

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



A Unified Performance Model for Best-Effort Services in WiMAX Networks

Jianqing Liu¹, Sammy Chan¹ and Hai L. Vu²

¹*City University of Hong Kong*

²*Swinburne University of Technology*

¹*Hong Kong S.A.R.*

²*Australia*

1. Introduction

Based on the work from the IEEE Working Group 802.16 and ETSI HiperMAN Working Group, the WiMAX (Worldwide Interoperability for Microwave Access) technology is defined by the WiMAX Forum to support fixed and mobile broadband wireless access. In the standard (IEEE 802.16 standard, 2009), it defines several air interface variants, including WirelessMAN-SC, WirelessMAN-OFDM, WirelessMAN-OFDMA and WirelessMAN-HUMAN. WiMAX networks can be operated in two different modes: point to multi-point (PMP) mode and mesh mode. Under the PMP mode, all traffics from subscriber stations (SSs) are controlled by the base station. Mesh mode is a distributed architecture where traffics are allowed to route not only between SSs and the base station but also between SSs. In this chapter, we focus on the WirelessMAN-SC air interface operating in the PMP mode.

In WiMAX networks, quality of service (QoS) is provided through five different services classes in the MAC layer (Andrews et al., 2007):

1. Unsolicited grant service (UGS) is designed for real-time applications with constant data rate. These applications always have stringent delay requirement, such as T1/E1.
2. Real-time polling service (rtPS) is designed for real-time applications with variable data rate. These applications have less stringent delay requirement, such as MPEG and VoIP without silence suppression.
3. Extended real-time polling service (ertPS) builds on the efficiency of both UGS and rtPS. It is designed for the applications with variable data rate such as VoIP with silence suppression.
4. Non-real-time polling service (nrtPS) is designed to support variable bit rate non-real-time applications with certain bandwidth guarantee, such as high bandwidth FTP.
5. Best effort service (BE) is designed for best effort applications such as HTTP.

To meet the requirements of different service classes, several bandwidth request mechanisms have been defined, namely, unsolicited granting, unicast polling, broadcast polling and piggybacking. In this chapter, we present a performance model for services, such as BE service, based on the broadcast polling mechanism which is contention based and requires

the SSs to use the truncated binary exponential backoff (TBEB) algorithm (Kwak et al., 2005) to resolve contention. There is some previous research work on the contention free and contention based bandwidth request mechanisms. Delay analysis of contention free unicast polling request mechanism is proposed in (Iyengar et al., 2005). In (Vinel et al., 2005), average delay of random access with broadcast polling in saturation IEEE 802.16 networks is studied. An analytical model of contention based bandwidth request for IEEE 802.16 networks is proposed in (He et al., 2007), in which bandwidth efficiency and channel access delay are obtained. In (Vu et al., 2010), the throughput and delay performances of best-effort services in IEEE 802.16 networks is analysed. Both (He et al., 2007; Vu et al., 2010) consider the saturated case that each SS always has traffics to send. In (Ni & Hu, 2010), the authors propose a model for the unsaturated case of the request mechanisms in WiMAX. Fallah *et al.* propose a 2-dimensional Markov chain (MC) model to evaluate the average access delay and the capacity of the contention slots in delivering bandwidth request (Fallah et al., 2008). Fattah *et al.* extend (Fallah et al., 2008) to analyze the IEEE 802.16 networks with subchannelization (Fattah & Alnuweiri, 2009). Chuck *et al.* also use the 2-dimensional MC model to obtain the performance of bandwidth utilization and delay (Chuck et al., 2010). However, (Fallah et al., 2008; Fattah & Alnuweiri, 2009; Chuck et al., 2010) assume that the probability of an SS sending a request is an input parameter of their models, instead being a function of the backoff process. Moreover, all existing works only explicitly model mean packet delay, but not the complete distribution.

This chapter significantly extends our work in (Vu et al., 2010) by proposing a unified model for the performance of the best-effort service of WiMAX networks. This model can capture the performances of both unsaturated and saturated cases, and derives the expressions for network throughput and packet delay distribution, rather than just mean packet delay. Each SS will be modeled as a M/G/1 queueing system, where the bandwidth request arrival follows a Poisson process, and the service time is determined by the broadcast polling mechanism. Since our model explicitly models the broadcast polling mechanism, it provides a more accurate estimate of the service time of bandwidth request and packet delay than (Fallah et al., 2008; Fattah & Alnuweiri, 2009; Chuck et al., 2010). The validity of our model will be evaluated by extensive simulations. Our model can be used by operators to configure the parameter settings at the MAC layer for performance optimization.

The rest of this chapter is organized as follows. In Section 2, we first briefly introduce the contention based broadcast polling mechanism. Section 3 proposes fixed point equations to analyze the system. Section 4 derives the expressions of some performance measures. Section 5 verifies the analytical results by simulations. Section 6 degenerates the unsaturated model to saturated networks. Finally, Section 7 concludes the chapter.

2. Broadcast polling

We consider an IEEE 802.16 network consisting of N SSs operating in the PMP mode through WirelessMAN-SC air interface. The SSs access the network through the time division multiple access technology. The MAC frame structure defined in the IEEE 802.16 standard for TDD in PMP mode is shown in Fig. 1. Each frame has a duration of Δ and is divided into uplink and downlink subframes. At the beginning of a downlink subframe, which has a duration T_{DL} , there are two important messages called downlink map (DL-Map) and uplink map (UL-Map)

messages. They specify the control information for the downlink and uplink subframes respectively. In the UL-Map, there is data or information element indicating whether there are transmission opportunities for bandwidth requests (REQs) and data packets. The uplink subframe is composed of bandwidth request bursts with duration T_{RE} and data bursts with duration T_{DA} , respectively. At frame i , when an SS has a data packet to send, it first sends a bandwidth request for transmitting its data in one of the transmission opportunities within the request interval of the uplink subframe. Upon receiving the bandwidth requests, the BS then allocates bandwidth and data slots for data transmission in the uplink data interval of frame $i + 1$ based on its scheduler.

