

EVALUATION AND ENHANCEMENT OF PERFORMANCE METRICS FOR IR AND BWR IN MAC 802.16

R. Bhakthavathsalam¹ and Khurram J.Mohammed²

¹Supercomputer Education and Research Centre
Indian Institute of Science, Bangalore, India.

bhaktha@serc.iisc.ernet.in

²Ghousia College of Engineering, Ramanagaram, India.

khurramashrafi@gmail.com

ABSTRACT

In the IEEE 802.16 standard, Initial Ranging (IR) is defined as the mechanism of acquiring the correct timing offsets and power adjustments such that the Subscriber Station (SS) is co-located with the Base Station (BS). Bandwidth Request (BWR) is the process by which the communication is established between BS and SS requesting the uplink bandwidth allocation. In this, we evaluate the performance of these two schemes based on the metrics of delay and throughput. Then we enhance their performance by incorporating circularity. Circularity is a paradigm that allows the identification of specific groups of packets or events. Using this, we introduce delay control and backoff window control in the case of IR and selective dropping of BWR packets. This new paradigm reduces the collisions among request packets and thereby, improves both these mechanisms. The evaluation and enhancement are performed through extensive simulated studies.

KEYWORDS

IEEE 802.16, MAC sub layer, Contention Resolution, Initial Ranging, Bandwidth Request, Circularity, Network Simulator 2

1. INTRODUCTION

The Institute of Electrical and Electronics Engineers (IEEE) 802.16 standard for Wireless Metropolitan Area Networks specifies the most recent technical features of wireless technology. Originally intended for Fixed Broadband Wireless Access (FBWA) networks and as a wireless competitor for wire-line DSL and cable modem access in particular in rural and low-infrastructure areas, the most recent stage of the IEEE 802.16 standard also provides mobility support mainly intended for nomadic users or users with little mobility. Worldwide Interoperability for Microwave Access (WiMAX) is a group established to empower the interoperability and promote the commercialization of products based on the IEEE 802.16 standard. The current IEEE 802.16-2004 standard with the extensions for mobility support amended in the IEEE 802.16e-2005 standard is the basis for two classes of WiMAX certified products [1]. The Orthogonal Frequency Division Multiplexing (OFDM) part of IEEE 802.16-2004 is known as Fixed WiMAX and the Orthogonal Frequency Division Multiple Access (OFDMA) part of IEEE 802.16e-2005 is known as Mobile WiMAX [2], [3].

1.1. Overview of 802.16 PHY and MAC layers

The IEEE 802.16 standard specifies four physical layers namely Single Carrier (SC), Single Carrier-a (SCa), single carrier transmission in line-of-sight and non-line-of-sight environments. A common Medium Access Control (MAC) sub layer is defined for all physical layers with only small adaptations to the different physical layers. The standard stipulates two types of

operation point-to-multi-point (P2MP) and mesh mode. In the following we focus on point-to-multipoint communication and the SC, SCa and OFDM physical layers. In P2MP mode of operation, the connection is established between a base station and subscriber stations/base stations or both. The MAC sub layer is present within Data Link layer of Open Systems Interconnection (OSI) model. MAC sub layer has a great role in OSI model. It serves as a link between lower hardware oriented PHY layer and other upper software driven layers. IEEE 802.16 MAC sub layer is further divided into three parts as Service specific Convergent Sub layer (CS), Common Part Sub layer (CPS), and Privacy Sub layer (PS). The MAC layer functions are Transmission Scheduling, Admission Control, Link Initialization, Fragmentation, Retransmission and support for integrated Voice/Data connections [4].

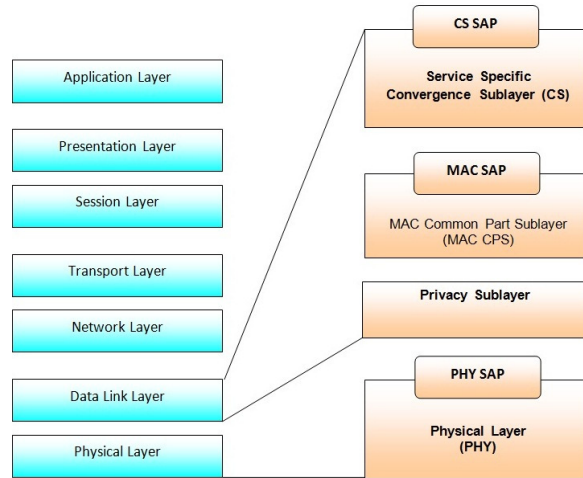


Figure 1. Protocol Layering

Figure 1 shows the positions of the different layers and sub layers mentioned above. Three types of MAC management connections are possible in MAC 802.16 to exchange control messages between the BS and SS. These connections specify different Quality of Service (QoS) requirements, needed at different management levels [5]. The BS is identified by a 48-bit Base Station Identifier (ID) and each SS is identified by a 48-bit MAC address. The 16-bit Connection Identifier (CID) identifies each connection with BS and SS in a session. In P2MP type of operation, the connection between BS and SS is established based on Request/Grant mechanism. The frame structure shown in Figure 2 best explains the operation of MAC layer [6]. Data units between the BS and SSs are exchanged in units of fixed-length frames. The frame length consists of a downlink and an uplink sub frame.

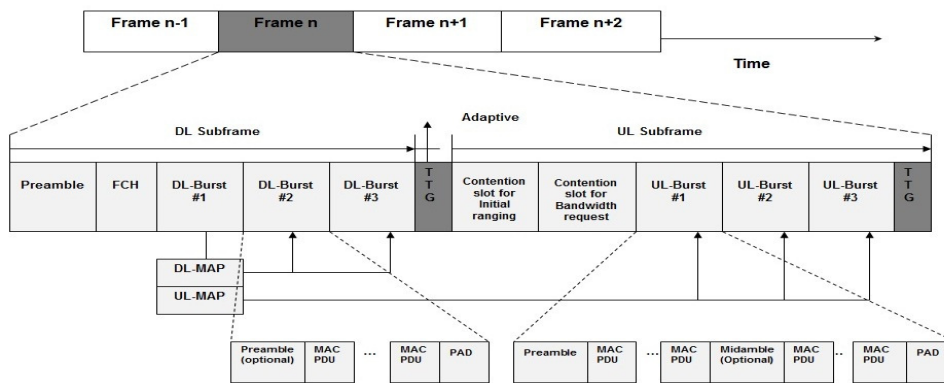


Figure 2. Frame Structure for IEEE 802.16

The transfer of data from BS to SSs takes place in the downlink sub frame and the uplink transfer occurs in opposite direction. The border between two sub frames is adaptive according to the data amount in downlink and uplink. An integer number of OFDM symbols carrying MAC datagrams are contained in each frame. These OFDM symbols comprise MAC resources available for data transmissions. Sharing of MAC resources between SSs is centralized in the BS. Downlink map (DL-MAP) and uplink map (UL-MAP) signaling messages are used to inform SSs about bandwidth allocations in downlink and uplink. At the beginning of the downlink sub frame a preamble for synchronization is transmitted, followed by the broadcast MAC signaling messages. Thereafter the bursts of user data are transmitted to the corresponding SS switch the use of different modulation types and coding rates. The Transmit/Receive (Tx/Rx) transition gaps between the sub frames allow the stations to switch between transmission and reception modes. The uplink sub frame begins with contention intervals scheduled for initial ranging and bandwidth request opportunities. IR during network entry accommodates the adjustments of time and power parameters by the new SSs. Meanwhile, BWR fulfills the needs of a SS for necessary resources for its transmission to the BS.

1.2. Network Entry Procedure

The SS needs to complete the network entry process for establishing its connections in the entire network. The different stages in the network entry procedure are shown in the Figure below [7]. The first stage of network entry is downlink (DL) channel synchronization. To begin with the SS scans for the available channels in the prescribed frequency list for its operations in the WiMAX network. The SS synchronizes at PHY level with the use of periodic frame preamble after identifying a DL channel. The DL channel descriptor (DCD) provides the information on modulation and other DL parameters, while the UL channel descriptor (UCD) contains the similar information for uplink. In Initial Ranging, the SS acquires the timing offsets and power adjustments from the BS.

