Performance Analysis of Contention Based Bandwidth Request Mechanisms in WiMAX Networks

Qiang Ni, *Senior Member, IEEE*, Ling Hu, Alexey Vinel, *Member, IEEE*, Yang Xiao, *Senior Member, IEEE*, and Marios Hadjinicolaou, *Member, IEEE*

Abstract—**WiMAX networks have received wide attention as they support high data rate access and amazing ubiquitous connectivity with great quality-of-service (QoS) capabilities. In order to support QoS, bandwidth request (BW-REQ) mechanisms are suggested in the WiMAX standard for resource reservation, in which subscriber stations send BW-REQs to a base station which can grant or reject the requests according to the available radio resources. In this paper we propose a new analytical model for the performance analysis of various contention based bandwidth request mechanisms, including grouping and no-grouping schemes, as suggested in the WiMAX standard. Our analytical model covers both unsaturated and saturated traffic load conditions in both error-free and error-prone wireless channels. The accuracy of this model is verified by various simulation results. Our results show that the grouping mechanism outperforms the no-grouping mechanism when the system load is high, but it is not preferable when the system load is light. The channel noise degrades the performance of both throughput and delay.**

Index Terms—**Bandwidth request (BW-REQ), broadband wireless access (BWA), IEEE 802.16, Markov chain, medium access control (MAC), performance analysis, quality-of-service (QoS), WiMAX.**

I. INTRODUCTION

IN recent years broadband wireless access (BWA) has gained special attention [1]. While the IEEE 802.11 technologies were very successful in indoor wireless LAN (WLAN) applications, it was concluded that the overall design and feature set of 802.11 WLANs were not well suited for outdoor BWA applications. In order to meet the outdoor need, the IEEE 802 committee set up a working group to develop a new standard for BWA applications, namely IEEE 802.16 [2]. To promote the 802.16 standard, an industrial association, namely Worldwide interoperability for microwave access (WiMAX) forum [3]–[5], was formed to define the interoperability specifications between 802.16 products from different manufacturers. A device which

passes the WiMAX conformance and interoperability testing is regarded as a WiMAX-certified device. IEEE 802.16 networks are also often referred to as WiMAX networks.

A. Overview of WiMAX Standard

Two operational modes are specified in WiMAX: a mandatory point-to-multipoint (PMP) mode and an optional mesh mode. In the PMP mode, a centralized base station (BS) serves a set of subscriber stations (SSs) within the same antenna sector in a broadcast manner. The transmissions between the BS and the SSs are realized in a frame structure by means of time division multiple access (TDMA). To support duplexing, WiMAX defines both time division duplexing (TDD) and frequency division duplexing (FDD). On the other hand, in the mesh mode, the SSs are organized in an ad hoc manner.

In this paper, we focus on the PMP mode, as it is the main choice preferred by the WiMAX operators. We analyze the TDMA/TDD mode, and similar analysis can also be applied to the FDD configuration. The TDMA/TDD frame structure is illustrated in Fig. 1. It consists of a downlink subframe for transmission from the BS to the SSs and an uplink sub-frame for transmission in the reverse direction. A Tx/Rx transition gap (TTG) and an Rx/Tx transition gap (RTG) are specified between the downlink and uplink sub-frames and between the uplink and the following downlink sub-frames in the next frame duration, respectively, to allow the SS terminals to switch from reception to transmission and vice versa. In the downlink sub-frame, two management messages (downlink MAP (DL-MAP) and uplink MAP (UL-MAP)) are transmitted which indicate the bandwidth allocation for data transmission in both the downlink and uplink directions, respectively. The uplink transmissions from the SSs to the BS are controlled by the BS through UL-MAP messages transmitted in a prior downlink subframe. At the medium access control (MAC) layer, the BS schedules the resources of the uplink channel for initial ranging, bandwidth request (BW-REQ), and data transmissions.

