Project	**IEEE 802.16 Broadband Wireless Access Working Group <http://ieee802.org/16>**
Title	**Contention Schemes For OFDM Mode A_L, Revision 0**
Date Submitted	**2002-01-04**
Source(s)	Jerry Krinock, Manoneet Singh, Mike Paff, Arvind Lonkar, Lawrence Fung Radia Communications, Inc. 275 N. Mathilda, Suite A Sunnyvale, CA 94086 USA Voice: 408-830-9726 extension 239 Fax: 408-245-099 mailto:jkrinock@radiacommunications.com
Re:	IEEE 802.16 Working Group, Letter Ballot #4, IEEE P802.16a/D1-2001
Abstract	This document addresses our concerns with the Bandwidth Request method in IEEE P802.16a/D1-2001, and contains revised text to improve it.
Purpose	The information should be considered with the comments to IEEE P802.16a/D1-2001.
Notice	
Release	
Patent Policy and Procedures	

Contention Schemes For OFDM Mode A_L

Jerry Krinock
Radia Communications, Inc.

1. Background

This contribution addresses improving the bandwidth-request method for mode OFDM A_L which is very briefly described in section 8.3.5.3.3.7.2, titled “Bandwidth requesting”. That section presently reads as follows [1]:

> *Bandwidth requests in OFDM are contention based, wherein regular uplink bursts shall be used for bandwidth requests. Bandwidth requests are further provisioned by a piggy-back mechanism provided by the MAC.*
>
> *The base station shall allocate a number of symbols every frame for bandwidth requests. This number of symbols shall be large enough to contain one or a multiple of long preamble uplink bursts with one OFDM symbol in data. SSs requiring bandwidth may, using a backoff mechanism, use these slots to request bandwidth*

With Internet Protocol (IP) traffic, many bandwidth requests may result from, say, a mouse click on a web page, which will request a very small uplink allocation, and to which piggy-backing is not well applicable, since web-surfing is not well predictable. Thus, it is expected that the average amount of bandwidth requested may be quite small, and that bandwidth requests may comprise a significant fraction of UL traffic. Under this scenario, a fast, efficient and robust bandwidth request mechanism will be important to the success of any mode in 802.16a.

We recognize the existing method, described above, as *slotted ALOHA*. Collision-free transmissions are received as valid requests by the BS, which can then allocate reserved slots for transmission during subsequent uplink frames. A similar method is used in the DOCSIS standard [6].

For maximum utililization of the contention slots, the MAC’s policies would attempt to allocate enough contention slots so that the probability of a successful contention (no collision), is $1/e = .368$. However, this will result in a rather high average delay, including an average 2.71 attempts for a successful contention. Also, a system tuned for this maximum utilization will quickly choke on a sudden, modest burst of increased contentions. Therefore, in practice the number of contention slots allocated must be increased by several times over this maximum-utilization value.

Based on the foregoing considerations, the contention allocations are expected to comprise a significant fraction of the UL resource. For reasonable efficiency therefore, it is important that each contention signal use a small bandwidth-time resource. In a single-carrier system such as DOCSIS, with its small minislots, each contention slot need be no larger than necessary to send the bandwidth-request MAC header, and slotted ALOHA is appropriate.

In OFDM, however, we do not have this fine granularity. Indeed, as presently described in sec. 8.3.5.3.3.7.2, the contention signal is a “long preamble uplink burst”, which means a PHY preamble (at least one OFDM symbol) followed by another OFDM symbol containing only a 48-bit bandwidth-request MAC header. The preamble is necessary since the BS has no knowledge of the channel distortion from the SS. Thus, the allocation for each contention slot is two OFDM symbols, which is a huge resource of bandwidth-time and transmitter power to use for such a small message. Others have noticed the noted this problem with the current draft, and there was a previous attempt to fix it. [6]

In this contribution, we propose a method which we have found to be faster, more efficient and robust than the *slotted ALOHA* method now existing in sec. 8.3.5.3.3.7.2. For purposes of comparison, we will refer to the proposed method as *focused contention.*

2. Focused Contention

2.1 Overview

In multi-carrier transmission, besides the time dimension, the frequency dimension is also available for designing a contention method. This is used in the OFDMA mode to design a contention method based on CDMA codes [3]. Similarly, the proposed *focused contention* method seeks to accommodate more than one contending user per symbol by partitioning the 200 available subcarriers into a number of separate *contention channels*. To maximize the probability of detection, instead of a complete bandwidth-request MAC header, the SS transmits an anonymous *contention code* chosen from a small codebook.

A disadvantage of focused contention is that, after a successful contention, the BS must subsequently allocate two OFDM symbols during which the successful contender shall identify itself and its needs by transmitting a complete bandwidth-request MAC header. This additional step is not necessary in slotted ALOHA, because the successful contention signal *contains* this very MAC header. Thus, although the access delay will be smaller with slotted ALOHA under conditions of low traffic when many contention slots are available and the probability of collision is low, our simulations show that focused contention will provide a lower average delay for a given contention allocation, and, more importantly, is less likely to choke on sudden, modest bursts of increased contention.

To allow the BS receiver to easily compensate for the unknown channel, focused contention modulates its code using time-differential modulation (DPSK) over two consecutive OFDM symbols, achieving an effect similar to the preamble used in slotted ALOHA.

Because multiple SS transmit on different subcarriers simultaneously as in OFDMA, to maintain orthogonality, it is necessary for the SS time and carrier frequency synchronization to be closely locked to the BS reference. However, the present draft of 802.16a already requires this. [2]

2.2 Partitioning In Time And Frequency

The proposed mechanism operates by setting up a time-frequency partition of the contention window within any OFDM uplink frame.

The available contention period in each frame is partitioned into pairs of consecutive symbols. Each such pair is called a *contention slot.* For purposes of description, the first symbol in each contention slot is indexed *0* and the second is indexed *1*. To contend, the SS differentially keys the two symbols in a the contention slots.

The 200 available used subcarriers are partitioned into sets called *contention channels* that carry the “bits” of a differentially keyed code. For purposes of description, the total number of subcarriers in a bandwidth request channel is labeled *K*, and the index *k* is used to refer to a particular subcarrier in the set. (K must be fixed in a given system design, and the simulation results shown subsequently suggest that K=4 is optimum for our purposes.)