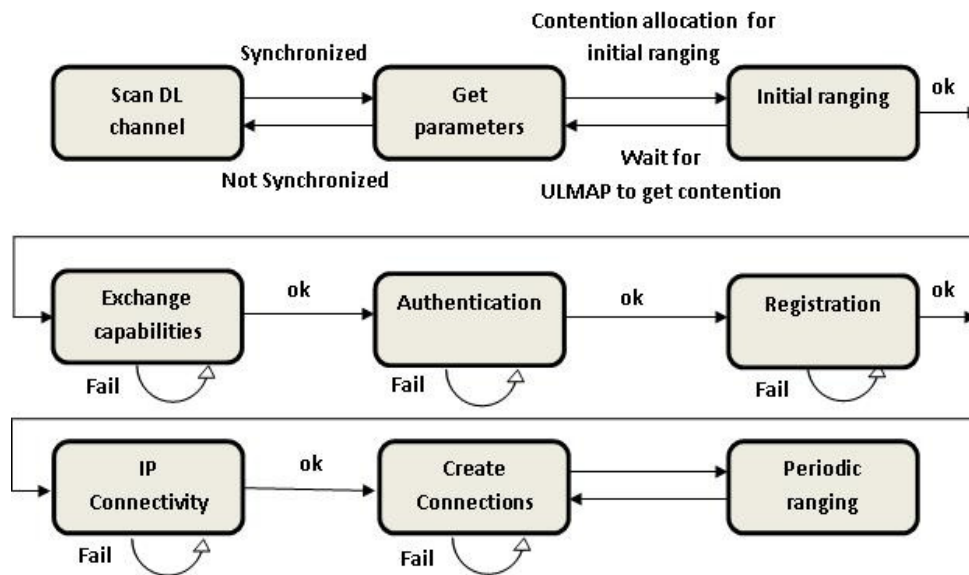


Figure 3. Network Entry Process

This enables the SS to properly communicate with the BS. IR is a very important part of the network entry procedure and is dealt with in more detail in the next section. The step of Exchanging Capabilities follows the IR completion. The Exchanging Capabilities involve rates and duplexing methods, modulation level, and coding scheme. Next follows Authentication

wherein, the BS validates the SS, determines the ciphering algorithm to be used, and sends an authentication response to the SS. In the following Registration step, the SS exchanges a registration request message to the BS and the BS in return sends the registration response. The registration response message includes the secondary management CID of the SS. Using this, a SS is allowed entry into the network and the SS is said to be manageable. Next in Internet Protocol (IP) Connectivity, the SS gets the IP address via DHCP. The SS also fetches the additional operational parameters using TFTP. After completing the IP connectivity step, transport connections are formed in the final Connection Creation.

2. IR AND BWR MECHANISMS

Initial Ranging is an important part of the Network Entry procedure performed by the SSs, upon power up, in IEEE 802.16 networks. In the IR procedure, the correct timing offsets and power adjustments are obtained from the BS so that the SS can successfully transmit data to the BS. It occurs after the SS has synchronized with a downlink channel from the BS and has obtained the uplink transmit parameters from the UCD MAC management message. After this the SS will scan the UL-MAP message to find an Initial Ranging Interval, consisting of one or more transmission opportunities allocated by the BS. Then SS initiates the IR procedure and assembles a Ranging Request (RNG-REQ) message and sends to the BS.

The SS sends this message as if it is co-located with the BS. This is done by setting the initial timing offset to the internal fixed delay equivalent to co-locating the SS next to the BS. The SS calculates the maximum transmit signal strength for initial ranging and transmits the RNG-REQ

message at a power level below this measured at the antenna connector. In case a response is not received from the BS, the next RNG-REQ message is sent at the next higher power level in the next appropriate Initial Ranging Interval. In case it receives a response from the BS, depending the contents of the response the SS does the following. If the Ranging Response (RNG-RSP) message contains the frame number in which the RNG-REQ message was sent, the SS will consider the previous attempt to be unsuccessful. Nevertheless, it will make the adjustments specified in the RNG-RSP message. If the RNG-RSP message contains the MAC address of the SS, then the request attempt will be considered successful. When the RNG-REQ message is successfully received by the BS, it will send an RNG-RSP message using the Initial Ranging CID. At this point the BS starts using Invited Initial Ranging Intervals addressed to the Basic Connection Identifier of the SS to complete the process of IR. But if the status in the RNG-RSP is success, the IR procedure will be completed. On receiving an RNG-RSP message with continue status, the SS first makes the power level and timing adjustments. Then it retransmits another RNG-REQ using the Basic CID assigned to it. The BS yet again sends an RNG-RSP message containing additional fine tuning, if required. This exchange of RNG-REQ and RNG-RSP messages continues till an RNG-RSP message with status success is received by the SS or the BS aborts the IR procedure.

Whenever the SS has to transmit the request packets it performs the Truncated Binary Exponential Backoff procedure. This method is the contention resolution procedure used in IEEE 802.16 networks [8], [9]. The minimum backoff window and the maximum backoff window are both controlled by the BS. Initially the SS set its backoff window to the minimum possible backoff window. Now the SS randomly selects a number from this backoff window. This number selected indicates the number of IR transmission opportunities that the SS must defer before transmitting the request packet. After the selected number of transmission opportunities is deferred, the SS transmits the RNG-REQ message.

After transmitting the request message, the SS waits for a response message from the BS. If the RNG-RSP message is received from the BS before the specified timeout then the contention resolution is considered to be a success. If not, the SS doubles its backoff window until the maximum backoff window is reached. It then randomly selects another number from this new

window and the deferring process is repeated. This may happen due to the collision of RNG-REQ packets or due to the loss of RNG-RSP messages. There exists a maximum limit for the number of such IR retries allowed. If this limit is reached by an SS, then the particular downlink channel being used is marked as unusable and it begins scanning for a new downlink channel. The backoff windows always increase in terms of powers of two. For example, the backoff window at a particular instant for an SS is lying between 0 and 31 (0 to $2^5 - 1$) and the random number picked is 11. The SS has to defer a total of 11 Initial Ranging Intervals before transmitting the RNG-REQ packet. This may require the SS to defer IR intervals over multiple frames. In case a collision is detected, the backoff window is increased twofold. Next a random number is picked between 0 and 63 and the deferring process is continued repeatedly for a maximum of 16 times after which the uplink channel is restarted from the beginning. These basic steps of the contention resolution procedure are identical to both Initial Ranging and Bandwidth Request. These steps are shown in Figure 4.