To meet the quality-of-service (QoS) requirements of multimedia applications, the WiMAX standard provides four basic scheduling services: unsolicited grant service (UGS), real-time polling service (rtPS), non-real-time polling service (nrtPS) and best effort (BE) [2], [6]. As for the UGS, when a UGS connection is set up, a fixed bandwidth is allocated to this flow. Thus, an SS is not allowed to request bandwidth for an established UGS flow. For other types of services, the SSs need to send requests to reserve bandwidth resources.
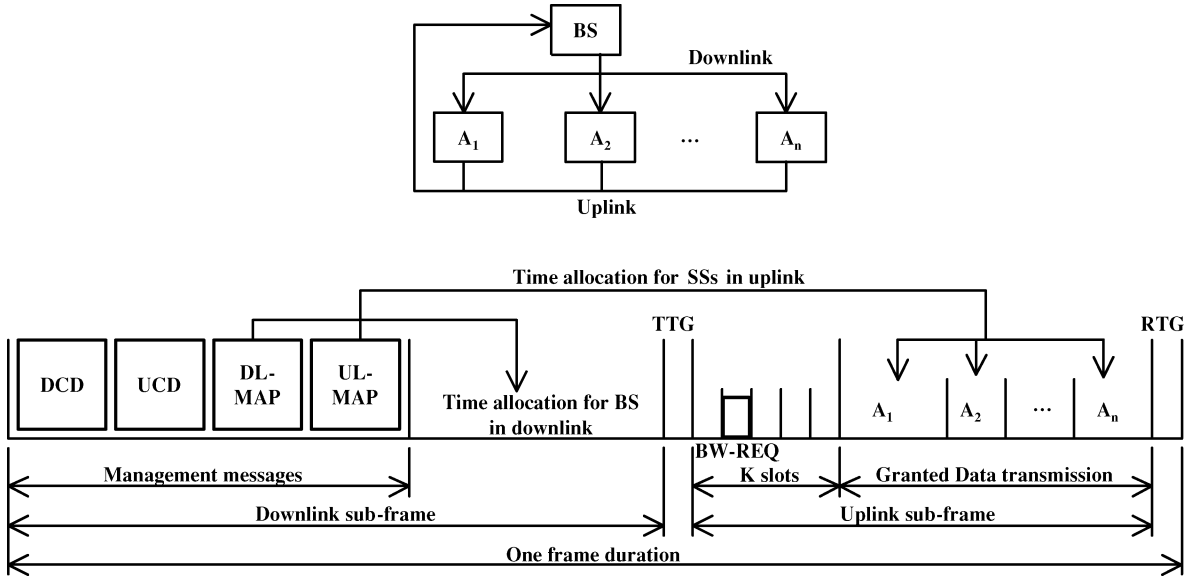
Fig. 1. TDMA/TDD operation of WiMAX networks in PMP mode [2], [7].

In this paper, we focus on the performance analysis of BW-REQ, which allows us to focus our attention on improving the efficiency of the BW-REQ algorithms as it is a fundamental component for the complete analysis of the WiMAX medium access control (MAC) protocols. In the next subsection, we will describe BW-REQ mechanisms as specified in the WiMAX standard.

B. BW-REQ Mechanisms in WiMAX

When the SSs want to request some uplink bandwidth from the BS, they can send either a stand-alone BW-REQ or a piggybacked one in an uplink data packet. This BW-REQ message may be aggregate or incremental. In the first case, an aggregate request indicates the total amount of bandwidth that one SS requires. On the other hand, if an SS sends an incremental BW-REQ, it means that it needs some additional amount of bandwidth to its existing allocation. An SS may establish multiple connections with a BS, where each BW-REQ request should be per-connection based [2], [7].

Two basic mechanisms are suggested in the WiMAX standard for BW-REQ transmission: contention-based random access and contention-free based polling access (unicast polling).

As shown in Fig. 1, in the case of contention-based random access, an SS transmits a BW-REQ during a slotted time duration known as a contention period, whose duration within a repetitive WiMAX frame consists of K slots[1]. A random backoff mechanism is used to resolve contention among multiple SSs. The mandatory method of random access used in WiMAX is called a truncated binary exponential backoff (TBEB) scheme without carrier sensing, in contrast to the widely used carrier sense multiple access with collision avoidance (CSMA/CA) mechanism in IEEE 802.11 WLANs.

[1]As specified in the WiMAX standard, the frame duration and the length of a contention period (K slots) are determined by the base station. In this paper, we do not consider downlink sub-frame and uplink data packet transmission in our model analysis. Only uplink BW-REQ contention period is to be considered. Hence, the frame duration is equal to the length of contention period, i.e., K slots.

Before each transmission attempt of a BW-REQ, an SS uniformly chooses a random integer from the interval of $[0, W_j - 1]$, where W_j denotes the current value of its contention window [2]. The chosen value, also referred to as a backoff counter, indicates the number of slots the SS has to wait before the transmission of this request. In contrast to WLAN, the backoff counter in WiMAX can be decremented every slot over the contention period in a frame. Since no carrier sensing is used, the SS does not freeze its backoff decrement if another station is transmitting. When the backoff counter reduces to zero, the SS is authorized to transmit its BW-REQ. If the backoff counter of one SS does not reach zero at the end of the contention period, the SS freezes its backoff counter and resumes decreasing it in the next contention period. If the backoff counters of two or more SSs reach zero at the same time, they will transmit BW-REQs simultaneously and cause a collision to occur. For the first transmission attempt, the contention window size starts with a minimal value which is denoted by W_0 in this paper. Upon each transmission failure, the station doubles its contention window value according to the standard. Hence, the backoff window after the j-th collision, W_j, becomes $2^j W_0$, until reaching a maximum value denoted by W_m. In this paper, W_m is equal to $2^f W_0$, where f denotes the maximum number of doubling the backoff window from W_0 to W_m. After the f-th attempt, while the number of retransmissions keeps increasing, the backoff window remains constant and equal to W_m until the maximum number of retransmissions m is reached. At this point, this BW-REQ will be discarded. On the other hand, if a transmission is successful, the backoff window size should be set back to the minimum value W_0, assuming that the channel is free again. Both W_0 and W_m are specified by the BS in a management field of the downlink sub-frame, called the uplink channel descriptor (UCD), as illustrated in Fig. 1.