2.3 Contention Codes

A codebook of *2K* codes is used in focused contention. The first *K* codes are the rows of a *KxK* Hadamard-Walsh matrix; i.e. Walsh codes. The remaining *K* codes are their bitwise complements. Simulation results have shown this design to give slightly better performance than a smaller number of Walsh codes, or a larger or smaller number of pseudonoise (PN) codes.

2.4 Contention Transmissions

To contend, a SS must choose four parameters:

1) An upcoming frame, with a nonzero contention allocation, in which to contend.
2) A contention slot within this frame.
3) A contentional channel (i.e., a set of subcarriers), and
4) A K-bit contention code, to be transmitted differentially across the subcarriers in the given contention slot.

The latter three parameters 2), 3) and 4) are chosen at random, with equal probability to each available choice.

The code bits differentially phase-modulate the subcarriers, forming DBPSK. The value of the bit modulating the *k*th subcarrier in the first OFDM symbol in the *n*th contention slot, $b_0(k, n)$, may be chosen arbitrarily by the SS as either +1 or -1. Doing the differential modulation, the corresponding value during the second OFDM symbol is

$$b_1(k, n) = C_m(k) \cdot b_0(k, n) \qquad (1)$$

where *m* is the index of the chosen contention code and $C_m(k)$ is the value of the *k*th bit of the *m*th contention code.

The contention transmission is shown in Figure 1.

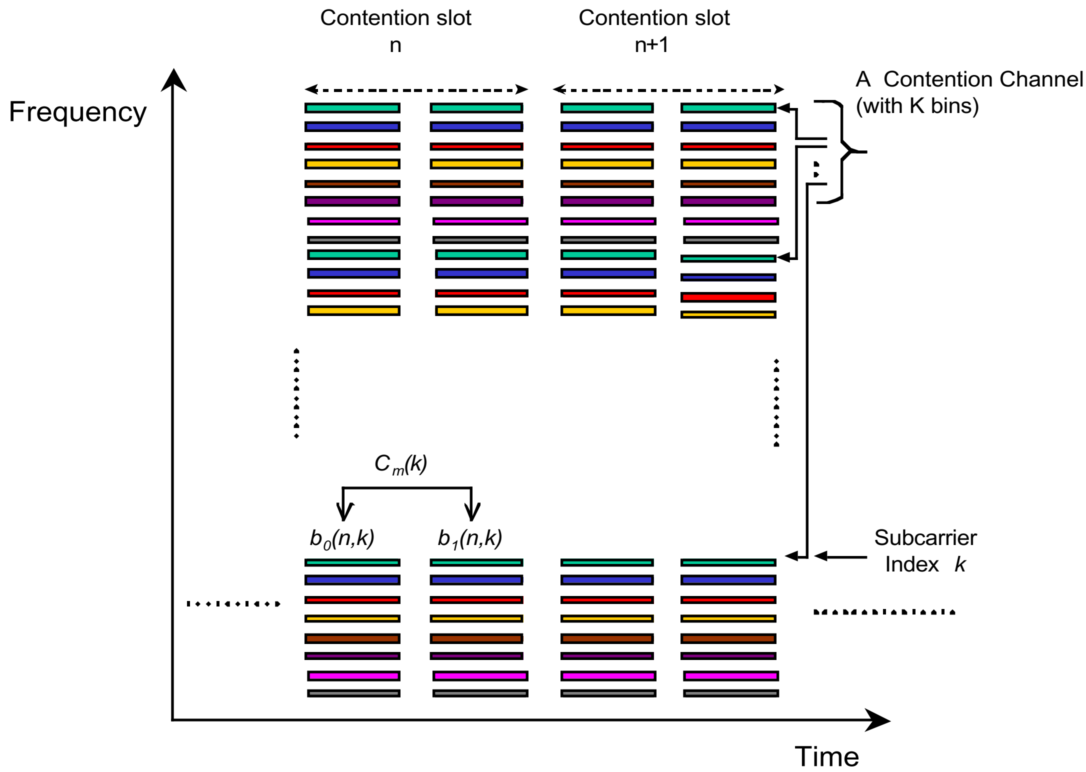

Figure 1 Contention Slots and Contention Channels

2.5 Power Boosting

Because the SS is designed to transmit 200 subcarriers, the power of each may be boosted when transmitting a contention signal. We are familiar with this in OFDMA, and where it has been called *power concentration* or *boosting.* The amount of the boost is important, and is discussed in section 4.2.

2.6 Received Signal

The transmitted bits b-0(k,n) and b1(k,n) are then received by the BS (across the frequency-selective channel) as

$$r_0(k, n) = b_0(k, n) \cdot H(k, n) + \eta(k) \quad (2)$$

$$r_1(k, n) = b_1(k, n) \cdot H(k, n) + \eta'(k) \quad (3)$$

where $\eta(k)$ and $\eta'(k)$ denote uncorrelated additive white Gaussian noise, and $H(k, n)$ is the complex channel transfer function of the *k*th subcarrier in the *n*th contention slot, which we assume does not vary between these two consecutive OFDM symbols.

More formally, modelling the channel as consisting of *L* paths, and denoting that during the *n*th contention slot, the *l*th path has delay $\tau_{l,n}$ and complex gain $h_{l,n} = h_{l,n}^{I} + h_{l,n}^{Q}$, then [7]

$$H(k, n) = \sum_{l=1}^{L} h_{l \cdot n} \cdot e^{-j2\pi f \tau_{l,n}} \Big|_{f = k/T_s} \qquad (4)$$

2.7 Detection

The BS must process the received signal during each contention slot to produce sufficient statistics for the likelihood that each of the contention codes was transmitted on each of the contention channels. (Since there are *2K* contention codes and *200/K* contention channels, the number of sufficient statistics is 400.) System implementers are free to “invent” their own algorithms. Each sufficient statistic must then be applied to a decision threshold.

2.7.1 Design Considerations

2.7.1.1 Consequences of Wrong Decisions

Table 1 is a matrix showing all possible contention stimuli vs. detector decisions, which is useful in designing the detection algorithm.

Table 1: Detector Performance Matrix

Number of SS which actually sent the given code in the given contention channel	Detector Output	
	Positive (Code Detected)	Negative (Code Not Detected)
0	*Undesired: No-Signal False Alarm*, which will cause the BS to allocate slots for a bandwidth-request MAC header, in which no SS will respond.	*Desired: non-detection of no signal*
1	*Desired: Successful detection*	*Undesired: Single Miss*, which will cause the SS to retry in a subsequent frame, depending on the backoff.
2 or more	*Undesired: Collision False Alarm*, which will cause the BS to allocate slots for a bandwidth-request MAC header, in which two SS will respond and collide, causing all of them to retry in a subsequent frame, depending on backoff.	*Desired: Multiple Miss*, which will cause the 2 or more involved SS to retry in a subsequent frame, depending on the backoff. This is the best answer to a collision!

