Broadcast Polling in IEEE 802.16 Networks with Average-Buffered Subscriber Stations

Chandan Singh Ujarari[1], Arun Kumar Pal[2]

Department of Mathematics Statistics and Computer Science, G. B. Pant University of Agriculture & Technology, Pantnagar, India

Abstract: ***We proposed a model to analyze the performance of the contention based services via broadcast polling in unsaturated IEEE 802.16 networks with channel errors. Subscriber station with finite buffer capacity K can be treated as a M/G/1/K queue. Service time can be calculated by the back-off process of broadcast polling. Using this model, the normalized network throughput, Average-Buffered size and the distribution of the packet delay are derived. For performance estimation of best effort or contention-based non-real time polling services this proposed analytical model is used. Additionally, we show that the model gives good approximations for network performance with a more realistic bursty arrival process at light load, while providing conservative performance measures at medium and high loads.***

Keywords: M/G/1/K, Laplace-Stieltjes transform, fixed point equations, WiMAX

1. Introduction

The Worldwide Interoperability for Microwave Access (WiMAX) Forum promotes THE IEEE 802.16 standard [1]. THE standard [1] defines different air interfaces for various physical layers: WirelessMAN-SC, WirelessMAN-OFDM, WirelessMAN-OFDMA and WirelessMAN-HUMAN. WirelessMAN-OFDMA supports mobility, while WirelessMAN-SC and WirelessMAN-OFDM support fixed connections. A key feature of 802.16 is that it is a connection-oriented technology. The subscriber station (SS) cannot transmit data until it has been allocated a channel by the base station(BS). This allows 802.16 to provide strong support for quality of service (QoS), such as unsolicited grant service (UGS), real-time polling service (rtPS), extended real-time polling service (ertPS), non-realtime polling service (nrtPS), and best-effort service (BE). Simultaneously, some bandwidth request mechanisms are employed to meet the demand of different schedule services, like unsolicited granting, unicast polling, broadcast polling and piggybacking. Performance of the network operating in the point-to-multipoint (PMP) mode through WirelessMAN-SC or WirelessMAN-OFDM air interfaces with no mobility and time division duplexing (TDD) mode is focused here. Our proposed model is applicable for both the SC and OFDM systems as the WirelessMAN-OFDM air interface also uses the TDMA approach to access the channel, but with different parameter values [1]. From the point of view of medium access layer, a subscriber station (SS) still follows the broadcast polling protocol to access the channel, irrespective of the carrier system used in the physical layer. Here we will study and model the performance of an unsaturated IEEE 802.16 network with contention based services which request bandwidth via broadcast polling.

There are several previous researches based on the performance evaluation of bandwidth request mechanisms in WiMAX networks has been done. Performance of contention-free bandwidth request based on unicast polling is studied in [2]-[3], while performance models for the contention based bandwidth request mechanisms in saturated or unsaturated WiMAX networks is proposed in [4]-[10]. In [4], the bandwidth efficiency and channel access delay with contention based bandwidth request scheme in saturated IEEE 802.16 networks was analyzed by the authors. In [5], the authors develop a range of performance metrics for the contention based bandwidth request scheme in saturated IEEE 802.16 networks, such as collision probability, normalized throughput and mean packet delay. To evaluate the average access delay and the capacity of the contention slots in delivering bandwidth request, a 2-D Markov chain (MC) model was proposed in [6] by Fallah *et al.* To analyze the IEEE 802.16 networks with sub channelization in [7], Fattah *et al.* extend the work in [6]. Chuck *et al.* also develop a 2-D MC model to obtain the performance of bandwidth utilization and delay [8]. On the other hand, [6], [7] and [8] assume that the probability of an SS sending a request (REQ) is an input parameter of their models, instead of being a function of the arrival and backoff processes. To evaluate the performance of the request mechanisms with grouping and no-grouping modes in WiMAX the authors suggest a 2-D MC model in [9]. However, they only analyze normalized throughput and collision probability. This model is significantly extended to include delay analysis and take into account channel errors [10] afterward. Moreover, all these mechanism assume the availability of infinite buffers at an SS. The major limitation of the previous results is buffer sizes are always finite. In this paper, we study the performance of nrtPS or BE services using the broadcast polling mechanism, which is contention based and requires SSs to use the truncated binary exponential backoff (TBEB) algorithm for contention resolution. We consider the circumstances where each SS has a finite buffer. Major contributions of this paper can be summarized as follows:

- A general queueing model is developed to analyze the performance of contention based services which use broadcast polling for bandwidth request. Our model also takes into account the random channel noise.
- The distribution of REQ service time and packet delay are well derived by Laplace-Stieltjes transform (LST).
- We derive the average buffer size and average time period frm the arrival of a packet until the start of the next request for this packet is initiated.

