A Novel Mechanism for Contention-based Initial Ranging in IEEE 802.16e Networks

Jing Chi
ENST & CNRS LTCI
46, rue Barrault 75013 Paris, France
Email: jing@enst.fr

Philippe Martins
ENST & CNRS LTCI
46, rue Barrault 75013 Paris, France
Email: martins@enst.fr

Marceau Coupechoux
ENST & CNRS LTCI
46, rue Barrault 75013 Paris, France
Email: coupechoux@enst.fr

Abstract—The paper proposes a Waiting-time Dependent Increasing rate Adapted Backoff (WDIA) algorithm to improve the contention efficiency in the IEEE 802.16e (OFDMA). A contention model is first proposed to analyze the frame-based behavior of contention resolution in 802.16e network. We introduce the concept of reduced overlapping between back-off windows of contending stations and we adopt adaptable increasing rate. An initial ranging scenario is constructed to evaluate the effectiveness of WDIA. The simulation results show that the WDIA algorithm achieves better performance than Truncated Binary Exponential Backoff (TBEB) algorithm in terms of number of retransmissions, access delay and resource utilization.

I. INTRODUCTION

In 802.16e, the uplink frame consists in three parts: contention interval for initial ranging (IR), contention interval for bandwidth request (BR), and data transmission interval for user traffic. These parts are allocated by the base station (BS) in UL-MAP. At network entry, subscriber stations (SS) contend for resource on the IR interval in order to access the network. As in many cellular networks, this random access is based on a traditional slotted Aloha in conjunction with a back-off algorithm.

Given more resources to the contention interval surely reduces the collision probability of user access, but also reduces the amount of resources dedicated to user data traffic. On the other hand, improving the contention efficiency within a relative stable contention interval allows better performance, while maintaining sufficient resources for data traffic.

In this paper, a Waiting-time Dependent and Increasing rate Adapted (WDIA) backoff algorithm is proposed to improve the contention performance (collision rate, access delay and transmission opportunity utilization) of IEEE 802.16e networks. WDIA has been originally proposed for WLAN [1] and is extended in this paper to the IR process in WiMAX. Focusing on contention performance, the Request-response mechanism based initial ranging is considered and the CDMA ranging code is not used. Although presented here for IR, WDIA can also be used for other types of contention-based uplink accesses, including periodic ranging, handoff ranging and bandwidth request.

Bianchi [2] has given a simple and accurate framework for the analysis of back-off based MAC protocols but focused on IEEE 802.11 DCF networks. He concluded that maximum performance can be achieved by adaptively tuning the value of the back-off window (BW) size based on the network size. In [3], a distributed backoff algorithm is proposed to maximize the throughput by computing the optimal value of CW for each station. However, the optimization is based on estimating the number of active stations that is inaccurate under a distributed contention manner. In [4], collision history based BW adjustment is proposed to randomize the transmission timing and minimize the collision probability. Authors focus on periodic ranging. However periodic ranging is performed using periodically granted bandwidth and not using contention-based ranging unless no grant is received when T4 (35s as maximum defined in 802.16e) expires. In [5], authors propose a Markov chain based analytical model to study initial ranging performance. They point out that a fixed size of initial BW, as proposed in the standard, neglects the effect of varying network load and define a new adaptive algorithm based on the number of contending SSs. Again, this number is difficult to estimate in practice.

In our work, we propose a contention model of initial ranging in 802.16e and give the direct relation between contention performance and backoff window overlapping. Based on analysis, we introduce the concept of reduced overlapping between BW of contending SS for improving contention performance. The way BW are adjusted is based on the time duration requests have already waited, the so called waiting time (WT). And the load of contention is used for optimizing the BW instead of the contending number of SS since it can be estimated under our separated BW mechanism.

The rest of the paper is organized as follows. In Section II, we summarize the procedure of initial ranging and point out the limitations of contention resolution in 802.16e. Section III presents the contention model and gives direct relation between contention performance and backoff window overlapping. Section IV proposes our algorithm WDIA and Section V evaluates the improvement of WDIA. Finally, conclusions are drawn in Section VI.