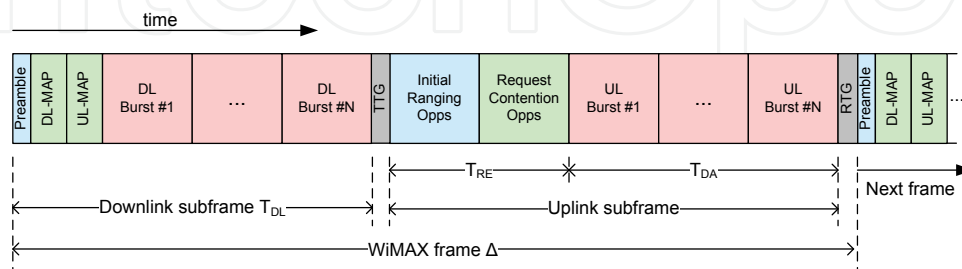


Fig. 1. IEEE 802.16 MAC frame structure with times division duplexing (TDD).

Let us consider a scenario where broadcast polling is used by the BS with m (fixed) transmission opportunities for bandwidth requests which are referred to as *request slots*. In this case, if there is only one request submitted to a request slot, the request is successful. On the other hand, if there are two or more SSs sending their requests in the same request slot, collision will happen and TBEB is used to solve this contention problem. Let W_i be the contention window for backoff state i , and each SS randomly selects a backoff time in the range $[0, W_i - 1]$. With TBEB, W_i is given by:

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

where r is referred to as the truncation value, W is the initial contention window and R is the maximum allowable number of attempts. If the request still fails after R attempts, the packet will be discarded. Then if there are other packets queueing in the buffer, the packet at the head of the queue will send bandwidth request in the next frame.

In this chapter, the SSs are only allowed to request bandwidth to transmit one packet per request, and all packets are assumed to have the same length. Let t_{RE} be the length of a request (or backoff) slot. Furthermore, we assume that the BS always allocates the same amount of uplink capacity consisting of $d \leq m$ data slots in every uplink subframe for uplink traffic. Each data slot is of length T ($T \gg t_{RE}$) which is the transmission time of a packet. As the standard does not define scheduling algorithms for both BS and SSs, we assume here that the BS uplink scheduler will uniformly allocate bandwidth to SSs whose bandwidth request is successful in the previous frame. Let j be the number of requests that do not collide. 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 slots available in the next frame; those $(j - d)$ requests are also considered unsuccessful.

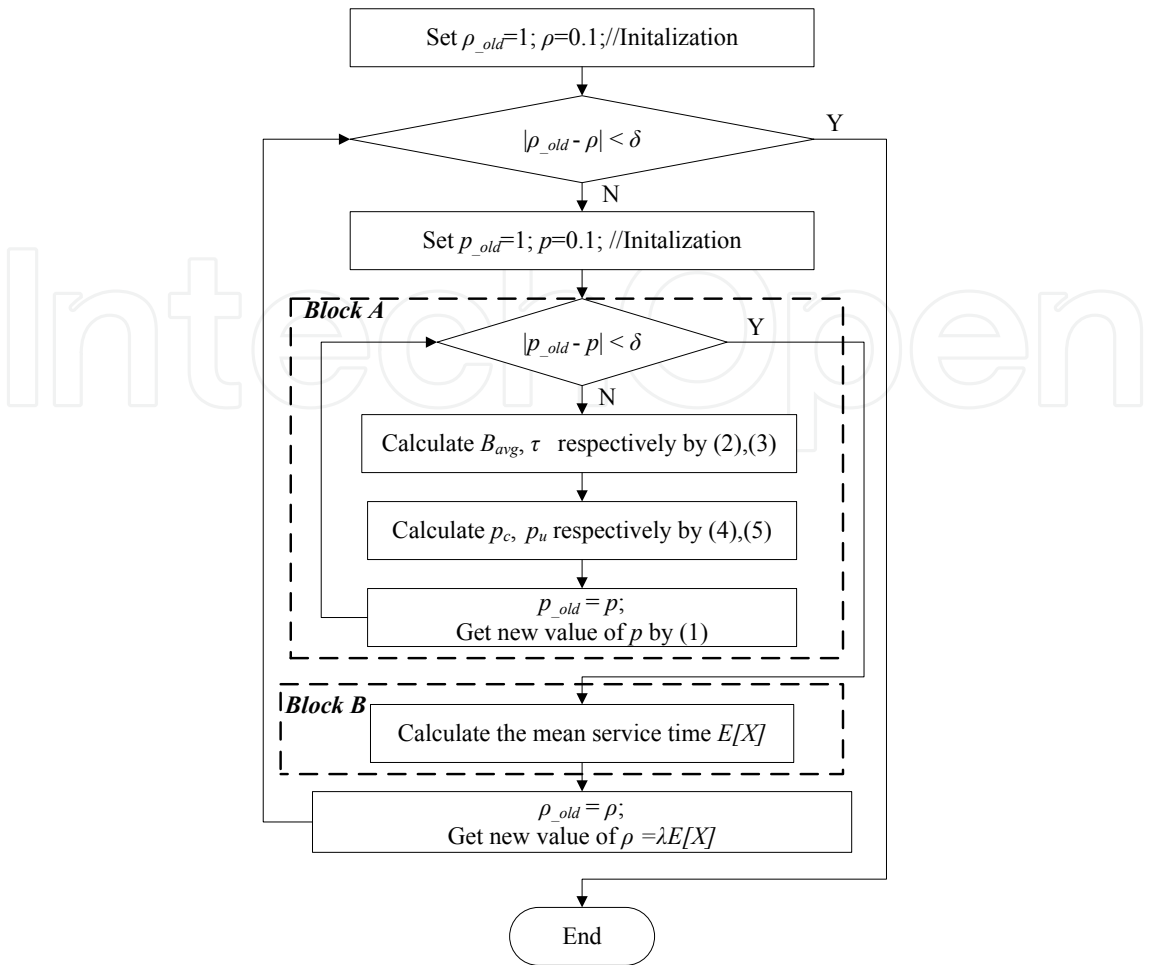


Fig. 2. An overview of the nested fixed point equations.

3. Fixed point equations

In this section, we will use the fixed-point method (Agarwal et al., 2001) to analyze the queueing behaviour at an SS. We assume that packets arrive at an SS according to a Poisson process with rate λ and each SS has an infinite buffer. An SS can therefore be modelled as a $M/G/1$ queueing system. We develop two sets of fixed point equations, one nested by the other, to calculate the failure probability p of an REQ and the offered load to the queue ρ , respectively. The relationship between these two sets of fixed point equations is illustrated by the flow chart shown in Fig. 2. The inner set, labelled as *Block A*, calculates the p for a given ρ . The outer set includes one more block, labelled as *Block B*, and calculates ρ which is relevant to the mean service time of an REQ.

