Simultaneous Transmission of MAP IE and Data for Minimizing MAC Overhead in the IEEE 802.16e OFDMA Systems

Juyeop Kim, Student Member, IEEE, and Dong-Ho Cho, Senior Member, IEEE

Abstract—An advanced MAP transmission scheme for improving Media Access Control (MAC) overhead in the IEEE 802.16e systems is proposed. In the IEEE 802.16e system, when a Base Station (BS) broadcasts a MAP message, which is a control message about scheduling information, it applies a robust Modulation and Coding Scheme (MCS) level and allocates a large amount of radio resources, which induce a huge MAP overhead. Our proposed scheme utilizes piggybacked MAP IEs, in which control messages are concatenated with data packets and transmitted with the MCS level applied to data transmission. Due to the fact that the rate at which data is transmitted is generally higher than the rate at which broadcasting messages are transmitted, the proposed scheme can increase average data rate of a MAP transmission and consequently reduce the amount of resources allocated to the MAP transmission. Numerical analysis and simulations are presented to show that the MAP overhead is critical to system performance and can be improved by the proposed scheme.

Index Terms—IEEE 802.16e, MAC overhead, MAP IE, piggyback.

I. INTRODUCTION

The IEEE 802.16e system, which is the standard for fixed and mobile broadband wireless access systems, aims to provide always-connected services to mobile users. Once a mobile user has entered the network initially, it can always access the network and receive packets without performing any further procedure to connect. This type of service provides great convenience to mobile users. However, the system has a burden of keeping in constant touch with all the mobile users. This burden includes additional computations and allocation of radio resources for control signaling.

As a result, the IEEE 802.16e system suffers a problem of huge MAC overhead. The additional control signaling for always-connected service causes a significant increase in the MAC overhead and limits the number of users that can be supported by the system. Such a system allocates a large amount of radio resources for supporting mobile users and may suffer a shortage of resources when the number of users in a cell exceeds a certain limit. The simulation results in [1] revealed that the MAC overhead of the IEEE 802.16e system is critical to the system performance.

There are many research works which design MAC protocol to minimize the MAC overhead in other communication systems. [2]-[4] However, those works cannot help to reduce the MAC overhead of the IEEE 802.16e system, because most of the huge MAC overhead in the IEEE 802.16e system is caused by MAP overhead, i.e., overhead due to the transmission of MAP messages. A MAP message is transmitted at the beginning of each frame to provide scheduling information to receivers. In order to ensure constant connection, all the mobile users in a cell should keep in touch with BS, which requires frequent indication of scheduling information. In addition, the IEEE 802.16e system employs various schemes to handle radio resource dynamically with respect to time and frequency, which contributes to increasing the amount of scheduling information in the MAP messages.

A certain amount of research has been done on improving the MAC overhead in the IEEE 802.16e system. Some works apply AMC to control messages for resource allocation to increase the spectral efficiency of control signaling. [5]-[9] These works are based on the assumption that it is possible to apply AMC to the transmission of broadcasting messages, which is not true in the case of the IEEE 802.16e system. In addition, they require additional signaling for indicating modulation and coding scheme of the control message transmission. In [9], a new control message is proposed for periodic assignment of resources to avoid signaling for subsequent allocation, but can only be applied to specific services whose traffic arrives periodically.

In this letter, we firstly evaluate the MAP overhead that is due to broadcasting messages and how it affects the system performance of the IEEE 802.16e system. Then, we propose a practical scheme that reduces the MAC overhead and that can be applied to the IEEE 802.16e system.

II. SYSTEM MODEL

We consider the IEEE 802.16e system that has $K$ downlink users with the same traffic patterns. An OFDM symbol contains $N_u$ available subcarriers for transmitting data and control messages. $c_{u}(t)$, the Signal to Interference and Noise Ratio (SINR) of subcarrier $l$ in OFDM symbol $t$ with respect to user $i$, can be expressed as following

$$c_{u}(t) = \frac{h_{ul}(t)}{I_{ul}(t) + N_0},$$  (1)

where $h_{ul}(t)$ and $I_{ul}(t)$ are the channel gain and interference level of subcarrier $l$ in OFDM symbol $t$ for user $i$, respectively, and $N_0$ is a noise level.