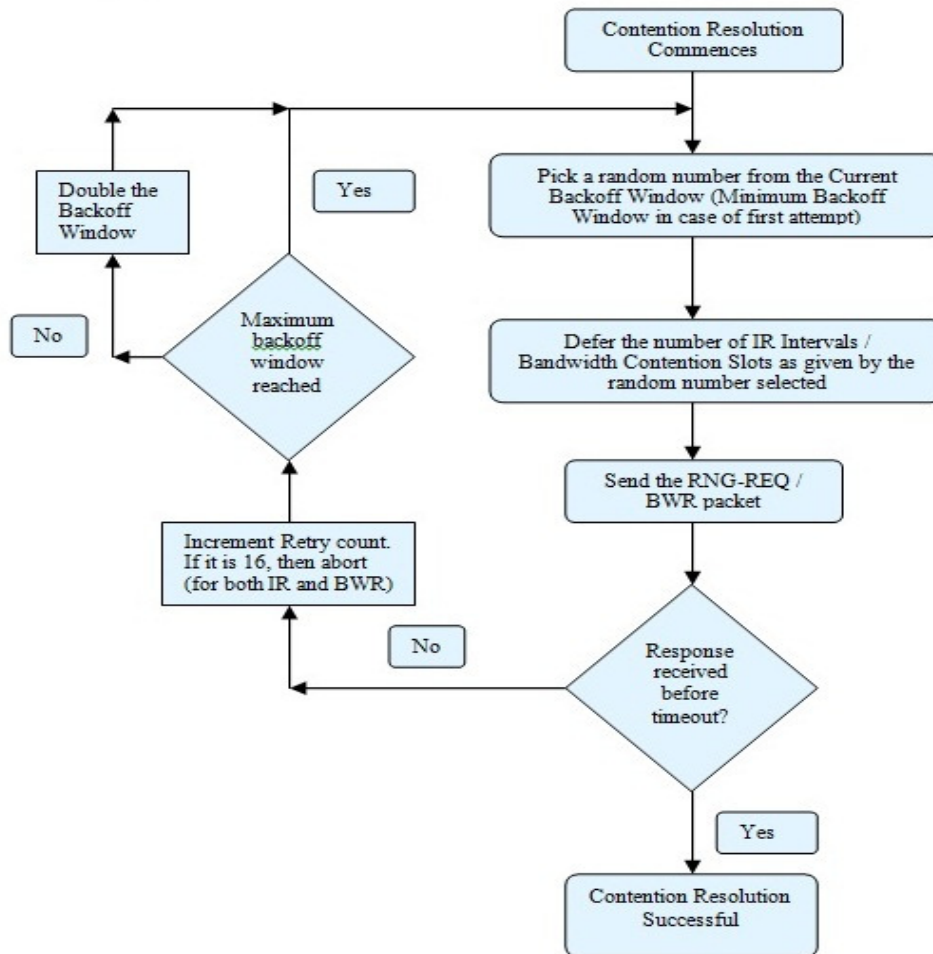


Figure 4. Contention Resolution Process

Bandwidth Request deals with the process through which the SS demands the bandwidth requirements for the uplink channel for data transmission [10], [11]. In Initial Ranging, the RNG-REQ comprises only single type of request, whereas BWR packets contain multiple types such as stand-alone, piggybacked BWR packets and also a either incremental or aggregate BWR messages. In all cases, the data requests are made in terms of the number of bytes required to be sent, but not in terms of the channel capacity. This is due to fact that the uplink burst profile

keeps changing dynamically. In the uplink sub frame of the TDD frame structure of 802.16, there are a number of Contention slots present for the purposes of BWR [12].

In contention based BWR mechanism, an SS tries to transmit its BWR packet to the BS in one of the contention slots found in the uplink sub frame. Under contention free condition, if a BWR packet is successfully received by the BS, the requested bandwidth is fully available. Once the bandwidth is reserved for the SS, it can freely transmit its data without contention in the next frame. In case of multiple SSs sending their request messages to the same BS simultaneously, it gives rise to collisions. So, here again the contention is resolved by using the truncated binary exponential backoff procedure, which is the same as in the case of IR. The BS decides the range of backoff windows from minimum to maximum numbers.

Also, the number of bandwidth request trials is limited to 16. One can measure the performance of a network in many ways. This is mainly due to the variety of networks that are present. Some of the metrics that can be used to measure network performance are delay, throughput, availability, bandwidth, etc. We have selected delay and success-ratio of the Ranging Request packets as the metrics for measuring the performance of the Initial Ranging mechanism. These two metrics are also used to compare the performances of IR before and after incorporating the paradigm of circularity. Also, the delay metric is used for the BWR mechanism to show the effect of incorporating circularity.

3. ANALYSIS OF IR AND EVALUATION OF TOTAL DELAY

In this section we analyze the IR scheme and evaluate mathematically the delay involved in the procedure. First we structure the IR mechanism as a set of distinct states with transitions among these states. Then this information is used to model IR as a Markov process. In a Markov process, the probability of the system making a transition to a particular state depends only on the state the system is currently in [13]. Also, in this Markov process, we calculate the delays associated with the transitions between the states. Finally, by making use of the delays and probabilities associated with each of the transitions, mathematical equation describing the total IR delay is obtained.

3.1. Modeling IR as a Markov Process

Markov processes offer very flexible, powerful, and efficient tools for describing and analyzing the dynamic system properties. Thus they help us to easily estimate the performance and dependability measures. Markov processes unify the fundamental theory underlying the concept of queuing systems. The notation of queuing systems has been considered as a high-level specification technique for (a sub-class of) Markov processes. The queuing systems can be mathematically evaluated by modeling them as Markov process. But besides highlighting the computational relation between Markov processes and queuing systems, it is worthwhile pointing out also that fundamental properties of queuing systems are commonly proved in terms of the underlying Markov processes. This type of Markov processes is also possible even when queuing systems exhibit properties such as non-exponential distributions that cannot be represented directly by discrete-state processes. Markovian models, such as embedding techniques or supplementary variables, can be used in such cases. These models serve as the theoretical framework for determining the correctness of computational methods applied directly to the analysis and evaluation of queuing systems [14].

A stochastic process is defined as a family of random variables $\{X_t; t \in T\}$ where each random variable X_t is indexed by parameter t belonging to T , which is usually called the time parameter if T is a subset of $R^+ = [0, \text{infinity})$. The set of all possible values of X_t (for each $t \in T$) is known as the state space S of the stochastic process. Many numbers of stochastic processes belong to an important class of Markov models. Thus, a Markov model represents a stochastic process

where the future state of the process depends only the current state and not on the history of the system. A Markov process is a memory-less stochastic process [13]. In the reference [15], the Backoff procedure is modeled as a Markov process. Here, we extend this to incorporate the entire Initial Ranging mechanism. After analyzing IR, we enumerate the following states as well as transitions needed for modeling the procedure.

1. Waiting for UL-MAP. This is also the start state.
2. SS is performing Backoff procedure.
3. Waiting for an RNG-RSP message from BS.
4. Continue
5. Success State – Wait for CDMA Allocation IE.
6. Abort – Start network entry procedure at a different DL channel
7. Waits for RNG-RSP again.
8. Proceed to next phase of network entry
9. Commence Periodic Ranging

The transitions among the states are as follows. In State 1, the SS waits for a UL-MAP. After receiving this message it makes a transition to State 2. Transmission of CDMA code occurs at end of State 2. Also a timer is set for waiting for RNG-RSP message. This transition leaves the system in State 3. When in State 3, if the timer for RNG-RSP expires then SS increments the power level and goes back to State 1. When in State 3, if RNG-RSP is obtained with Ranging code as well as the Ranging slot, then it makes a transition to State 4. Here the necessary adjustments specified in RNG-RSP are made and system moves to State 1. When in State 3, if RNG-RSP is obtained with success status, then the system transits to State 5. Here it waits for CDMA Allocation IE. After reception it sends RNG-REQ message on the allocated bandwidth and moves to State 7. When in State 7, on reception of RNG-RSP with success status it moves to State 8. On reception of RNG-RSP with continue status it moves to State 9. Else on reception of RNG-RSP with abort status, it goes to State 6 and SS starts the network entry procedure again. When in State 3, if RNG-RSP is obtained with abort status then the system goes to State 6 and SS starts the network entry procedure again. The following matrix diagram shows the Markov process that represents the Initial Ranging procedure of IEEE 802.16 network standard. The Transition Probability Matrix corresponding to the states in this process is shown in table 1.

Table 1. Probability matrix corresponding to the states

	1	2	3	4	5	6	7	8	9
1	0	1	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0
3	a1	0	0	a2	a3	a4	0	0	0
4	1	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0
7	0	0	0	0	0	b3	0	b1	b2
8	0	0	0	0	0	0	0	1	0
9	0	0	0	0	0	0	0	0	1

The transitions out of states 3 and 7 are non-deterministic and therefore we have used algebraic symbols to represent the probabilities associated with outgoing transitions. Using these probabilities, we design the Markov process representation of IR as shown in Figure 5. The states 6, 8 and 9 lead out of the IR mechanism and are the absorbing states. Next, we use the

transition matrix obtained above to obtain the overall delay formula. For this, we first need to tabulate the delays involved in the individual states. The states 6, 8 and 9 are absorbing states. For these states, the probability of transition to themselves is 1.

