

RESEARCH

Open Access

Cross-layer design for radio resource allocation based on priority scheduling in OFDMA wireless access network

Yen-Wen Chen*, Chang-Wu Chen and Yi-Shiou Lin

Abstract

The orthogonal frequency-division multiple access (OFDMA) system has the advantages of flexible subcarrier allocation and adaptive modulation with respect to channel conditions. However, transmission overhead is required in each frame to broadcast the arrangement of radio resources to all mobile stations within the coverage of the same base station. This overhead greatly affects the utilization of valuable radio resources. In this paper, a cross layer scheme is proposed to reduce the number of traffic bursts at the downlink of an OFDMA wireless access network so that the overhead of the media access protocol (MAP) field can be minimized. The proposed scheme considers the priorities and the channel conditions of quality of service (QoS) traffic streams to arrange for them to be sent with minimum bursts in a heuristic manner. In addition, the trade-off between the degradation of the modulation level and the reduction of traffic bursts is investigated. Simulation results show that the proposed scheme can effectively reduce the traffic bursts and, therefore, increase resource utilization.

Keywords: scheduling, mapping, OFDMA, overhead, QoS, WiMAX

1. Introduction

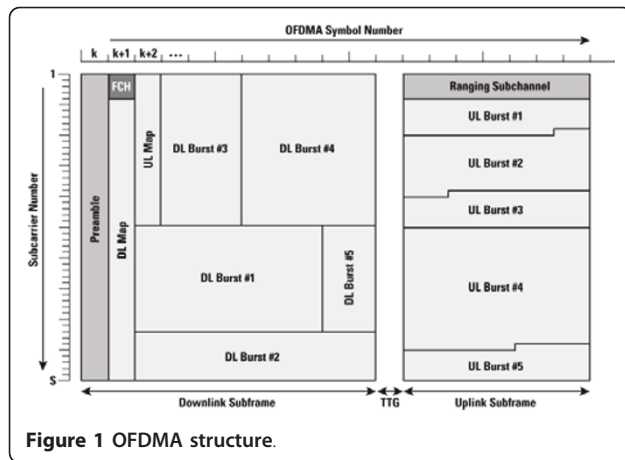
Channel quality is the basis of radio resource allocation for QoS traffic streams in OFDMA systems. The radio resources allocated and the modulation scheme adopted for downlink and uplink transmissions are adaptively adjusted by the base station (BS) in accordance with the required bandwidth and the channel condition of each receiving station [1,2]. The use of adaptive modulation can improve the transmission performance and throughput, especially when the channel quality is unstable. Generally, the issues of QoS scheduling and resource allocation are separated in their functions but tightly correlated in performance. The scheduling algorithm decides which traffic has the higher priority to use the network resources, while the resource allocation algorithm deals with the distribution of network resources. In the case of OFDMA, because the available resources will be affected by the channel conditions and the overhead of the control and management information, base

stations must deal with these two issues in a cooperative way.

The OFDMA system divides the transmit channels into several orthogonal subchannels, and each subchannel is composed of subcarriers. Three basic kinds of subcarrier allocation schemes, partial usage of subchannel (PUSC), full usage of subchannel (FUSC), and adaptive modulation and coding (AMC), are defined in IEEE 802.16 [3,4]. The PUSC and FUSC are diversity (or distributed) type subcarrier permutation schemes and AMC is a contiguous (or adjacent) type subcarrier permutation scheme. Generally, the diversity subcarrier permutation performs well in a high speed mobile environment while the contiguous subcarrier permutation is suitable for fixed or low speed applications. The radio resources of the OFDMA system can be constructed as a two-dimensional matrix as shown in Figure 1: the number of subchannels by the number of symbols. Both uplink and downlink subframes include data bursts of different types from multiple users.

This matrix can be referred to for the resource allocation of traffic streams with various kinds of QoS. Recently, based on the standard of IEEE 802.16/802.16e

* Correspondence: ywchen@ce.ncu.edu.tw
Department of Communication Engineering, National Central University,
Taiwan



[3,4], Worldwide interoperability for microwave access (WiMax) has been regarded as one of the most appropriate technologies for the next generation of broadband wireless access, using OFDMA for efficient transmission between the BS and mobile stations (MS). In order to provide QoS, WiMax adopts a connection-oriented approach at its link layer. The establishment of each connection between the MS and BS is admitted by the BS, and the BS takes care of the resource allocation for each connection in a centralized manner [5,6]. The BS arranges radio resources in accordance with the QoS of each traffic stream and the channel conditions. Several schemes have been proposed to study the scheduling efficiency of QoS traffic in OFDMA based networks [7-10]. Because the channel condition is time-varying, the BS must choose the proper subchannels and a suitable modulation scheme for each MS. Best channel first (BCF) scheduling [10] with the best channel first scheme selects the user who has the best average received SNR among the available subchannels to transmit data. Although this scheme can achieve better total throughput, the QoS of connections may not be satisfied. In [9,11], a cross-layer approach was proposed to assign priority to each connection, and the priority factors were calculated according to the QoS requirement and channel condition of each connection. After the arrangement of radio resources in accordance with these priorities, the information of resource arrangements for connections in each frame is broadcasted by the BS through the downlink MAP (DL_MAP) and uplink MAP (UL_MAP) fields of the frame. The information in the DL_MAP and UL_MAP is required to be referenced by each MS for receiving and transmitting its data frames. However, the transmission of the MAP information may introduce large overhead of the downlink channel if the traffic bursts for each MS are not properly mapped into subchannels [12,13]. It was indicated in [13] that the throughput behavior of an OFDMA