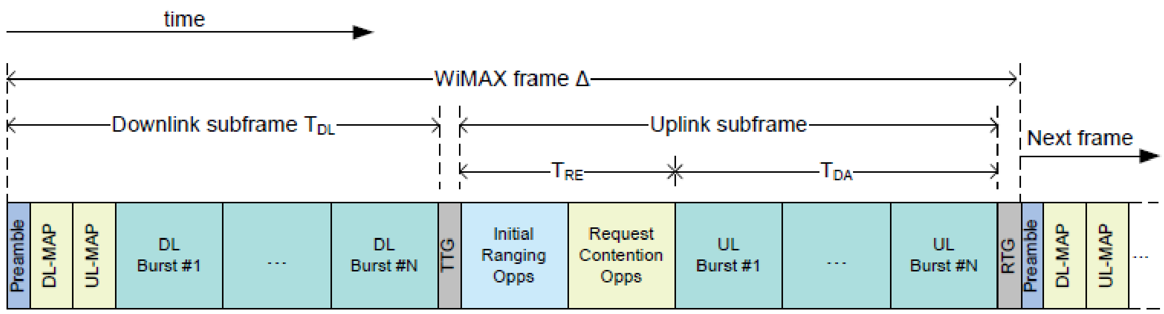


Figure 1: IEEE 802.16 MAC frame structure with time division duplexing

The rest of this paper is organized as follows. Section II first describes the operation of broadcast polling, and then derives the LST of the service time of REQs. Section III develops the fixed point equations to obtain numerical results of some important parameters of our model. Section IV derives the expressions of network throughput and the distribution of packet delay. Section V the analytical results evaluates the impact of different parameters on performance metrics. Results with a more realistic arrival process where packet inter-arrival time follows a Pareto distribution are also presented in this section. Section VI provides an algorithm to achieve the optimum network throughput. Finally, Section VII concludes the paper.

2. Broadcast Polling and Mean Service Time of Reqs

The MAC frame structure defined in the IEEE 802.16 standard in TDD mode is shown in Fig. 1. Each frame is divided into downlink (DL) subframe and uplink (UL) subframe with fixed duration Δ. The duration of DL subframe is *TDL* which begins with preamble for synchronization and equalization. For frame control the DL subframe is followed by DL-MAP and UL-MAP, then downlink burst begins. If there are transmission opportunities for bandwidth requests and data packets then UL-MAP informs the SSs. Bandwidth request interval with duration *TRE*, including initial ranging opportunities and request contention opportunities, and UL bursts with duration *TDA* composed UL subframe. The two gaps called Transmit/receive transition gap (TTG) and Receive/transmit transition gap (RTG) respectively are there in Fig. 1. If an SS has a data packet to send, it will send an REQ in a request slot according to the TBEB algorithm within the request contention opportunities interval of the UL subframe. The contention window *Wi* for backoff state *i* with TBEB is given by as follows:

$$W_i = \begin{cases} 2^i W, & 0 \le i \le r \\ 2^r W, & r < i < R \end{cases}$$

Where *W*, r and R denotes the initial contention window, truncation value and maximum allowable number of attempts. At state *i*, the number of request slots that an REQ has to backoff is a number randomly chosen between 0 and *Wi* −1. UL burst of the next frame is allocated by the base station (BS). If an REQ is successful in state *i*, the packet will be sent in an UL burst of the next frame. Otherwise, the SS will backoff according to state i+1 in the next frame. When the backoff state is altered from *i* to i+1, *Wi* is doubled if $i < r$, or else *Wi* remains constant. If an REQ fails up to *R* transmissions, the corresponding packet will be discarded. We suppose that the SSs are only allowed to request bandwidth to transmit one packet by each REQ because the standard does not state how many packets each REQ represents. This assumption is the default behaviour of an SS. Also, all packets are assumed to have the same length, which is realistic in the MAC layer. Suppose length of a request slot be t_{RE}, *m* be the number of request slots in each uplink subframe. For uplink traffic we assume additionally that the BS always allocates the same amount of uplink capacity consisting of $d \le m$ data slots in every uplink subframe. The transmission time of a packet is t_{DA} ($t_{DA} \ge t_{RE}$) which is length of each data slot. If bandwidth request is successful in the previous frame then BS uplink scheduler will uniformly allocate bandwidth to SSs since standard does not define scheduling algorithms for both BS and SSs. Let *j* be the number of successfully transmitted REQs in an UL subframe. If $j < d$ then in the next frame there will be $(d - j) > 0$ unused data slots, which are wasted. However, if $j > d$ then $(j - d) > 0$ requests must be declined because there are only *d* data slots available in the next frame; so $(j - d)$ requests are also considered unsuccessful.