II. IEEE 802.16e CONTENTION-BASED INITIAL RANGING

The network entry and initialization process for entering and registering a new SS to the network, is specified in IEEE 802.16e [6]. The procedure can be divided into the following phases: a) Scan for downlink channel and establish...
synchronization with the BS b) Obtain transmit parameters (from UCD message) c) Perform ranging.

Initial ranging is a primary and important procedure for a SS to access the network and setup connection. It is representative since the periodic and handover ranging or bandwidth request have similar process.

We give some basic notations in initial ranging:

- BW: Back-off window with a range of type \([0,W_r-1]\), \(W_r\) is reset after each transmission. The initial/maximum value is \(BO_x\) and \(BO_e\).
- \(BO_x\) and \(BO_e\): back-off start/end defined in IEEE 802.16e standard. It is the initial/maximum back-off window size for initial ranging contention, specified by BS and expressed as a power of 2.
- \(BO\): back-off number randomly chosen from its window range \([0,W_r-1]\) whenever a backoff process begins.
- \(r\): REQ retry counter.
- \(r_{\text{max}}\): the maximum number of contention retries defined in IEEE 802.16e.
- \(T_{\text{UCD}}\): The time interval of two consecutive UCD message. The UCD message is periodically broadcasted by the BS.

The ranging process is shown in Fig. 1, and the cases of three SSs experiencing success or collisions are illustrated. Initial ranging is started once a SS that tries to enter the network, receives a UCD message to be synchronized with the BS for the first time. The UCD Interval is defined in 802.16 with a maximum value of 10s, and this value is not changed in 802.16e. A SS enters the contention process by setting its internal BW before sending a ranging request RNG-REQ. The initial BW value is specified as \(BO_x\) in the UCD message sent by BS. The SS shall randomly select the backoff \((BO)\) value within its BW, and defer that number of contention transmission opportunities \((TO)\) before transmitting the RNG-REQ [6] [7]. A TO provides the required bandwidth for RNG-REQ transmission.

After the transmission, the SS waits for a RNG-RSP message. Once received, the contention resolution completes as the case of SS1 shown in Fig. 1. If no response has been received within \(T_3\) response reception timeout, the SS shall consider the contention transmission as lost. The SS shall then increase its BW by a factor of two, as long as it is less than the maximum BW \((BO_e)\). The SS shall randomly select a number within its new BW and repeat the deferring process described above until a successful transmission or the retry counter \(r\) reaches the maximum number allowed [6] [7]. This is the case of SS2 and SS3.

This mandatory contention resolution is based on a truncated binary exponential backoff (TBEB) to avoid possible re-collision, with the \(BO_x\) and \(BO_e\) controlled by the BS and broadcasted in UCD message. The TBEB algorithm has been well investigated and various modifications have been proposed. TBEB limitations in terms of efficiency and latency have been extensively studied for IEEE 802.3 and 802.11 systems in the literature [2], [3]. IEEE 802.16 networks have however the particularity that the backoff is triggered by T3 timeout (much bigger than the ACK timeout in IEEE 802.11) and the contention process is frame-based. We now analyze the contention resolution for initial ranging in 802.16e and summarize its limitations.

A. Deteriorated delay caused by timeout-based retry

As shown in Fig. 2, in 802.16e, the contention process is not started until a UCD message is received. The backoff for retransmission is not started until \(T_3\) timer expires. \(T_3\) can be set between 20 and 200 ms according to the standard: this causes a much longer waiting time than the ACK timeout in IEEE 802.11. A typical \(T_3\) value of 200 ms, which equals at least 10 frames (for 20 ms frame duration). The backoff counter \((BO)\) is decremented by the number of TO. In each frame, \(N_{\text{TO}}\) TOs is defined by BS. If \(BO < N_{\text{TO}}\), the SS transmits the request in the current frame, else it has to wait for TO in subsequent frames and try to transmit when its BO reaches 0.