When polling-based BW-REQ allocation is chosen, the BS will maintain a list of registered SSs and poll them according to this list. Each SS is only allowed to transmit the BW-REQ message after it is polled. Note that scheduling algorithms for

polling are vendor-dependent and are not specified in the standard. One may choose a simple round-robin scheduler to poll each SS sequentially in the polling list. Or vendors can choose priority-based polling mechanisms for BW-REQ scheduling.

Furthermore, random access may work in combination with polling, which is referred to as a grouping mode. When a group is polled, contention is used for those SSs within the group to compete for BW-REQ transmission. This grouping mechanism is suggested when available bandwidth is insufficient for a BS to individually poll many SSs.

In WiMAX, no explicit acknowledgement (ACK) frame is sent back to indicate whether a BW-REQ message is successfully transmitted or distorted. If a grant is not awarded within a special time duration, called T16 timeout [2], the SS decides that the BW-REQ was corrupted and starts a resolution process. On the other hand, on receiving a grant within the T16 timeout, the SS uses the allocated bandwidth for uplink transmission of data packets or to piggyback additional requests if necessary. Furthermore, the SS might know how much bandwidth is awarded by the BS by observing the following grant.

C. Related Work

Analytical models are powerful and versatile tools that can help us to understand the performance of protocols under different scenarios and therefore to explore possible methods in improving the mechanisms.

In [8], Giuseppe Bianchi proposed a stochastic Markov chain model to analyze the performance of the IEEE 802.11 distributed coordination function (DCF) under the saturated assumption, where stations always have data to send and persist in contention to access the channel. Using the saturation assumption, some other researchers extended his work to the performance analysis of service differentiation for 802.11 and 802.11e (e.g., [9] and [10]). Although the saturation assumption allows queuing dynamics to be neglected and avoids the need for detailed modeling of traffic characteristics, networks do not typically operate in saturated conditions [11]. Hence, some modified Markov chain models were proposed to model the 802.11 DCF in unsaturated conditions, e.g., [12] and [13].

To the best of our knowledge, until now most Markov chain models were proposed for IEEE 802.11 WLAN, while only a few papers were presented to model WiMAX networks. Those models proposed for WLAN cannot be directly used for WiMAX, since the random access method suggested in WiMAX is based on a truncated binary exponential backoff mechanism without carrier sensing, which is different from WLAN, as detailed in Section I-B.

In [14]–[16], Markov chain models under the saturated conditions were proposed for the performance analysis of the contention-based BW-REQ mechanism in WiMAX. Since the sizes of BW-REQ packets are generally small and the generation of BW-REQs is typically bursty, unsaturated conditions are more common in WiMAX than in 802.11 WLAN.

Recently, an unsaturated Markov chain model for WiMAX was proposed in [17] for the contention-based BW-REQ mechanism. However, their model assumes the channel is error-free.

Furthermore, the grouping mode is not addressed in their analysis. We presented an unsaturated model under error-free channel for BW-REQ analysis in [11].

In this paper, we derive a new unified analytical model for both grouping and no-grouping modes under various traffic and channel conditions. The main contributions of this paper are as follows.

- We first derive a simple and unified analytical model that can investigate various contention-based BW-REQ mechanisms in WiMAX with grouping and no-grouping modes. Various traffic and channel conditions are covered by our model, including saturated and unsaturated traffic conditions as well as error-free and error-prone channel conditions.
- Our analytical model can be easily utilized to investigate how the performance of the protocol is affected by a variety of important parameters, such as the number of groups, the number of stations in each group, the traffic load per station, and channel conditions.
- Our analytical model can be used to ensure efficient operation of BW-REQ mechanisms and to optimize the system performance under varying network conditions.

The rest of this paper is organized as follows: Section II presents in detail our analytical model. In Section III we validate our model by comprehensive comparisons with simulation results. Finally, Section IV draws the conclusion and highlights the future work.

II. A UNIFIED MODEL FOR CONTENTION BASED BW-REQ IN WIMAX: NO-GROUPING AND GROUPING

We present our analytical model in this section. Our aim is to derive a unified analytical model that can investigate both grouping and no-grouping modes. Our main performance metrics are collision probability, mean delay and throughput.

A. Model Description

Let us assume that there are one BS and n SSs in a WiMAX system, with each SS having one connection to the BS. In the case of multiple connections per SS, n refers to the total number of connections in the system.