For reducing signaling overhead, the system defines a subchannel as a set of $N_u/N_s$ subcarriers collected by a certain subchannelization rule and treats it as a minimal unit for resource allocation and link adaptation. Consequently,
there are \( N_s \) subchannels in an OFDM symbol. Let \( b_{jp} \) denote the index of the \( p \)th subcarrier in subchannel \( j \) and let \( \mathbf{c}_{ij}(t) = [c_{ib_1}(t) \ c_{ib_2}(t) \ \ldots \ c_{ib_{(N_u/N_s)}(t)}]^T \) be the channel state vector of subchannel \( j \) in OFDM symbol \( t \) for user \( i \).

Fig. 1 shows the frame structure of the system. A frame is composed of \( N_u \) subchannels and \( N_t \) time slots in the downlink subframe, where a time slot is composed of \( N_t \) OFDM symbols. The BS performs scheduling and allocates resources at the beginning of each frame. We denote a resource unit as rectangular resource region for the minimum unit of link adaptation and resource allocation. Consequently, a frame contains \( N_uN_t \) resource units. We assume that channel state varies from frame to frame and is static within a frame.

For link adaptation, the BS chooses an appropriate Modulation and Coding Scheme (MCS) level for each data transmission. Let \( M = \{1, 2, \ldots, M\} \) be the index set of the available MCS levels and \( r_m \) be the number of bits that are transmitted using a single resource unit when using MCS level \( m \). Then, \( d_{ij}(t) \), the number of resource units needed to transmit one bit with respect to subchannel \( j \) in OFDM symbol \( t \) and user \( i \), can be evaluated as following,

\[
d_{ij}(t) = \left[ \max_{m \in M, (c_{ij}(t))} r_m \right]^{-1}
\]

(2)

where \( M_a(c) \) is the set of available MCS levels under channel state \( c \). There are several practical ways to define \( M_a(c) \). We use the Exponential Effective SINR Mapping (Exp-ESM) method, in which the MCS level is chosen on the basis of a scalar value called effective SINR [10]. With this method, \( M_a(c_{ij}(t)) \) can be evaluated as following,

\[
M_a(c_{ij}(t)) = \{ m \in M \mid \frac{N_u}{N_s} \sum_{p=1}^{N_u/N_s} c_{ib_{jp}} > c_{l,m} \} \tag{3}
\]

where \( c_{l,m} \) is the lower bound of the average SINR to which MCS level \( m \) can be applied and can be determined by the method in [11]. Here, \( \frac{N_u}{N_s} \sum_{p=1}^{N_u/N_s} c_{ib_{jp}} \) is the effective SINR.

In addition, we consider diversity subchannelization, in which subcarriers in a subchannel are distributed over the frequency domain. Fig. 2 shows the structure of diversity subchannels. The subcarriers are divided into \( N_u/N_s \) subgroups and a subcarrier is picked from each subgroup according to a permutation rule. Assuming that the channel state is flat within a subgroup, \( c_{ij}(n) \) becomes constant for all \( j \in \{1, 2, \ldots, N_s\} \) and for all OFDM symbols within a frame, and \( d_{ij}(t) = d_{ij}(n) = \ldots = d_{iN_s}(t) \). Thus, we further express \( d_{ij}(n) \) as \( d_{i}(n) \) for all \( j \).

A. MAP Overhead Analysis

We assume that a MAP message is transmitted with fixed MCS level \( m_{map} \). It is composed of multiple MAP IEs and each MAP IE describes the resource region of a burst, which is a collection of data packets whose MCS levels are the same. Consequently, the number of MAP IEs in a MAP message is equal to the number of bursts transmitted in the frame. We assume that each burst is composed of data packets which will be delivered to a user so that the number of MAP IEs is equal to the number of scheduled users in the frame.

In addition, we further assume heavy traffic environment that \( K \) is sufficiently large and that most of the resource units in each frame are utilized. These assumptions are reasonable when analyzing MAC overhead, because MAC overhead becomes critical to system performance when the system suffers from a shortage of radio resources. To quantify the effect of MAP message on system performance, we define MAP overhead as the average ratio of the number of resource units allocated to MAP transmission to the total amount of resource units used in the frame. MAP overhead \( V \) is defined mathematically as following,

\[
V = E \left[ \frac{R_m(n)}{R_m(n) + R_d(n)} \right]
\]