3.1 Failure probability of an REQ

As in (He et al., 2007), a request is regarded as unsuccessful either when the request experiences collision during transmission (with probability p_c) or when the request is successfully transmitted but the BS could not allocate bandwidth to it due to insufficient data slots (with probability p_u). For simplicity, these two events are assumed to be independent. Then p can be expressed as

$$p = 1 - (1 - p_c)(1 - p_u).$$

(1)

Based on TBEB, we can derive the average number of backoff slots B_{avg} an SS has to wait before sending requests as

$$B_{avg} = \frac{m}{2} + \eta \sum_{i=0}^{r-1} p^i \left(\frac{2^i W - 1}{2} \right) + \eta \left(\frac{2^r W - 1}{2} \right) \sum_{i=r}^{R-1} p^i, \quad (2)$$

where $\eta = (1 - p)(1 - p^R)^{-1}$, and $(1 - p^R)$ is a normalization factor.

Knowing B_{avg} , the probability that an SS attempts to send the requests in a slot can be written as

$$\tau = \rho / (B_{avg} + 1), \quad (3)$$

where $\rho = \lambda E[X]$ and $E[X]$ is the average REQ service time, which will be derived in next subsection.

Given that there are N SSs in the system, the probability p_c can be expressed as

$$p_c = 1 - (1 - \tau)^{N-1}. \quad (4)$$

Let ξ be the probability that a collision-free request is made in a given slot, given that there are N SSs, each attempting to send requests with probability τ . Under the assumption that requests are independent, we have

$$\xi = N\tau(1 - \tau)^{N-1}.$$

The probability that there are j collision-free requests among m request slots, $0 \leq j \leq n = \min(m, N)$, is then given by a truncated binomial distribution

$$Q(j) = \frac{\binom{m}{j} \xi^j (1 - \xi)^{m-j}}{\sum_{i=0}^n \binom{m}{i} \xi^i (1 - \xi)^{m-i}}.$$

The probability that a collision-free request is unsuccessful due to lack of bandwidth in the subsequent frame can be expressed as

$$p_u = \frac{\sum_{j=d+1}^n (j - d) Q(j)}{\sum_{j=0}^n j Q(j)}. \quad (5)$$

Equations (1) to (5) form the inner set of fixed point formulations for p . As shown in **Block A** of Fig. 2, for a given ρ , p can be obtained by repeatedly solving these equations until p converges. The resultant p obtained is subsequently used in the outer set of fixed point equations evolving around the traffic load of an SS, ρ . In the following, we will develop the outer set of fixed point equations for ρ .

3.2 Mean service time of an REQ

This subsection presents the details of **Block B** of Fig. 2, which calculates the mean service time of REQs.

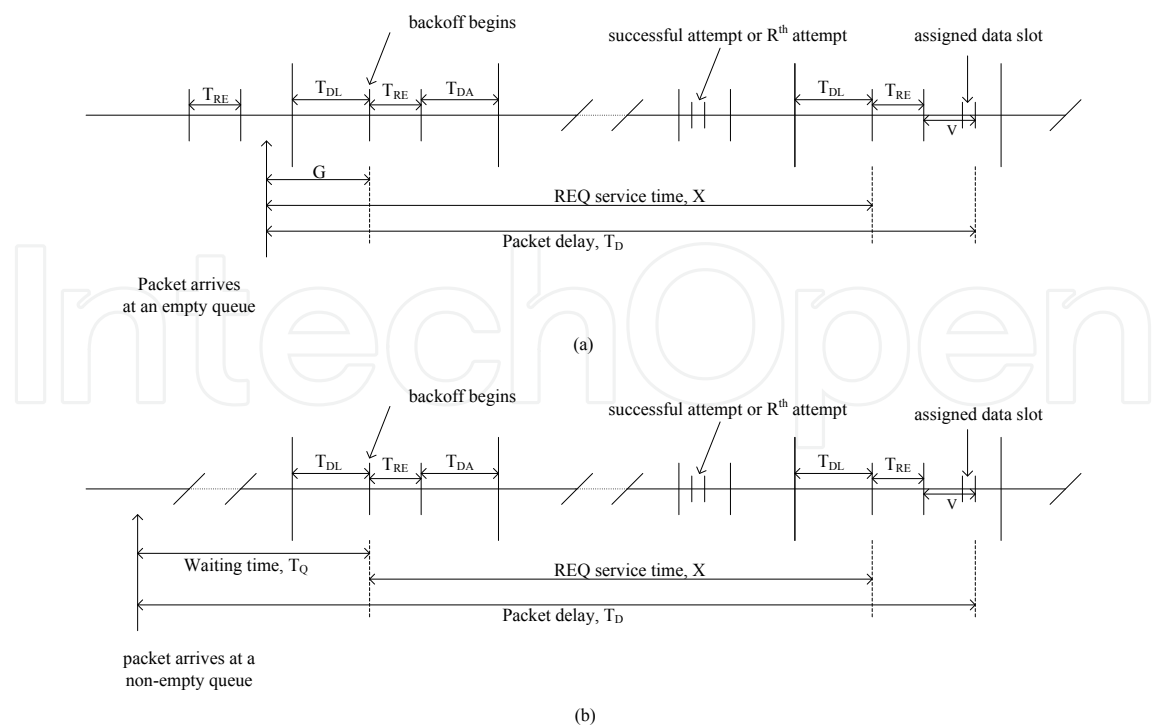


Fig. 3. The service time of an REQ when (a) its packet arrives at an empty queue, (b) its packet arrives at a non-empty queue

Referring to Fig. 3, the definition of REQ’s service time depends on whether the queue is empty or not upon the arrival of a new packet at an SS. We specify below separately these two cases:

- S0:** The queue is empty (with probability $1 - \rho$, Fig. 3(a)). If a packet arrives at an empty queue, its REQ’s service time will include 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 , Fig. 3(b)). If a packet arrivals 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 R^{th} request attempt is made.