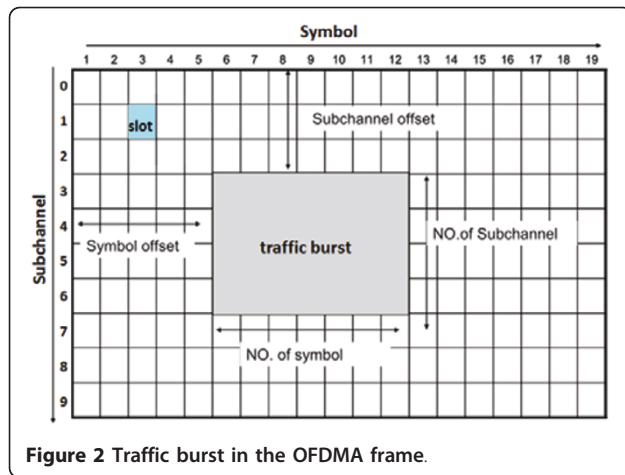
system is significantly influenced by the signaling overhead and that neglecting the signaling overhead leads to wrong performance conclusions. Furthermore, it was shown that the MAP messages occupy up to 20-60% of downlink resources [12]. Therefore, the mapping of traffic into bursts is a crucial issue for resource utilization in OFDMA systems.

In this paper, a novel burst mapping algorithm for downlink traffic, which considers the channel quality, coding and modulation, and the traffic priority, is proposed to reduce the size of MAP. The proposed scheme deals with the burst mapping in a cross layer manner for the purpose of improving resource utilization. In order to reduce the size of the MAP message, the proposed scheme utilizes the concept of "target side" with a flexible boundary adaptation to effectively fit the traffic in rectangular blocks so that the number of traffic bursts can be minimized. In addition, it is known that degrading the modulation level will exhaust more subchannels. However, in some cases, it may be more helpful to fit the downlink traffic of MS into a rectangular subchannel block so that the number of traffic bursts can also be minimized. It is also possible to increase the resource utilization if the modulation level is properly degraded. This trade-off issue is also analyzed.

This paper is organized as follows. The overview of WiMax access technology and the overhead analysis of MAP are described in the following section. In Section 3, the burst mapping algorithm is proposed. The influence of the radio resource utilization for the degradation of the modulation level is also analyzed. The simulation results of the proposed algorithm are illustrated and discussed in Section 4. Finally, the conclusions are provided in the last section.

2. MAP overhead of WiMax access

Each WiMax connection obtains a connection identification (CID) from the BS when it is admitted to the network. The BS then allocates appropriate resources for each connection in accordance with its desired QoS. Resource allocation can be divided into uplink and downlink. The BS informs the MS using the fields of UL_MAP and DL_MAP, for which a traffic burst is allocated for the transmission and receipt of each MS. In OFDMA, although the subcarrier allocation schemes may be different, the radio resources allocated in one frame can be conceptually regarded as the collection of a number of slots, where each slot is formed by subchannels and OFDMA symbols. According to [3,4], the numbers of symbols accommodated by one slot can have different arrangements for PUSC, FUSC, and AMC. For the example shown in Figure 2, there are one symbols included in one slot because DL FUSC is divided into slots of one symbol by one subchannel.



Each traffic burst, depending on its number of bits to be delivered and the modulation scheme adopted, may consist of one or more than one slot. However, these slots must be represented by a rectangular shape so that the BS can easily specify the range of the traffic burst in the DL_MAP. In WiMax specifications [3,4], each traffic burst is determined by the symbol offset, subchannel offset, number of symbols, and number of subchannels, as shown in Figure 2.

For the resource allocation of the uplink, the BS periodically polls mobile stations for the bandwidth request of each connection, except for the connections with unsolicited grant service (UGS) because UGS is a constant bit rate service; therefore, the BS reserves the bandwidth of UGS connections in advance. Each connection issues the bandwidth request (if it demands uplink bandwidth) to the BS when receiving the polling message. Based on the bandwidth requests, the BS allocates the radio resource for each connection according to the priority of each connection and the channel condition of the MS. Also, one MS may establish more than one connection for different services simultaneously. For efficiency, the BS aggregates the bandwidth allocated for the connections of the same MS into a traffic burst for transmission because the connections of the same MS get the same channel condition. Thus, for uplink transmission, the BS allocates the radio resource via each mobile station basis, and the resource allocation for connections within the same MS is the responsibility of the MS.

For the downlink transmission, because the current traffic condition of each connection, e.g., buffered packets and quality of service, is known by the BS, the BS can dominate the resource allocation of each connection. In order to satisfy the QoS desired by each connection and to optimize the utilization of radio resources, more than one traffic burst may be arranged. Thus, for

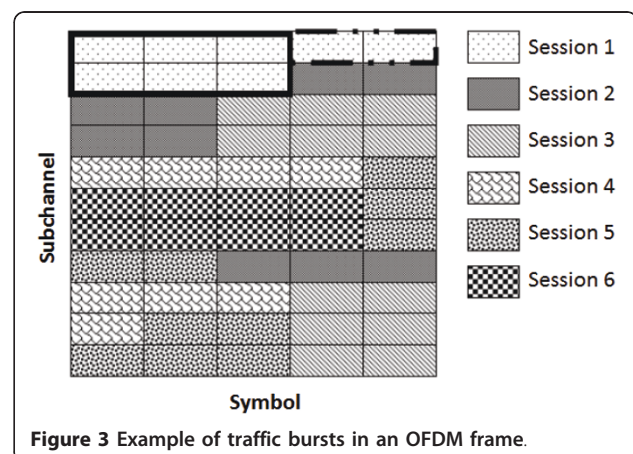
downlink transmission, the BS allocates the radio resources on a per connection basis. If more than one connection (CID) exists in a single MS, ideally, it would be possible to aggregate the traffic of connections belonging to the same MS into one traffic burst. The advantage of aggregating traffic into one traffic burst is to reduce the number of traffic bursts so that the overhead in DL_MAP can be minimized.