For the simplicity of analysis, we assume that the total number of SSs, n, can be exactly divided into $g(\geq 1)$ groups with each group having exactly *n/g* stations. Non-even group allocation will be our future work. Actually, the no-grouping (pure contention) scheme is a special type of grouping scheme, i.e., the group number $g = 1$. The BS polls each group in sequence. At the beginning of each slot, only the stations within the group that is currently being polled are allowed to decrement their backoff counters. Stations from other groups have to wait until their group is polled. When a station's backoff counter reaches zero, it can transmit its BW-REQ. Hence, for a given station in any group, we can model its backoff behavior with a discrete-time 2-D Markov chain $\{l(t), c(t)\}$, as depicted in Fig. 2, where the item $l(t)$ denotes the stochastic process of backoff stage at different contention window levels and the
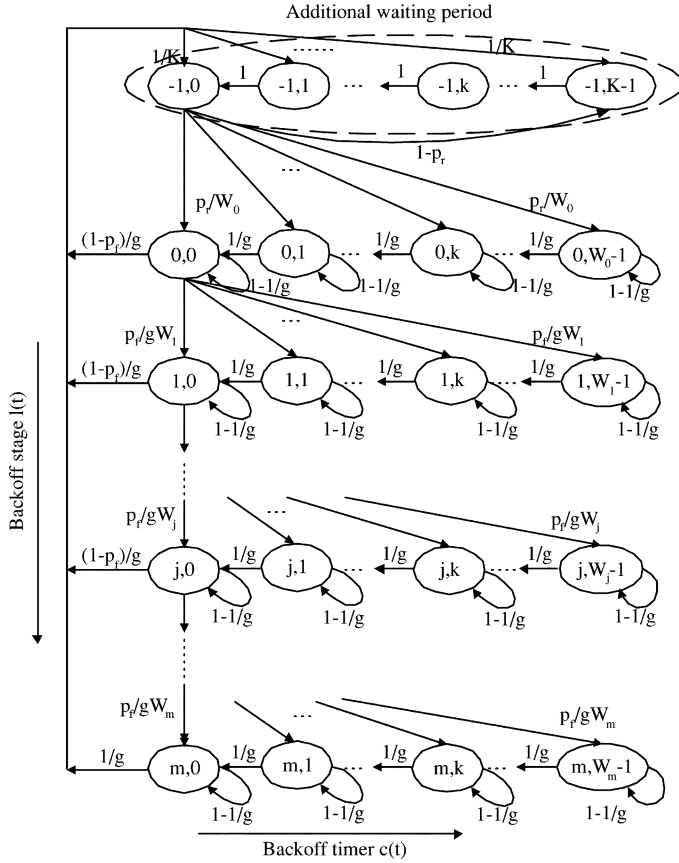
Fig. 2. Unified model for no-grouping and grouping contention-based BW-REQ Access.

item $c(t)$ represents the stochastic process of backoff timer at a given time slot t.

According to the standard specifications, WiMAX utilizes a TDMA based frame structure for BW-REQ and data transmissions. After each BW-REQ transmission attempt, the SS does not immediately start the backoff process for the next transmission attempt; instead, it has to wait until the beginning of the next frame. To model this behavior, we consider each WiMAX frame consists of a fixed number (denoted by K) of equally-sized slots for contention access. The duration of a slot is sufficient for one BW-REQ transmission. As shown in Fig. 2, we introduce in our Markov chain model the additional waiting states $\{-1, k\}, k = 0, 1, \ldots, K-1$ to represent the above case when a station just completes a BW-REQ transmission (either successfully transmitted or discarded upon reaching the retry limit m). In this case, the number of additional waiting slots that an SS should wait is uniformly distributed over $[0, K-1]$. On reaching the $\{-1, 0\}$ state, if a BW-REQ is queued, the SS goes to any state of the $\{0, k\}$ row starting the backoff; otherwise, it goes to the $\{-1, K-1\}$ state (i.e., waiting until the next frame).

In order to model both saturated and unsaturated conditions, the following behaviors are captured. On completion of a previous additional waiting period, if a BW-REQ is generated at the MAC queue (which occurs with probability p_r), a station can enter the first row of the backoff stage; otherwise (i.e., no request is generated) it has to wait for the next frame with the probability $1-p_r$. When $p_r \rightarrow 1$, it models saturated conditions. In other cases, it models unsaturated conditions. Without loss of

generality, we assume that BW-REQs arriving at the MAC of each SS follow Poisson distribution[2] with arrival rate λ. Denote the average duration when a BW-REQ is generated till it leaves the MAC queue (either successfully transmitted or dropped after reaching the retry limit) by $E(T)$, we can obtain the relation between p_r and the arrival rate λ as follows:

$$p_r = 1 - e^{(-\lambda E(T))}. \tag{1}$$

Here, $E(T)$ also refers to the average delay of BW-REQs, which will be derived in Section II.B.

Let p_c denote the probability that each BW-REQ collides, and p_e denote the error probability of each BW-REQ caused by random channel noise. We assume these probabilities are independent of the number of retransmissions which a request might have suffered. Hence, the total failure probability p_f is expressed as

$$p_f = p_c + p_e - p_c p_e. \tag{2}$$

In the backoff stage, whenever the value of the backoff counter reaches zero the station makes a transmission attempt. If this transmission attempt is successful, the state will move to any of the $\{-1, k\}$ waiting states and wait for the beginning of the next frame. On the other hand, when a collision occurs at the stage $j - 1$, the backoff stage will increase and the new backoff counter will be uniformly chosen in the range $[0, W_j - 1]$. At state $\{m, 0\}$, the BW-REQ will either be successfully transmitted or discarded by the station. After this state, the Markov chain goes back to the first row of the model.

In the following, we will explain in detail the transition probabilities between the states.