Considering the various consequences, it is apparent that the detection system design should place a high cost of an *Undesired Collision False Alarm,* because it affects both average access delay as well as wasting allocations. Although it only causes wasted allocations, the *No-Signal False Alarm* rate should be negligible also.

2.7.1.2 Effect of SNR on Threshold Selection

In the real world, selecting the receiver threshold is further complicated by the range of signal-to-noise ratios (SNR) which must be accomodated. The system operators' goal will initially be to stretch all cells to the limit, while using puniest power amplifiers available in the SS, requiring those at the fringe to run at maximum power. However, in time, as cell sizes are reduced, the SS at the reduced fringes will have excess power, and the operator may set the power control setpoint to take advantage of a higher system SNR. At the higher SNR, there will be fewer missed detections but more false alarms. In an ideal world, the operator or system designer could lower the contention-code error threshold to compensate for the higher SNR, but we are uncertain whether this preferred design would actually be done in practice. System SNR could also vary some with atmospheric conditions. For these reasons, to be conservative, we have required in our simulations that the contention-code error threshold be fixed with respect to SNR, and we have run simulations with 15 dB system SNR, and at high SNR, using this same fixed threshold.

To verify the performance possible with focused contention, we developed an ad-hoc algorithm which produces an *error* magnitude as each sufficient statistic; i.e. the lower a given error magnitude, ther greater is the likelihood that a given code was transmitted on the given contention channel. This algorithm is detailed in section 3.5.

2.8 Threshold Selection and SNR

As in any signal-detection system, in order to determine the optimum threshold setting, one must plot the so-called *receiver operating characteristic* which shows missed detections and false alarms as a function of the sufficient statistic. As our error threshold is increased, fewer transmissions are missed but more false alarms are received. As the error threshold is decreased, the false alarms are reduced but more transmissions are missed.

3. Description of the Simulation

The simulations used to obtain the results in this document are an extension of our previous work[], with the addition of a rudimentary MAC layer model, and calculation of access delay times. Contentions are now generated as the simulation progresses, so that the retransmissions could be added to the model.

All signal generation, distortion, reception, and detection is simulated with complex numbers in the frequency domain, assumed constant for each OFDM symbol.

3.1 Subcarrier and Frame Structure

The 256 subcarriers (200 used subcarriers) were a modelled in a 6 MHz channel, with guard interval 1/8 of the reciprocal of the subcarrier spacing, and frame period 5 milliseconds. These values were “pulled out of the air” for convenience and are felt to be inconsequential to our results.

3.2 Contention Generation

At the beginning of each frame, a time-domain simulation is run with timestep = (OFDM symbol duration) / (number of FFT points) for the duration of one frame. The probablity of a new contention transmission at each timestep is calculated form the desired traffic load (average number of contentions per OFDM symbol). At each timestep, this rather heavily-weighted coin is flipped and, sometimes, a new contention is ordered. Upon completion of this trial, a new contention is scheduled for each contention so ordered, and also one for each retransmission which had been scheduled from prior frames.

Note that this gives contentions with Poisson-distributed interarrival times, like the classical models of telephone call initiations or passengers arriving at a bus stop. This seemed to us to be sensible. The traffic model adopted by IEEE 802.16 [8] does not seem to address contention traffic, so we made up our own.

A contention is scheduled by generating the three latter random parameters given in section 2.4. In addition, a SS transmitter power level is randomly selected from the range of ± 1.0 dB from nominal, to model real-life power control.

3.3 Modulation

For each contention slot, the active bandwidth-request signals are picked out of the schedule, differentially-modulated as explained in section 2.4 and multiplied by the randomly-selected tx power level. The power is also boosted above the normal subcarrier power level by the desired number of dB.

3.4 Channel Distortion and Multiaccess Interference

A random channel is realized from the SUI-4 model [7], and then the modulation in each tone, for both the first and second signals, is multiplied by the complex frequency-domain channel gain from the random channel realization. The received signal due to each active bandwidth-request signal is thus two vectors of complex numbers, one for the first OFDM symbol and one for the second.

We did run simulations using other SUI channel models, but all results shown in this document were made using a SUI-4 channel model, since it is the most severe model applicable to mode OFDM-A_L. (The delay spreads in SUI-5 and SUI-6 exceed the largest guard interval available in mode OFDM-A_L for most channel bandwidths).

All simulations were run using the 256-point default FFT mode with 28 lower-frequency guard tones and 27 higher-frequency guard tones as currently specified in our draft standard. The channel width used was 6 MHz; this number was pulled out of the air for convenience and are not felt to influence our comparative results.

All such active bandwidth-request signals in the instant contention slot all signals are added together to form a composite receive signal. Note that if all active bandwidth-request signals in a given contention slot have chosen different bandwidth-request channels, the additions will occur on different tones, and assuming no intercarrier interference in the BS receiver, there will therefore be no interference among these bandwidth-request signals.

3.5 Detection

After so generating the received signal for a given contention slot, a BS receiver is simulated for each available contention channel, for each available contention code in each contention slot. The algorithm explained below was developed ad hoc to satisfy the considerations given in section 2.7.1.

The receiver computes two metrics. The first metric measures the error in the received signal correlated to the expected code:

$$errCode(m) = (1/K) \cdot \sum_{k} \left| r_0(k, n) - C_m(k) r_1(k, n) \right| \quad (5)$$

where

K = code length = number of subcarriers in a contention channel
k = index on subcarriers in a contention channel
$r_0(k,n)$ = BS rx FFT output, first OFDM symbol, contention slot n, carrier k
$r_1(k,n)$ = BS rx FFT output, second OFDM symbol, contention slot n, carrier k
$C_m(k)$ = code bit, code m, bit k

The second metric measures the error in the power level relative to the expected power level of 1.0. This helps to reduce the rate of multiple-collision false alarms, since if more than one SS transmits on a given contention channel, the received power is usually higher than if only one transmits.

$$errPower(m) = |1/K| \cdot \left| \left| \sum_{k} \left| r_0(k, n) \right|^2 \right| - 1 \right| \quad (6)$$

Note that, in the above expression, $\left| r_1(k, n) \right|^2$ could have been used as well, since both signals suffer the same tx power variation and same channel instance.