We consider an IEEE 802.16 network with one BS serving *N* SSs working in the PMP manner. At each SS arrival of data packets are assumed to be Poisson process with rate *λ* packets per second. Let the probability that the buffer of an SS has at least one packet be ρ_c. The buffer size of each SS is denoted as *K*. Hence an SS can be modeled as a *M/G/*1*/K* queueing system. Referring to Fig. 2, the definition of the service time of an REQ depends on whether the queue is empty or not upon the arrival of a new packet at an SS. We identify below separately these two cases:

S0: The empty queue (with probability $1-\rho_c$, Fig. 2(a)). In this case the service time of its REQ includes the time period from its arrival until the start of the request interval where the backoff of the first attempt is initiated, and its backoff process from the beginning of the first request interval until the beginning of the request interval prior to which a successful request or the R^{th} request attempt is made.

S1: The queue is non-empty (with probability ρ_c, Fig. 2(b)). If a packet arrives at a non-empty queue, it will be placed in the buffer until it becomes the head-of-the-line (HOL) packet. The REQ service time of this packet is defined as the time duration from the beginning of the request interval where the backoff of the first attempt is initiated until the beginning of the request interval prior to which a successful request or the *Rth* request attempt is made.

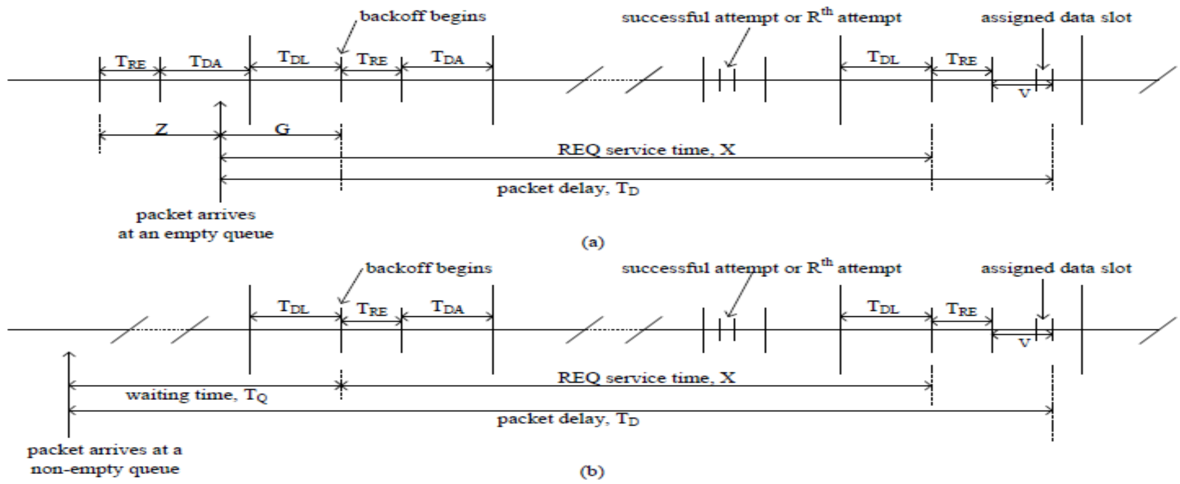


Figure 2: The service time of an REQ when (a) its packet arrives at an empty queue, (b) its packet arrives at a non-empty queue

We analyze the distribution of the service time of REQs as follows. In Fig. 2(a), when a arrival comes at an empty queue then for a certain amount of time to start the TBEB backoff process packet arrival will wait. Let G be a random variable (r.v.) representing the time period from the arrival of a packet until the start of the next request interval where the backoff of the first request for this packet is started. Note that $G = \Delta - Z$ where Z is an exponential r.v. representing the time elapsed from the beginning of the bandwidth request interval in the current WiMAX frame till the arrival of the first packet from the Poisson process.

The cumulative distribution function of G is expressed as $F_G(g) = 1 - F_Z(\Delta - g)$, and after conditioning on $Z \le \Delta$ can then be given by as follows.

$$F_G(g) = \begin{cases} \frac{e^{-\lambda\Delta}(e^{\lambda g}-1)}{1-e^{-\lambda\Delta}}, & 0 \le g \le \Delta \\ 1, & g \ge \Delta \end{cases} \quad (1)$$

The probability density function (*pdf*) of G is given as

$$f_G(g) = \begin{cases} \frac{\lambda e^{\lambda g}}{e^{\lambda\Delta}-1} & 0 \le g \le \Delta \\ 0 & g > \Delta \end{cases} \quad (2)$$

The LST of $f_G(g)$ is given by $L_G(g) = \frac{\lambda(e^{-s\Delta} - e^{-\lambda\Delta})}{(\lambda - s)(1-e^{-\lambda\Delta})}$