In accordance with the frame format of WiMax specifications [3,4], the number of bits, b , required in a DL_MAP to specify the assignment of traffic bursts can be stated as

$$b = 104 + \sum_{i=1}^n (44 + 16C_i) \quad (1)$$

where n is the number of traffic bursts within a frame and C_i is the number of CIDs associated with the traffic burst i . It is easy to understand that at least 60 bits of overhead are required for each additional traffic burst. Inappropriate allocation of time slots for the required bandwidth of each connection leads to more traffic bursts within the OFDMA frame and introduces more overhead in the DL_MAP field. For example, as shown in Figure 3, slots are allocated to six traffic sessions according to their channel conditions and bandwidth needs. The ideal scheme would allocate one burst for each traffic session; however, in this case, there are a total of 15 traffic bursts formed due to inappropriate allocation. Note that those slots which are not rectangle block are viewed as different traffic bursts.

In accordance with WiMax specifications [3,4], and assuming each burst contains the traffic of only one connection, it will require 1,004 bits to specify the 15 traffic bursts in the DL_MAP. However, only 464 bits are needed if six traffic bursts are used. The difference of the DL_MAP between these two assignments is 600 bits. Note that the information of DL_MAP is conveyed



using broadcasted CID, and the lowest modulation scheme, e.g., BPSK, would be adopted so that all mobile stations could successfully receive it. As a result, more radio resources would be exhausted in this scheme as compared to transport CID. To remedy this problem, it is the objective of this paper to study the efficient allocation algorithm in a cross-layer manner so that the overhead can be minimized.

3. Target side-based resource allocation scheme

As mentioned in the previous section, reducing the number of traffic bursts can minimize the overhead introduced in the DL_MAP. In addition to effective resource allocation, the resources should be allocated in a prioritized manner so that the QoS connections can receive their desired quality. In [11], the scheduling priorities of real time polling service (rtPs), non-real time polling service (nrtPs), and best effort (BE) traffic were derived by considering the expected delay, channel condition, and fairness. However, the arrangement of traffic bursts, or block mapping, was not considered. In this paper, we focus on the issue of block mapping, and the above scheme is adopted to decide the scheduling order of traffic flows in the WiMax frame. As the radio resource can be allocated by subchannels in the OFDMA system, the subchannels with highest modulation level will be considered to be allocated for that mobile station.

The "target side"-based allocation (TSA) scheme for the OFDMA system is proposed to satisfy the above objective. In addition, a more heuristic scheme, TSA with flexible modulation (TSA-FM), which considers the trade-off of overhead caused by the increase in traffic burst number and the bandwidth loss caused by the degradation of modulation level, is provided to further improve the utilization of radio resources.

3.1. TSA scheme

The radio resource to be allocated in one frame can be formed into a two-dimensional array of slots. In order to increase the resource utilization, the BS decides which subchannel(s) could support the highest modulation level for the MS with the highest scheduling priority by referring to its channel condition. Then, the allocation of slots is performed from left to right of the selected subchannels within the two dimensional slots map. Some slots of a subchannel may have been allocated to other MS with higher priority when an MS is allocated for the same subchannel. The residual slots of a subchannel may not be sufficient to provide enough bandwidth for a given MS. Without the appropriate arrangement, this would require more traffic bursts for a specific session. The most common scheme, or normal scheme, is to allocate slots in the sequence of the

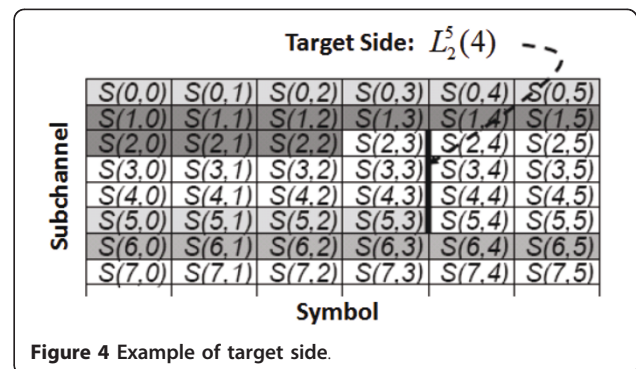
selected subchannels. For the example of session 1 shown in Figure 3, eight slots are needed to convey the data with subchannels 1 and 2 as preferences in accordance with the channel condition. Thus, five slots are allocated in subchannel 1 first, and the other three slots are allocated in subchannel 2. This introduces two traffic bursts. If the first four slots are allocated in both of subchannels 1 and 2, then only one traffic burst will be required. In order to arrange the slots of an MS with a rectangular shape, instead of allocating the slots in a per subchannel basis, the target side is applied as a reference boundary of consecutive subchannels for the allocation of slots. Consider a two dimensional array of slots where $S(i, j)$ denotes the slot located at the i th row (or the i th subchannel) and the j th column (or the j th symbol). The target side is defined as the leftmost vertical line with a number of consecutive subchannels of the two-dimensional array so that the slots to the right of target boundary of those consecutive subchannels are all available for allocation. Let $S(i, j) = 0$ denote an available slot, and let $S(i, j) = 1$ mean an allocated slot. Then, for the set of consecutive subchannels from i_1 to i_2 , it represents as $SUB(i_1, i_2) = \{i_1, i_1+1, \dots, i_2-1, i_2\}$, the leftmost position x can be defined as

$$x \equiv \cap[SUB(i_1, i_2)] \quad (2)$$

where the operator \cap on $SUB(i_1, i_2)$ finds the leftmost common position of the consecutive subchannels such that

$$S(j, k) = 0, \quad \forall(j, k) \quad i_1 \leq j \leq i_2 \quad \text{and} \quad x \leq k \quad (3)$$