Finally the two errors are added together to develop a sufficient statistic:

$$errComposite(m) = errCode(m) + \alpha \cdot errPower(m)$$

which is then compared with a decision threshold to determine whether or not a bandwidth-request signal is declared to be detected or not.

For each such receiver (contention slot, contention channel, contention code), if the errComposite exceeds an empirically-determined error threshold, the result is ignored. If it does not exceed the allowed threshold, the actual signals transmitted are examined and finally the result is executed per Table 1.

3.6 Detection Consequences and Access Delay Calculations

3.6.1 MAC Policies

We tried applying several exponentially-weighted backoff schemes before to the retransmission model, but found that this just increased the access delays by the average backoff. Dynamic backoff is an important tool for regulating system operation under high traffic loads and smoothing out bursty traffic, but is just another variable (which should be held constant) in comparing PHY system performance.

The behaviors described below also assume that the needed allocations are always immediately available.

Thus, our access delay results are “best-case” MAC numbers. The idea is to show the effect of the PHY layer.

3.6.2 Slotted ALOHA

In simulating slotted ALOHA, it was assumed that all bandwidth requests would be successfully detected, unless there was a collision, in which case the request would fail. It was also assumed that no false alarms would ever occur. A successful request is always followed by a payload allocation 2 frames later. If a collision occurred, a new request was generated 2 frames + backoff later. Thus, if no collision occurred, the access delay would be 2 frames. If the one retry was needed and the backoff was zero, the access delay would be 4 frames. If two retries were needed with no backoff, the access delay would be 6 frames, etc., etc.

3.6.3 Focused Contention

For the focused contention method, all of the possibilities in Table 1 were simulated. It was assumed that a successful request would be responded with a bandwidth-request MAC header allocation two frames later, and a payload uplink allocation two more frames later. Thus, if no collision occurred, the access delay would be 4 frames, and, again ignoring backoff, each retry will again add two more frames, resulting in access delays of 6 frames, 8 frames, 10 frames, etc.

3.7 Simulation Length and Stability

A simulation of this type takes some time to arrive at a steady-state value of retransmissions because, as retranmissions begin, this increases the traffic which increases the collision frequency, which increases the retranmission rate even more.

It was verified that all simulations at least ran long enough for the retransmission frequency to stabilize to its steady-state value. Most simulations were several hundred frames; those with less traffic density were run longer.

Because we wanted to measure PHY layer performance, we assumed a dumb MAC which did not take any measures to restrict traffic. In fact, as the raw detection probability fell below 60% or so, often the system would run away with retransmissions and break down. This explains the abrupt endpoints on some of the curves in the following section.

4. Simulation Results

Figure 2 shows shows average access delay for slotted ALOHA and for focused contention using various code lengths. The “system” AWGN (for a normal OFDM symbol) is 15 dB, which is about the threshold of operation for the most robust modulation and coding. A power boost of 6 dB is assumed; therefore the S/N ratio on these contention tones is 21 dB.

With our chosen 6 MHz channel bandwidth and 5 millsecond frame time, the frame in these simulations consists of 136 OFDM symbols. In these first figures, 20% of each frame, 26 symbols, are allocated to the contention period.

Note that focused contention with longer code lengths starts to be shaped like slotted ALOHA; the access delay increases with increasing traffic load. This is because of the increased collisions. Also note that with a code length = 2, focused contention exhibits a “floor” of missed detections, degrading access delay slightly, independent of traffic load. This is because, although collisions are rare, AWGN causes missed detections without any traffic.

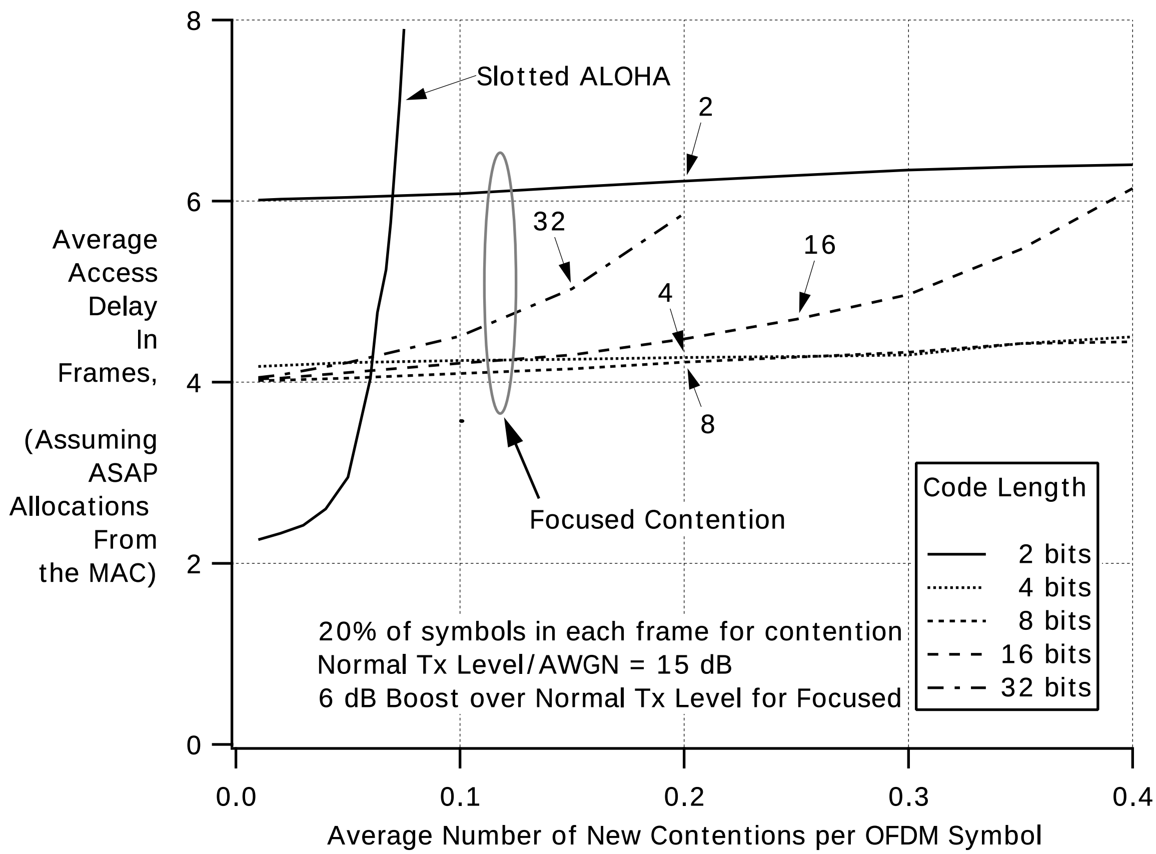

Figure 2 Average Access Delays with 15 dB AWGN, 20% of UL symbols for contention, power boost = 6 dB

A plot of false alarm rates from the same simulation runs is shown in Figure 3. Note that these are the sum of both false alarm types shown in Table 1.