B. Mathematical Analysis

If we adopt the short notation $P\{(j_1, k_1)|(j_0, k_0)\} = P\{(l(t+1) = j_1, c(t+1) = k_1)|(l(t) = j_0, c(t) = k_0)\}$, the one-step non-null transition probabilities of the above Markov chain model are:

- after a successful transmission, the SS waits for the end of the current frame: $P\{(-1, k)|(j, 0)\} = (1 - p_f)/g$, $j \in [0, m]$, $k \in [0, K-1]$;
- when the retransmission limit is reached, the SS waits for the end of the current frame: $P\{(-1, k)|(m, 0)\} = 1/g$, $k \in [0, K-1]$;
- a new BW-REQ starts at backoff stage 0 with the probability of p_r, and the backoff is initially uniformly distributed in the range $[0, W_0 - 1]$: $P\{(0, k)|(-1, 0)\} = p_r/W_0$, $k \in [0, W_0 - 1]$;
- the SS remains in the (-1) row and waits for the end of the next frame as long as its queue is empty: $P\{(-1, K-1)|(-1, 0)\} = 1 - p_r$;
- the backoff counter decrements only if the group which the SS belongs to is currently polled: $P\{(j, k)|(j, k+1)\} = 1/g$, $j \in [0, m], k \in [0, W_j - 2]$;
- the backoff counter does not decrement when the group which the SS belongs to is not polled: $P\{(j, k)|(j, k)\} = 1 - 1/g, j \in [0, m]$, $k \in [0, W_j - 1]$;

- the SS enters the next backoff stage when a collision occurs: $P\{(j,k)|(j-1,0)\} = p_f/(gW_j), j \in [0,m], k \in [0, W_j - 1]$.

Let $b_{j,k} = \lim_{t \to \infty} P\{(l(t)=j, c(t)=k)\}$ be the stationary distribution of the Markov chain, with $j \in [0,m], k \in [0, W_j - 1]$. Hence, in the steady state, the following relations are valid.

If $f < m$

$$W_j = \begin{cases} W_0 2^j, & j \in [0, f] \\ W_0 2^f, & j \in [f, m]. \end{cases} \quad (3)$$

Otherwise

$$W_j = W_0 2^j, j \in [0, m]. \quad (4)$$

The normalization condition of this chain is

$$\sum_{j=0}^{m} \sum_{k=0}^{W_j-1} b_{j,k} + \sum_{k=0}^{K-1} b_{-1,k} = 1. \quad (5)$$

In the following, we write all probabilities in (5) in terms of $b_{0,0}$ and use the normalization condition to determine $b_{0,0}$.

1) Backoff States of the Markov Chain $(0 < j \leq m)$: From this part of the Markov chain, we can derive (6), as shown at the bottom of the page.

Through the iteration, (6) can be simplified as

$$b_{j,k} = (W_j - k)\frac{p_f}{W_j}b_{j-1,0}, j \in [1,m], k \in [0, W_j - 1]. \quad (7)$$

In particular, when $k = 0$, we get

$$b_{j,0} = p_f b_{j-1,0}, \quad j \in [1, m] \quad (8)$$

Dividing (7) by (8), we get the equation expressed by $b_{j,0}$ as follows:

$$b_{j,k} = \frac{W_j - k}{W_j}b_{j,0}, \quad j \in [1,m], k \in [0, W_j - 1]. \quad (9)$$

2) BW-REQ Arrival Row of the Markov Chain $(j = 0)$: From this part of the Markov chain, we can derive (10), as shown at the bottom of the page.

Through iteration, (10) can be simplified as

$$b_{0,k} = (W_0 - k)\frac{gp_r}{W_0}b_{-1,0}, \quad k \in [0, W_0 - 1]. \quad (11)$$

When $k = 0$, we get

$$b_{0,0} = gp_r b_{-1,0}. \quad (12)$$

Dividing (11) by (12), we get the equation expressed by $b_{0,0}$

$$b_{0,k} = \frac{W_0 - k}{W_0}b_{0,0}, \quad k \in [0, W_0 - 1]. \quad (13)$$

Combining (9) and (13) gives a unified expression for $b_{j,k}$

$$b_{j,k} = \frac{W_j - k}{W_j}b_{j,0}, \quad j \in [0, m], k \in [0, W_j - 1]. \quad (14)$$

Through iteration of (8), we get

$$b_{j,0} = p_f^j b_{0,0}, \quad j \in [0, m]. \quad (15)$$

Using (14) and (15), we get

$$\begin{aligned}
\sum_{j=0}^{m} \sum_{k=0}^{W_j-1} b_{j,k} &= \sum_{j=0}^{m} \sum_{k=0}^{W_j-1} \frac{W_j - k}{W_j}b_{j,0} \\
&= \sum_{j=0}^{m} \sum_{k=0}^{W_j-1} \frac{W_j - k}{W_j}p_f^j b_{0,0} \\
&= \frac{b_{0,0}}{2}\left[W_0 Z + \frac{p_f^{m+1} - 1}{p_f - 1}\right]
\end{aligned} \quad (16)$$

where, if $f < m$

$$Z = \frac{[(2p_f)^{f+1} - 1](p_f - 1) + 2^f\left(p_f^{m+1} - p_f^{f+1}\right)(2p_f - 1)}{(2p_f - 1)(p_f - 1)} \quad (17)$$