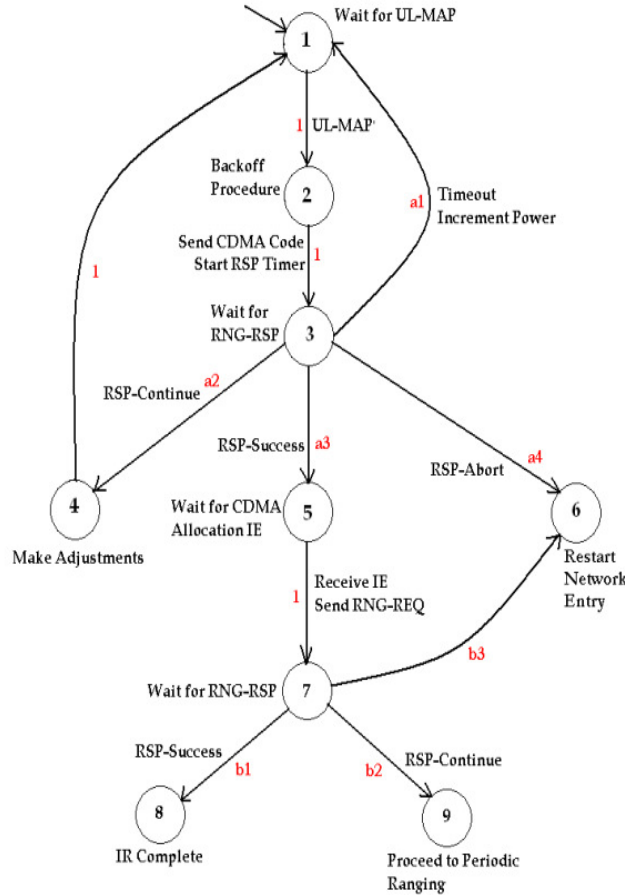


Figure 5. IR Procedure as a Markov Process

Table 2. Details of the delays involved along with the associated probabilities

Delay Involved	Probabilities
UL-MAP Reception (1 to 2)	1
Backoff Delay + Sending CDMA (2 to 3)	1
RNG-RSP Timeout (3 to 1)	a1
RNG-RSP Reception + Processing (3 to 4, 5 or 6)	a2,a3,a4
IE Allocation Delay + Sending RNG-REQ (5 to 7)	1
RNG-RSP Reception + Processing (7 to 8, 9 or 6)	b1,b2,b3

The details of the delays involved along with the associated probabilities are given in the table below. In states 3 and 7, the outgoing probabilities are marked with algebraic symbols a1 to a4 and b1 to b3. This is because the probabilities of the transitions originating from these states are non-deterministic in nature. Nevertheless, the sum of probabilities of all transitions originating from states 3 and 6 are still equal to 1.

The numerical values of the delays involved are expressed below where the symbol \rightarrow indicates the transition from one state to other as depicted above in Figure 5 ($i \rightarrow j$, $1 \leq i, j \leq 9$):

1. UL-MAP Reception (UL_M) = 5ms (Maximum of one complete frame length)[1 \rightarrow 2]
2. CDMA Sending Time = $CDMA_{ST}$ = Transmission Time = $5ms/2 = 2.5ms$
[Length of UL subframe= (Frame Length/2) where frame length is 5ms][2 \rightarrow 3]
3. RNG-RSP Timeout = $T_3 = 200$ ms[3 \rightarrow 1]
4. RNG-RSP Reception (RSP) + Processing (average value)
= $T_3/2 + \text{Max. RNG-RSP Processing Time}/2$
= 100 ms + $10ms/2 = 105$ ms[3 \rightarrow 4, 5, or 6]
5. CDMA Allocation IE delay = $CDMA_{IE} = 5ms$ (same as 1)[5 \rightarrow 7]
6. Sending RNG-REQ (Same as 2) = $2.5ms$ [5 \rightarrow 7]
7. RNG-RSP Reception + Processing (average value) = $105ms$ [7 \rightarrow 8 or 9]

We assume that the delay involved for making changes at SS is negligible compared to the other delays involved.

3.2. Mathematical Derivation of the Backoff delay

Consider the first time an SS enters Backoff procedure. Let the initial Contention Window be w_0 . The random number will be picked in the range $[0, w_0-1]$. Let this random number be called 'k'. The SS has to defer a total of k contention slots. Let the number of CS's in a frame be n_{cs} . The number of frames that have to be deferred is k/n_{cs} . The delay involved here will be $(k/n_{cs}) * \text{frame length}$. After k/n_{cs} frames have passed the SS defers a further $(k \text{ modulo } n_{cs})$ Contention Slots. The delay involved here is equal to $(k \text{ mod } n_{cs}) * T_{cs}$, where T_{cs} is the length of one CS.

$$\text{Total delay so far} = (k/n_{cs}) * \text{frame length} + (k \text{ mod } n_{cs}) * T_{cs} \quad (1)$$

Here the value of k can vary from 0 to w_0-1 . Thus, we take an average of the delay (AD_0) over the random number k.

$$AD_0 = (1/w_0) * \sum_{k=0}^{(w_0-1)} \left(\frac{k}{n_{cs}} \right) * \text{frame length} + (k \text{ mod } n_{cs}) * T_{cs} \quad (2)$$

Next we make an assumption that the probability of a successful transmission in a CS is 'p'. Thus, probability of failure will be '1-p'. In case of a failure the contention window is doubled in size. Let the new window be $[0, w_1-1]$. Similar to equation (1a), the delay involved will be

$$AD_1 = (1/w_1) * \sum_{k=0}^{(w_1-1)} \left(\frac{k}{n_{cs}} \right) * \text{frame length} + (k \text{ mod } n_{cs}) * T_{cs} \quad (3)$$

Here $w_1 = 2 * w_0$

Again there could be success or failure. So, it will enter the third Backoff window phase $[0, w_2-1]$. Continuing in this fashion, we get the following delays for the next three phases.

$$AD_2 = (1/w_2) * \sum_{k=0}^{(w_2-1)} \left(\frac{k}{n_{cs}} \right) * \text{frame length} + (k \text{ mod } n_{cs}) * T_{cs} \quad (4)$$

Here $w_2 = 2 * w_1$

$$AD_3 = (1/w_3) * \sum_{k=0}^{(W_3-1)} \binom{k}{n_{cs}} * \text{frame length} + (k \bmod n_{cs}) * T_{cs} \quad \text{Here } w_3 = 2 * w_2 \quad (5)$$

$$AD_4 = (1/w_4) * \sum_{k=0}^{(W_4-1)} \binom{k}{n_{cs}} * \text{frame length} + (k \bmod n_{cs}) * T_{cs} \quad \text{Here } w_4 = 2 * w_3 \quad (6)$$

We make another assumption at this point. The SS is assumed to complete successful transmission of its CDMA code, in a maximum of 5 Backoff phases. Thus, the worst case of transmission will be four failures followed by a success. The final formula for the Backoff delay (BD) will be as follows:

$$\begin{aligned} BD &= p * \{AD_0 + t/2\} \\ &+ ((1-p) * p * \{[AD_0 + t] + [AD_1 + t/2]\}) \\ &+ ((1-p)^2 * p * \{[AD_0 + AD_1 + 2t] + [AD_2 + t/2]\}) \\ &\quad + ((1-p)^3 * p * \{[AD_0 + AD_1 + AD_2 + 3t] + [AD_3 + t/2]\}) \\ &+ ((1-p)^4 * p * \{[AD_0 + AD_1 + AD_2 + AD_3 + 4t] + [AD_4 + t/2]\}) \end{aligned} \quad (7)$$

Here t is the time-out after which failure is assumed. So, we take half that value for success i.e. $t/2$.

3.3. Mathematical Derivation of the Overall IR Delay

By traversing the transition diagram (Figure 5) and multiplying the probabilities with the corresponding delays (Table 2), the total delay can be calculated. The resulting formula is as follows. The first part of the delay is in the loops 1-2-3-1 and 1-2-3-4-1. We call this D_{loop} . Then either success or abort occurs which is added to this part to get the final formula.

$$\begin{aligned} D_{loop} &= 1 * UL_M + 1 * (BD + CDMA_{ST}) \\ &\quad + a1 * (T3 + D_{loop}) \\ &\quad + a2 * (RSP + D_{loop}) \end{aligned} \quad (8)$$

Simplifying we get,

$$D_{loop} = \frac{UL_M + BD + CDMA_{ST} + a1 * T3 + a2 * RSP}{1 - (a1 + a2)} \quad (9)$$

Now, the total delay involved can be represented using the formula given below.

$$\begin{aligned} D_{total} &= D_{loop} \\ &\quad + a3 * [RSP + CDMA_{IE} + RNG-REQ + (b1 + b2 + b3) * RSP] \\ &\quad + a4 * RSP \end{aligned} \quad \text{(Here } b1 + b2 + b3 = 1) \quad (10)$$

Substituting the expression for the delay in the loop into the formula for overall delay in IR, we get the following final formula.