The MAC layer packet delay is decomposed into four types of defer time: the defer time \(T_{\text{preTO}}\) before a TO comes (the back-off process is not yet started); the defer time \(T_3\) before having a negative response from the BS (and so entering another BO process); the defer time of a backoff; the defer time from the time instant request is successfully transmitted until the instant RNG-RSP is received (if the RNG-RSP is received in the next frame after a successful request transmission, this duration is approximately the frame duration \(T_f\)). In Fig. 2, we can see in case of SS1, the request is successfully transmitted without collision. The delay consists of \(T_{\text{preTO}}\) and \(T_f\). In case of SS2, we can see after each collision the SS experience a \(T_3\) and a backoff defer before next transmissions. It is clear that the collision has a huge impact on the global access delay.

B. Lower bandwidth utilization caused by contention interval

In IEEE 802.16e, the SS is allowed to send RNG-REQ only during the initial ranging interval with limited number of TO in each frame. Thus, this constraint induces higher collision rate and a fast deterioration as the number of contending SSs increases. In worst case, due to the contention failure,
the BS cannot allocate the dedicated resource to unregistered SS although it has enough bandwidth. So, the bandwidth allocation depends heavily on the performance of contention.

C. Collision caused by random back-off within overlapped BW

The upper bound of BW is controlled by the TBEB but the lower one is always zero. So, overlapping between back-off intervals of contending SSs can be quite high. The random BO selection in the overlapped BW induces a bigger collision rate since the same BO may always be selected even from a bigger BW range. The effect of BW overlapping on collision and delay is modeled and simulated in context of WLAN [1].

III. MODEL OF CONTENTION RESOLUTION IN SINGLE CHANNEL

Most of the existing studies on contention use a Markov model and calculate the transition probability diagram to every steady-state following the analysis in IEEE 802.11. Their model are based on an assumption of constant request arrival and related with the distribution of ranging request arrivals. It is not realistic in initial ranging process of IEEE 802.16e. And it is complicated to derive the primary factors that lower contention performance [4] [8]. Our model gives the direct relation between successful transmission probability and backoff window overlapping. It is then easier to be applied to improve contention performance by optimizing the backoff window.

Concerning the timeout triggered, frame-based characteristic of TBEB in 802.16e, we assume that:

1. N SSs have arrived in the network during a UCD interval. The user arrival rate is below 20/s so \( n < 100 \).
2. During one UCD interval all the n SSs can finish IR.
3. Collisions happen between the SSs in the same retry stage. The collisions happen between the SSs in different retry stages occurs with a small probability that can be omitted.

These assumptions are useful for deriving the primary factors that lower contention performance. In Section V we present simulation results to show that these assumptions are reasonable.

As shown in Fig. 2, the n SSs arrived shall not transmit RNG-REQ until receiving a UCD message for the first time. Thus all the SSs arriving in the same UCD interval start IR process at the same time and set their retry counter \( r \) as 0. So the distribution of ranging request arrivals has no effect on the contention. It also indicates that contention within a UCD interval only happens among the n SSs and does not involve the new arrived SSs in this UCD interval since their contention process is not started until receiving a UCD message. With assumption (2) contention does not involve the contending SSs in prior UCD interval since their IR is finished. Thus within one UCD interval the number of contending SSs is not constant but decreasing from \( n \). The decrement of the number of contending SSs is simulated and shown in Fig. 4. It is not proper to model the contention as a stochastic process with a constant SS arrival rate and steady transition probability diagram, while most existing work do. In our model the successful transmission probability increases as the number of contending SSs and the contention intensity decreases.

In each retry stage \( r \geq 0 \), the SSs are grouped by their prior collisions and the SSs collided in the same TO are assigned to a same group. The SSs in the same group start their next backoff at the same time, while the SSs from different groups start the backoff with a time differ related to their last contention. As shown in Fig. 2, the backoff for retry is not started until T3 expires. With assumption (3), in our model, omitting a small-probability collision type we analyze the successful transmission probability considering only the impacts from SS in the same retry stage.