otherwise

$$Z = \frac{(2p_f)^{m+1} - 1}{2p_f - 1}. \quad (18)$$

3) K Waiting States After a Successful Transmission Attempt or the Packet is Discarded on Reaching Retry Limit $(j = -1)$: From this part of the Markov chain, we can derive that

$$\begin{cases} b_{-1,k} = \frac{1}{K} + b_{-1,k+1}, & k \in [0, K-2] \\ b_{-1,K-1} = \frac{1}{K} + (1 - p_r)b_{-1,0} \end{cases}. \quad (19)$$

$$\begin{cases} b_{j,k} = \left(1 - \frac{1}{g}\right)b_{j,k} + \frac{1}{g}b_{j,k+1} + \frac{p_f}{gW_j}b_{j-1,0}, & j \in [1,m], k \in [0, W_j - 2] \\ b_{j,W_j-1} = \left(1 - \frac{1}{g}\right)b_{j,W_j-1} + \frac{p_f}{gW_j}b_{j-1,0}, & j \in [1, m] \end{cases}. \quad (6)$$

$$\begin{cases} b_{0,k} = \left(1 - \frac{1}{g}\right)b_{0,k} + \frac{1}{g}b_{0,k+1} + \frac{p_r}{W_0}b_{-1,0}, & k \in [0, W_0 - 2] \\ b_{0,W_0-1} = \left(1 - \frac{1}{g}\right)b_{0,W_0-1} + \frac{p_r}{W_0}b_{-1,0} \end{cases}. \quad (10)$$

Through iteration, (19) can be simplified as

$$b_{-1,k} = (K - k)\frac{1}{K} + (1 - p_r)b_{-1,0}, \quad k \in [0, K - 1] \quad (20)$$

and when $k = 0$, we get

$$b_{-1,0} = \frac{1}{p_r}. \quad (21)$$

Dividing (20) by (21), we get the equation expressed by $b_{-1,0}$

$$b_{-1,k} = b_{-1,0}\left(1 - p_r\frac{k}{K}\right), \quad k \in [0, K - 1]. \quad (22)$$

Using (22) and (12), we can obtain the sum probability of the (-1) row

$$\sum_{k=0}^{K-1} b_{-1,k} = \sum_{k=0}^{K-1} b_{-1,0}\left(1 - p_r\frac{k}{K}\right)$$
$$= \frac{b_{0,0}}{g}\left[\frac{K}{p_r} - \frac{1}{2}(K - 1)\right]. \quad (23)$$

Then, by applying (16) and (23) to the normalization condition (5), we finally obtain $b_{0,0}$, as shown in (24), at the bottom of the page.

By utilizing the Markov chain, the probability τ that an SS transmits a BW-REQ in a randomly chosen slot time is equal to (25), also shown at the bottom of the page.

According to the Markov chain, considering the probability weight of each backoff delay (for the backoff stage j, the average time for the backoff counter reaches 0 is $(W_j - 1)/(2)$), taking into account the average waiting delay in K slots as well as the grouping effect, the average delay of BW-REQ in the slots can be calculated by

$$E(T) = g\sum_{j=0}^{m} p_f^j \frac{W_j - 1}{2} + \frac{K}{2}$$
$$= \frac{1}{2}g\left(W_0 Z - \frac{p_f^{m+1} - 1}{p_f - 1}\right) + \frac{K}{2}. \quad (26)$$

Using (1) and (26), we obtain

$$p_r = 1 - e^{-\lambda\left[\frac{1}{2}g\left(W_0 Z - \frac{p_f^{m+1} - 1}{p_f - 1}\right) + \frac{K}{2}\right]}. \quad (27)$$

In (25) and (27), there are three unknown variables: τ, p_r and p_f. In order to solve such a non-linear system, we also use another condition: each station can transmit BW-REQ with the probability $(1 - p_c)$ only when none of the remaining stations in the same group attempt to transmit during that slot. Therefore, it holds that

$$1 - p_c = (1 - \tau)^{\frac{n}{g} - 1} \quad (28)$$

which can be written as

$$1 - \left(\frac{p_f - p_e}{1 - p_e}\right) = (1 - \tau)^{\frac{n}{g} - 1}. \quad (29)$$

Hence, we obtain three equations, (25), (27) and (29) containing the aforementioned three unknown variables, which can be solved using numerical techniques. Therefore, we are able to obtain the collision probability and the mean delay performance. Due to space limitation, we omit the details of numerical techniques here.