$$\begin{aligned} D_{Total} &= \frac{UL_M + BD + CDMA_{ST} + a1 * T3 + a2 * RSP}{1 - (a1 + a2)} \\ &\quad + a3 * [RSP + CDMA_{IE} + RNG-REQ + RSP] \\ &\quad + a4 * RSP \end{aligned} \quad (11)$$

We define the Initial Ranging delay as the time taken by an SS to complete the IR scheme. Therefore, this is the time elapsed from the moment when an SS finds an IR opportunity to the moment when it receives an RNG-RSP message with a success status from the BS. This is done for the purposes of comparing the IR delay before and after the application of circularity. The IR delay process consists of the transmission delays of the request and response messages, the time needed for contention resolution, and the time needed by SS and BS to process the messages received.

$$\text{IR Delay} = \{ \text{Time at which RNG-RSP with Success status is received} \} \quad (12)$$

$$- \{ \text{Time at which the first IR opportunity is found} \}$$

During the IR procedure, multiple SSs can send their request messages to the same BS simultaneously giving way to the occurrence of packet collisions. This leads to the contention resolution procedure being restarted with double the size of the backoff window. This leads to an increased IR delay. Thus, the contention resolution phase is the most affected by collisions among packets. We are interested in reducing the time spent by the SSs in competing with each other to send their request messages to the BS. This directly reduces the overall IR delay as well. With respect to the IR delay, we consider two approaches. Firstly, its reduction upon implementing circularity in IR is demonstrated. Then, the effect of applying circularity on BWR, IR and both BWR and IR upon the IR delay is studied.

4. ANALYSIS AND EVALUATION OF IR PERFORMANCE METRICS: SUCCESS RATIO

We define the IR success ratio as the ratio of the number of successfully completed IR procedures of various SSs to the sum of successfully completed IR procedures of various SSs and the number of retransmissions needed to be done as a result of the request packets timing out. This ratio is also directly affected by the collisions between RNG-REQ packets. This metric can be mathematically expressed as shown in the equations below:

$$\text{Success ratio} = \frac{\text{Successful attempts}}{\text{Total number of attempts}} \quad (13)$$

$$\text{Success ratio} = \frac{\text{RNG-RSP with Success Status}}{\text{RNG-RSP with Success Status} + \text{RNG-REQ Expire}} \quad (14)$$

Initially the success ratio is measured under 2 cases namely under the existing IR scheme and after implementing circularity for IR. Here, the effect of changing circularity values for delay and window control is also shown. Then, the success ratio is measured under 4 cases. They are existing IR scheme, with circularity for IR, with circularity for BWR and with circularity for both IR and BWR. In the next section we explain the paradigm of circularity that aims to reduce the number of collisions between request packets sent by various SSs.

5. PERFORMANCE ENHANCEMENT OF IR AND BWR

In IEEE 802.16 networks, the IR scheme is used by the SSs in order to acquire the timing offsets and the power adjustments from the BS, so that it can successfully transmit data packets. Although the mechanism is completely defined in the 802.16 standard, the performance of this mechanism is affected by the collisions between the RNG-REQ packets sent by different SSs.

In this section we propose an enhanced mechanism for IR, which incorporates the principle of circularity. Circularity is a novel technique whose objective is to reduce the number of collisions between the request packets in contention scenarios. It is defined as a number that allows us to identify specific groups of events or packets in the network [16].

The number of packets or events in one such group is equal to the circularity value. In each group, one of the packets or events is said to be circularity-satisfied. Here, we introduce certain control measures in case of circularity-satisfied packets and events. By doing this we achieve a decrease in the IR delay as well as an increase in the IR throughput. The circularity value is a positive integer. In order to identify the circularity-satisfied packets or events, we keep a count of the number of such packets or events. This count is common for all the SSs since we do not keep an individual counter for each SS. Whenever the value of this counter is a multiple of the circularity value, the packet or event is said to be circularity satisfied. If the counter is represented by k and the circularity value by c , then the mathematical representation for satisfying circularity is as follows

$$k \text{ modulo } c = 0 \quad (15)$$

The control measures taken are the following. Before sending the first RNG-REQ message or after sending its RNG-REQ packet if the SS does not receive a RNG-RSP message before a timeout, the RNG-REQ is said to have timed out. Then the SS doubles its backoff window and restarts the contention resolution procedure. The number of such expire events is counted. When an 'expire' event is circularity satisfied, the backoff window is doubled an extra time. By setting the backoff windows this way, in case of circularity satisfied expire events, the random numbers chosen by the different SSs will have lesser probabilities of being equal. This would mean that the backoff counters of the SSs would also have lesser probabilities of reaching zero at the same instant. Hence, the likelihood of collisions among the request packets decreases. After the requisite number of IR intervals is deferred, the SS is ready to send its RNG-REQ packet. We keep a count of such RNG-REQ packets as well. In the case of circularity satisfied RNG-REQ packets, we introduce a certain finite delay before the RNG-REQ packet is sent on the Initial Ranging Interval. Due to the delay introduced a particular request packet is sent a little later than it should have been. So, this sacrifice allows another SS to send its request packet in the meantime. This further reduces the probability of these packets colliding with each other.

In case of the BWR mechanism, the BWR packets are also kept count of. When the packet count of a BWR packet is a multiple of circularity, the packet is said to be circularity-satisfied. Such a BWR packet is discarded [17]. The events that occur in the IR intervals and the BWR contention slots of a particular frame affect the IR intervals and BWR contention slots of not only the next frame but also the frames after that. Consider the BWR contention slots of a particular frame 'U'. If there are many BWR packets sent successfully in a frame, then in the next frame 'U+1' there will be a larger number of uplink data bursts allocated for the data to be sent. This would of course be dependent on the availability of bandwidth in the next frame. A larger amount of data being sent implies fewer resources available on the uplink for contention based processes namely IR and BWR. Next consider the events in the IR intervals. After the successful transmission of RNG-REQ packets, the SS will complete the other stages of network entry and commence their BWR mechanisms. Thus, the IR process affects the activities of the BWR contention slots of the subsequent frames. Therefore, the activities of IR and BWR in a particular frame affect each other in the future. Hence, we try to establish a more efficient usage of the resources assigned by the base station for the contention-based processes. This is what circularity aims to do. Although our main focus remains the enhancement of the IR scheme, the implementation of circularity for BWR and its effect on IR is also dealt with in the next section.