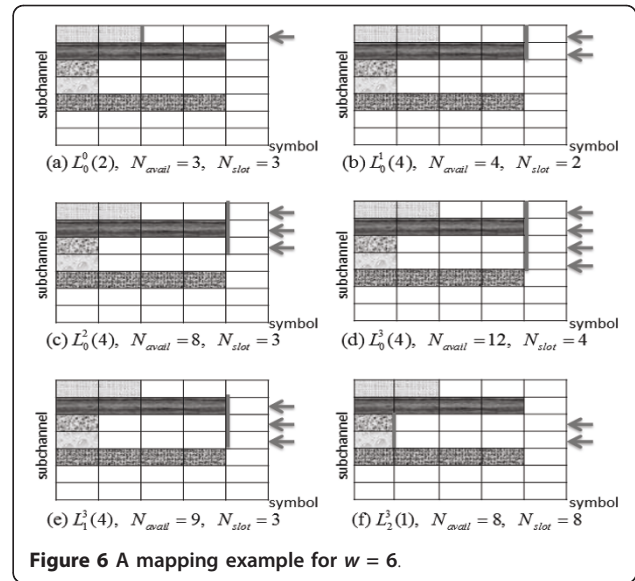
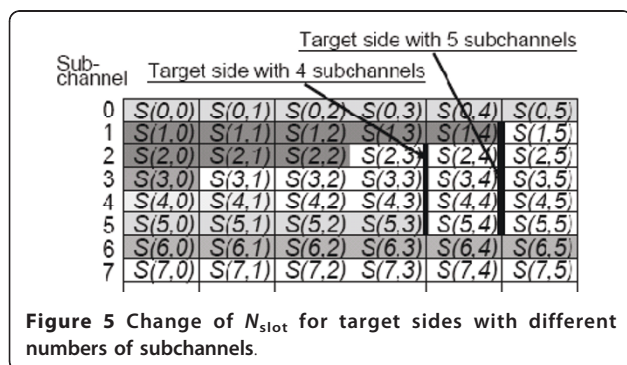
k is the rightmost position of the column. The target side is then denoted $L_{i_1}^{i_2}(x)$ over consecutive subchannels i_1 and i_2 at the position x . For the example shown in Figure 4, where the blank (or white) slots represent the available slots, the leftmost position x from subchannels 2 to 5, $\cap[SUB(2,5)]$, is equal to four. Hence the target side $L_2^5(4)$ indicates that slots (2, 4), (2, 5), (3, 4), (3, 5), (4, 4), (4, 5), (5, 4), and (5, 5) are available for



allocation. These slots form a 4×2 rectangular area that is the mapping of a traffic burst.

The target side is flexible to allow the subchannels of the MS to be allocated, and which subchannels are appropriate for the transmission of the MS is dependent on its channel condition. As mentioned above, the scheduling priorities of each session are determined by the expected delay, channel condition, and fairness, as proposed in [11], and this paper focuses on the allocation of slots of traffic bursts. Assume that the bandwidth required of the session, which will be scheduled, is w slots with respect to the modulation level it will use for transmission. And let M be the set of subchannels that are applicable for the use of the modulation level decided for that session according to the channel condition of the associated MS. Then, for a $m \times n$ (the number of subchannels by the number of symbols) slots matrix, the basic concept of the proposed TSA scheme is stated as follows.

First, the proposed algorithm in line 1 determines whether the traffic burst for the desired bandwidth w is found or not by examining the number of available slots bounded by the target side (N_{slot}) and the factor relationship between the required bandwidth w and i th subchannel (N_{sub}). It is not always true that w slots with a rectangular shape can be found when N_{slot} is greater than w . There are two procedures, *re_target_side* and *normal_mapping*, in the algorithm to allocate available slots and to re-adjust the target side. The *normal_mapping* procedure in line 12 of the TSA scheme is a straightforward slot mapping scheme that allocates the scheduled session with slots of the appropriate subchannel(s) in sequence [11]. This procedure is only applied when the proposed scheme cannot find available slots formed by a rectangular shape for that session. The procedure in line 15 of *re_target_side* is designed to back down some subchannels with less available symbols so that the position of the target side x can be smaller and the value of N_{slot} can be larger. Thus, the total number of available slots are not bounded by the target side is examined to search for this possibility. For the example shown in Figure 5, it is assumed that the subchannels



from 1 to 5 are suitable for the session. The value of N_{slot} is five for the target side with five subchannels (i.e., $L_1^5(5)$); while it becomes eight if subchannel 1 can be backed down (i.e., $L_2^5(4)$).

In order to judge whether the abandonment of a subchannel is worthwhile, a heuristic approach is applied. The procedure of *re_target_side* backs down subchannel i_1 and sets i_1 to be i_1+1 if the residual number of slots, after the abandonment of this subchannel, is greater than w . For the example in Figure 5, the total number of available slots which are blank for subchannels 1 to 5 is 15, and it is 14 after the abandonment of subchannel 1. Subchannel 1 will be discarded if the required number of slots is less than 14 in our approach. This arrangement will increase the value of N_{slot} from 5 to 8.

An illustrative example of the mapping procedure is shown in Figure 6. It is assumed that the required bandwidth, w , is six slots. The mapping starts from subchannel 0, and the total number of available slots which are blank is 3 as indicated. As the number of slots in subchannel 0 is not sufficient for allocation, subchannel 1 is included. Although the total number of available slots of subchannels 0 and 1 is 4, N_{slot} becomes two because $\cap[\text{SUB}(0,1)]$ is equal to 4 and the target side is $L_0^1(4)$. When subchannel 2 is included, $\cap[\text{SUB}(0,2)]$ is also equal to 4, and it still cannot allocate the six slots in one rectangular block. Although the total number of available slots from subchannel 0 to 2 is 8, the *re_target_side* procedure is not invoked. The reason is that the abandonment of subchannel 0 would result in the total number of available slots being 5, which is less than w . The *re_target_side* procedure is performed when subchannel 3 is included. After the abandonment

of subchannels 0 and 1, $\cap[\text{SUB}(2,3)]$ becomes one and the target side moves back to $L_2^3(1)$. Then the values of N_{slot} and N_{sub} are 8 and 2, respectively, and finding available slots is satisfied so the required bandwidth can be allocated in one traffic burst.