The normalized throughput S is defined as the fraction of time the channel is used to successfully transmit BW-REQs. Let P_S, P_I and P_F denote the probability of successfully sending a REQ, being idle and being a failure, respectively. T_S, T_I and T_F are the corresponding slot durations of sending a REQ, being idle and being a failure. Hence, the throughput S can be calculated by

$$S = \frac{P_S T_S}{P_S T_S + P_I T_I + P_F T_F}. \quad (30)$$

Since $T_S = T_I = T_F = 1$ (one slot), we obtain

$$S = \frac{P_S}{P_S + P_I + P_F} = P_S = \frac{n}{g}(1 - p_e)\tau(1 - \tau)^{\frac{n}{g} - 1}. \quad (31)$$

$$b_{0,0} = \frac{2gp_r(p_f - 1)}{gp_r\left[W_0 Z(p_f - 1) + p_f^{m+1} - 1\right] + (p_f - 1)[2K - p_r(K - 1)]} \quad (24)$$

$$\tau = \sum_{j=0}^{m} b_{j,0} = \sum_{j=0}^{m} p_f^j b_{0,0}$$
$$= \frac{2gp_r\left(p_f^{m+1} - 1\right)}{gp_r\left[W_0 Z(p_f - 1) + p_f^{m+1} - 1\right] + (p_f - 1)[2K - p_r(K - 1)]}. \quad (25)$$
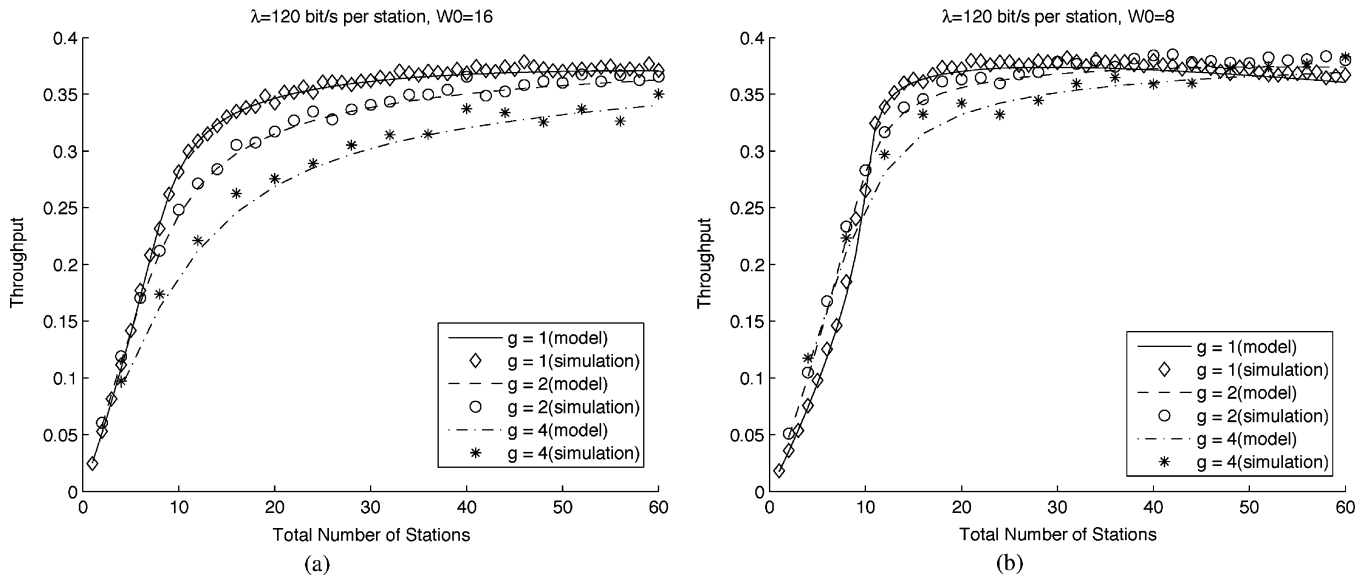
Fig. 3.   Normalized throughput (grouping versus no-grouping). (a) $W_0 = 16$. (b) $W_0 = 8$.
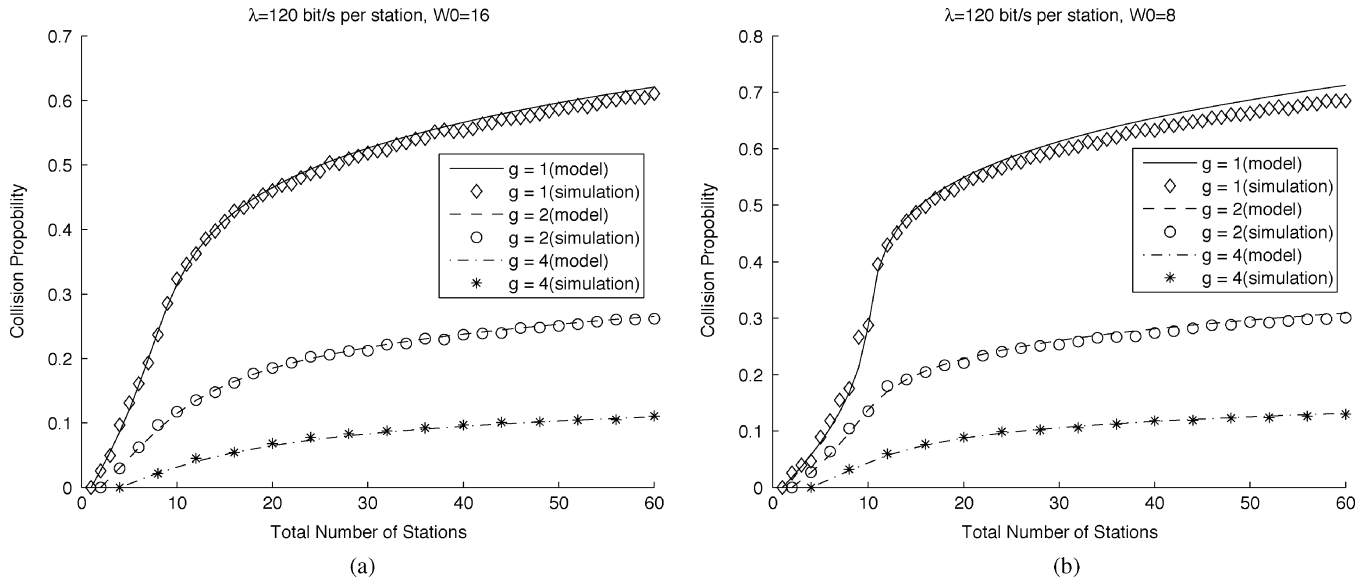