A. Successful Transmission Probability

The collision occurs when more than one SS transmit the request in the same TO. The collision probability for a random SS \( i \) in \( r \)th retry is defined as:

\[
P_{c,i}^r = 1 - P_{s,i}^r,
\]

where \( P_{s,i}^r \) is the successful transmission probability for a random SS \( i \) in \( r \)th retry, if the backoff is started at the same TO, \( P_{s,r}^i \) is the probability the other SSs choose a different BO value than SS \( i \). As in 802.16e the n SSs started contention simultaneously at the instant a TO arrives, thus the SSs with the same BO value collide. [0, \( W_{r} \)-1] is the backoff window for an SS in \( r \) transmission stage. The probability that SS \( i \) choose value \( k \) from [0, \( W_{r} \)-1] is derived as:

\[
P_{s,i}^r(BO = k) = \frac{1}{W_{r}},
\]

The successful transmission probability for first transmission is derived as:

\[
P_{s,0}^i = \sum_{k=0}^{W_0-1} \frac{1}{W_0} (1 - \frac{1}{W_0})^{n-1}.
\]

From equation (3), it is clear \( P_{s,0} \) is the minimum value of \( P_{s,r} \), not only because \( W_0 \) is the minimum value of BW, but also because the overlap of the BW between the SSs is the whole BW (\( W_0 \)) and the number of contending SSs (\( n \)) is the biggest. These in all increases the collision probability. It also indicates that the worst case of contention happens in the beginning part, which particularly needs improvements.
When \( r > 0 \), for the retransmission backoff, the number of contending SSs in \( r \)th contention is \( n_r \). The SSs are grouped by their prior collisions and the SSs collided in the same TO are assigned to a same group. We use \( m_r \) to represent the number of groups in \( r \)th contention, and \( n_r \) to represent the number of SSs in group \( j, j \in (1, m_r) \), \( n_r = \sum_{j} n_r \). Only the SS in the same group enters the next backoff simultaneously with the same BW \( W_r \). While the SSs from different group start the backoff with a time differ of \( \Delta_r \) (which is related to their last contention), if \( W_r - \Delta_r > 0 \) their BW overlap is \( W_r - \Delta_r \) and else the overlap is 0. Their relations is shown in Fig. 3. To simplify the formula, we use group 1 to denote the group SS i belongs. The time differ of SS i and a SS from group j is derived as:

\[
\Delta^{j,1}_{r-1} \in (1, W_{r-1} - 1), (j \neq 1),
\]

The successful transmission probability in \( r \)th retry is derived as:

\[
P_{s,r} = \frac{1}{W_r} \left( 1 - \frac{1}{W_r} \right)^{n_r - 1}
\]

\[
\prod_{j \neq 1} \left( \frac{1}{W_r} \right)^{n_r} = (1 - \frac{1}{W_r})^{n_r - 1} \prod_{j \neq 1} \frac{\Delta^{j,1}_{r-1}}{W_r} + \frac{W_r - \Delta^{j,1}_{r-1}}{W_r} (1 - \frac{1}{W_r})^{n_r}
\]

We use \( P_{s,r} \) to denote the average value of successful transmission probability of \( n_r \) SSs, \( n_r = n \prod_{r=1}^{r} (1 - P_{s,r}) \) is monotonic decreasing. We assume the average value of \( n_r \) is \( n_r = \frac{n}{W_{r-1} - 1} \), as \( \Delta^{j,1}_{r-1} \) conforms a uniform distribution within \( [1, W_{r-1} - 1] \).

From equation (4) we can see in 802.16e, the overlap of BW affects the collision and the successful transmission probability. This inspired us to design an algorithm to improve contention performance by reducing the overlap.

Our model gives a simple method to analyze the contention process in initial ranging. The direct relation between successful transmission probability and backoff window overlapping is given. In Fig. 4, we shows the decrement of contending SSs. In Fig. 5, the dash line shows the \( P_{s,r} \) derived from our model.