(4)

where \( R_d(n) \) and \( R_m(n) \) are the number of resource units allocated to data and MAP transmission, respectively, in the \( n \)-th frame. Note that (4) is quite different from the definition of control overhead in [7], where it is equal to the ratio of expected amount of resources allocated to control signaling to the expected total amount of resources used. Letting \( b_{nu} \) be the size of a MAP IE, \( N_k(n) \) be the number of users scheduled in the \( n \)-th frame, and \( d_m \) be \( m_{map} \), \( R_m(n) \) can be evaluated as following

\[
R_m(n) = N_k(n)b_{nu}d_m.
\]

(5)

Given that the system has a sufficiently large number of users, \( N_k(n) \) should be also large. Using the central limit
\[ R_d(n) = \sum_{k=1}^{N_k(n)} b_i(n)d_i(n) \]
\[ \simeq N_k(n)E[b_i(n)d_i(n)] \simeq N_k(n)E[b_i(n)]E[d_i(n)] \]

where \( b_i(n) \) is the size of the data transmitted to user \( i \) in the \( n \)-th frame. Note that approximation in (7) is based on the assumption that \( b_i(n) \) and \( d_i(n) \) are independent. In general, \( b_i(n) \) and \( d_i(n) \) are not always independent, because when an original data packet is large and the available data rate for user \( i \) is low, \( b_i(n) \) becomes the size of a fragmented data packet. However, when the original data packet is sufficiently small, \( b_i(t) \) is equal to the size of the original data packet, which does not depend on \( d_i(n) \).

Let \( \overline{b} \) and \( \overline{d} \) be \( E[b_i(n)] \) and \( E[d_i(n)] \), respectively. Applying (5) and (7) to (4), \( V \) can be evaluated as following

\[ V = E\left[ \frac{N_k(n)b_m\overline{d}}{N_k(n)b_m\overline{d} + N_k(n)b\overline{d}} \right] \]
\[ = E\left[ \frac{b_m\overline{d}}{b_m\overline{d} + \overline{b}\overline{d}} \right] = 1 - \frac{\overline{b}}{b_m\overline{d} + \overline{b}\overline{d}}. \]

(9) reveals that \( V \) is a function of \( \overline{b}\overline{d} \) and independent of \( N_k(n) \) in heavy traffic. \( V \) increases as \( \overline{b}\overline{d} \) decreases, because the portion of the resources allocated to data transmission becomes smaller. In addition, \( V \) depends on the average of \( b_i(n) \) and is not affected by probability distribution of \( b_i(n) \). This means that the MAP overhead will be the same for any traffic whose average packet size is the same. Assuming that the average size of data packet is small and that fragmentation rarely happens, \( \overline{b} \) is equal to the sum of the header size and the average packet size. In addition, \( \overline{d} \) can be evaluated as following

\[ \overline{d} = \sum_{m=1}^{M} \frac{1}{r_m} Pr(d_i(n) = 1/r_m), \]

where \( Pr(B) \) is the probability that \( B \) will occur. In the case of a Rayleigh fading channel with mean \( \lambda \), \( \overline{d} \) can be evaluated as following

\[ \overline{d} = \sum_{m=1}^{M} \frac{1}{r_m} (\exp(e_{i,m}/\lambda) - \exp(e_{i,m+1}/\lambda)). \]

Using \( V \), we can evaluate the average system throughput, denoted by \( T \), as following

\[ T = E\left[ N_sN_t(1 - \frac{R_m(n)}{R_d(n) + R_m(n)})E[d_i(n)] \right] \]
\[ \simeq N_sN_t(1 - V)\overline{d}. \]

When \( \overline{b}\overline{d} \ll b_m\overline{d}_m \), \( V \) approaches to 1 and \( T \) consequently approaches to 0. In practice, \( b_m\overline{d}_m \) is large because the BS tries to send MAP message to all users and applies a very robust MCS level to MAP transmission, which induces large \( d_m \). Thus, in case of traffic whose average packet size is small, such as Voice over IP (VoIP), system capacity will be limited due to large MAP overhead.

III. SIMULTANEOUS TRANSMISSION OF MAP IE AND DATA

To increase the average data rate of MAP transmission, our proposed scheme combines MAP IEs with data packets and transmits them to users. When a BS transmits a data packet to a user, it constructs a MAP IE that describes the resource allocation in future frames and piggybacks the MAP IE on the data packet. (We further call this type of MAP IE as a piggybacked MAP IE and a conventional MAP IE as a broadcasting MAP IE.) Given that the data rate of the piggybacked MAP IEs is equal to \( d_i(n) \), which is usually smaller than \( d_m \) under the proposed scheme, the average data rate of MAP transmission is increased and MAP overhead is reduced.