6. ANALYSIS AND ENHANCEMENT OF BWR

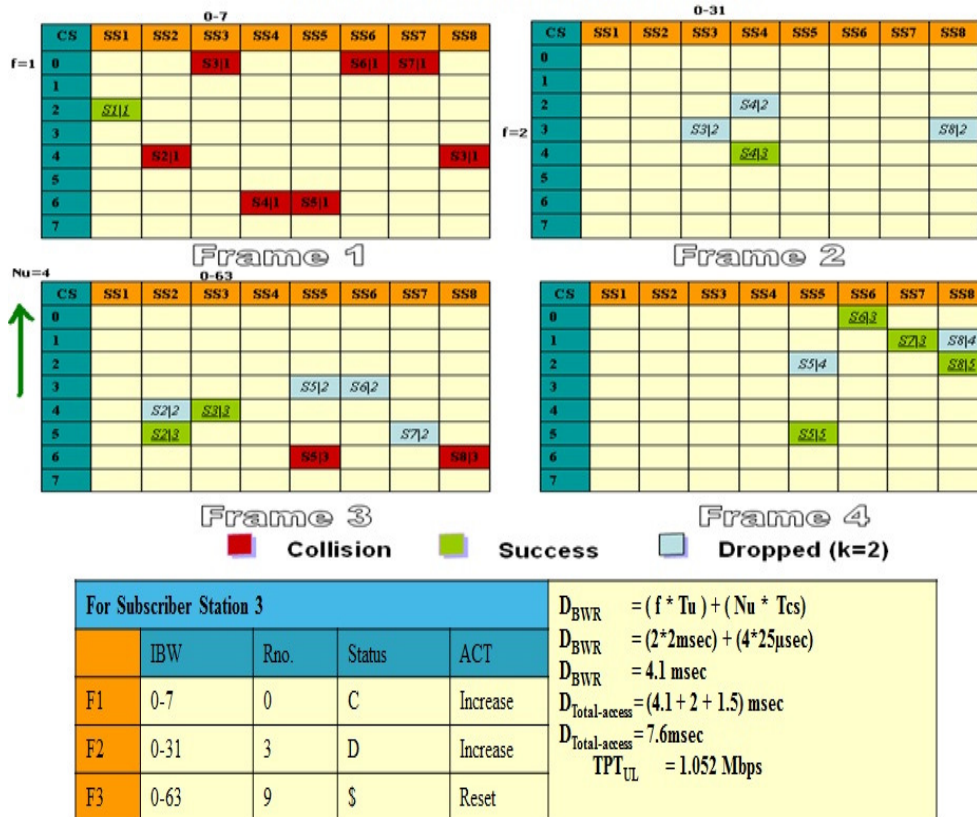


Figure 6. Frame Analysis of BWR with Circularity = 2

In order to explain the effects of the principle of circularity on the bandwidth request mechanism, we carry out a frame analysis that shows the events that occur during the mechanism. The frame analysis of BWR with the principle of circularity is as shown in the Figure 6. Let us explain the analysis taking SS1 as an example. First, SS1 tries to transmit its BWR in the first frame i.e. (F1) in the Contention Slot 2. Here, no other SS is transmitting its BWR and SS1 is able to transmit its request message without any collision. Next we consider SS2. SS2 tries to transmit its BWR in the first frame in CS4. But in this CS, SS8 is also sending its BWR. Hence these packets will collide which is denoted by C.

After the collision each SS, doubles its backoff window (if it has not reached a maximum already). Then they will select a number randomly from the new backoff window. The requisite number of BWR CS is deferred. The second BWR packets of SS2 and SS8 have been dropped, since they are circularity satisfied. This is denoted by D. The backoff window is doubled again. SS2 selects the number 10 and SS8 selects the number 11. With this value selected SS has a successful transmission of its BWR packet. But, SS8 again experiences a collision. The standard backoff procedure is continued in this case. Since the circularity value chosen is k = 2 for simplicity, every alternate request packet is dropped by each SS. Another form of frame analysis is shown in the following figure illustrating the events associated with each SS on a frame-by-frame basis. The performance metrics used to compare the BWR scheme with and without circularity are the access delay and throughput.

We define the access delay of the BWR scheme as follows [17].

$$D_{\text{Total-access}} = D_{\text{BWR}} + D_{\text{gr}} + T_{\text{data}} \quad (16)$$

Where D_{BWR} is the time spent in the contention process. This can be expressed mathematically as follows:

$$D_{\text{BWR}} = (f \cdot T_U) + (N_U \cdot T_{\text{CS}}) \quad (17)$$

After calculating the total access delay we compute the uplink throughput, for transmitting data of 1000 bytes, using the formula given below.

$$\text{TPT}_{\text{UL}} = \frac{\text{Bytes sent} \cdot 8}{D_{\text{Total-access}}} \quad (18)$$

7. SIMULATION STUDIES

The following sections describe in detail the simulation scenarios of various cases.

7.1. Simulation Setup

The simulations have been carried out using the Network Simulator 2 (ns-2) which is a discrete event simulator. We have added the WiMAX patch developed by the Advanced Network Technologies Division of the National Institute of Standards and Technologies [18]. The simulation script is written in the Tool Command Language (Tcl) [19], [20]. The model currently implemented is based on the IEEE standard 802.16-2004 and the mobility extension 802.16e-2005 [21]. The model currently supports TDD. In this model, the uplink transmission follows only after the downlink occurs in each frame. The DL_MAP and UL_MAP messages determine the bursts of allocation and transmission opportunities for each station. The BS allocates slots that involve contention in the uplink direction. These slots are used in two cases:

- Initial Ranging request
- Bandwidth request

This model resolves contention by using a truncated binary exponential backoff procedure. The BS broadcasts the UCD messages that determine the window sizes. Also the BS decides the number of contention slots allocated in each frame.

The WiMAX control agent is used in the Tcl script in order to produce a detailed account of the activities in the network [21], [22]. The network configuration used is as follows. A single base station (BS) is considered. A sink node is considered that is attached through a wired link to the BS. The number of subscriber stations is simulated with BS. The performance metrics used are the delay, collision and throughput. The values for these metrics are calculated from the output file generated by the WiMAX control agent.

7.2. Implementation of Circularity

The basic network component objects in the data path are written in C++. The event scheduler is also scripted and compiled using C++. These compiled objects are made available to the OTcl interpreter through an OTcl linkage that creates a matching OTcl object for each of the C++ objects and makes the control functions and the configurable variables specified by the C++ object act as member functions and member variables of the corresponding OTcl object [23]. The backend files of the WiMAX module are coded in C++ are modified to implement the principle of circularity [24].

Table 3. List of important parameters that have been used during the simulation

channel Type	WirelessChannel
Radio Propagation Model	TwoRayGround
Network Interface Type	Phy/WirelessPhy/OFDM
MAC Type	802_16
Interface Queue Type	DropTail Priority Queue
Link Layer Type	LL
Antenna Model	OmniAntenna
Maximum Packets in Interface Queue	50
Routing Protocol	DSDV
BS coverage	20 meters
Simulation Time	50 seconds

In the case of Initial Ranging, we introduce two modifications namely delay control for the request packets and window control for the backoff Window. Firstly, a counter is kept for the RNG-REQ packets that are scheduled to be sent from the different SS. Whenever the counter value is an integer multiple of the circularity value, a finite delay is introduced. Secondly, a counter is kept of the number of RNG-REQ expire events that occur. Whenever this counter is a multiple of the circularity value, the Backoff Window is quadrupled (instead of the original doubling). This is done provided the Backoff Window does not exceed the designated maximum.

In the case of Bandwidth request, we selectively drop the request packets. A counter is kept for the BWR packets scheduled to be sent by the SSs. When this counter value is a multiple of the circularity value, the corresponding BWR packet is dropped.

7.3. Simulation Results

The following sections describe the simulation analysis and comparison of various cases for IR, BWR and both combined IR and BWR modules using ns-2 with different circularity values.

7.3.1. Simulation with IR

We describe the results of the simulations conducted using ns-2 [25], [26] in this section. In the first graph shown in Figure 7, we compare the delay incurred in the IR mechanism in the existing and enhanced scenarios. The circularity value used in selectively delaying the RNG-REQ packets is 3. The circularity value used in selectively doubling the backoff window an extra time is 5. The second graph compares the success-ratio of IR between the existing and improved IR schemes. There are three sets of Circularity values that are plotted in the graph shown in Figure 8 along with the existing IR scheme.

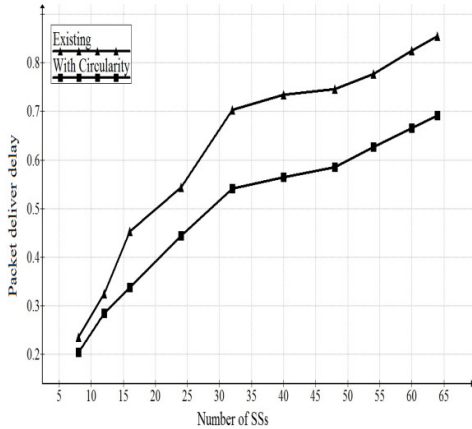


Figure 7. IR Delay comparison

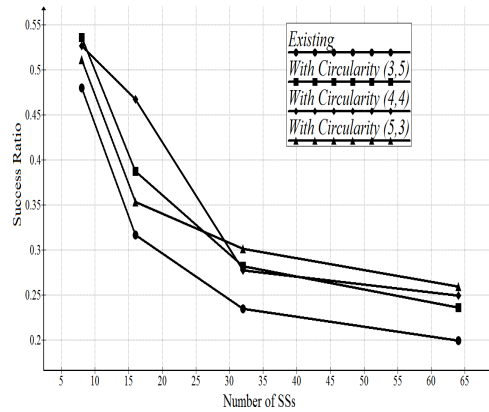


Figure 8. IR Success-Ratio comparison across three different sets of circularity valuepairs

Three circularity value pairs, used for delay control and window control, are considered. They are (3, 5), (4, 4) and (5, 3). The implementation of circularity in the IR scheme has decreased the delay incurred and increased the throughput. Within the improved scenario we observe the following. Around the 8 node scenario, for small network sizes, the circularity value pair (3, 5) gives higher throughput. Around 16 node scenario the circularity value pair (4, 4) gives higher throughput. For higher numbers of SSs (32 and 64) the circularity value pair (5, 3) gives higher throughput. From the above observation the following conclusions can also be made. With increasing number of SS, the circularity value controlling the delay must be increased. This implies that for increasing network sizes, the introduction of delay must be less frequent. With increasing number of SS, the circularity value controlling the window size must be decreased. This implies that for increasing network sizes, the doubling of window size must be more frequent.