Consider case **S0**, let G be a random variable representing the time period from packet’s arrival until the start of the request interval where the backoff of the first request for that packet is initiated. The cumulative distribution function of G is written as

$$F_G(g) = \begin{cases} \frac{e^{-\lambda\Delta}(e^{\lambda g}-1)}{1-e^{-\lambda\Delta}} & 0 \leq g \leq \Delta, \\ 1 & g \geq \Delta. \end{cases} \tag{6}$$

The probability density function (pdf) of G is written as

$$f_G(g) = \begin{cases} \frac{\lambda e^{\lambda g}}{e^{\lambda \Delta} - 1} & 0 \leq g \leq \Delta, \\ 0 & g > \Delta. \end{cases}$$

Based on (6), the average of G can be obtained as

$$E[G] = \frac{\Delta}{1 - e^{-\lambda \Delta}} - \frac{1}{\lambda}, \quad (7)$$

where $E[\cdot]$ is the average operator. And, we can obtain the Laplace-Stieltjes transform of $f_G(g)$

$$\mathcal{L}_G(s) = \frac{\lambda(e^{-s\Delta} - e^{-\lambda\Delta})}{(\lambda - s)(1 - e^{-\lambda\Delta})}.$$

Next, we need to analyze the collision resolution process by TBEB. Let $H^{(i)}, 0 \leq i < R$, be a discrete random variable 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

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

where w.p. stands for “with probability” and $A_i = \lceil W_i/m \rceil$, which is the smallest integer greater than or equal to W_i/m . Hence, the average number of backoff frames incurred by the i^{th} attempt of an REQ can be expressed as

$$E[H^{(i)}] = A_i - A_i(A_i - 1) \frac{m}{2W_i} \quad i = 0, 1, \dots, R - 1.$$

Then, the Laplace-Stieltjes transform of $H^{(i)}$ can be obtained as follows

$$\mathcal{L}_{H^{(i)}}(s) = \sum_{j=1}^{A_i-1} \frac{m}{W_i} e^{-js} + \left(1 - \frac{(A_i - 1)m}{W_i}\right) e^{-A_i s}. \quad (8)$$

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

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

So, the Laplace-Stieltjes transform of $Y^{(i)}$ can be given as

$$\mathcal{L}_{Y^{(i)}}(s) = \prod_{j=0}^i \mathcal{L}_{H^{(j)}}(\Delta s). \quad (9)$$

Therefore, the accumulated backoff time Y for an arbitrary REQ is given as

$$Y = \begin{cases} Y^{(i)} & \text{w.p. } (1-p)p^i, i = 0, 1, \dots, R-2 \\ Y^{(R-1)} & \text{w.p. } p^{R-1}. \end{cases} \quad (10)$$

From (10), the pdf of Y , denoted by $f_Y(y)$, can be obtained, and $E[Y]$ can be written as

$$E[Y] = (1-p) \sum_{i=0}^{R-2} p^i E[Y^{(i)}] + p^{R-1} E[Y^{(R-1)}], \quad (11)$$

where

$$E[Y^{(i)}] = \Delta \sum_{j=0}^i E[H^{(j)}].$$

And the Laplace-Stieltjes transform of Y can be written as

$$\mathcal{L}_Y(s) = \sum_{i=0}^{R-2} (1-p)p^i \mathcal{L}_{Y^{(i)}}(s) + p^{R-1} \mathcal{L}_{Y^{(R-1)}}(s). \quad (12)$$

Note that $\mathcal{L}_G(s)$, $\mathcal{L}_{H^{(i)}}(s)$, $\mathcal{L}_{Y^{(i)}}(s)$ and $\mathcal{L}_Y(s)$ will be used in Section 4.2 where the distribution of packet delay is derived.

At the instant of packet arrival, the queue at the SS may be in one of two cases: **S0** or **S1**.

For case **S0**, the service time of an REQ is $X_0 = G + Y$, noting that G and Y are independent, so the pdf of X_0 can be written as

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

So, $E[X_0] = E[G] + E[Y]$, and the Laplace-Stieltjes transform of X_0 can be written as

$$\mathcal{L}_{X_0}(s) = \mathcal{L}_G(s)\mathcal{L}_Y(s). \quad (13)$$

For case **S1**, the service time of an REQ is $X_1 = Y$, and the Laplace-Stieltjes transform of X_1 is therefore given by that of Y .

Thus the service time of an REQ is given by

$$X = \begin{cases} X_0 & \text{w.p. } 1-\rho \\ Y & \text{w.p. } \rho \end{cases} \quad (14)$$

and the mean service time can be written as

$$\begin{aligned} E[X] &= (1-\rho)(E[G] + E[Y]) + \rho E[Y] \\ &= E[Y] + (1-\rho)E[G]. \end{aligned} \quad (15)$$

Hence, the outer set of fixed point equations is completed by updating ρ as in Fig. 2.

4. Performance metrics

4.1 Throughput

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

$$\Gamma = \frac{N\lambda(1 - p^R)}{d/\Delta}. \quad (16)$$

4.2 Distribution of packet delay

Recall that X is a random variable representing the service time experienced by an REQ, irrespective whether the REQ will be successful or unsuccessful. Let us define a related random variable X' , which represents the service time experienced by a successful REQ. In addition, referring to Fig. 3, we define another random variable V which represents the time from the beginning of a data subframe to the end of a packet transmission. Hence, for a successful REQ, the corresponding packet delay $D(t)$ is comprised of the waiting time of the REQ in the queue $W_q(t)$, X' of the REQ, T_{RE} and V , which can be written as

$$D(t) = W_q(t) + X' + T_{RE} + V. \quad (17)$$

So, the Laplace-Stieltjes transform of $D(t)$ can be written as

$$\mathcal{L}_D(s) = \mathcal{L}_{W_q}(s)\mathcal{L}_{X'}(s)\mathcal{L}_V(s)e^{-sT_{RE}}. \quad (18)$$

To calculate $\mathcal{L}_D(s)$, we first need to derive $\mathcal{L}_{W_q}(s)$. In (Welch, 1964), the waiting time distribution has been derived for the generalized $M/G/1$ queueing process. Hence, we can apply this result for our model. The waiting time distribution for our model can be rewritten as

$$\mathcal{L}_{W_q}(s) = \frac{(1 - \lambda E[Y])\{\lambda[\mathcal{L}_{X0}(s) - \mathcal{L}_Y(s)] - s\}}{[1 - \lambda(E[Y] - E[X_0])][\lambda - s - \lambda\mathcal{L}_Y(s)]}. \quad (19)$$

The service time experienced by successful REQs X' is given by

$$X' = \begin{cases} Y' & \text{w.p. } \rho \\ Y' + G & \text{w.p. } 1 - \rho \end{cases} \quad (20)$$

where the random variable Y' is the accumulated backoff time for successful REQs only, and is given by

$$Y' = Y^{(i)} \quad \text{w.p. } \eta p^i. \quad (21)$$

Hence, the Laplace-Stieltjes transform of Y' is expressed as

$$\mathcal{L}_{Y'}(s) = \sum_{i=0}^{R-1} \eta p^i \mathcal{L}_{Y^{(i)}}(s) = \eta \sum_{i=0}^{R-1} p^i \prod_{j=0}^i \mathcal{L}_{H^{(j)}}(\Delta s) \quad (22)$$