To describe the operation of the proposed scheme, let \( MPDU(n) \) be a data packet transmitted in frame \( n \), \( IE(n) \) be the MAP IE that describes the resource allocation for \( MPDU(n) \) transmission, \( N_{pk}(n) \) be the number of packets in the queue in the \( n \)-th frame and \( \delta_{est} \) be the time interval, measured in frames, between the current moment and the moment of next data transmission. The operation of the proposed scheme is presented in detail as following,

1) Initialize. \( n := 0 \).
2) if (A BS transmits \( MPDU(n) \) to a user)
   - \{ go to step 3 \}; else \{ go to step 7 \};
3) if (\( d_i(n) > d_m \&\& N_{pk}(n + \delta_{est}) > 0 \))
   - \{ The BS piggybacks \( IE(n+\delta_{est}) \) to \( MPDU(n) \); \}
4) if (the user already has \( IE(n) \))
   - \{ The BS excludes \( IE(n) \) from broadcasting MAP; \}
5) The BS broadcasts a MAP message.
6) if (the user successfully decodes \( IE(n + \delta_{est}) \))
   - \{ The user sends an ACK; \}
   - else, \{ The user sends a NACK; \}
7) \( n := n + 1 \), go to step 2

Fig. 3 shows an example of the operation of the proposed scheme. The BS piggybacks the MAP IE that describes the resource region of data packet 2 when transmitting data packet 1. The user successfully decodes the piggybacked MAP IE with data packet 1 and sends an ACK. The BS then transmits data packet 2 without sending the MAP IE in the MAP message. When the user fails to receive the data packet and piggybacked MAP IE and sends a NACK, the BS broadcasts a MAP IE for the retransmission of data packet 2, like conventional operation.
The proposed scheme can effectively reduce the amount of resources allocated to MAP transmission without any additional control signaling. The proposed scheme may result in additional retransmissions of the MAP IE when a data packet is not delivered or becomes corrupted and the piggybacked MAP IE in the data packet is lost. However, this rarely happens because the system generally keeps Packet Error Rate (PER) very small. Note that the piggybacked MAP IE describes the resources that are to be allocated for next data transmission. After obtaining a piggybacked MAP IE from a current data packet, the user can receive the next data packet without decoding the broadcasting MAP IE. Thus, the proposed scheme can be applied to services whose packet arrival is predictable, such as VoIP, video streaming, and uplink polling.

In addition, it is simple to apply the proposed scheme to the IEEE 802.16e system. The proposed scheme does not increase complexity of a conventional MAP transmission process, because it only requires additional decision with respect to step 3, which is a linear process. Also, feedback information needed for the proposed scheme can be obtained from the feedback mechanisms performed for conventional schemes. Channel state, which is needed for the decision in step 3, can be obtained from AMC feedback of conventional schemes, and ACK/NACK information used for the decision in step 4 can be obtained from Hybrid ARQ feedback of conventional schemes.

A. MAP Overhead Analysis

Let \( R_{m,b}(n) \) and \( R_{m,p}(n) \) be the number of resource units allocated to broadcasting MAP IEs and a piggyback MAP IEs, respectively, in the \( n \)-th frame. \( R_m(n) \) is expressed as following,

\[
R_m(n) = R_{m,b}(n) + R_{m,p}(n)
\]

(14)

\[
R_{m,b}(n) = N_{k,b}(n) b_m d_m
\]

(15)

\[
R_{m,p}(n) = \sum_{i=1}^{N_{k,p}(n)} b_m d_i(n)
\]

(16)

where \( N_{k,b}(n) \) and \( N_{k,p}(n) \) are the number of users whose burst is indicated by a broadcasting MAP IE and piggyback MAP IE, respectively, in the \( n \)-th frame. Note that \( R_{m,p}(n) \) depends on \( d_i(n) \), because the same MCS level that is applied to data transmission is applied to the transmission of the piggybacked MAP IE. Then, \( N_{k,b}(n) \) and \( N_{k,p}(n) \) satisfy following

\[
N_{k,b}(n) + N_{k,p}(n) = N_k(n)
\]

(17)

\[
N_{k,p}(n) = \alpha(n) N_k(n)
\]

(18)

where \( \alpha(n) \) denotes the portion of users whose data bursts are indicated by piggyback MAP IEs in the \( n \)-th frame, which generally changes in time because it depends on the time-varying channel state. However, when \( N_k(n) \) is sufficiently large, \( \alpha(n) \) asymptotically approaches \( E[\alpha(n)] \) by the central limit theorem. We further express \( E[\alpha(n)] \) as \( \alpha \).