False alarms in focused contention have two bad effects. The *multiple collision false alarms* cause the colliding SS to both respond in the same UL allocation, and by the time they recover from this 4 frames have gone by. We have, however, with our detection algorithm and threshold settings, forced this to be a very rare event, so that its effect on average access delay is negligible. Nevertheless, it is correctly modelled in the simulation so the delay-increasing effect of these false alarms is included in our figures such as Figure 2.

The second bad effect of false alarms, of either type, is the wasted 2-symbol UL allocations for bandwidth-request MAC headers. From Figure 3, note that, for the best code length choices, there will be about 0.01 false alarm for every bona fide contention. Thus, the bandwidth-time resource required for RNG-RSP and bandwidth-requests MAC headers will be increased by 1% due to bogus RNG-RSP to false alarms.

Thus, both of the effects of false alarms are negligible.

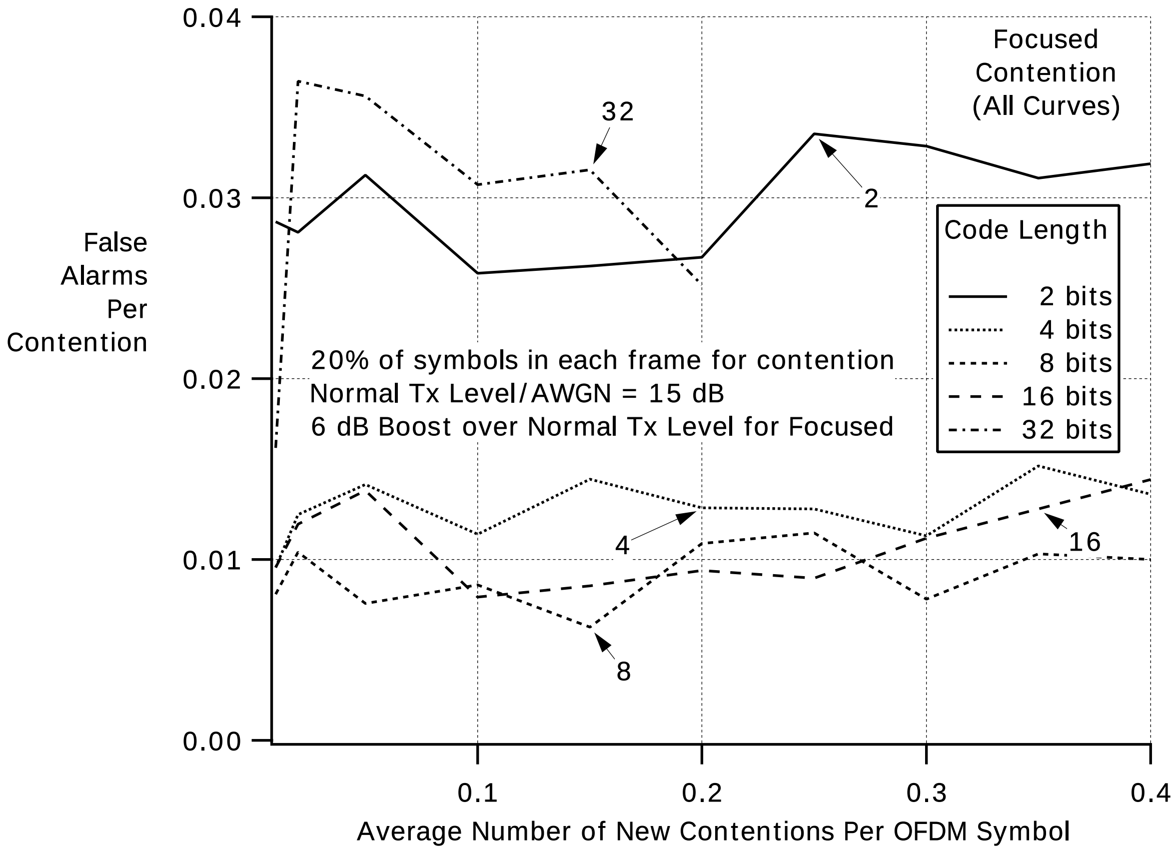

Figure 3 False Alarms (total, all types) with 15 dB AWGN, 20% of UL symbols for contention, power boost = 6 dB

Note that the false alarm rates shown in Figure 2 are constant with traffic; this means that the false alarms are proportional to traffic. We therefore averaged each line in Figure 2 and plotted the false alarm rate as a function of code length. The result is Figure 4