Therefore, the Laplace-Stieltjes transform of X' can be written as

$$\mathcal{L}_{X'}(s) = \rho \mathcal{L}_{Y'}(s) + (1 - \rho) \mathcal{L}_{Y'}(s) \mathcal{L}_G(s). \quad (23)$$

Using (19) and (23), the remaining term in (18) that needs to be determined is $\mathcal{L}_V(s)$. Let $q(j)$ be the probability that there are j successful requests other than the tagged SS in a frame. The probability $q(j)$ follows a truncated binomial distribution

$$q(j) = \frac{Q(j+1)}{1 - Q(0)}, \quad 0 \leq j \leq n - 1. \quad (24)$$

Using the assumption that the BS randomly allocates data slots to successful requests, the pmf of V can be expressed as

$$V = iT \quad \text{w.p.} \quad \sum_{j=i}^{n'} \frac{q(j-1)}{j}, \quad i = 1, 2, \dots, n', \quad (25)$$

where $n' = \min(n, d)$. Then, the Laplace-Stieltjes transform of V can be written as

$$\mathcal{L}_V(s) = \sum_{i=1}^{n'} e^{-iT_s} \sum_{j=i}^{n'} \frac{q(j-1)}{j}. \quad (26)$$

From (19), (23) and (26), $\mathcal{L}_D(s)$ can be determined. Hence, by the properties of Laplace-Stieltjes transform, any moments of the delay distribution can be derived straightforwardly. In particular, the mean packet delay \overline{D} is given by

$$\overline{D} = -\frac{d\mathcal{L}_D(s)}{ds} \Big|_{s=0},$$

and the variance of packet delay is given by

$$\sigma_D^2 = \frac{d^2\mathcal{L}_D(s)}{ds^2} \Big|_{s=0} - \overline{D}^2.$$

5. Model validation and numerical results

In this section, we verify our analytical model using computer simulation and investigate the performances under various configurations of N , W and λ . To this end, we have developed an event-driven simulation program to simulate the broadcast polling mechanism of IEEE 802.16. The simulator was written in C++. In the simulation model, the channel is operated in TDD mode, in which a frame is divided into a downlink and uplink subframe. The MAC and physical layer parameters were configured in accordance with default parameters taken from the standard (IEEE 802.16 standard, 2009). In particular, the frame duration is 1 msec consisting of 2500 mini slots each of 0.4 μ sec length. Each bandwidth request consists of 6 mini slots including 3 mini slots for subscriber station transition gap (SSTG), 2 mini slots for preamble and one mini slot for a bandwidth request message of 48 bits. The length of a data slot including the preamble and transition gap is 37.6 μ sec (i.e. 94 mini slots). Each SS has an

infinite buffer fed by a Poisson traffic source with mean arrival rate λ packet per msec. The head-of-queue packet of each SS makes bandwidth request and follows the TBEB mechanism. Based on the contention result, the processes of bandwidth allocation and packet transmission are then carried out. The duration of each simulation is 5000 seconds long, with an initial transient period of 300 seconds. For the analytical results, we set δ of Fig. 2 equal to 10^{-8} . As shown in the following figures, the numerical results match well with values obtained from simulation.

Therefore, our model is suitable for studying the impact of different parameters on the performance of contention-based services of IEEE 802.16.

We evaluate the impact of the number of SSs (N) and the initial backoff window (W) on various performance metrics. We set $r = 4, R = 8, m = 10, d = 8, \lambda = 0.1$. The results are shown in Fig. 4(a) to Fig. 4(f). The failure probability of REQ (p) under different N with $W = 8, 16, 32$ are plotted in Fig. 4(a). As expected, larger N leads to more request contentions and thus larger p . On the other hand, p decreases as W increases. This is because when W increases, there are more choices of a request slot in each backoff stage. As a result, the probability that an SS transmits a request in a request slot (τ) becomes smaller. So, p_c and p decrease.

Fig. 4(b) plots the mean service time of REQs against N with $W = 8, 16, 32$, respectively. Since p increases with N , it means that larger N increases the average number of attempts of a successful REQ. This results in a larger mean service time. Similarly, larger W leads to larger backoff time which constitutes the service time of REQs. Therefore, the mean service time also increases with W .

Fig. 4(c) and Fig. 4(d) plot the mean and variance of packet delay against N for various W , respectively. Since the mean service time contributes part of the mean packet delay, as expected from Fig. 4(b), the mean packet delay also increases with both N and W .

Fig. 4(e) also indicates that larger W results in higher traffic load for a given N . However, increasing W does not increase the net throughput when N is fixed. Therefore, it is actually better to choose small W and tolerate a slightly higher REQ unsuccessful probability.

Next, we evaluate the impact of the packet arrival rates (λ) on the performance metrics. We set $r = 4, R = 8, m = 10, d = 8, N = 30, W = 8, 16, 32$. The results are shown in Fig. 5(a) to Fig. 5(f). Essentially, increase in λ means increasing the offered traffic load ρ . Therefore, this set of results would resemble to that of varying N . The failure probability of REQ under different λ and W are plotted in Fig. 5(a). As packet arrival rate increases, each node is more likely to make requests and hence p also increases.

At last, we also consider how d influences the performance of the mean packet delay and normalized network throughput. As shown in Fig. 6(a), mean packet delay does not change too much against d for a given N . On the other hand, the normalized network throughput varies greatly, so it is important to choose suitable values of m and d .