Mean time period from the arrival of a packet until the start of the next request interval where the backoff of the first request for this packet is initiated is given by

$$E_G(g) = \int_0^{\infty} g f_G(g) dg = \frac{(1-e^{\lambda\Delta}+\lambda\Delta e^{\lambda\Delta})}{\lambda(1-e^{-\lambda\Delta})} \quad (3)$$

Let $H^{(i)}$, $0 \le i < R$, be a discrete r.v. representing the number of backoff frames incurred by the i^{th} attempt of an REQ. Since the backoff period is uniformly chosen from $[0, W_i - 1]$ in the i^{th} attempt, the probability mass function (pmf) of $H^{(i)}$ is given by

$$H^{(i)} = \begin{cases} j & w.p.\ {}^{m}/_{W_i},\ j = 1, 2, \dots, A_i - 1 \\ A_i & w.p.\ 1 - \frac{(A_i - 1)m}{W_i} \end{cases} \quad (4)$$

where w.p. is for "with probability" and A_i is given by

$$A_i = \left\lceil {}^{W_i}/_{m} \right\rceil$$

Then, the LST of the pmf of $H(i)$ can be obtained as follows

$$L_{H^{(i)}}(s) = \sum_{j=1}^{A_i - 1} \frac{m}{W_i} e^{-js} + \left(1 - \frac{(A_i - 1)m}{W_i}\right) e^{-A_i s}.$$

Let $Y^{(i)}$, $0 \le i < R$, be a discrete r.v. representing the accumulated backoff time that an SS has spent from backoff state 0 to backoff state i,

$$Y^{(i)} = \sum_{j=0}^{i} H^{(j)} \Delta. \quad (5)$$

So, the LST of the pmf of $Y^{(i)}$ can be given as

$$L_{Y^{(i)}}(s) = \prod_{j=0}^{i} L_{H^{(j)}}(\Delta s).$$

Case S0: In this case the service time of an REQ is $X_0 = G + Y$, noting that G and Y are independent, so the pdf of X_0 is given as follows

$$f_{X_0}(x) = \int_{-\infty}^{\infty} f_G(x-y) f_Y(y) dy$$

And its LST is given simply by

$$L_{X0}(s) = L_G(s)\, L_y(s)$$

Case S1: In this case ,the service time of the REQ is $X_1 = Y$. so the pmf of X_1 is $f_y(y)$. Hence we can defined service time of an REQ's as follows.

$$X = \begin{cases} G + Y & w.p.\ 1 - \rho_c \\ Y & w.p.\ \rho_c \end{cases} \quad (6)$$

And the LST of the pdf of an REQ's service time can be written as

$$L_X(s) = (1 - \rho_c) L_{X_0}(s) + \rho_c L_Y(s)$$

3. Fixed Point Equations

In this section, to calculate the service time distribution of REQs we will discuss two nested sets of fixed point equations to obtain ρ_c and p.

3.1 Outer Set

The service time distribution of an REQ based on whether the queue is empty or not upon the arrival of its packet at an SS, and thus it is a function of ρ_c. ρ_c is just the utilization of the $M/G/1/K$ queue modeling an SS, it cannot be directly obtained from the standard expressions for $M/G/1/K$ model. We define the system state of an SS at time t to be the number of requests in its buffer at that instant. Consider the imbedded Markov Chain of the system states just after the departure instants of the requests that leave the queue after obtaining service.

Suppose the time instant at which the i_{th} request departs from the SS after obtaining service is denoted by t_i, $i = 1, 2, 3, \ldots,$. Let n_i is the number of requests left behind when the i_{th} request departs at t_i state of the SS & n_i has the range between 0 and $K-1$ since the departure of the request cannot leave the SS completely, i.e. with the SS in state K. Let a_i be the number of arrivals (from the Poisson arrival process) in the i^{th} service time. The corresponding equations for the Markov Chain can then be given as

$$n_{i+1} = \begin{cases} \min(a_{i+1}, K-1) & for\ n_i = 0 \\ \min(n_i - 1 + a_{i+1}, K-1) & for\ n_i = 1,2,\ldots,K-1 \end{cases}$$

Now the equilibrium state probabilities $p_{d,k}$ at the departure time of requests from the queue can be easily can be easily compute. Where $k = 0, 1, 2, \ldots, K\text{-}1$. The transition probabilities of the imbedded Markov Chain is required for the $p_{d,k}$ which is given by

$$p_{d,jk} = P\{n_{i+1} = k / n_i = j\} \quad 0 \le j, k \le K-1 .$$

The probability of k job arrivals to the queue during the service time of a packet in the empty queue is given by

$$\alpha_k = \int_{t=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} f_{X_{s0}}(t)dt. \qquad (7)$$