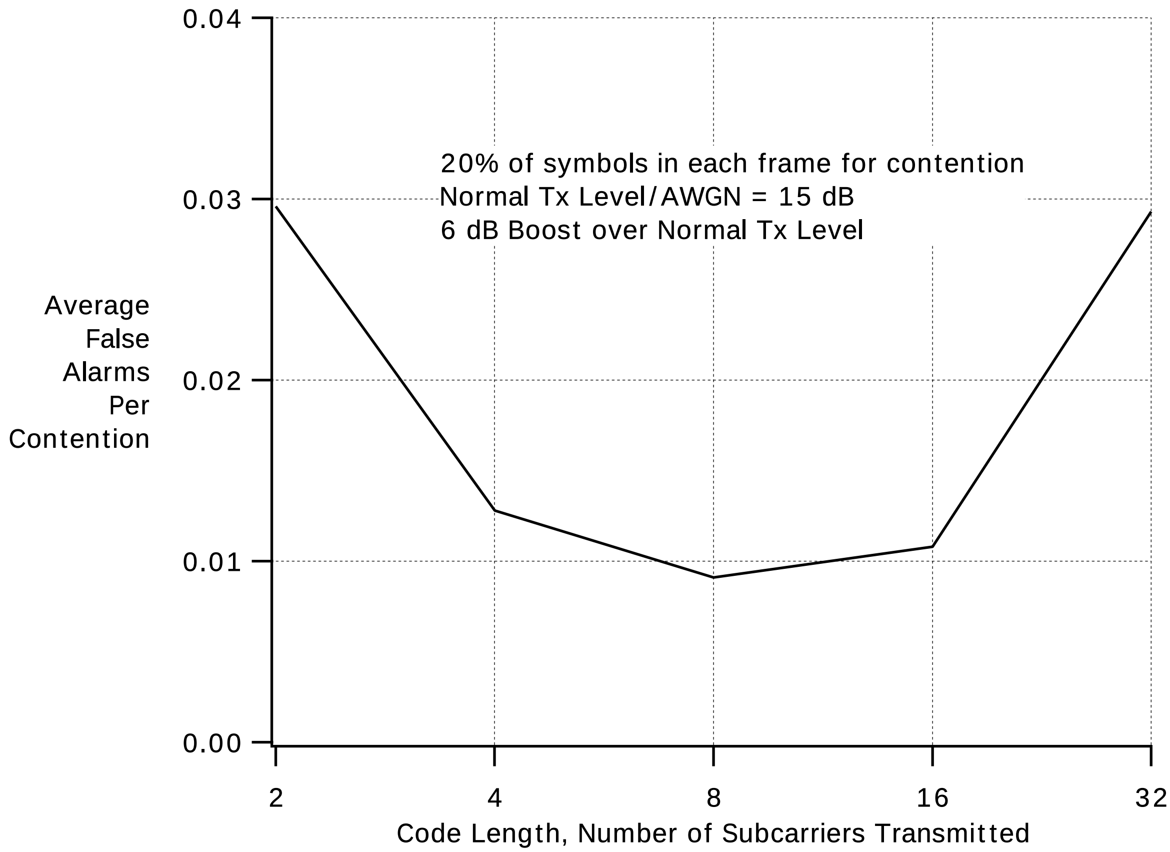

Figure 4 False Alarms (total, all types) with 15 dB AWGN, 20% of UL symbols for contention, power boost = 6 dB, as a function of Code Length for Focused Contention

4.1 Choice of Optimum Code Length

As we have already alluded, in order to minimize the probability of collisions and thus the access delay, one would want to partition the 200 available subcarriers into as many request channels as possible. However, this has the disadvantage of reducing the code length (since each code bit requires a subcarrier), which will decrease the “coding gain”, decreasing the detection probability and increasing false alarms.

From Figure 2, Figure 3and Figure 4 we conclude that the optimum code length for focused contention is probably 4 or 8 bits. *Since we don’t want any more options in the standard,* we ran additional simulations to help us decide on one answer.

As our results so far have shown, one advantage of focused contention over slotted ALOHA is the ability to accomodate many more contenders with the same allocation, which thus far has been 20%. In practice, one might want to reduce the allocation. Our next figure was run with only 10% of the UL symbols allocated to contention. This worked out to be the first 12 symbols in the 136-symbol frame. The resulting average access delay is shown in Figure 6.

,

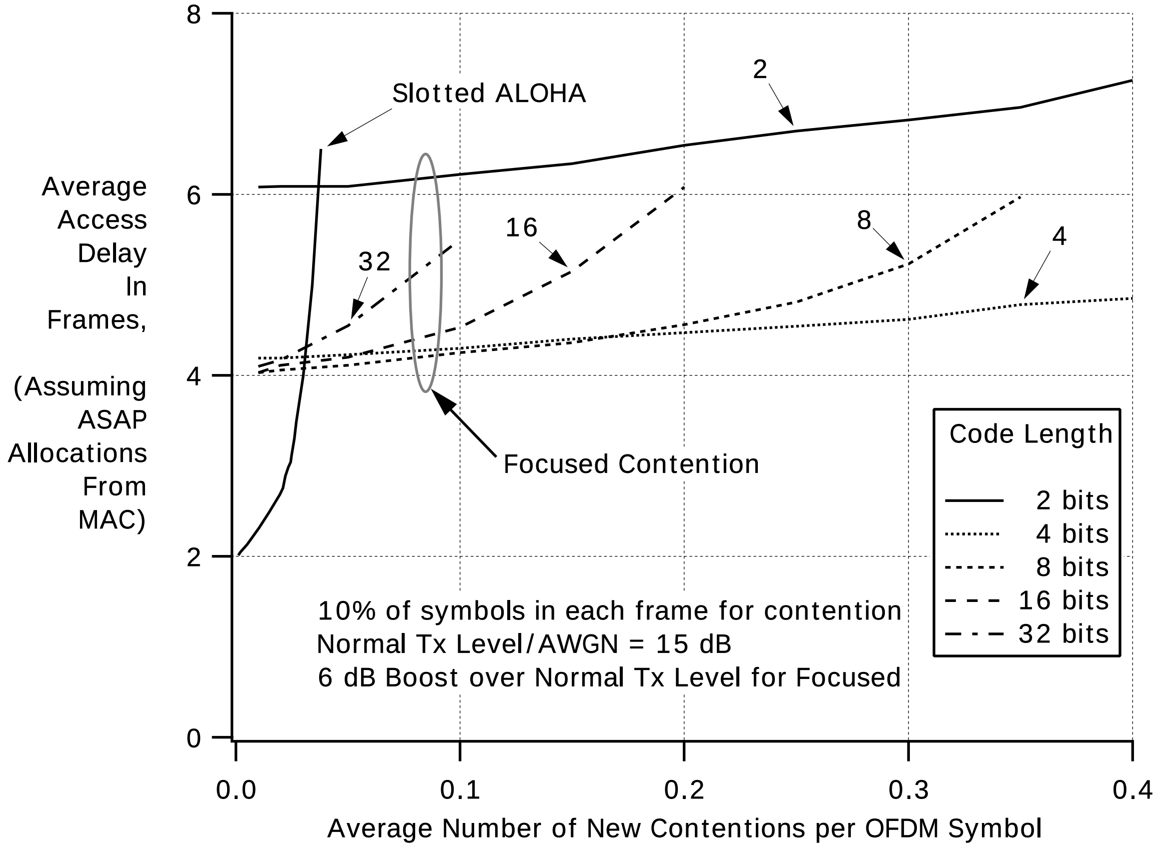

Figure 6 Average Access Delays with 15 dB AWGN, 10% of UL symbols for contention, power boost = 6 dB

By reducing the allocation to contention, but still plotting over the same x-axis of traffic load, we have in effect doubled the load on the system. Note that the slotted ALOHA system, way on the left side there, is already “broken” (going to inifinite access delay) before focused contention even starts to bend. However, now we can see the limits of focused contention, as the longer code lengths (with smaller numbers of contention channels available) begin to show increased access delay with increasing traffic, due to collisions. However, looking at code length 4, we see that it is still degrading gracefully while handling over ten times the traffic which brought the slotted ALOHA system to its knees. (That is called “robust”.)