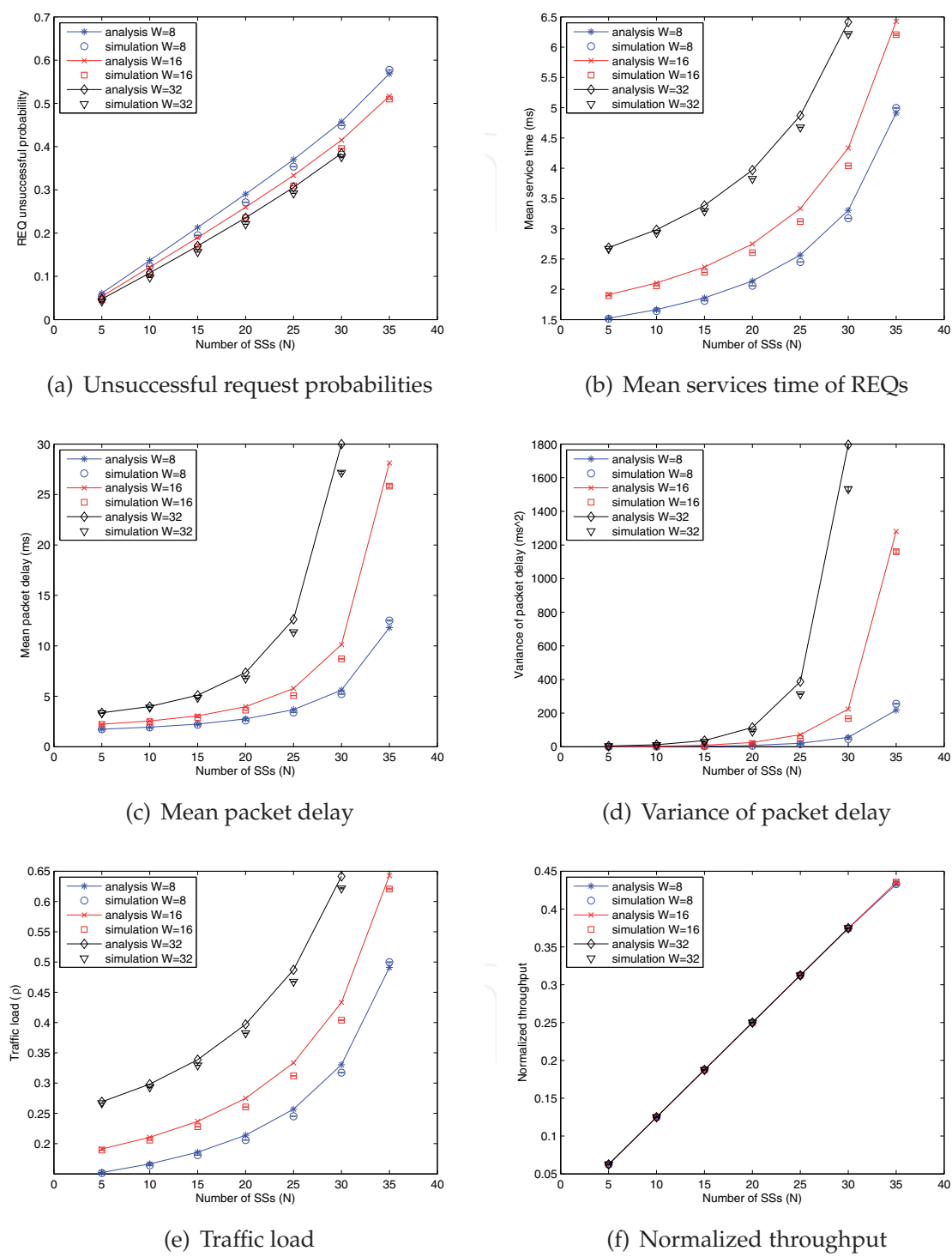


Fig. 4. Results for varying N and W , when $r = 4, R = 8, m = 10, d = 8, \lambda = 0.1$.

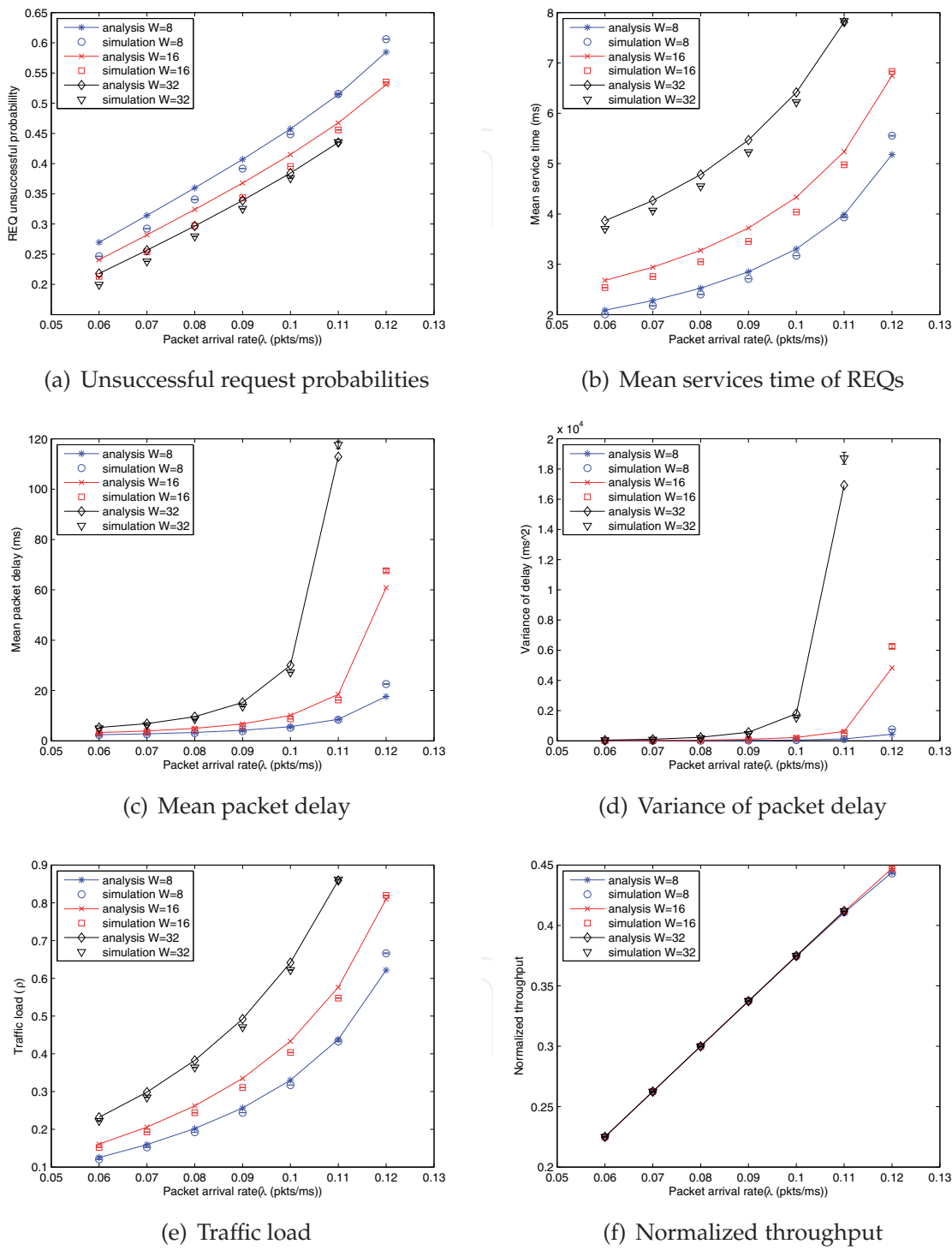


Fig. 5. Results for varying λ and W , when $r = 4, R = 8, m = 10, d = 8, N = 30$.