The probability of k job arrivals to the queue during the service time of a packet in the non-empty queue is given by

$$\beta_k = \int_{t=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} f_Y(t)dt. \qquad (8)$$

The transition probability in terms of α_k and β_k for empty and non-empty respectively is given by

$$p_{d,0k} = \begin{cases} \alpha_k & 0 \le k \le K-2 \\ \sum_{m=K-1}^{\infty} \alpha_m & k = K-1 \end{cases} \qquad (9)$$

&

$$p_{d,jk} = \begin{cases} \beta_{k-j+1} & j-1 \le k \le K-2 \\ \sum_{m=K-j}^{\infty} \beta_m & k = K-1 \end{cases} \qquad (10)$$

Where $j = 1, 2, \ldots, K\text{-}1$.

Hence the transition probability matrix is given by

$$\mathbf{P_d} = [p_{d,jk}] = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{K-2} & \sum_{m=K-1}^{\infty} \alpha_m \\ \beta_0 & \beta_1 & \beta_2 & \cdots & \beta_{K-2} & \sum_{m=K-1}^{\infty} \beta_m \\ 0 & \beta_0 & \beta_1 & \cdots & \beta_{K-3} & \sum_{m=K-2}^{\infty} \beta_m \\ 0 & 0 & \beta_0 & \cdots & \beta_{K-4} & \sum_{m=K-3}^{\infty} \beta_m \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta_0 & \sum_{m=1}^{\infty} \beta_m \end{bmatrix}$$

The equilibrium state probabilities $p_{d,k}$, $k = 0, 1, 2, \ldots, K\text{-}1$ at the departure instants can be expressed in terms of K -1 balance equations along with the normalization condition, which are given as follows.

$$p_{d,k} = \sum_{j=0}^{K-1} p_{d,j}\, p_{d,jk} \qquad (11)$$

The normalized condition of this is as follows

$$\sum_{k=0}^{K-1} p_{d,k} = 1 \qquad (12)$$

Now we can substitute the transition probabilities $p_{d,jk}$ in equations (11) & (12) by which we get a set of linear equations that may help to solve the corresponding state probabilities. Since there are only K unknowns ($p_{d,k}$, $k = 0, 1, \ldots, K\text{-}1$) to be found this implies only K-1 equations are required from K equations of (11) apart from the normalized condition of (12). The set of K equations can be summarized as in following equation (13).

$$p_{d,k} = p_{d,0\alpha_k} + \sum_{j=1}^{k+1} p_{d,j}\, \beta_{k-j+1} \qquad k = 0,1,\ldots,K-2 \qquad (13)$$

$$p_{d,K-1} = 1 - \sum_{k=0}^{K-2} p_{d,k}$$

Let $p_{a,k}$, $k = 0, 1, \ldots, K$ be a newly arriving packet probability, irrespective of whether it finally joins the queue or not and finds k packets waiting in the queue (where an SS is an equilibrium state). Let p_k, $k = 0, 1, \ldots, K$ be the steady state probability that buffer has k packets in it. Then by using PASTA property we can then write that

$$P_k = p_{a,k} \ :, k = 0, 1, \ldots, K$$

Now the average buffer size is given as

$$E(k) = \sum_{k=0}^{K-1} k\, p_k = \sum_{k=0}^{K-1} k\, p_{a,k} = K_{avg} \qquad (14)$$

If the equilibrium buffer overflow probability (SS is in state K) is P_B. Therefore $P_B = p_K$ then we have

$$p_k = p_{a,k} = (1 - P_B) p_{d,k}, \ k = 0, 1, \ldots, K-1. \qquad (15)$$

Where $\sum_{k=0}^{K} p_{a,k} = 1$ and $\sum_{k=0}^{K-1} p_{d,k} = 1$. Hence the average buffer size is given by

$$E(k) = \sum_{k=0}^{K-1} k\,(1 - P_B)p_{d,k} = K_{avg} \quad (16)$$

For the case k=0, using (15) we can write

$$1 - \rho(1 - P_B) = 1 - \rho_c = (1 - P_B)p_{d,0} \quad (17)$$

Where ρ is the traffic load of an SS & the relation between ρ_c and P_B is given by as follows

$$\rho_c = (1 - P_B)\rho = (1 - P_B)\lambda E[X] \quad (18)$$

Where $E[X] = \frac{dL_X(s)}{ds}$ at s=0. After substituting value of P_B in equation (18) we can find the value of ρ_c and $p_{d,0}$.

3.2 Inner Set

If there is no collision then request is successful with probability (1-p_c) and base station has sufficient bandwidth to serve it with probability 1-p_d. If request is not corrupted by random channel noise (with probability 1-p_e where p_e is the frame error rate) then it can be another condition for a request to be successful. Then we can write p as follows

$$p = 1 - (1 - p_c)(1 - p_u)(1 - p_e) \quad (19)$$

Where p_u is the probability that a collision free request is unsuccessful due to the lack of bandwidth in the subsequent frame [1].