Algorithm: TSA scheme

Input: a session that requires w slots in m -by- n slots matrix

Output: the allocation of w slots in m -by- n slots matrix

Initialize (preparation):

Set $i_1 = i_2 = 0$, $i_1, i_2 \in M$ $N_{\text{avail}_{i_1}^{i_2}}$ is the total number of available slots from subchannel i_1 to i_2 . N_{sub} is the number of successive subchannels. N_{slot} is the number of available slots from subchannel i_1 to i_2 based on target side.

Procedure TSA(w)

1. **if** ($N_{\text{slot}} \geq w$ && $w \bmod N_{\text{sub}} = 0$)
2. $i' \in [i_1, i_1 + N_{\text{sub}} - 1]$
3. $x' \in [x, x + w/N_{\text{sub}} - 1]$
4. **else**
5. **if** ($N_{\text{avail}_{i_1}^{i_2}} \leq w$)
6. **if** ($i_2 + 1 \in M$)
7. Set $i_2 = i_2 + 1$, $x = \cap[\text{SUB}(i_1, i_2)]$
8. $N_{\text{slot}} = N_{\text{sub}} \cdot (n - x)$
9. **return** TSA(w)
10. **elseif** ($i_2 + 1 \notin M$)
11. There is no appropriate subchannel in M
12. **return** normal_mapping
13. **end**
14. **else**
15. **return** re_target_side(w)
16. **end**
17. **End**

Algorithm: re_target_side(w)

Input: a session that requires w slots in m -by- n slots matrix

Output: adjusted target side $L_{i_1}^{i_2}(x')$

i_1' is the update of i_1 , x' is new target side.

Procedure re_target_side(w)

1. **while** ($i_1 \leq i_2$)
2. $i_1 = i_1 + 1$, $L_{i_1}^{i_2}(x')$, $N'_{\text{slot}} = N_{\text{sub}} \cdot (n - x')$
3. **if** ($N_{\text{avail}_{i_1}^{i_2}} \geq w$)
4. **if** ($N'_{\text{slot}} > N_{\text{slot}}$)
5. abandon the subchannel $i_1 - 1$
6. Set $i_1' = i_1$, $L_{i_1}^{i_2}(x')$
7. **break**
8. **end**
9. **end**
10. **end**

3.2. TSA-FM scheme

It is obvious that if more subchannels could be adopted for the allocation, the possibility of arranging one traffic burst for the session under scheduling would increase. One way to increase the number of subchannels for allocation is to decrease the modulation level. For example, in accordance with the channel condition, there are ten subchannels for allocation using 64 QAM. And, if 32 QAM is adopted, five more subchannels might be available for this session, and the total number of appropriate subchannels for allocation would increase to 15. However, the number of bits conveyed by one slot would be decreased when the modulation is downgraded from 64 QAM to 32 QAM. More slots are required to convey the data of this session because of the decrease of spectral efficiency. Although the overhead of the DL_MAP field decreases as the number of traffic bursts decreases when 32 QAM is adopted, more radio resources are required for this session when compared with a session with a higher modulation scheme. The objective of the proposed TSA-FM scheme is to consider whether it is possible to gain further benefit of resource utilization through the degradation of modulation level based on the above phenomenon.

From the resource utilization point of view, the adjustment of modulation level is a trade-off issue. In order to finely compare the sacrificed bandwidth caused by the degradation of modulation level and the extra overhead of DL_MAP introduced by additional traffic burst, the analysis of resource utilization was performed. Let $\text{Cost}_{\text{DL_MAP}}$ and $\text{Cost}_{\text{modulation}}$ be the extra bandwidth needed in DL_MAP, due to the additional traffic burst(s), and the decreased bandwidth, due to the degradation of modulation level, respectively. It is the objective for the degradation of modulation level to have $\text{Cost}_{\text{modulation}}$ be less than $\text{Cost}_{\text{DL_MAP}}$. The $\text{Cost}_{\text{modulation}}$ can be calculated from

$$\text{Cost}_{\text{modulation}} = (b_{\text{before}} - b_{\text{after}})w_{\text{after}} \quad (4)$$

where b_{before} and b_{after} denote the numbers of bits that can be accommodated by one slot of the original modulation scheme and the modulation scheme to be degraded, respectively. The number w_{after} indicates the number of slots required to convey the traffic of the session when the degraded modulation scheme is adopted. For example, by assuming each slot consists of 48 subcarriers and one symbol, the number of bits carried on a slot with 64 QAM_{3/4} modulation scheme is 216 bits/slot, and it would be 192 bits/slot if 64 QAM_{2/3} is used. Then, for the transmission of 2160 bits, 10 slots are required for 64 QAM_{3/4} modulation scheme; however, it needs 12 slots for 64 QAM_{2/3} modulation scheme. The cost, due to the degradation of the

modulation level, is 288 (i.e., $(216 - 192) \times 12$) bits. The value of $\text{Cost}_{\text{DL_MAP}}$ is derived from

$$\text{Cost}_{\text{DL_MAP}} = (b_{h\text{-modu}} - b_{\text{MAP-modu}})(w_{\text{MAP-before}} - w_{\text{MAP-1}}) \quad (5)$$