4.2 Power Boosting and Code Length

The 6 dB power boost which has been used in the foregoing results is an important component of focused contention. Without it, the performance of the 4-bit code system is quite poor, similar to what the 2-bit code system does in Figure 2, and the optimum code length is moved out to 8 or 16 bits. (This is what we had proposed in our earlier work, before we tried the power boost). As can be seen from Figure 4 however, if we did have to use 8 or 16 bit codes, the performance would be not nearly as good.

The power boost *available* from the SS power amplifier is 10*log10(200/K), which would be 17 dB for K=4. However, since we are dealing with a digital radio, all that is really needed is to get the receiver “off the threshold” ; returns diminish rapidly after that. To verify this, we compared the result of a focused contention system with 4-bit codes using power boosts of 0, 6 and 12 dB. The result is shown in Figure 7.

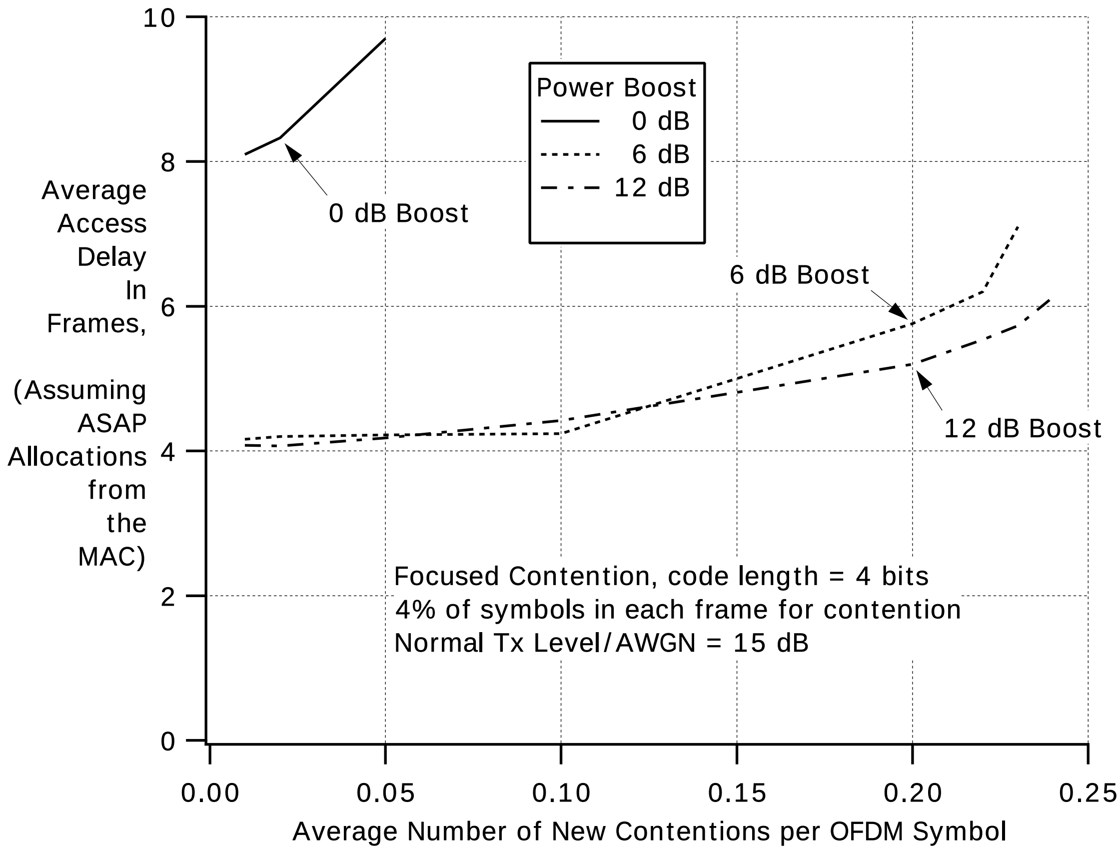

Figure 7 Effect of Power Boost on the Average Access Delay of a Focused Contention system with 15 dB AWGN, 4% of UL symbols for contention

Note that the performance with 0 dB power boost is quite poor, but the performance with 6 dB boost is almost as good as with 12 dB boost.

The advantage of the power boost is that all bandwidth-request transmissions are effectively received at high SNR, because the system must be designed to allow this same SS to be received by the same base station when it is transmitting a normal, non-contention OFDM symbol of all 200 subcarriers. Another way to look at this is to consider that many SS are now transmitting simultaneously, and the BS is receiving their aggregated power. Fortunately, however, they are easily separable since each SS is *focused* on a *contention channel* consisting of only four subcarriers.

We considered several possible questions regarding power boosting:

Question 1. Will this require an awkward SS transmitter design, and possibly a power control in the upconverter of the SS? Answer: No. The 6 dB gain may be realized by shifting the SS transmitter's IFFT output one bit to the left.

Question 2. What about peak power? Answer: With fewer tones combined, the peak:average power ratio is actually less than when transmitting the normal 200 tones.

Question 3. How will the BS receiver handle 6 dB more power during the contention periods? Answer: Because the BS knows when the contention periods are coming, the receiver IF or RF gain can be neatly reduced by 6 dB during the contention periods.