3.3 Relationship between p and ρ_c

We presume initial value of ρ_c to solve p and ρ_c, which is input into the outer set of fixed point equations .This input ρ_c is first used in the inner set of fixed point equations to solve p. Then, they are used together in the outer set of fixed point equations to calculate $E[X]$, PB and thus a new ρ_c. The above process repeats if the new ρ_c has not converged and it is fed back to the outer set of fixed point equations as the input. Otherwise, we have solved ρ_c and p.

4. Performance Measures

4.1. Throughput

A packet is discarded if its request has failed after R attempts, the throughput of each SS is given by $\lambda(1-P_B)(1-p^R)$. Since the network provides a capacity of d data slots in each frame with duration Δ, then the normalized network throughput Γ is thus given by

$$\Gamma = \frac{N\lambda(1-P_B)(1-p^R)}{d/\Delta} \quad (20)$$

4.2. Packet Delay

The r.v. X is service time of an REQ, it does not matter REQ is successful or unsuccessful. Let the service time of a successful REQ is define by a r.v. X'. If waiting time of an REQ is represented by W_q and V be a discrete r.v. representing the time from the beginning of a data subframe to the end of a packet transmission, as shown in Fig. 2. TRE is the duration of the bandwidth request opportunities depicted Fig. 1. Therefore for a successful REQ, the corresponding packet delay D can be written as

$$D = W_q + X' + TRE + V \quad (21)$$

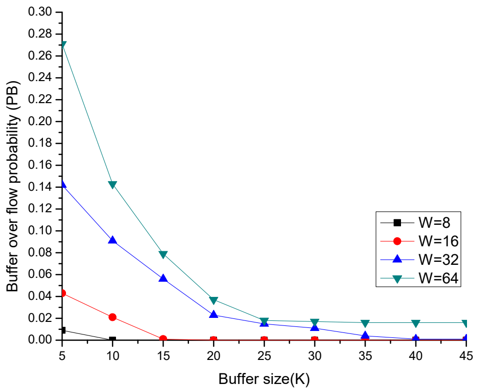
as a result, the LST of the pdf of D can be written as

$$L_D(s) = L_{W_q}(s)L_{X'}(s)L_V(s)e^{-sTRE} \quad (22)$$

5. Numerical Results

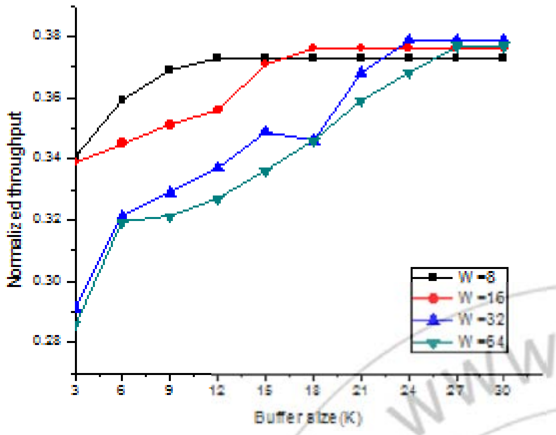
In our analytical model we analyze the relative study of buffer overflow probability, mean packet delay and normalized throughput under various factor N, m, d, λ, K and W. The channel is operated in TDD mode with the frame structure as shown in Fig. 1. The data rate at the physical layer is assumed to be 130 Mbps. In accordance with default parameters taken from [1] MAC and physical layer parameters are configured. The duration of frame is 1 μsec an it consist of 5000 physical slots or 2500 mini slots each of 0.4 μsec. Each bandwidth request consists of 6 mini slots where 3 mini slots are for subscriber station transition gap (SSTG), 2 mini slots for preamble and one mini slot for a bandwidth request message of 48 bits. the transmission of an approximately 0.5 KB packet per data slot is allowed by the preamble and transition gap which length is 37.6 μsec (i.e. 94 mini slots). Each SS has a finite buffer. Our model is suitable for studying the impact of different parameters on the performance metrics of contention-based services of the IEEE 802.16 networks. Let us have a network setting of $N = 45$, $r = 3$, $R = 5$, $m = d = 10$, $\lambda = 130$, $p_e = 0$, and $W = 8$, 16, 32, 64 to calculate the impact of the buffer size K on buffer overflow probability, mean packet delay and normalized throughput.