where the value of $w_{\text{MAP-1}}$ represents the number of slots needed for broadcasting the resource allocation information in DL_MAP by assuming that only one traffic burst is required after the degradation of modulation level. The value of $w_{\text{MAP-before}}$ is the number of slots required in the DL_MAP when the modulation level is not degraded. $b_{h\text{-modu}}$ and $b_{\text{MAP-modu}}$ denote the numbers of bits that can be carried in one slot for the highest modulation level adopted by the session and the modulation scheme used in transmitting MAP information, respectively. Smaller number of subchannels can be used for allocation if the degradation of the modulation level is not performed, but more traffic bursts will be required. The number of bits required in the DL_MAP can be calculated according to Equation 1. For example, one traffic burst with three CIDs needs 196 (i.e., $104 + (44 + 16 \times 3)$) bits. It is noted that lower modulation level must be applied to guarantee the DL_MAP information can be broadcasted to all mobile stations successfully. Therefore the number of bits conveyed by one slot is limited. If the QPSK1/2 modulation is applied, only 48 bits can be transmitted in one slot. It requires five (i.e., $\lceil 196/48 \rceil$) slots to carry the resource information in DL_MAP for one traffic burst. If it needs two traffic bursts without degrading the modulation level, then the total number of bits required is 288 (i.e., $104 + 2 \times (44 + 16 \times 3)$) bits. The number of required slots in DL_MAP is 6 (i.e., $\lceil 288/48 \rceil$) slots. The cost of DL_MAP, $\text{Cost}_{\text{DL_MAP}}$, is 168 (i.e., $(216 - 48) \times (6 - 5)$) bits. Note that the increase of bits in the DL_MAP not only depends on the number of traffic bursts, but also the number of CIDs accommodated in one traffic burst. If there are 5 CIDs in the traffic burst, an increment of 124 bits is required for 1 additional traffic burst. Thus, for the above example with five CIDs, 320 bits are required to carry the resource allocation information and the number of required slots in DL_MAP becomes 7.

As mentioned above, the degradation of the modulation level has the advantage of decreasing the number of traffic bursts at the expense of spectral utilization. An appropriate degradation of modulation level shall be under the constraint of $\text{Cost}_{\text{modulation}} > \text{cost}_{\text{DL_MAP}}$. In Equation 5, the value of $w_{\text{MAP-before}}$ is determined by knowing the number of traffic bursts for the session under scheduling using the original modulation level. However, it is noted that the proposed TSA scheme is designed for mapping the required bandwidth into single traffic burst; otherwise, the procedure of *normal-mapping* is performed. In order to reduce the computing complexity, the concept of backtracking is not considered

in the proposed scheme, and it is not possible to know the number of traffic bursts in advance. Therefore, it is also not easy to predict which modulation level should be degraded for an optimal solution. A heuristic approach is to assume the maximum number of traffic bursts as the reference bound for degradation. Thus, for a session with the demand of w slots using the original modulation scheme, the maximum number of traffic bursts, which occurs when each slot is arranged as one traffic burst, is w . The value of $\text{Cost}_{\text{DL_MAP}}$ of Equation 5 can then be obtained accordingly. The degradation of modulation level could be controlled subject to

$$\text{Cost}_{\text{modulation}} < \text{Cost}_{\text{DL_MAP}} \quad (6)$$

For the $a \times b$ slots matrix with a subchannels, the mapping of the proposed algorithm starts from the first appropriate subchannel in the sequence without backtracking, and whether a complete traffic burst can be found is determined after all appropriate subchannels are examined. In contrast, the *re_target_side* procedure searches for a appropriate target side after the abandonment of the subchannel. Therefore, the computing complexity of the proposed TSA algorithm is $O(a^2)$. And for the TSA-FM scheme, the degradation of modulation level will be considered when the required bandwidth cannot be mapped to a traffic burst successfully. The worst case is that it would try all of the modulation levels that are lower than its original modulation. Since the number of modulation levels is fixed, its computing complexity is also $O(a^2)$.

4. Experimental results

In order to investigate the performance of the proposed scheme, simulations were performed to compare the efficiency of the TSA scheme and the traditional best channel first (normal) mapping scheme. The OFDMA parameters applied during the simulation is listed in Table 1. Both of 12-subchannel and 48-subchannel configuration types were considered, and each slot was assumed to consist of three symbols. These arrangements form the 12×5 slots and 48×5 slots in one OFDMA frame. Each slot of the 2 configurations consists of 192 and 48 subcarriers.

The traffic sources generated for simulations consist of three kinds of delivery classes: rtPs, nrtPs, and BE, with different QoS parameters. Each delivery class and its associated QoS parameters are stated in two scenarios, as shown in Table 2. Scenario 1 was applied to

Table 1 OFDMA parameters applied for simulations

System	FFT	Frame	DL/UL ratio	CP	BW	Tx Power
TDD	1024	5 ms	50%/50%	1/8	7 MHz	22 dBm

Table 2 Traffic sources adopted for simulation

Scenarios	Delivery class	QoS parameters (number of sources)
Scenario 1 (ten traffic sources)	rtPs	0.64 Mbps with 50 ms max. delay (2)
	rtPs	0.32 Mbps with 20 ms max. delay (2)
	nrtPs	0.3 Mbps (1); 0.5 Mbps (1); 0.7 Mbps (1)
	BE	0.2 Mbps (1); 0.4 Mbps (1); 0.6 Mbps (1)
Scenario 2 (20 traffic sources)	rtPs	0.64 Mbps with 50 ms max. delay (2)
	rtPs	0.32 Mbps with 30 ms max. delay (2)
	rtPs	0.16 Mbps with 20 ms max. delay (2)
	nrtPs	0.5 Mbps (2); 0.7 Mbps (2); 0.9 Mbps (2); 1.1 Mbps (2)
	BE	0.6 Mbps (2); 0.8 Mbps (2); 1.0 Mbps (2)

examine the allocation efficiency of traffic bursts, and scenario 2, which generates much heavier traffic load than that of scenario 1, was performed to measure the performance of resource utilization.