The solid line shows the mean value of \( P_{s,r} \) calculated from simulation of TEBEB. The comparison of numerical results and simulation results show the accuracy of model. The maximum differ of the two is 5.3%.

IV. WDIA ALGORITHM

A. Method description

The basic ideas of WDIA algorithm is illustrated in Fig. 6. We introduce waiting time WT as a control parameter in the backoff algorithm. WT is tracked for each SS. It starts from the instant the SS arrives at the network. BW is shown as a function of the waiting time WT and the number of retransmissions. For each retransmissions, the SS slides the lower bound of the BW range \( W_r \) according to its own WT with a deterministic function. The size of BW \( W_r \) itself increases with a tunable multiplicative factor. It is also shown on the figure, the overlapping part between two BW. By changing the lower bound \( W_r \) of the BW, we can see that this overlapping is reduced compared to the classical back-off algorithm. It is so expected that the collision probability is reduced.

B. Mathematic model and parameter setting

Inheriting the essence of WDB mathematic model [1], BO function contains two addends: a deterministic function \( W_r(WT) \) and a random function \( RandInt(0, W_r^0) \). In Fig. 6, the deterministic function is the curve of lower bounds and the random function is represented by the successive intervals named \( W_r^0 \). \( RandInt(0, W_r^0) \) is a random integer generated
from uniform distribution between 0 and \( W_r^{\delta} \). The functions are given by:

\[
W_r^i(WT) = WT (\text{mod} \frac{T_3}{T_f}).
\]  
(5)

\[
W_r^{\delta} = \begin{cases} 
BO_s & \text{if } r = 0 \\
\alpha W_r^{\delta}_{r-1} & \text{if } r \neq 0 
\end{cases}
\]  
(6)

\( \alpha \) is a tunable integer for the increasing rate of BW after each collision. It provides flexible adjustment for the mathematical model to be applied under different performance requirements and network configuration. As a consequence, the back-off value after \( r \) retries is obtained as follows:

\[
BO = W_r^i(WT) + \text{RandInt}[0, W_r^{\delta}].
\]  
(7)

Let us now make two remarks on the above model.

1) WT-dependent overlap-restricted BW: WT is used to differentiate the lower bounds of BW for each SS. Each SS slides its own BW to the offset \( W_r^i(WT) \), which is calculated before starting a backoff. The offset value is set to be smaller than the number of TO within T3 so that \( W_r^i(WT) \) will not be too big to induce time waste. This overlap-restricted BW reduces the probability of choosing the same BO thus the collision rate is reduced.

2) Adaptive BW increasing rate: As analyzed in [8], the exponential increasing of BO is effective when there is a medium number of contending SSs. When the number is small or large, it is time wasting. In WDIA, the BW increasing rate \( \alpha \) is adapted to the intensity of contention. This is different from the fixed value chosen by binary exponential back-off. Under the overlap restricted BW, a big intensity of contention is indicated if an SS encountering a collision with relative bigger BO; thus a bigger rate is set. Otherwise, the rate is set to a smaller value to avoid unnecessary time waste caused by a big BW. The estimation of the number of contending SSs is based on probability of collisions happen at a SS with big BO. \( P_{c,r}^i(BO > k|n) \) is the probability of collisions of SS \( i \) when its BO value is bigger than \( k \) and the number of contending SSs is \( n \). It is derived as:

\[
P_{c,r}^i(BO > k|n) = \sum_{u=1}^{T_3/T_f-k+W_r^{\delta}-2} \sum_{m=1}^{n-1} c_{m-1}^{m-1} \left( \frac{T_f \cdot u}{T_3 \cdot W_r^{\delta}} \right)^{m+1}.
\]  
(8)

From this equation, we can get the confidence interval of \( n \) if SS \( i \) experienced a collision when its BO value is bigger than \( k \).

V. PERFORMANCE EVALUATION

In order to evaluate the efficiency of WDIA in the initial ranging period in 802.16e system, we construct a simulation environment using QualNet version 4.0.