Given that \( K \) is sufficiently large, \( N_{k,p}(n) \) is also sufficiently large. By the central limit theorem, (16) can be approximated as following

\[
R_{m,p}(n) \simeq N_{k,p}(n) E[b_m d_i(n)] = N_{k,p}(n) b_m \overline{d}.
\]

(19)

Applying (14)-(19) to (4), \( V \) is evaluated as following

\[
V = E \left[ \frac{N_{k,p}(n) b_m \overline{d} + N_{k,b}(n) b_m d_m}{N_{k,p}(n) b_m \overline{d} + N_{k,b}(n) b_m d_m + N_k(n) b_d} \right]
\]

(20)

\[
= E \left[ \frac{\alpha b_m \overline{d} + (1 - \alpha) b_m d_m}{\alpha b_m \overline{d} + (1 - \alpha) b_m d_m + b_d} \right]
\]

(21)

\[
= 1 - \frac{\alpha b_m \overline{d} + (1 - \alpha) b_m d_m}{\alpha b_m \overline{d} + (1 - \alpha) b_m d_m + b_d}
\]

(22)

Here, \( \alpha b_m \overline{d} + (1 - \alpha) b_m d_m \) in (22) represents the average portion of the resources that is allocated to MAP transmission. As we have evaluated in (9), the \( b_m d_m \) is the average portion of the resources allocated to MAP transmission in case of the conventional scheme. Note that \( b_m d_m > \alpha b_m \overline{d} + (1 - \alpha) b_m d_m \), because \( d_m \) is generally greater than \( \overline{d} \). Thus, \( V \) in the proposed scheme is smaller compared with \( V \) in case of the conventional scheme.

IV. PERFORMANCE EVALUATION

We performed simulations based on [12] and the parameters, which are given in Table I. The channel model is set as a Pedestrian B 3km/h model, whose details are described in [13].

Fig. 4 shows the results for MAP overhead for various average SINRs and average packet sizes. The numerical results are almost the same as the simulation results. In all cases, the MAP overhead decreases as the average size of data packets increases, because the number of resource units allocated to data transmission increases and the portion of the resources allocated to MAP transmission consequently decreases. In case of a conventional scheme, the MAP overhead increases as the average SINR increases, because the number of scheduled users increases and the number of MAP IEs consequently increases. In addition, the results show that the MAP overhead is always smaller when the proposed scheme is applied compared to the case when the conventional scheme is used. Note that the MAP overhead does not change much with respect to the average SINR. This is because when the average SINR is sufficiently high, \( \alpha \approx 1 \) and \( V \) approaches \( 1 - \frac{\overline{d}}{(b_m + \overline{d})} \). In this case, the MAP overhead tends to be independent of the average SINR.

Fig. 5 shows maximal throughput with respect to various average SINRs and average packet sizes. Because the numerical results are almost the same as the simulation results, we can conclude that our MAP overhead model is well-defined to represent system throughput. The results show that the proposed scheme can improve system capacity by up to 50% and 75% when the average SINR is 8dB and 15dB, respectively. The performance improvement is greater in case of small packet size, because MAP overhead is sufficiently large in this case and can be reduced much more by applying the proposed scheme. Thus, it is effective to apply the proposed scheme when traffic generates small packets, such as VoIP or video streaming. In fact, it is not needed to suppress MAP overhead in case of traffic with large packet size, because as referred
in the last paragraph of section II, MAP overhead problem is not critical in this case.

V. CONCLUSIONS

Reducing MAP overhead is an important issue in the IEEE 802.16e system and can be a critical factor in view of system performance. To determine the effect of MAP overhead, we analyzed how MAP overhead affects the system performance and showed that a low rate of transmission for MAP messages is the main factor that causes huge MAP overhead. To reduce the MAP overhead, we proposed a piggybacked MAP IE that is combined with data packets and transmitted at a higher data rate. We showed that the proposed scheme effectively reduces the MAP overhead and improves system throughput.

In addition, MAP overhead will be more improved when the proposed scheme is combined with a certain data compression scheme. Meanwhile the proposed scheme reduces the amount of resource assigned for MAP transmission, the data compression scheme can reduce amount of bits for MAP message. Combining the data compression schemes to our proposed scheme can be a further work.

REFERENCES