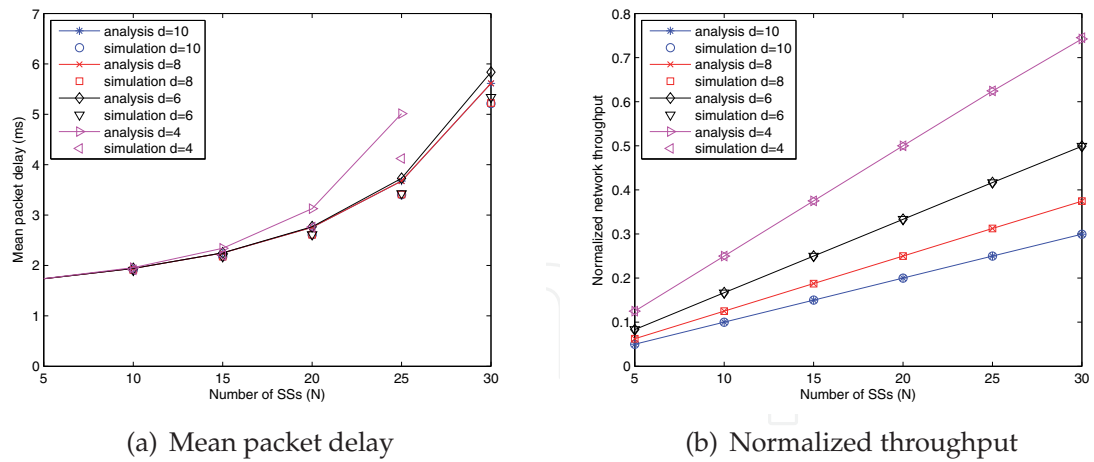


Fig. 6. Results for varying N and d , when $r = 4, R = 8, m = 10, W = 8, \lambda = 0.1$.

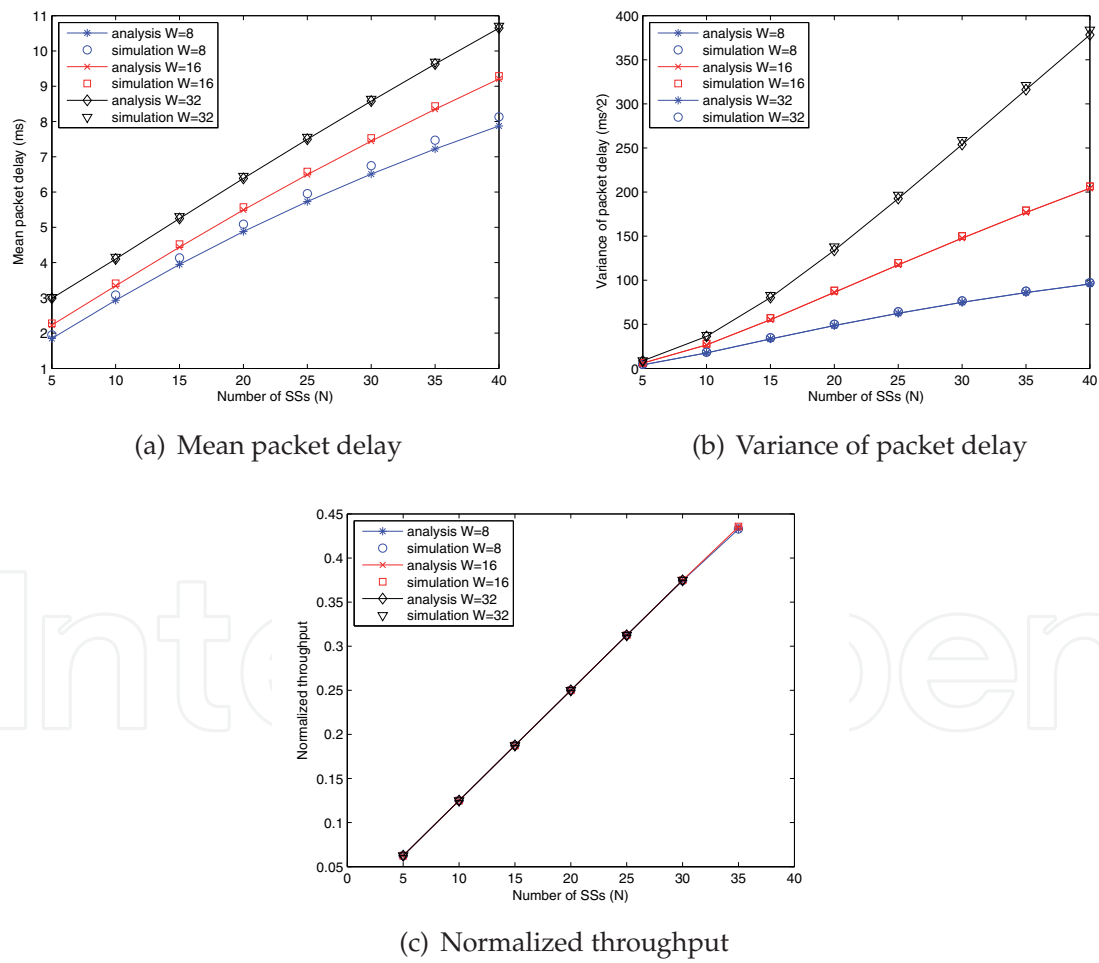


Fig. 7. Results for saturated networks, when $r = 4, R = 8, m = 10, d = 8, \lambda = 0.1$.