During the simulations, the jakes model was adopted to emulate the channel environment. The average number of traffic bursts and the overhead of DL_MAP in one frame for the proposed TSA scheme, which does not consider the flexible modulation level adjustment, and the normal mapping schemes are compared in Table 3. The scenario 1 traffic load was offered for simulations.

As expected, the simulation results show that the proposed scheme utilizes lower average numbers of traffic bursts than that of the normal mapping scheme and the overhead in the DL_MAP of the proposed scheme is also smaller. It is noted that scenario 1 generates ten traffic sources for simulation. Hence the minimum (or optimal) number of traffic bursts is 10 in one frame. According to the simulation results, the average traffic burst numbers of the proposed scheme are 10.61 and 10.56 for the 12 and 48 subchannels, respectively. They are very close to the above minimum number. It is also worth mentioning that the normal mapping scheme of the 48-subchannel case requires much more traffic bursts than the others. As indicated in the 48-subchannel case of Table 3, the average overhead of the normal mapping scheme is about 85% higher than the proposed TSA scheme. The reason is that each slot of the 48-subchannel case conveys less data than that of the 12-

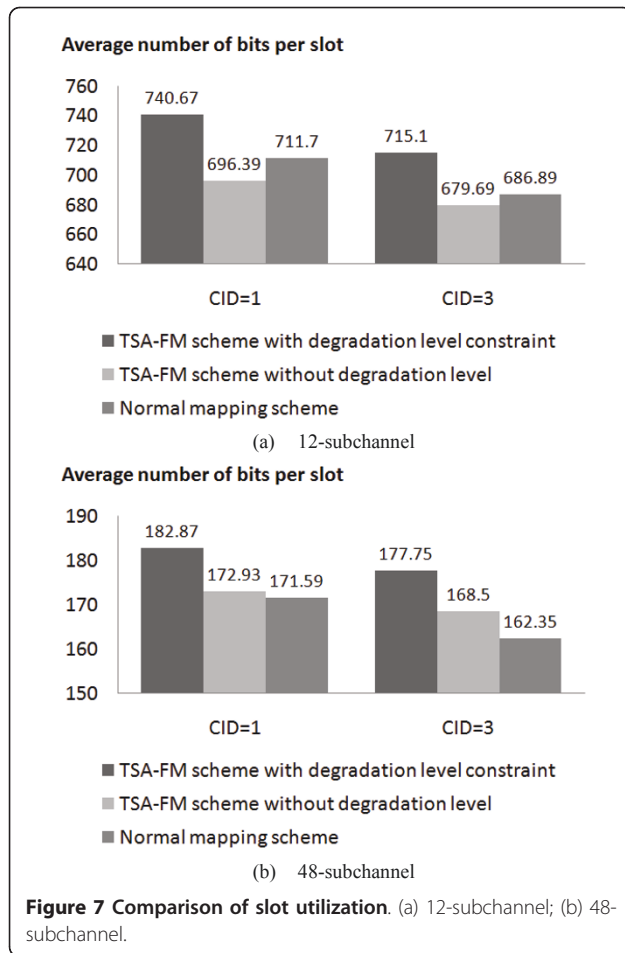
subchannel case. Also, more slots are required for the same bandwidth requirement. The normal mapping scheme always allocates the slots with the best channel of the session to be scheduled subchannel by subchannel without considering the proper mapping of the traffic burst. It tends to introduce fragmental slots and, as a result, more traffic bursts are required.

The effectiveness of the TSA-FM scheme is examined by providing a heavier traffic load (scenario 2) for simulation so that some sessions need to reduce the modulation level to achieve fewer traffic bursts. In addition to the comparison with the normal mapping scheme, the effect of constraining the modulation level using Equation 6 is also analyzed.

Figure 7 shows that the average number of bits can be accommodated by one slot for the TSA-FM with and without degradation level constraint approaches and the normal mapping scheme under different numbers of subchannels and CID. The average number of bits per slot is calculated by the division of the total number of bits, including data and the DL_MAP, and the number of slots for downlink. It is obvious that the utilization of the proposed TSA-FM scheme with degradation level constraint is superior to that of the normal mapping scheme. Thus, an appropriate decrease of modulation level and proper traffic burst allocation are helpful for the optimization of overall resource utilization. However, it is noted that, when compared to the normal mapping scheme, there is no benefit if there is no degradation level constraint. The utilization of the TSA-FM

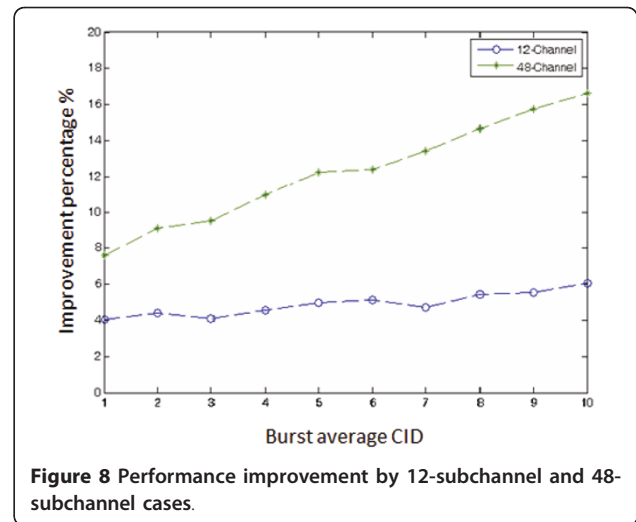
Table 3 Traffic bursts and DL_MAP overhead comparison

No. of subchannels	Allocation schemes	Average number of traffic bursts per frame	Average overhead in DL_MAP (obtained from eq.(1))
12	Proposed TSA scheme	10.61	740.65 bits
	Normal mapping scheme	12.25	839.17 bits
48	Proposed TSA scheme	10.56	737.59 bits
	Normal mapping scheme	24.04	1366.54 bits



scheme without the degradation constraint is even worse than the normal mapping scheme for the case of 12-subchannel. This coincides with the concern mentioned that the benefit gained from the proper burst arrangement may not compensate for the loss of utilization caused by the decrease of modulation level. For the case of 48-subchannel, the average number of bits per slot of the proposed TSA-FM scheme is higher than that of the normal mapping scheme, regardless of with or without degradation level constraint. The reason, which has been explained in Table 3, is that relatively large numbers of traffic bursts are generated due to the fragmental slots of the normal mapping scheme, and the overhead increased in DL_MAP is also comparatively high.