Fig. 4.1 shows the buffer overflow probability against K for various W. For fix K, the buffer overflow probability increases when we increase W. Since average number of backoff slots increases when W increases and this increases the mean service time and hence there is a higher buffer overflow probability. When W is fixed, as expected, the buffer overflow probability decreases as K increases. We define a *critical buffer size* beyond which the buffer overflow probability becomes negligible for a particular W. For example, when $W = 8$, the critical buffer size is about 10, for $W = 16$ it is 15, for $W = 32$ it is about 36 and for $W = 64$ it is infinity as shown in Fig. 4.1. Here we observe that as the load increases (i.e. W increases), the critical buffer size increases. When the network is closed to saturation, the critical buffer size could be very large, i.e., approaching infinity, as depicted by the curve for $W = 64$ in Fig. 4.1. Also, the decreasing rate of buffer overflow probability with respect to K is larger at higher load.



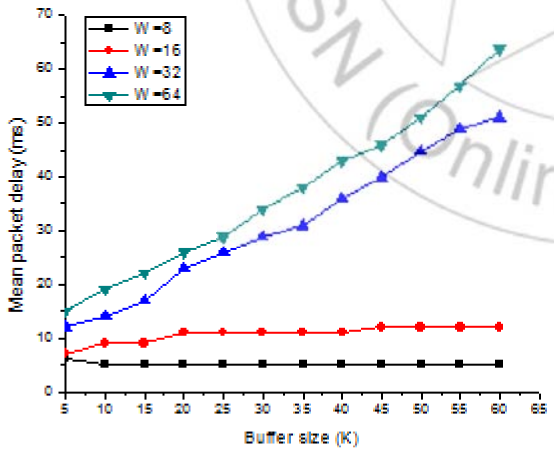
(4.1) Buffer overflow probabilities

The normalized throughput against K for different W (i.e., loads) is shown in Fig. 4.2. For a given W, increase of K up to the critical buffer size results in the increase in throughput. Since buffer overflow probability decreases faster for a higher load (e.g., $W = 32$, $W = 64$) hence the rate of increase in throughput with respect to K is larger. On the other hand, for further increase of K beyond the critical buffer size, the throughput remains the same because either all the input load has been accommodated and admitted, or the network is already saturated.



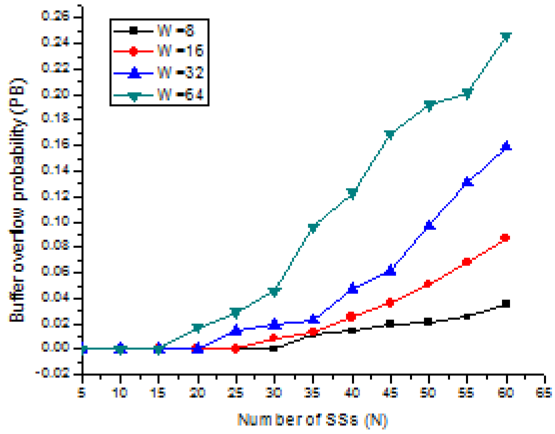
(4.2) Normalized throughput

Fig. 4.3 plots the mean packet delay against K for various W. For a given K when we increase W it causes larger mean packet delay due to the larger mean service time. Due to larger queues for a fix W, increase of K up to the critical buffer size increases the mean packet delay. Again for $W = 32$, 64 the network is closed to saturation, and the mean delay is seen to be fast increasing. There is no advantage in setting K to be larger than the critical value. For a given buffer size, if the network is closed to or already saturated, then increasing K would only increase packet delay.