5. Conclusions

An efficient and robust bandwidth request mechanism is presented for the OFDM mode of the IEEE802.16a standard. Extensive PHY-level simulations were performed to study the performance of the proposed scheme. The proposed method is compared with the slotted ALOHA method now in IEEE 802.16a by default, and found to show substantially better performance. In particular, the throughput obtained with the proposed method is over an order of magnitude greater than that obtained using slotted ALOHA, with the same per-frame overhead

Slotted ALOHA does perform better under extremely light traffic loads or, equivalently, a relatively high allocation of the UL time for contention periods. This is because the proposed method requires two extra frames for sending the bandwidth-request MAC header after successful contention. However, these are usually going to be negligible when compared to the queing delays which will occur in actual operation

The more important considerations are that,

1. For systems supporting many simultaneous, part-time users, contention is expected to be significant, and therefore the overhead imposed by allocating enough contention slots to allow slotted ALOHA to work will cause a substantial penalty to system efficiency.

2. The proposed method is more robust than slotted ALOHA.

The proposed method may be implemented in IEEE 802.16a by adopting the text given in the next section to revision D1-2001.

6. Proposed Changes to Text in IEEE 802.16a/D1-2001 Draft Standard

This section contains the text which is referenced in our comments.

6.1 Proposed Additional Text in sec. 6.2.2.3.6

The following parameters may be included in the RNG-RSP message:

Contention Slot:
Index number of the the contention slot that was used by the SS in a contention.

Contention Channel:
Index number of the the contention slot that was used by the SS in a contention.

Contention Code:
Index number of the the contention slot that was used by the SS in a contention.

6.2 Proposed Additional Text in IEEE 802.16 sec. 8.3.5.3.3.1

The *Contention Period REQ* allocated by the BS in the UL-MAP, shown in Fig. 211, shall be an even number of OFDM symbols. Each consecutive pair of OFDM symbols in the contention period is referred to as a *contention slot*. Contention slots are indexed consecutively in time by a *contention slot number*, starting at the beginning of the contention period with contention slot number 0.

6.3 Proposed Revision of IEEE 802.16 sec. 8.3.5.3.3.7.2

8.3.5.3.3.7.2 Bandwidth Requests

This section describes the procedure to be followed by a SS in order to effect a bandwidth request.

8.3.5.3.3.7.2.1 Parameter Selection

The SS shall first choose an upcoming frame, with one or more allocated contention slots, during which to make its request.

The SS shall also choose, at random with a uniform distribution, a valid contention slot number allocated during the chosen upcoming frame.

The SS shall also choose, at random with equal probability, a *contention code* from Table 2.

Table 2: Contention Codes

Bandwidth-Request Code	bit0	bit1	bit2	bit3
0	1	1	1	1
1	1	-1	1	-1
2	1	1	-1	-1
3	1	-1	-1	1
4	-1	-1	-1	-1
5	-1	1	-1	1
6	-1	-1	1	1
7	-1	1	1	-1

The SS shall also choose, at random with equal probability, a *contention channel* from Table 3. The indices {-100 to +100} in the body of Table 3 refer to the subcarrier indices as defined in sec. 8.3.5.3.3.5.

Table 3: Contention Channels

bandwidth-request channel index	subcarr0	subcarr1	subcarr2	subcarr3
0	-100	-50	+1	+51
1	-99	-49	+2	+52
2	-98	-48	+3	+53
...	...	...	...	...
k	k-100	k-50	k+1	k+51
...	...	...	...	...
48	-52	-2	+49	+99
49	-51	-1	+50	+100

8.3.5.3.3.7.2.1 Contention Transmission

After choosing its four parameters, the SS shall transmit, during the chosen contention slot in the chosen frame, only the four subcarriers {subcarr0, subcarr1, subcarr2, subcarr3} which comprise the chosen contention channel.

The phase and amplitude (modulation) of each subcarrier shall be constant during each of the two OFDM symbols which comprise the contention slot, as in a normal OFDM symbol.

During the first OFDM symbol, each of the four subcarriers may be transmitted with any arbitrary phase.

During the second OFDM symbol, the phase shall depend on the corresponding bit in the chosen contention code, and the (arbitrary) phase transmitted during the first OFDM symbol on the same subcarrier. If the code bit is +1, the phase shall be the same as that transmitted during the first OFDM symbol. If the code bit is -1, the phase shall be inverted, 180 degrees with respect to the phase transmitted during the first OFDM symbol.

During both OFDM symbols, the amplitude of each of the four subcarriers shall be boosted by 6 dB above its *normal* amplitude, that used during a non-contention OFDM symbol, including the current power-control correction.

6.4 Proposed Additions to Section 11.1.4:

Table 4—Additions to Table 233

Name	Type	Length	Value
Contention Slot	18	2	Used to indicate the Contention Slot number of the message which is being responded to. This TLV is used in conjunction with the Contention Channel, Contention Frame Number and Contention Code values to identify the SS to which this response is directed.
Contention Channel	19	1	Used to indicate the Contention Channel number of the message which is being responded to. This TLV is used in conjunction with the Contention Slot, Contention Frame Number and Contention Code values to identify the SS to which this response is directed
Contention Code	20	1	Used to indicate the Contention Code number of the message which is being responded to. This TLV is used in conjunction with the Contention Slot, Contention Frame Number and Contention Channel values to identify the SS to which this response is directed
Contention Frame Number	21	1	The eight least significant bits of the frame number in which the message being responded to was received. This TLV is used in conjunction with the Contention Slot, Contention Code and Contention Channel values to identify the SS to which this response is directed

7. References

[B1] "The current draft of our standard": IEEE Draft Standard for Local and Metropolitan Area Networks – Part 16; Document number IEEE P802.16a/D1-2001, dated 2001 Nov 30.

[B2] "Data-Over-Cable Service Interface Specifications, Radio frequency Interface specifications", SP-RFIv1.1-130991105, Interim Specification, Cable Labs.

[B3] I. Kitroser, Y. Segal and Z. Hadad, "Bandwidth Request Using CDMA Codes in OFDMA(OFDM) Base PHY for TG3 and TG4", IEEE802.16.3c-01/55, April 2001.

[B4] In [1], sec. 8.3.5.2.5.1.

[B5] D. Bertsekas and R. Gallager, "Data Networks", Prentice Hall, New Jersey, 1992.

[B6] T. Kaitz and R. Halfon, "Subcarrier Based Polling for 802.16ab OFDM PHY", IEEE802.16abc-01/30, Sept. 2001.

[B7] V. Erceg et al., "Channel Models for Fixed Wireless Applications", IEEE802.16.3c-01/29r4 {accepted by TG3}, Sept. 2001.

[B8] C.R. Baugh et al., "Traffic Model for 802.16 TG3 MAC/PHY Simulations", IEEE 802.16.3c-01/30r1, March 2001.