7.3.2. Simulation with BWR

Now, we digress a bit and show the reduction in the time taken by SSs to complete their BWR mechanisms. This is achieved by applying the principle of circularity to selectively drop the BWR packets.

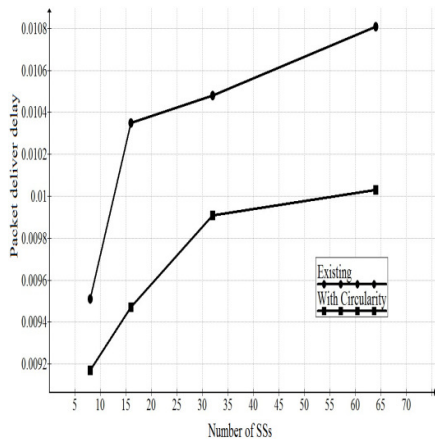


Figure 9. BWR Delay comparison with circularity using ns-2

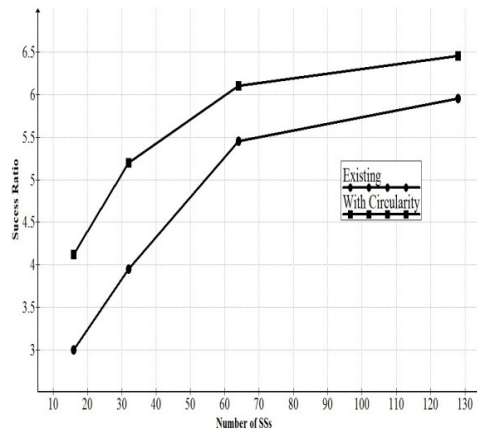


Figure 10. BWR Delay comparison with circularity using generic code (C++)

This achieves a significantly more efficient utilization of the BWR contention slots that are available on the uplink. The graph showing this improvement follows. By altering Tcl code of BWR file, we can observe significant changes in access delay [27] with circularity for different test cases. Let us consider three different test cases in the following:

Case 1: Placing all Mobile Nodes (SS's) crowded.

Case 2: Placing Mobile Nodes (SS's) equidistant.

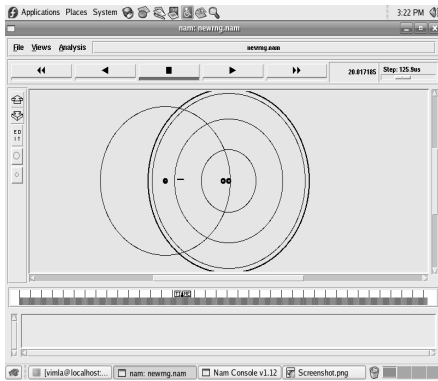


Figure 11.Placement of all mobile nodes accessing from same point

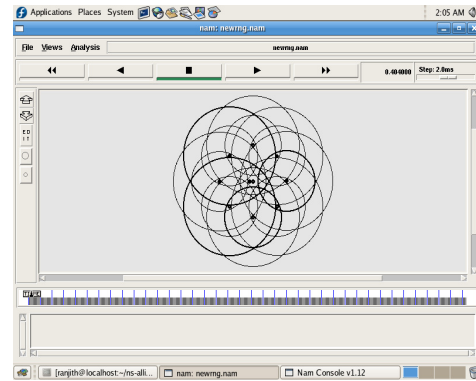


Figure 12.Placement of all mobile nodes accessing from equidistant point

Case 3: Mobile Nodes (SS's) on the circumference.

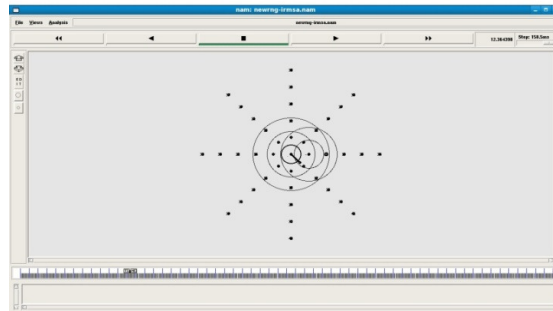


Fig.13.Circumferential Topology with no motion

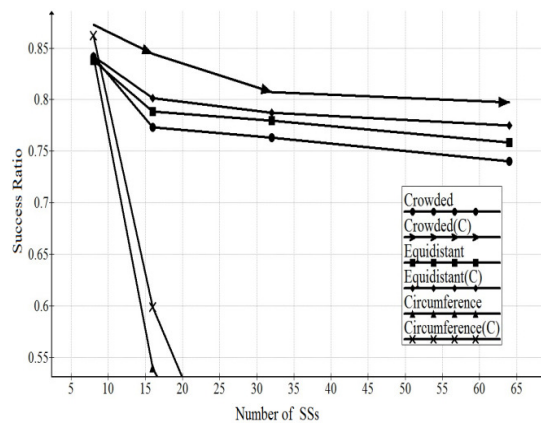


Figure 14.Comparison of various topological scenarios.

Simulation results of delay and throughput for the Case 1 wherein the SSs are densely placed in the same location are depicted in the Figures 9 and 10. The Figures 11, 12 and 13 illustrate different topological scenarios consider for further simulation. The comparative results are presented in Figure 14. The enhanced performance is evinced in all cases while adopting the principle of circularity for improving the contention resolution process. In circumferential topology, the coverage area between BS and SSs expands farther away as network size increases. This results in drastic drop in the throughput.

7.3.3.Simulation with IR and BWR

Due to the influence of contention in the IR and BWR mechanisms on each other, a more efficient collision avoidance mechanism will result in a lesser delay and a higher success-ratio for the IR mechanism. This is proved by the following graphs. Four cases are considered here. Firstly in Figure 15, the delays for the existing IR scheme and the IR scheme with circularity are plotted. Then the delay with circularity implemented for BWR and both IR and BWR are shown.

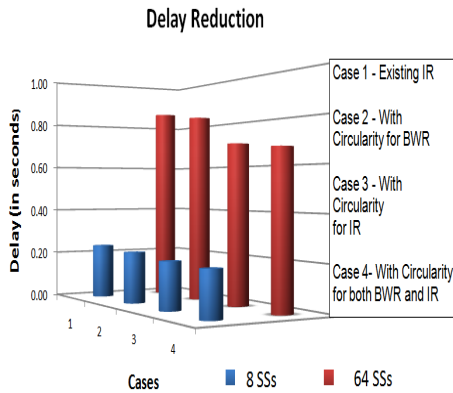


Figure 15. Variation of IR delay across four cases of circularity implementation

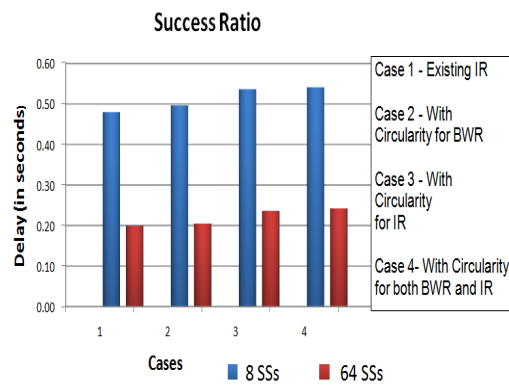


Figure 16. Variation of IR success ratio across four cases of circularity implementation

In Figure 16, the same four cases are considered for the success-ratio of the IR mechanism. From the third and fourth set of results, the influence of the efficient use of IR intervals and the BWR contention slots on the IR delay and the success ratio of RNG-REQ packets are established. Irrespective of whether circularity is applied to IR or BWR, the IR performance improves. Also better IR performance is observed when circularity is applied to IR mechanism. Finally, the performance of IR is best when we implement circularity for both IR and BWR.