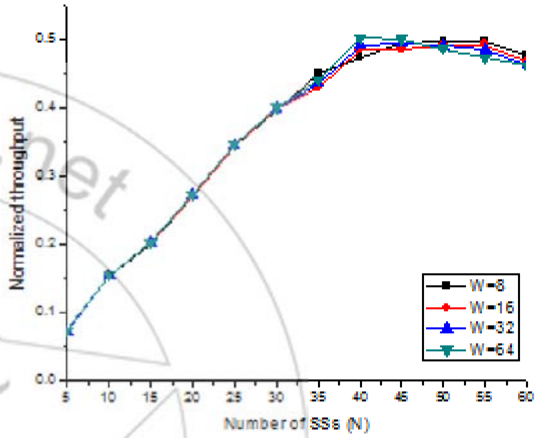


(4.3) Mean packet delay

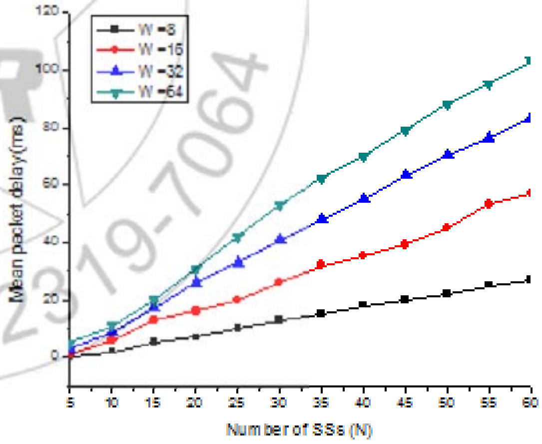
Now we evaluate the impact of N and W on various performance metrics. Similarly, we set $r = 3$, $R = 5$, $m = d = 10$, $\lambda = 130$ and fix $K = 10$. The results are shown in Fig. (5.1) & in Fig. (5.3). For a given W and fixed K, a larger N leads to higher buffer overflow probability and delay. The normalized throughput against N for various W is plotted in Fig. (5.2). When N is small, say $N < 30$, the normalized throughput for different W is about the same. However, for a given larger N, the normalized throughput increases and then decreases as W increases from 8 to 64. This indicates that we can vary W to maximize the throughput.



(5.1) Buffer overflow probabilities



(5.2) Normalized throughput



(5.3) Mean packet delay

6. Conclusion

We develop an analytical model to evaluate the performance metrics of the contention-based services in unsaturated IEEE 802.16 networks over imperfect channel where subscriber stations have only finite buffers. Expressions for network throughput, carried traffic of each subscriber station and the distribution of packet delay are derived. Using the model, we have been able to investigate the impact of various parameters on the performance metrics of the 802.16 networks. The average buffer size and Mean time period from the arrival of a packet until the start of the next request interval where the backoff of the first request for this packet is initiated are also defined here. We observe that there is no benefit to provide large buffers to subscriber stations.

References

[1] IEEE 802.16-2009, IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems, May 2009.
[2] R. Iyengar, P. Iyer, and B. Sikdar, "Analysis of 802.16 based last mile wireless networks," in *Proc. 2005 IEEE GLOBECOM*, pp. 3123–3127.
[3] B. Sikdar, "Queueing analysis of polled service classes in the IEEE 802.16 MAC protocol," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5767–5772, Dec. 2009.
[4] J. He, K. Guild, K. Yang, and H. H. Chen, "Modeling contention based bandwidth request scheme for IEEE 802.16 networks," *IEEE Commun. Lett.*, vol. 11, no. 8, pp. 698–700, Aug. 2007.
[5] H. L. Vu, S. Chan, and L. Andrew, "Performance analysis of best-effort service in saturated IEEE 802.16 networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 460–472, 2010.
[6] Y. P. Fallah, F. Agharebparast, M. R. Minhas, H. M. Alnuweiri, and V. C. M. Leung, "Analytical modeling of contention-based bandwidth request mechanism in IEEE 802.16 wireless networks," *IEEE Trans. Veh. Technol.*, vol. 5, no. 5, pp. 3094–3107, 2008.
[7] H. Fattah and H. Alnuweiri, "Performance evaluation of contention-based access in IEEE 802.16 networks with subchannelizaion," in *Proc. 2009 IEEE ICC*, pp. 1–6.
[8] D. Chuck, K. Chen, and J. M. Chang, "A comprehensive analysis of bandwidth request mechanisms in IEEE 802.16 networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 2046–2056, May 2010.
[9] Q. Ni and L. Hu, "An unsaturated model for request mechanisms in WiMAX," *IEEE Commun. Lett.*, vol. 14, no. 1, pp. 45–47, Jan. 2010.
[10] Q. Ni, L. Hu, A. Vinel, Y. Xiao, and M. Hadjinicolaou, "Performance analysis of contention based bandwidth request mechanisms in WiMAX networks," *IEEE Systems J.*, vol. 4, no. 4, pp. 477–486, Dec. 2010.
[11] Y. Wang, S. Chan, M. Zukerman, and R. J. Harris, "Priority-based fair scheduling for multimedia WiMAX uplink traffic," in *Proc. 2008 IEEE ICC*.
[12] A. Arcia, N. Montavont, A. Paltrineri, and D. Ros, "Improving uplink bandwidth management in 802.16 networks," in *Proc. 2009 LCN*.
[13] A. Vinel, Y. Zhang, M. Lott, and A. Tuilikov, "Performance analysis of the random access in IEEE 802.16," in *Proc. 2005 IEEE PIMRC*.
[14] http://home.iitk.ac.in/~skb/qbook/qbook.html
[15] S. K. Bose, *An Introduction to Queueing Systems*. Kluwer Academic/ Plenum Publishers, 2002.
[16] A. Baiocchi and N. Blefari-Melazzi, "Steady-state analysis of the MMPP/G/1/K queue," *IEEE Trans. Commun.*, vol. 41, no. 4, pp. 531–534, Apr. 1993.
[17] D. Veitch, "Novel models of broadband traffic," in *Proc. 1993 IEEE GLOBECOM*, vol. 2, pp. 1057–1061