6. Saturated networks

As defined in Section 1, saturated networks mean that each SS always has a packet to send. In other words, $\rho = 1$. Hence, the outer set in Fig. 2 is not required for the saturation case and (3) becomes

$$\tau = 1/(B_{avg} + 1) \quad (27)$$

Meanwhile, the case **S0** in Section 3 does not exist. Therefore, the service time of an REQ X is equal to Y . For the same reason, the service time of an successful REQ X' is equal to Y' . Obviously, there is no need to calculate the waiting time in the queue of an REQ for saturated networks. So the delay of a packet can be changed to packet access delay as the time duration from the beginning of the request interval in which a request initiates the TBEB process till the end of the transmission of the packet, which is given by

$$D_{sat} = Y' + T_{RE} + V. \quad (28)$$

So, the Laplace-Stieltjes transform of D_{sat} can be written as

$$\mathcal{L}_{D_{sat}}(s) = \mathcal{L}_{Y'}(s)\mathcal{L}_V(s)e^{-sT_{RE}}. \quad (29)$$

And the normalized network throughput for saturated works is given by

$$\Gamma_{sat} = \frac{\sum_{j=1}^d jQ(j) + \sum_{j=d+1}^k dQ(j)}{d}. \quad (30)$$

In order to verify this degenerated model for the saturated network, the mean and variance of packet access delay and throughput against N with different W are plotted as Fig. 7(a) to Fig. 7(c). It can be seen that the analytical and simulation results again match very well.

7. Conclusion

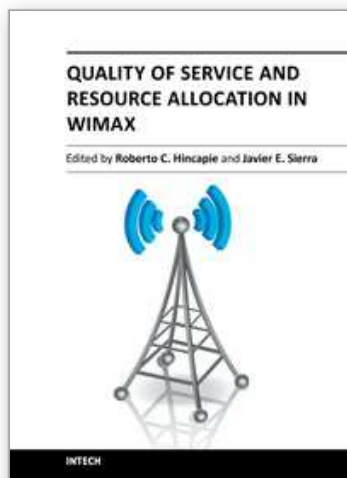
In this chapter, we have developed a unified performance model to evaluate the performances of the contention-based services in both saturated and unsaturated IEEE 802.16 networks. Different from some related works which assume that the probability of an SS sending a bandwidth request is an input parameter, our model takes into account the details of the backoff process to evaluate this probability. By solving two nested sets of fixed point equations, we have obtained the failure probability of a bandwidth request and the probability that a subscriber station has at least one REQ to transmit. Based on these two probabilities, the network throughput and the distribution of packet delay are derived. The model has been validated by simulations and shown to be accurate. Using the model, we have been able to investigate the impact of various parameters on the performance metrics of the 802.16 network.

8. References

IEEE 802.16-2009. IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems, *IEEE*, May 2009.

- J. G. Andrews; A. Ghosh & R. Muhamed (2007). *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*, Prentice Hall, ISBN 0-13-222552-2.
- B. Kwak; N. Song & L. E. Miller. Performance Analysis of Exponential Backoff. *IEEE/ACM Trans. on Networking*, vol. 13, no. 2, 2005, pp. 343-355.
- R. Iyengar; P. Iyer & B. Sikdar. Delay Analysis of 802.16 based Last Mile Wireless Networks. *Proceedings, IEEE Globecom'05*, 2005, pp. 3123-3127.
- A. Vinel; Y. Zhang; M. Lott & A. Tiurlikov. Performance Analysis of the random access in IEEE 802.16. *Proceedings, IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Berlin, September 2005.
- J. He; K. Guild; K. Yang & H. H. Chen. Modeling Contention Based Bandwidth Request Scheme for IEEE 802.16 Networks. *IEEE Communications Letters*, vol. 11, no. 8, August 2007, pp. 698-700.
- H. L. Vu; S. Chan & L. Andrew. Performance Analysis of Best-Effort Service in Saturated IEEE 802.16 Networks. *IEEE Trans. on Vehicular Technology*, vol. 59, no. 1, 2010, pp. 460-472.
- Q. Ni & L. Hu. An Unsaturated Model for Request Mechanisms in WiMAX. *IEEE Communications Letters*, vol. 14, no. 1, Jan. 2010, pp. 45-47.
- Y. P. Fallah; F. Agharebparast; M. R. Minhas; H. M. Alnuweiri & V. C. M. Leung. Analytical Modeling of Contention-Based bandwidth Request Mechanism in IEEE 802.16 Wireless Networks. *IEEE Trans. on Vehicular Technology*, vol. 5, no. 5, 2008, pp. 3094-3107.
- H. Fattah & H. Alnuweiri. Performance Evaluation of Contention-Based Access in IEEE 802.16 Networks with Subchannelization. *IEEE ICC on Communications*, 2009, pp. 1-6.
- D. Chuck; K. Chen & J. M. Chang. A Comprehensive Analysis of Bandwidth Request Mechanisms in IEEE 802.16 Networks. *IEEE Trans. on Vehicular Technology*, vol. 59, no. 4, 2010, pp. 2046-2056.
- R. P. Agarwal; M. Meehan & D. O'Regan. Fixed point theory and applications. Cambridge University Press, New York, ISBN 0-52-180250-4, 2001.
- Peter D. Welch. On a Generalized M/G/1 Queueing Process in Which the First Customer of Each Busy Period Receives Exceptional Service. *Operations Research*, vol. 12, no. 5, 1964, pp. 736-752.

IntechOpen



Quality of Service and Resource Allocation in WiMAX

Edited by Dr. Roberto Hincapie

ISBN 978-953-307-956-1

Hard cover, 376 pages

Publisher InTech

Published online 03, February, 2012

Published in print edition February, 2012

This book has been prepared to present state of the art on WiMAX Technology. It has been constructed with the support of many researchers around the world, working on resource allocation, quality of service and WiMAX applications. Such many different works on WiMAX, show the great worldwide importance of WiMAX as a wireless broadband access technology. This book is intended for readers interested in resource allocation and quality of service in wireless environments, which is known to be a complex problem. All chapters include both theoretical and technical information, which provides an in depth review of the most recent advances in the field for engineers and researchers, and other readers interested in WiMAX.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jianqing Liu, Sammy Chan and Hai L. Vu (2012). A Unified Performance Model for Best-Effort Services in WiMAX Networks, Quality of Service and Resource Allocation in WiMAX, Dr. Roberto Hincapie (Ed.), ISBN: 978-953-307-956-1, InTech, Available from: <http://www.intechopen.com/books/quality-of-service-and-resource-allocation-in-wimax/a-unified-performance-model-for-best-effort-services-in-wimax-networks>

INTech
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen