt (ST) bit in (Oh et al., 2008), and this has a bandwidth request (BR) field which can be specified as a required bandwidth in bytes. In the HV method, the SS transmit a BRUSC header by using the random access scheme when a packet to transmit is generated in a silent-period, or when the voice activity changes from being in a silent-period to a talk-spurt. At this time, the grant-size is the same with the bandwidth required by the BRUSC header.

Unfortunately, the HV algorithm can suffer from collisions when the EVRC is applied to the system. In case of the AMR speech codec and G.7xx with silence suppression, the collision cannot affect the QoS performance for the VoIP services, because the transmission rate of a BRUSC header is very low. However, a SS transmits a BRUSC header every 20 msec during a silent-period by the random access scheme when the EVRC is applied to the system as shown in Fig. 3 (c). For this reason, the message overhead required to transmit a packet rapidly increases because the transmission rate of a BRUSC header increases. For this problem, the HV algorithm may be inadequate for EVRC. Consequently, this chapter proposes the cross-layer VoIP scheduling algorithm to support the whole VoIP speech codecs with efficient use of radio



(a) ertPS for G.7xx with silence suppression

(b) ertPS for AMR speech codec

(c) HV algorithm for EVRC

Fig. 3. Resource inefficiency of the conventional VoIP scheduling algorithms

bandwidth.

4. Proposed cross-layer framework for VoIP services

In order to overcome the challenges of the VoIP services in IEEE 802.16e/m systems mentioned in section 3, we design the cross-layer framework for VoIP services which is shown in Fig. 4. It consists of the cross-layer QoS parameter mapping scheme and the new cross-layer VoIP scheduling algorithm. The description of the cross-layer QoS parameter mapping scheme and the cross-layer VoIP scheduling algorithm are as follows.

4.1 Cross-layer framework for VoIP services

We propose the cross-layer QoS parameter mapping scheme to compensate for the absence of the QoS parameter mapping scheme in IEEE 802.16e/m systems. The cross-layer QoS parameter mapping scheme consists of three functions such as the QoS parameter creation function, CID creation function, and CID mapping function as shown in Fig. 4.

4.1.1 QoS parameter creation function

The QoS parameter creation function is the main function in the cross-layer QoS parameter mapping scheme. It generates the QoS parameters using the session information in the application layer. When a VoIP session is opened in the application layer, the session initiation function activates a session initiation protocol (SIP) to connect a session between the end devices. At this time, the SIP message includes a SDP to deliver the session information, e.g. media type, transport protocol, media format, and so on, for guaranteeing the required QoS. In SDP, a field 'm' presents the media information such as m= (media) (port) (transport) (format list). For example, 'm=audio 49170 RTP/AVP 0' means that media is audio, port number is 49170, transport protocol is real time protocol (RTP) with audio video profile (AVP), and

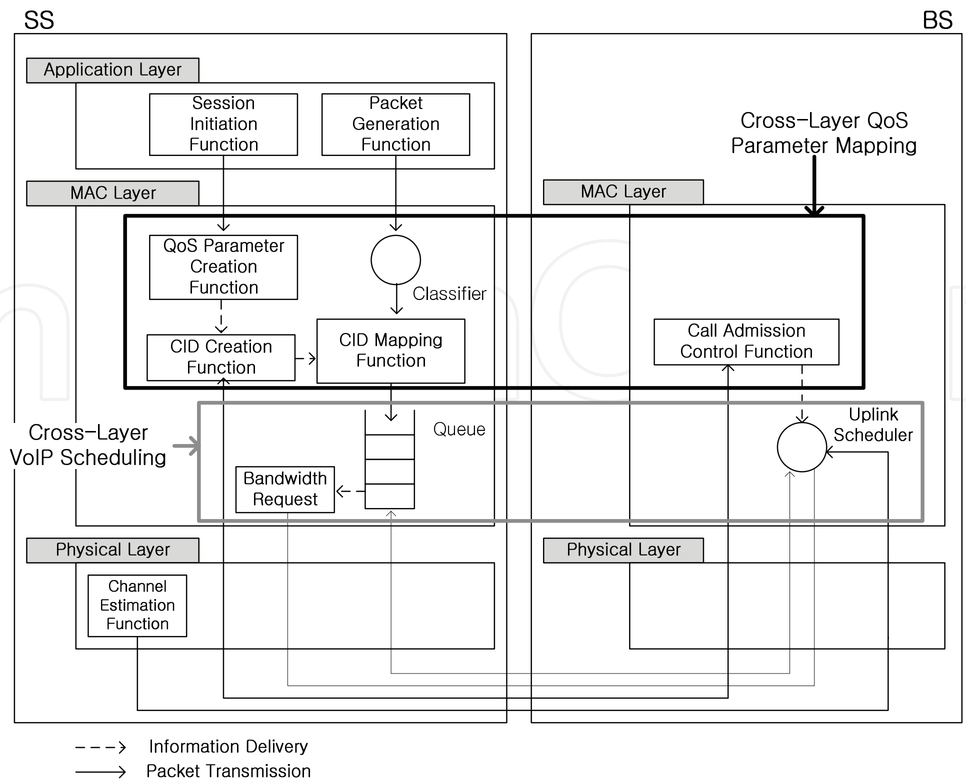


Fig. 4. Cross-layer framework for VoIP services

voice codec is G.711 (0) (Handley & Jacobson, 1998). In this chapter, the proposed scheme uses the field 'm' to identify the kinds of VoIP speech codec applied in the application layer. The features of VoIP services can be identified by the kinds of VoIP speech codec as shown in Table 1. For this reason, the QoS parameter creation function can obtain the features of the VoIP services such as the packet-size and packet-generation-interval from the SDP. Therefore, the QoS parameter creation function can generate the QoS parameters using the features of VoIP services as shown in Table 2.

4.1.2 CID creation function

The CID creation function generates a CID between a BS and a SS. It transmits a dynamic service addition request (DSA-REQ) message which includes the QoS parameter set, as shown in Table 2, to a call admission control function in a BS. The call admission control function decides whether the system supports the VoIP service or not based on the QoS parameter set

QoS parameter set	Values
Maximum sustained traffic rate	PS × PGI
Maximum traffic burst	PS
Minimum reserved traffic rate	PS × PGI
Minimum tolerable traffic rate	PS × PGI
Unsolicited grant interval	PGI
Unsolicited polling interval	PGI
SDU inter-arrival interval	PGI

Table 2. QoS Parameter Mapping Example for the VoIP Scheduling Algorithms

in the DSA-REQ message, and it sends a DSA response (DSA-RSP) message which includes a CID if the system can support the VoIP services. The CID creation function delivers the CID to the CID mapping function, when it receives the DSA-RSP message from the call admission control function.

4.1.3 CID mapping function

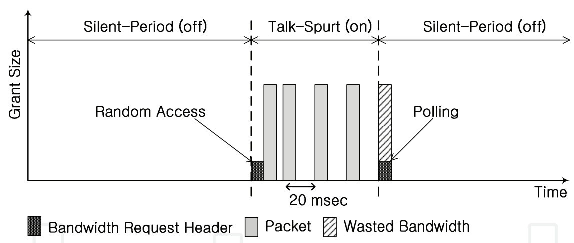
When the CID mapping function receives a CID from the CID creation function, it updates a CID table which consists of CID and the information of the user datagram protocol (UDP)/IP header such as the source/destination UDP port number, source/destination IP address, and protocol, and so on. The CID mapping function identifies a packet received from the IP layer using the information of the UDP/IP header, and it searches the CID which corresponds with the information of the UDP/IP header. For examples, the CID mapping function can identify a packet which includes a SIP message using the UDP port number because the UDP port number of SIP is 5060 or 5061. In addition, it can identify a VoIP packet using a source/destination IP address. The reason is that the source/destination IP addresses of the packets in a VoIP session are fixed. After the packet identification and CID mapping, the CID mapping function stores the packets in a queue which corresponds with the CID. The IEEE 802.16 systems transmit the packets stored in the queue by using VoIP scheduling algorithms.

4.2 Cross-layer VoIP scheduling algorithm

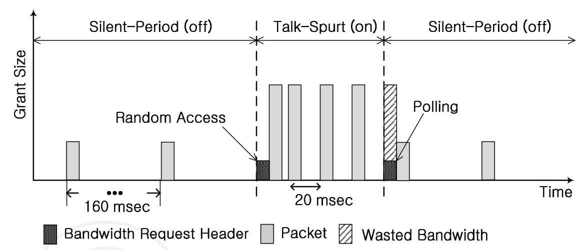
In order to solve the inefficiency of the conventional VoIP scheduling algorithms mentioned in section 3, we propose the new cross-layer VoIP scheduling algorithm. This proposed algorithm has three main features. Firstly, it exploits the QoS parameters, e.g. the grant-size and the grant-interval, generated by the cross-layer QoS parameter mapping scheme. Secondly, it adjusts the grant allocation policy according to the kinds of VoIP speech codec in the silent-period to save the uplink bandwidth. When the G.7xx with silence suppression is applied in the application layer, a BS stops the periodic grant allocation during the silent-periods in the proposed algorithm. When the EVRC or AMR speech codecs are applied in the application layer, a BS periodically allocates a grant every 20 msec or 160 msec during silent-periods, respectively. Thirdly, it adopts the random access scheme only when the voice activity changes from a silent-period to a talk-spurt. In addition, the proposed algorithm uses a BRUSC header to inform a BS of the SS's voice activity, as in the HV algorithm. In this chapter, we define that the ST bit '0' means a silent-period, whereas the ST bit '1' means a talk-spurt.

4.2.1 In case of silent-period

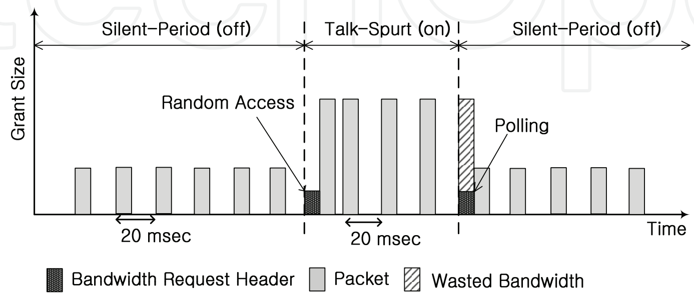
Figs. 5 (a), (b), and (c) represent the cross-layer VoIP scheduling algorithm for G.7xx, AMR speech codec, and EVRC, respectively. As shown in Fig. 5, a SS sends a BRUSC header with the ST bit '0' by using the polling scheme when the voice activity changes from a talk-spurt to a silent-period. When a BS receives a BRUSC header with the ST bit being '0', the BS stops the grant allocation or periodically allocates the grant. In case of G.7xx, the BS stops the periodic grant allocation in order to save radio bandwidth in the silent-period. In case of the AMR speech codec and EVRC, the BS periodically allocates a grant every 160 msec and 20 msec during the silent-period, respectively. The grant-size corresponds with the bandwidth specified in the BR field of the BRUSC header.



(a) for G.7xx with silence suppression



(b) for the AMR speech codec



(c) for EVRC

Fig. 5. Cross-layer VoIP scheduling algorithm

4.2.2 In case of talk-spurt

A BS periodically allocates a grant to a SS. The grant-size can be variable according to the data rate of the AMR speech codec. The proposed algorithm uses a BRUSC header or grant-management subheader for the variable data rate in the talk-spurt, similar to the HV algorithm. When the voice activity changes from a silent-period to a talk-spurt, a SS transmits a BRUSC header with the ST bit '1' by the random access scheme, as shown in Fig. 5. In the random access scheme, a SS transmits a ranging-request (RNG-REQ) message through a ranging subchannel to obtain the radio bandwidth in order to transmit a BRUSC header. A RNG-REQ message includes an orthogonal ranging code randomly selected by the SS. The grant-size is determined by the packet-size. When a BS receives the BRUSC header with the ST bit as '1', the BS allocates a grant to the SS at the next frame, and it periodically assigns a grant to the SS every grant-interval.

5. Performance evaluation

This section represents the performance evaluation results for the cross-layer QoS parameter mapping scheme and cross-layer VoIP scheduling algorithm. In order to compare the resource efficiency and QoS performance, we evaluate the system performance in terms of the average number of the allocated subchannel and average access delay. The average number of the allocated subchannel indicates the total number of subchannels, which is allocated by a BS per second. The average access delay means the average time to transmit a packet from a SS to a BS. In addition, we analyze the VoIP capacity according to the VoIP scheduling algorithms where the VoIP capacity means the maximum tolerable number of VoIP users.

VoIP speech codecs in application layer	Scenarios	Default values Grant-size	Default values Grant-interval
G.723.1 without	Scenario 1	188	20
silence suppression	Scenario 2	30	10
G.11 without	Scenario 1	40	30
silence suppression	Scenario 2	30	10
G.729 without	Scenario 1	188	20
silence suppression	Scenario 2	40	30

Table 3. Simulation Scenarios for the QoS Parameter Mapping Scheme

5.1 Simulation results for the cross-layer QoS parameter mapping scheme

The end-to-end performance evaluation simulator for the cross-layer framework has been built as shown in Fig. 4. In the end-to-end performance evaluation simulator, the whole functional blocks are modeled, and these are represented in Fig. 4. In addition, we considers the IEEE 802.16e/m orthogonal frequency division multiple access (OFDMA) system that uses 5msec time division duplex (TDD) frame size, 10 MHz bandwidth, and 1024 fast Fourier transform (FFT). In order to implement the channel variation, we consider the path loss, log-normal shadowing, and frequency-selective Rayleigh fading according to the user's mobility. To evaluate the performance for the QoS parameter mapping scheme, we consider one VoIP user in a cell. In addition, we assume that the IEEE 802.16 systems define the QoS parameters related to the VoIP services as the default values considering a specific VoIP speech codec, because the IEEE 802.16 standard does not mention the QoS parameters generation method. The default values are defined as shown in Table 3 and we consider the VoIP speech codecs as G.723.1, G.711, and G.729 without silence suppression, and defines two scenarios for each VoIP speech codec applied in the application layer, as shown in Table 3.
Fig. 6 presents the simulation results for the QoS parameter mapping scheme. Figs. 6 (a) and (b) show the simulation results when the UGS is applied to the system, whereas Figs. 6 (c) and (d) indicates the simulation results when the ertPS is applied to the system. As shown in Fig. 6 (a), the average access delay can go to infinity when the UGS is applied to the system, if the grant-size is smaller than the packet-size which is specified by the VoIP speech codec. The reason is that the access delay linearly increases when the number of transmitting packets increases because of a queuing delay of the whole VoIP packets, see Fig. 2 (a). On the other hand, the proposed algorithm can reduce the access delay to 3 msec. In case of G.723.1 and G.729, the average access delay of scenario 1 and 2 increases by 4 8 msec compared to that of the proposed algorithm. This increase of the access delay is caused by the mismatch of the QoS parameters and features of the VoIP services. However, the average access delay can not affect the QoS of the VoIP services because the maximum tolerable delay is defined, in (Srinivasan, 2007), as 50 msec. Unfortunately, these cases can suffer from resource inefficiency in term of the average number of allocated subchannel, as shown in Fig. 6 (b). Except for the G.729 with scenario 2, the average number of allocated subchannel increases by 400 $\sim$ 1200 subchannels per second compared to that obtained for the new proposed algorithm. In case of G.711, the average number of allocated subchannels for scenarios 1 and 2 are much smaller than that of the proposed algorithm. However, the SS experiences long access delays to transmit packet in scenarios 1 and 2. These cases can cause a serious deterioration of the QoS performance for VoIP services. Consequently, the system can waste wireless resources as well as increase the access delay, if the system uses the default values for the QoS parameters of VoIP services when the UGS is applied to the system. As shown in Figs. 6 (a) and (b), the

proposed algorithm can save the waste of wireless resources and as well as reduce the access delays.

Unlike the UGS, the ertPS can manage the grant-size according to the packet-size. For this reason, the ertPS can improve the system performance even though the system exploits the default values for the QoS parameters of VoIP services. As shown in Figs. 6 (c) and (d), the access delay and the average number of allocated subchannels when using the conventional algorithm with the ertPS decrease compared to those obtained for the conventional algorithm with the UGS. The average access delay can be reduced from "infinity" to less than 10 msec in the case of G.711. However, the waste of radio bandwidth and the increase of access delays still exist because of the mismatch of the grant-interval and the packet-generation-interval. In the case of G.723.1, the SS has to wait for a grant in scenario 1 because a BS periodically allocates a grant every 20 msec even though a packet is generated every 30 msec in the application layer. In scenario 2 of G.723.1, the SS does not need to wait for a grant because a BS allocates a grant every 10 msec. However, this case can waste two grants every 30 msec. For this inefficiency, the average number of allocated subchannel increases by about 200 % compared to the proposed algorithm as shown in Fig. 6 (d). In the case of G.729, a transmitting packet is delayed because the grant-interval is larger than the packet-generation-interval. For this reason, the average number of allocated subchannel decreases in scenarios 1 and 2 compared to that of the proposed algorithm whereas the average access delay increases by 10 16 msec. Therefore, the cross-layer QoS parameter mapping scheme can improve the system performance in terms of the number of allocated subchannels and access delays.

5.2 Numerical results for the cross-layer VoIP scheduling algorithm

This subsection represents the system performance for the new cross-layer VoIP scheduling algorithm in terms of the VoIP capacity. The VoIP capacity means the maximum supportable number of VoIP users. In order to analyze the system performance, the voice traffic has been modeled as an exponentially distributed ON-OFF system with mean ON-time $1/\lambda$ and mean OFF-time $1/\mu$. Fig. 7 represents the one-dimensional Markov chain for N independent VoIP users (Oh et al., 2008). In Fig. 7, each state indicates the number of VoIP users in the ON state. Since the sum of the whole steady-state probability is unit, the steady-state probability is derived as

$$\begin{aligned} p_N(k) &= \binom{N}{k}\left(\frac{\mu}{\lambda+\mu}\right)^k\left(\frac{\lambda}{\lambda+\mu}\right)^{(N-k)}, \\ & k=0,1,2,\cdots,N. \end{aligned} \tag{1}$$

The average number of VoIP users in the silent-period (N_{OFF}) is

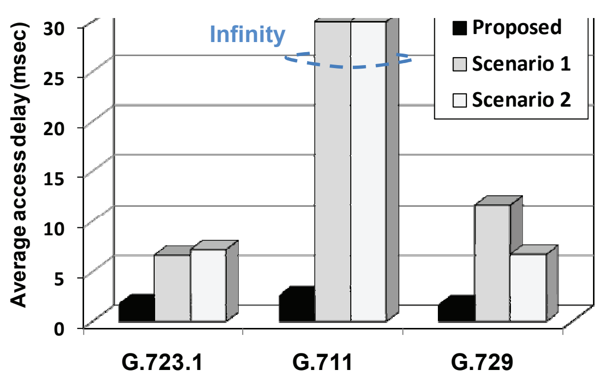
$$N_{OFF}(N)=\frac{N\lambda}{\lambda+\mu}, \tag{2}$$

where N is the number of VoIP users.

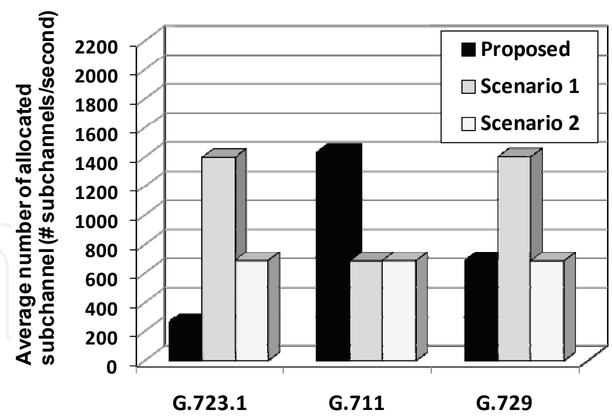
In this chapter, the unit of the grant-size is defined as the number of slots. The average number of uplink slots required every grant-interval for a VoIP user in each scheduler is given by

$$S_{UGS} = S_{ON\_max}, \tag{3}$$

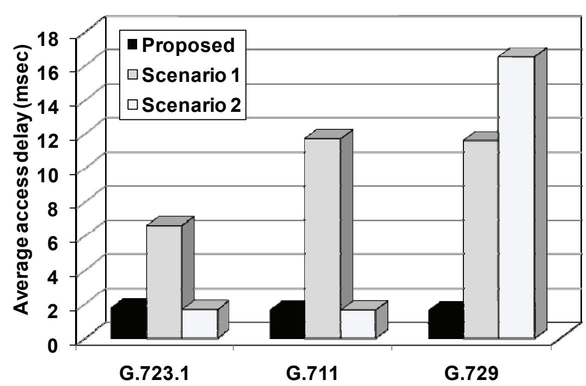
$$S_{ertPS} = \left(\frac{S_{ON}}{\lambda}+\frac{S_{GMH}}{\mu}\right), \tag{4}$$



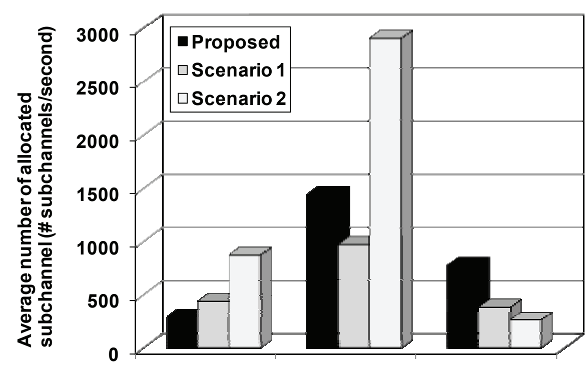
(a) average access delay with UGS



(b) average number of allocated subchannels with UGS



(c) average access delay with ertPS



(d) average number of allocated subchannels with ertPS

Fig. 6. Simulation results for the cross-layer QoS parameter mapping scheme

$$S_{HV}(N,F) = \left(\frac{S_{ON}}{\lambda} + \frac{S_{SI} + S_{BRUSC} R_{HV}(N,F)}{(T_{GIS}/T_{GIT})\mu} \right), \tag{5}$$

$$S_{pro}(N,F) = \left(\frac{S_{ON}}{\lambda} + \frac{S_{SI} + S_{BRUSC} R_{pro}(N,F)}{(T_{GIS}/T_{GIT})\mu} \right), \tag{6}$$

where S_{ON_max}, S_{ON}, S_{SI}, and S_{GMH} are the number of uplink slots required to send a maximum-size speech frame, variable-size speech frame, silence(or noise) frame, and generic-MAC header, respectively. F is the number of bandwidth request ranging codes. Note that the S_{GMH} in (4) can be changed to S_{SI} in the EVRC, because the EVRC generates a noise frame every packet-generation-interval. T_{GIT} (sec) and T_{GIS} (sec) indicate the grant-interval during the talk-spurts and the grant-interval during the silent-periods, respectively. In (5) and (6), R_{HV} and R_{pro} represent the average number of retransmissions for a BRUSC header in the HV algorithm and the new proposed algorithm, respectively.
The average number of uplink slots required every grant-interval for a VoIP user in the UGS and ertPS is independent on the number of VoIP users and the number of bandwidth

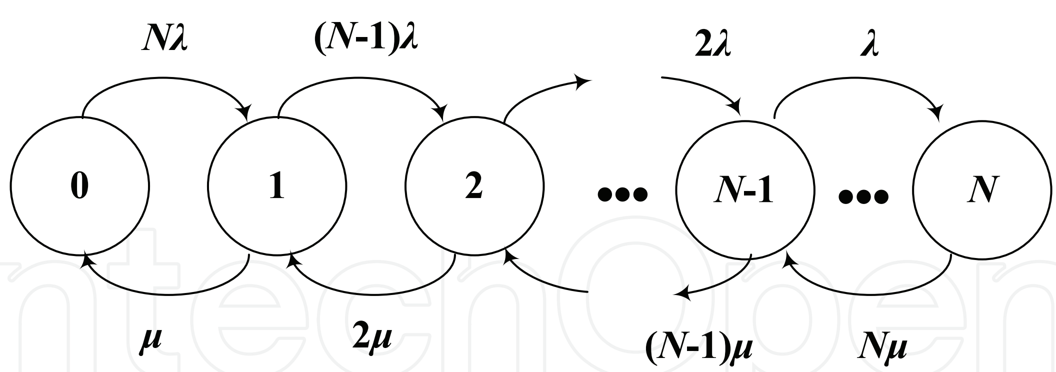


Fig. 7. Markov chain model for N independent VoIP users with exponentially distributed ON-OFF system

request ranging codes, because a BS periodically allocates a grant to a SS every grant-interval. However, in the HV, a SS sends a BRUSC header to transmit a SID frame every TGIS by the random access scheme in the silent-period. For this reason, the average number of contending users ($N_C(N)$) in a frame is

$$N_{C\_HV}(N) = N_{OFF}(N) \times \frac{T_{MF}}{T_{GIS}}, \tag{7}$$

where T_{MF} is the MAC frame size. In (7), the second term on the right side means the transmission rate of one user in a frame. In the random access scheme, the SS transmits a ranging-request (RNG-REQ) message through a ranging subchannel to obtain the radio bandwidth to transmit a BRUSC header. A RNG-REQ message includes an orthogonal ranging code randomly selected by the SS. When several SSs simultaneously choose the same orthogonal ranging code in a ranging subchannel, they experience a collision. In the random access scheme, other SSs should not select the ranging code which is already selected by a SS in a frame. Thus, the success probability ($P_{SUC}(N,F)$) in a frame is given by

$$P_{SUC}(N,F) = \left(1 - \frac{1}{F}\right)^{N_{C\_HV}(N)-1}. \tag{8}$$

The average number of retransmissions in the HV algorithm is given by

$$\begin{aligned} R_{HV}(F) &= \sum_{k=0}^{\infty} k P_{SUC}(N,F)(1 - P_{SUC}(N,F))^{k-1} \\ &= \frac{1}{P_{SUC}(N,F)}. \end{aligned} \tag{9}$$

By using (2), (7), and (8), the average number of retransmission in the HV can be derived as

$$R_{HV}(N,F) = \left(1 - \frac{1}{F}\right)^{1 - \frac{N\lambda}{\lambda+\mu} \times \frac{T_{MF}}{T_{GIS}}}. \tag{10}$$

In the proposed algorithm, a SS transmits a BRUSC header by the random access scheme only when a voice activity changes from a silent-period to a talk-spurt, unlike in the HV algorithm.

Thus, the average number of contending users in a frame is equal to $N_{OFF}(N) \times T_{MF}$. For this reason, the average number of retransmissions in the new proposed algorithm is

$$R_{pro}(N,F) = \left(1 - \frac{1}{F}\right)^{1-\frac{N\lambda}{\lambda+\mu} \times T_{MF}}. \tag{11}$$

At this time, the VoIP capacity (m) for each VoIP scheduling algorithm can be defined as follows.

$$m(N,F) = \frac{T_{GIT}}{T_M F} \times \frac{S_{TOT}}{S(N,F)}, \tag{12}$$

where S_{TOT} is the total number of uplink slots in a frame (Srinivasan, 2007) and $S(N,F)$ means the average number of uplink slots required every T_{GIT} for each VoIP scheduling algorithm such as S_{UGS}, S_{ertPS}, S_{HV}, and S_{pro}. In (12), the term on the right side represents the product of the number of frame during the grant-interval of the talk-spurt and the maximum supportable number of VoIP users in a frame. Unfortunately, the $S(N,F)$ of HV and proposed algorithm is given with respect to the number of VoIP users and the number of bandwidth request ranging codes as shown in (5), (6), (10), and (11). For this reason, it is difficult to analyze the VoIP capacity in the HV and proposed algorithms.

For the simple analysis process, we approximately analyze the average number of retransmission as follows.

$$\begin{aligned} R_{HV}(N,F) &= \left(1 - \frac{1}{F}\right)^{1-\frac{N\lambda}{\lambda+\mu} \times \frac{T_{MF}}{T_{GIS}}} \\ &= 1 + \frac{\lambda T_{MF} N}{(\lambda+\mu) T_{GIS}} \times \frac{1}{F} + \frac{\left(\frac{\lambda T_{MF} N}{(\lambda+\mu) T_{GIS}}\right) \times \left(\frac{\lambda T_{MF} N}{(\lambda+\mu) T_{GIS}} + 1\right)}{2} \times \left(\frac{1}{F}\right)^2 + \cdots \end{aligned} \tag{13}$$

By using MacLaurin series, $R_{HV}(N,F)$ can be written as (13). Here, IEEE 802.16 defines the number of orthogonal ranging codes as 256 where the ranging code consists of initial, handover, bandwidth request, and periodic ranging codes. However, the number of ranging codes in a frame can be above 256 because the number of ranging slots which consists of 256 ranging codes can be one or more. Thus, F can be a sufficiently large number. For this reason, $R_{HV}(N,F)$ can be approximately given by

$$R_{HV}(N,F) \approx 1 + \frac{\lambda}{\lambda+\mu} \times \frac{T_{MF}}{T_{GIS}} \times \frac{1}{F} \times N, \tag{14}$$

where $1/F$ is much less than one. Here, (14) is substituted for (5). Fig. 8 depicts the average number of uplink slots required every grant-interval for a VoIP user ($S_{HV}(N,F)$) according to the number of VoIP users and number of bandwidth request ranging code when VoIP speech codec is the EVRC. As shown in Fig. 8, N and F can be neglected in terms of $S_{HV}(N,F)$. In addition, this result is similar to the case of the AMR speech codec and G.7xx, because those speech codecs generate packets by using the lower transmission rate in the silent-period. Therefore, $S_{HV}(N,F)$ can be approximately represented as

$$S_{HV} \approx \left(\frac{S_{ON}}{\lambda} + \frac{S_{SID} + S_{BRUSC}}{(T_{GIS}/T_{GIT})}\right). \tag{15}$$

As in the HV algorithm, the new proposed algorithm can approximately analyze the $S_{pro}(N,F)$ as (16), because the proposed algorithm transmits a BRUSC header only when the

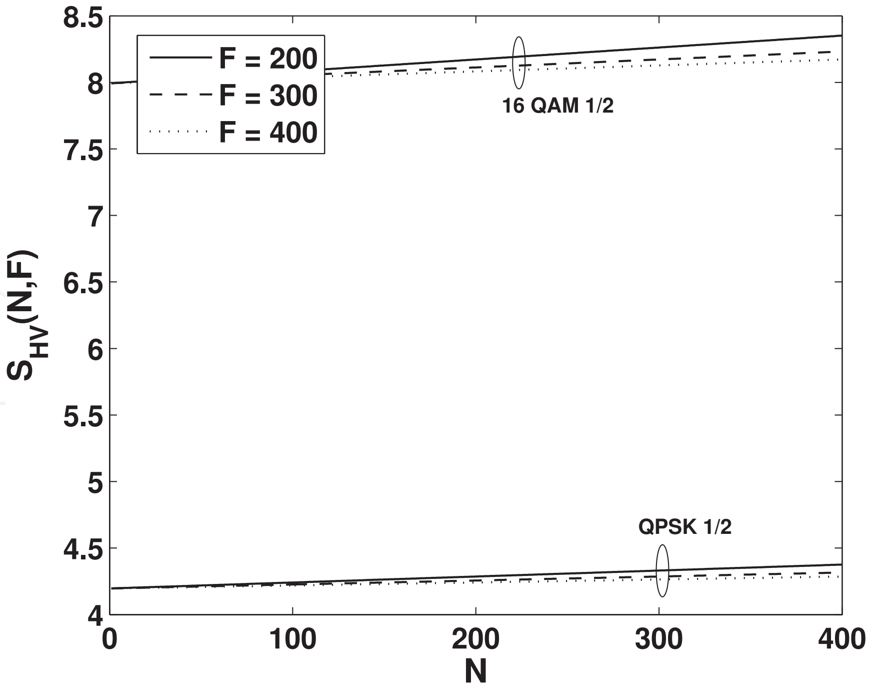


Fig. 8. $S_{HV}(N,F)$ vs. N and F (MCS level = QPSK 1/2 and 16 QAM 1/2, VoIP speech codec = EVRC, S_{TOT} = 144 slots, T_{MF} = 5 msec, FFT size = 1024, λ = 2.5, μ = 1.67, and bandwidth = 10 MHz)

voice activity changes from silent-period to talk-spurt.

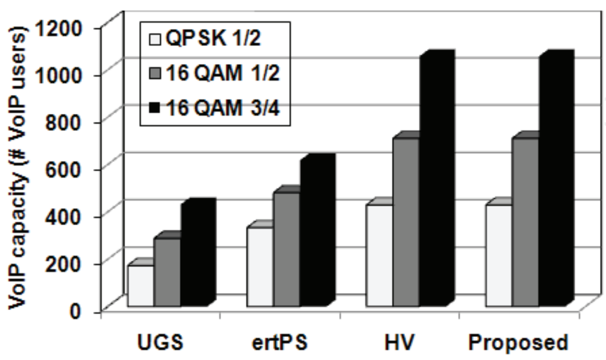
$$S_{pro} \approx \left(\frac{S_{ON}}{\lambda} + \frac{S_{SID} + S_{BRUSC}}{(T_{GIS}/T_{GIT})} \right). \tag{16}$$

By using (14) and (15), (12) can be derived as

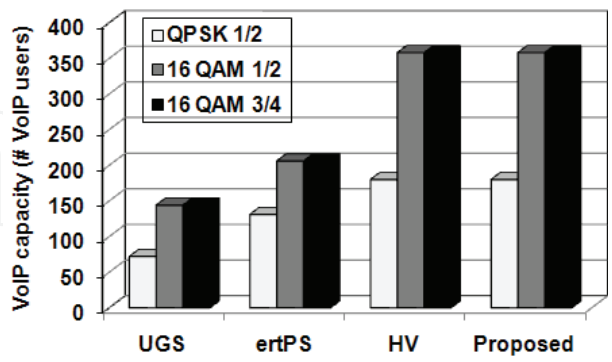
$$m = \frac{T_{GIT}}{T_{MF}} \times \frac{S_{TOT}}{S}, \tag{17}$$

where S is the average number of uplink slots required every T_{GIT} for each VoIP scheduling algorithm. In (17), the VoIP capacity can be easily analyzed, because m is not dependent on the N and F.

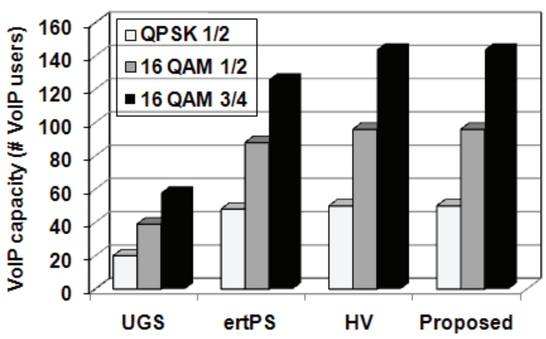
Fig. 9 presents numerical results for the VoIP capacity according to the modulation and coding scheme (MCS) levels. It can be seen that the HV and the proposed algorithm can increase the VoIP capacity except for the EVRC compared to the conventional ertPS and UGS, respectively. The reason is that the algorithms can save the uplink bandwidth in the silent-period by using the random access or the adaptation of the grant-interval. However, the HV and the proposed algorithm could not obtain the gain in terms of VoIP capacity when the VoIP speech codec is the EVRC, as shown in Fig. 9 (e). The HV is particularly inefficient in using the radio bandwidth compared to the ertPS when the VoIP speech codec is the EVRC, because the HV transmits a BRUCS header to send a noise frame of the EVRC every 20 msec. By using this feature of the HV, the VoIP capacity decreases by 29 % compared to when the ertPS is used. Unlike the HV, the proposed algorithm can efficiently use the radio bandwidth because of the adaptation of the grant-interval when the VoIP speech codec is the EVRC as well as the G.711 and AMR speech codec.



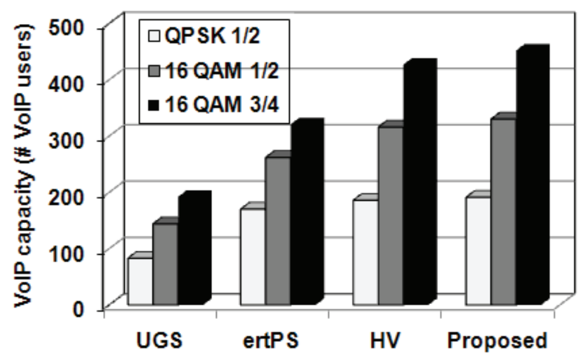
(a) G.723.1 with silence suppression



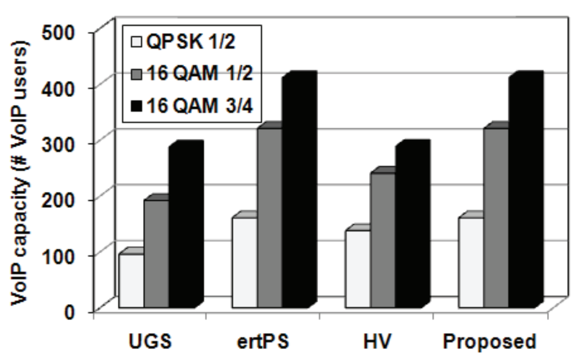
(b) G.729 with silence suppression



(c) G.711 with silence suppression



(d) AMR



(e) EVRC

Fig. 9. VoIP capacity vs. VoIP scheduling algorithms and MCS levels (S_{TOT} = 144 slots, T_{MF} = 5 msec, FFT size = 1024, λ = 2.5, μ = 1.67, compressed RTP/UDP/IP header size = 3 bytes and bandwidth = 10 MHz)