Note in Figure 7 that the number of CID accommodated by one traffic burst will affect the overall utilization. The utilization improvement by the proposed TSA-FM with the degradation level constraint scheme as compared to the normal mapping scheme for different numbers of CID is illustrated in Figure 8. It indicates that, even under the high traffic load, the improvements



for the 12-subchannel and 48-subchannel cases range from 4 to 6% and 8 to 16%, respectively.

5. Conclusions

In this paper, the influence of traffic burst allocation was studied, and a novel cross-layer design to improve the utilization of radio resource was proposed. The proposed TSA scheme decreases the transmission overhead by regularizing the radio resources for individual traffic bursts. The simulation results show that the required traffic bursts number of the proposed scheme is much less than that of the normal mapping scheme and is only a little higher than the optimal value when traffic load is not high. In addition, we introduced the concept of the adaptive decrease of modulation levels for better arrangement of traffic bursts to further improve resource utilization when traffic load is heavy. We also investigated the constraint of the degradation of modulation level. Experimental simulations were conducted to determine the performance improvement depending on the number of subchannels and the number of CID. The simulation results indicate that the influence of the traffic burst mapping is significant when the capacity of one slot is relatively much less than the desired bandwidth of the session to be allocated. This is because fragmental slots are more likely to occur in a normal mapping scheme, which requires more traffic bursts to be allocated for the same bandwidth. The simulation results also show that the overall utilization can be effectively increased if the modulation level decreases under the proposed constraint.

Abbreviations

AMC: adaptive modulation and coding; BCF: best channel first; BS: base station; CID: connection identification; DL_MAP: downlink MAP; FUSC: full usage of subchannel; MAP: media access protocol; MS: mobile stations;

OFDMA: orthogonal frequency-division multiple access; PUSC: partial usage of subchannel; QoS: quality of service; UGS: unsolicited grant service; UL_MAP: uplink MAP; TSA: target side-based allocation; TSA-FM: TSA with flexible modulation.

Acknowledgements

This research work was supported in part by the grants from the National Science Council (Grant numbers: NSC 97-2221-E-008-033, and NSC 98-2221-E-008-063).

Competing interests

The authors declare that they have no competing interests.

Received: 21 December 2010 Accepted: 5 July 2011

Published: 5 July 2011

References

1. M Sternad, D Aronsson, Channel estimation and prediction for adaptive OFDM downlinks. *IEEE VTC 2003-Fall*, Orlando, FL. (October, 2003)
2. KDG Li, H Liu, Computationally efficient bandwidth allocation and power control for OFDMA. *IEEE Trans Wirel Commun.* **2**(6), 1150–1158 (2003)
3. IEEE Std 802.16-2004, IEEE Computer Society, IEEE Microwave Theory and Techniques Society, IEEE Standard. (1 October, 2004)
4. IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor1-2005, IEEE Computer Society, IEEE Microwave Theory and Techniques Society, IEEE Standard. (28 February, 2006)
5. SH Ali, KD Lee, VCM Leung, Dynamic resource allocation in OFDMA wireless metropolitan area networks. *IEEE Wirel Commun.* **14**(5), 6–13 (2007)
6. TCH Alen, AS Madhukumar, F Chin, Capacity enhancement of a multi-user OFDM system using dynamic frequency allocation. *IEEE Trans Broadcast.* **49**(4), 344–353 (2003)
7. Y Ben-Shimol, I Kitroser, Y Dinitz, Two-dimensional mapping for wireless OFDMA systems. *IEEE Trans Broadcast.* **52**(3), 388–396 (2006)
8. C Cicconetti, C Eklund, L Lenzini, E Mingozzi, Quality of service support in IEEE 802.16 networks. *IEEE Netw Mag.* **20**(2), 50–55 (2006)
9. M Ergen, S Coleri, P Varaiya, QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems. *IEEE Trans Broadcast.* **49**(4), 362–370 (2003)
10. WS Jeon, DG Jeong, Combined connection admission control and packet transmission scheduling for mobile Internet services. *IEEE Trans Veh Technol.* **55**(5), 1582–1593 (2006)
11. Q Liu, X Wang, GB Giannakis, A cross-layer scheduling algorithm with QoS support in wireless networks. *IEEE Trans Veh Technol.* **55**(3), 839–847 (2006)
12. T Kwon, H Lee, S Choi, J Tim, D Cho, S Cho, S Yun, W Park, K Kim, Design and implementation of a simulator based on a cross-layer protocol between MAC and PHY layers in a WiBro compatible IEEE 802.16e OFDMA system. *IEEE Commun Mag.* **43**(12), 136–146 (2005)
13. J Gross, H Geerdes, H Karl, A Wolisz, Performance analysis of dynamic OFDMA systems with inband signaling. *IEEE J Sel Areas Commun.* **24**(3), 427–436 (2006)

doi:10.1186/1687-1499-2011-28

Cite this article as: Chen et al.: Cross-layer design for radio resource allocation based on priority scheduling in OFDMA wireless access network. *EURASIP Journal on Wireless Communications and Networking* 2011 **2011**:28.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com