8. CONCLUSION

The Initial Ranging scheme is a critical part of the network entry procedure. The delay incurred during the mechanism is mainly a function of the delay incurred in the backoff procedure that is at the center of IR. Due to collisions among RNG-REQ packets, the backoff procedure may be repeated many times leading to higher IR delay and thereby increased time to complete the network entry procedure. The concept of Markov process is used in analyzing the IR mechanism. From the above observation we can also conclude the following: 1) with increasing number of SS, the circularity value controlling the delay must be increased. This implies that for increasing network sizes, the introduction of delay control must be less frequent. Similarly, 2) with increasing number of SS, the circularity value controlling the window size must be decreased. This implies that for increasing network sizes, the doubling of window size must be more frequent.

Bandwidth Request (BWR) is the process wherein the communication takes place between SSs and BS for uplink bandwidth allocation. We introduce the concept of circularity aiming to reduce collisions among the BWR packets and make better use of the contention slots. We carry out a frame analysis that shows the events that occur during the mechanism. We have proved that the selective dropping of BWR packets decreases the number of collisions owing to improved mechanism of collision avoidance. This directly leads to decrease in the total access delay required by an SS to complete its BWR procedure. The throughput of the data sent is thereby considerably increased.

Furthermore, we introduce the circularity concept in both IR and BWR module combined at the same time. This will reduce the time required by the SSs to successfully transmit their request packets. Hence both IR and network entry get completed in lesser time and also the contention process in BWR mechanism is resolved faster. Thus, reduction in number of collisions increases the success ratio of request packets sent. We have also analyzed and compared the effects of the circularity value with IR, BWR and combined IR and BWR modules on the success ratio.

ACKNOWLEDGEMENTS

The authors sincerely thank the authorities of Supercomputer Education and Research Center, Indian Institute of Science for the encouragement and support.

REFERENCES

- [1] IEEE Standard for Local and Metropolitan area networks, (2004) Part 16: *Air Interface for Fixed Broadband Wireless Access Systems*; IEEE Computer Society and IEEE Microwave Theory and Techniques Society.
- [2] IEEE Standard for Local and metropolitan area networks, (2005) Part 16: *Air Interface for Fixed and Mobile Broadband Wireless Access Systems, Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*; IEEE Computer Society and IEEE Microwave Theory and Techniques Society.
- [3] Chakchai So-In, Raj Jain, and Abdel-Karim Al Tamimi, (2010) "Capacity Evaluation for IEEE 802.16e Mobile WiMAX," *Journal of Computer Systems, Networks, and Communications (JCSNC)*, Special issue on WiMAX, LTE, and WiFi Interworking, Vol. 1, No. 1.
- [4] Chakchai So-In, Raj Jain, Abdel Karim Al Tamimi, (2010) "A Scheduler for Unsolicited Grant Service (UGS) in IEEE 802.16e Mobile WiMAX Networks," *IEEE Systems Journal*, V-4, No. 4.
- [5] J. Y. Kim and D. H. Cho, (2007) "Piggybacking Scheme of MAP IE for Minimizing MAC Overhead in the IEEE 802.16e OFDMA Systems", In *IEEE 66th Vehicular Technology Conference (VTC-2007 Fall)*, pp284-288.
- [6] Bhandari, B. N., Ratnam, K., Raja Kumar, and Maskad, X. L, (2005) "Uplink Performance of the IEEE 802.16 Medium Access Control (MAC) Layer Protocol", *ICPWC, Hong Kong*.
- [7] Ahmed Younus, (2009) "WiMAX – Broadband Wireless Access", *Tech. Univ. of Munich, Germany*.
- [8] J. He, K. Guild, K. Yang, and H.-H. Chen, (2007) "Modelling Contention Based Bandwidth Request Scheme for IEEE 802.16 Networks", *IEEE Communications Letters*, 11(8):698-700.
- [9] S. M. Oh and J. H. Kim, (2005) "The Analysis of the Optimal Contention Period for Broadband Wireless Access Network", in *Third IEEE International Conference on Pervasive Computing and Communications Workshops*, pp215-219.
- [10] J.G. Andrews, A.Ghosh. and R. Muhamed, (2007) "Fundamentals of WiMAX Understanding Broadband Wireless Networking", *Pearson Education Inc., One Lake Street Upper Saddle River, NJ07458, USA*.
- [11] E. Nuaymi (2007) "WiMAX Technology for Broadband wireless Access" *John Wiley & Sons Ltd, the atrium. Southern gate, Chichester, West Sussex PO198SQ, England*.

- [12] J. Yan and G. S. Kuo, (2006) “Cross-layer Design of Optimal Contention Period for IEEE 802.16 BWA Systems”, In IEEE International Conference on Communications, Vol. 4, pp1807-1812.
- [13] Sheldon M. Ross, (2000) “Introduction to Probability Models”, 7th Edition. Academic Press, New York.
- [14] WiMAX Forum, Krishna Ramadas and Raj Jain, (2008) “WiMAX System Evaluation Methodology” No. 2.1.
- [15] Lidong Lin, WeijiaJia, Bo Han and Lizhuo Zhang, (2007) “Performance Improvement using Dynamic Contention Window Adjustment for Initial Ranging in IEEE 802.16 P2MP Networks”, Proceedings of the IEEE Wireless Communication and Networking Conference, Hong Kong.
- [16] Mohammad Z. Ahmad, DamlaTurgut, and R. Bhakthavathsalam, (2006) “Circularity based Medium Access Control in Mobile Ad hoc Networks”, 5th International Conference on AD-HOC Networks & Wireless (ADHOCNOW’06), Ottawa, Canada.
- [17] R. Bhakthavathsalam and Khurram J. Mohammed, (2009) “Analysis and Enhancement of the BWR mechanism in MAC 802.16 for WiMAX”; Proceedings of the 6th Annual IEEE Consumer Communications & Networking Conference, Las Vegas, Nevada, USA.
- [18] National Institute of Standards and Technologies, (2011)“*Seamless and Secure Mobility Tool Suite*”, Advanced Network Technologies, <http://www.antd.nist.gov/seamlessandsecure/doc.html>.
- [19] M. Greis, (2010) Tutorial for the Network Simulator NS2. [Online]. Available: <http://www.isi.edu/nsnam/ns/tutorial/>
- [20] D. Libes, (1993) “A debugger for Tcl applications,” in Tcl/Tk Workshop, [Online]. Available: <http://expect.nist.gov/tcl-debug/tcl-debug.ps.Z>
- [21] The Network Simulator – ns-2. (2012) [Online]. Available: <http://www.isi.edu/nsnam/ns/>
- [22] The Network Simulator Wiki (2010) [Online]. Available: <http://nsnam.isi.edu/nsnam/index.php/>
- [23] Berkeley Continuous Media Toolkit. *OTcl Tutorial*. (2012) [Online]. Available: <http://bmerc.berkeley.edu/research/cmt/cmtdoc/otcl/>
- [24] An AWK primer.(2011) [Online]. Available: <http://www.vectorsite.net/tsawk.html>
- [25] Bazzi, A. Leonardi, G. Pasolini, G. Andrisano, O. DEIS, (2009) “IEEE802.16e Simulation Issues” Napa Valley, CA.
- [26] Sanfilippo, (2012) “*An introduction to the Tcl programming language*” [Online]Available: <http://www.invece.org/tclwise>
- [27] The GDB Developers,(2012)“*GDB: The GNU project debugger*” [Online] Available:<http://www.gnu.org/software/gdb>

Authors

Dr. R.Bhakthavathsalam is presently working as a Senior Scientific Officer in SERC, Indian Institute of Science, Bangalore. His areas of interests are Pervasive Computing and Communication, Wireless Networks and Electromagnetics with a special reference to exterior differential forms. Author held the position of Fellow of Jawaharlal Nehru Centre for Advanced Scientific Research during 1993-95. He is a Member of IEEE Com. Soc., ACM and CSI.



Khurram J. Mohammed is currently an Assistant professor at the Department of Information Science and Engineering, Ghousia College of Engineering, Ramanagaram, India. He received his M.Tech (Computer Science and Engineering) from Dr. Ambedkar Institute of Technology, Bangalore and B.E from Ghousia College of Engineering, Ramanagaram. His research works include Wireless Communication, Formal Languages and .NET applications. He is a Member in several professional bodies.