A. Simulation Model

We set up the scenario with one central BS and several SS randomly located. In QualNet, the OFDMA PHY is supported and initial ranging is processed using RNG-REQ and RNG-RSP messages. To test the performance of initial ranging, the Poisson process (with parameter \( \lambda \)) is used for modeling the arrival of new SS. In the simulation scenario, the channel condition is ideal thus retransmission is only caused by collision and the RNG-RSP is received after successful transmission of RNG-REQ. The parameters of the simulation are shown in Table I referring to [7], [9].

B. Function setting

The setting of the deterministic and random functions is based on the analysis and adjusted through experiment.

\[
W_r^i(WT) = WT (\text{mod} \frac{T_3}{T_f}).
\]  
(9)

\[
W_r^{\delta} = \begin{cases} 
8 & \text{if } r = 0 \\
\alpha W_r^{\delta}_{r-1} & \text{if } r \neq 0 
\end{cases}
\]  
(10)

Through our simulation we get that setting a bigger increasing rate is effective when the BW is relatively small so we adapt the value of \( \alpha \) in first two transmission stages. To best benefit the contention performance we set \( \alpha = 3 \) if the first back-off value is more than 8 or the second back-off value is more than 16, otherwise \( \alpha = 2 \).

C. Simulation results

In Fig. 7, we show that the percent of collisions happens between the SSs with different retries is below 7.11%. It proved that under this timeout triggered backoff mechanism, most collisions happens between the SSs in the same retry stage. If we focus on the collision type with big probability in the analysis model, the type with a much smaller probability can be omitted. Thus assumption (3) is reasonable.
In the following figures, solid lines depict metrics of WDIA and dash lines for TBEB.

In Fig. 8 solid line shows the percent of SSs that finish IR within one UCD interval with average user arrival rate below 20/s. It proved that with the typical value of UCD interval (5s, which is \( \gg T_f \) 20ms), assumption (2) is reasonable. The dash line shows that WDIA achieves bigger value than TBEB so more SSs can be accepted to the network within one UCD interval.

Fig. 9 plots the total number of RNG-REQ transmitted by \( n \) SS to process network entry. It can be observed that the total retransmission is reduced using WDIA by 28.5% up to 37.5%. This improvement is gained from the reduced overlap of BW and the adaptive BW increasing rate, which effectively separates BW and dramatically reduces collisions.

In Fig. 10 the big impact of retransmission on delay verifies the analysis. With the concentration of contention performance, the delay is measured by the period from instant the first back-off begins to the instant the RNG-RSP is received. In consequence with the retransmission reduction, the delay caused by response timeout waiting and retry backoff waiting is greatly reduced. The ranging access delay is brought down by 30% down to 64%.

Fig. 11 shows the average number of TOs required by \( n \) SSs to finish initial ranging in one UCD interval. Less TO is required when WDIA is applied, thus the left TO may be saved for data transmission. It is a benefit for resource utilization.

**VI. CONCLUSION**

In this paper, we proposed a contention model of initial ranging and a Waiting-time Dependent and Increasing rate Adapted (WDIA) backoff algorithm for the IEEE 802.16e (OFDMA) wireless MAN. We introduce the concept of reduced overlapping between BW of contending SS for improving contention performance. The way BW are adjusted is based on the time duration requests have already waited, the so called waiting time (WT). And the contention intensity is used for optimizing the BW. The simulation results showed that in 802.16e initial ranging scenario, our WDIA outperforms the Truncated Binary Exponential Backoff (TBEB) algorithm in terms of retransmissions, access delay and resource utilization.

WDIA is efficient for all types of contention-based uplink access, including initial ranging, periodic ranging, handoff ranging and bandwidth request, under both light and heavy load conditions. In our future work, we aims to design a priority-based WDIA to achieve QoS guarantee for eRTS when contention based bandwidth request is used.

**REFERENCES**


Fig. 11. TO required for initial ranging in one UCD