As shown in Fig. 9, the gain of the HV and the proposed algorithm depends on the kinds of VoIP speech codec in the application layer. The gain increases by 70 % when the VoIP speech codec is G.723.1 or G.729, as shown in Figs. 9 (a) and (b). The G.723.1 and G.729 generate a small-size voice frame in talk-spurt, whose size is 19.88 bytes and 10 bytes, respectively. For this reason, the number of supportable VoIP users increases with respect to other VoIP speech codecs due to the saved bandwidth in the silent-periods. From these numerical results, the HV and the proposed algorithm can support 150 $\sim$ 400 VoIP users more than the other algorithms. Consequently, the HV and the proposed algorithm, which do not periodically allocate a grant in the silent-period, can increase the VoIP capacity and the proposed algorithm can in particular increase the VoIP capacity by 15 % $\sim$ 70 % regardless of the kinds of VoIP speech codec in the application layer.

6. Conclusion

VoIP traffic can have various features according to the kinds of VoIP speech codecs, hence wireless systems need to consider the features of VoIP speech codec. In this chapter, we have considered variable packet-size and packet-generation-interval for main VoIP speech codecs, and proposed a new cross-layer framework to efficiently support a VoIP service in IEEE 802.16 systems. The cross-layer framework for a VoIP service consists of a cross-layer QoS parameter mapping scheme and a cross-layer VoIP scheduling algorithm. The cross-layer QoS parameter mapping scheme directly obtains the QoS parameters for a VoIP service using the QoS information of the application layer. The cross-layer VoIP scheduling algorithm efficiently supports a VoIP service based on the QoS parameters generated by the proposed cross-layer QoS parameter mapping scheme. By the performance evaluation results, it has been shown that the new algorithm can efficiently support a VoIP service regardless of the kinds of VoIP codec in the application layer.

7. References

3GPP-TS-26071 (1999). 3GPP TS 26.071 v3.0.1: Mandatory speech codec speech processing functions AMR speech codec; general description.

3GPP-TS-26092 (2002). 3GPP TS 26.092 v5.0.0: AMR speech codec; comfort noise aspects.

3GPP-TS-26201 (2001). 3GPP TS 26.201 v5.0.0: AMR wideband speech codec; frame structure.

3GPP2-EVRC (2004). 3GPP2 C.S0014-A v1.0: Enhanced variable rate codec.

Handley, M. & Jacobson, V. (1998). RFC 2327 - SDP: Session description protocol.

Hong, S. E. & Kwon, O. H. (2006). Considerations for voip services in IEEE 802.16 broadband wireless access systems, *Proceedings of IEEE VTC spring*, pp. 1226–1230.

IEEE (2006). IEEE standard for local and metropolitan area networks - part 16: air interface for fixed and mobile broadband wireless access systems amendment 2.

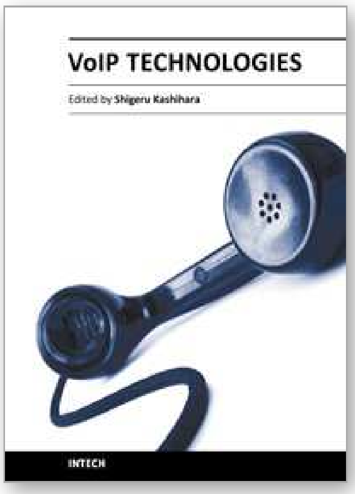
ITU-T-G711 (2000). ITU-T recommendation G.711 appendix II: A comfort noise payload definition for ITU-T G.711 use in packet-based multimedia communication systems.

ITU-T-G7231 (1996). ITU-T recommendation G.723.1 appendix A: Silence compression scheme.

ITU-T-G729 (2007). ITU-T recommendation G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction.

Lee, H. W., Kwon, T. S. & Cho, D. H. (2005). An enhanced uplink scheduling algorithm based on voice activity for voip services in IEEE 802.16d/e systems, *IEEE Commun. Lett.* Vol. 9: 691–692.

Oh, S. M., Cho, S. H., Kwun, J. H. & Kim, J. H. (2008). VoIP scheduling algorithm for AMR speech codec in IEEE 802.16e/m system, *IEEE Commun. Lett.* Vol. 12(No. 5).
Oh, S. M. & Kim, J. H. (2005). The analysis of the optimal contention period for broadband wireless access network, *Proceedings of IEEE PWN 05*, pp. 215–220.
Srinivasan, R. (2007). Draft IEEE 802.16m evaluation mothodology document.



VoIP Technologies

Edited by Dr Shigeru Kashihara

ISBN 978-953-307-549-5

Hard cover, 336 pages

Publisher InTech

Published online 14, February, 2011

Published in print edition February, 2011

This book provides a collection of 15 excellent studies of Voice over IP (VoIP) technologies. While VoIP is undoubtedly a powerful and innovative communication tool for everyone, voice communication over the Internet is inherently less reliable than the public switched telephone network, because the Internet functions as a best-effort network without Quality of Service guarantee and voice data cannot be retransmitted. This book introduces research strategies that address various issues with the aim of enhancing VoIP quality. We hope that you will enjoy reading these diverse studies, and that the book will provide you with a lot of useful information about current VoIP technology research.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821