Fig. 4.   Collision probability (grouping versus no-grouping). (a) $W_0 = 16$. (b) $W_0 = 8$.

III. MODEL VALIDATION AND PERFORMANCE ANALYSIS

To verify the analytical results, we write Matlab codes to simulate the WiMAX systems. The simulation model implements the contention access protocols (both grouping and no-grouping) as specified in the standard. Note that the standard neither provides optimal setting values for W_0, m, W_m and K nor defines any relation between these parameters. Consider if $W_0 < K$, then some slots of the total K slots per frame will not be used during the first transmission attempt. Hence, we conclude that a relation should hold, i.e., $W_0 \geq K$. Thus, in our following analysis, the frame length K is chosen to be 8 slots (i.e., $K = 8$) and the minimal contention window size W_0 is chosen to be either 8 or 16 for testing. Both m and f are set as 7.

We plot in Figs. 3–7 the analytical model results versus the simulation results, including performance criteria for throughput, collision probability and mean delay. The analytical model results are plotted with lines (solid and various dashed lines). The simulation results are plotted with diamond dots, circle dots and star dots, with each point the mean value of ten experiments[3]. As shown in the figures, the analytical model results match the simulation results well. This demonstrates that our model is accurate in capturing the behaviors of grouping and no-grouping under both unsaturated and saturated conditions.

Figs. 3, 4 and 5 show the performances under error-free channel conditions. For every criterion, we vary the total number of stations with different types of group settings $(g = 1, 2, 4)$.

Fig. 3 shows how the normalized throughput behaves under different configurations. Fig. 3(a) compares the throughput performance between grouping $(g = 2, 4)$ and no-grouping $(g =$

---

[3]In each experiment, the time duration of 1000 periodic frames is chosen.
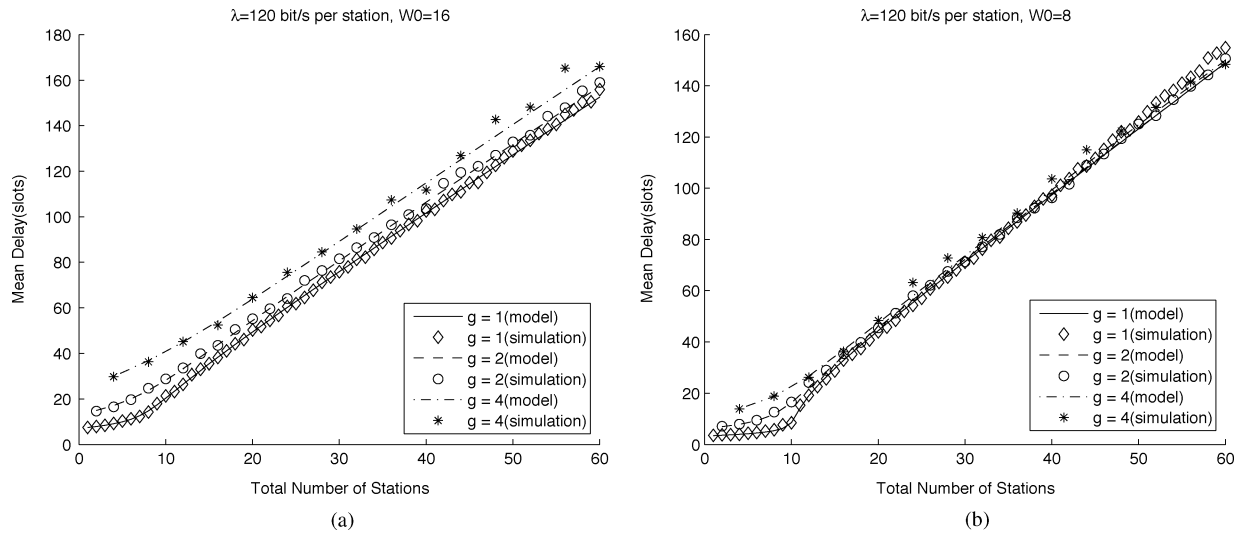
Fig. 5.   Mean delay (grouping versus no-grouping). (a) $W_0 = 16$. (b) $W_0 = 8$.
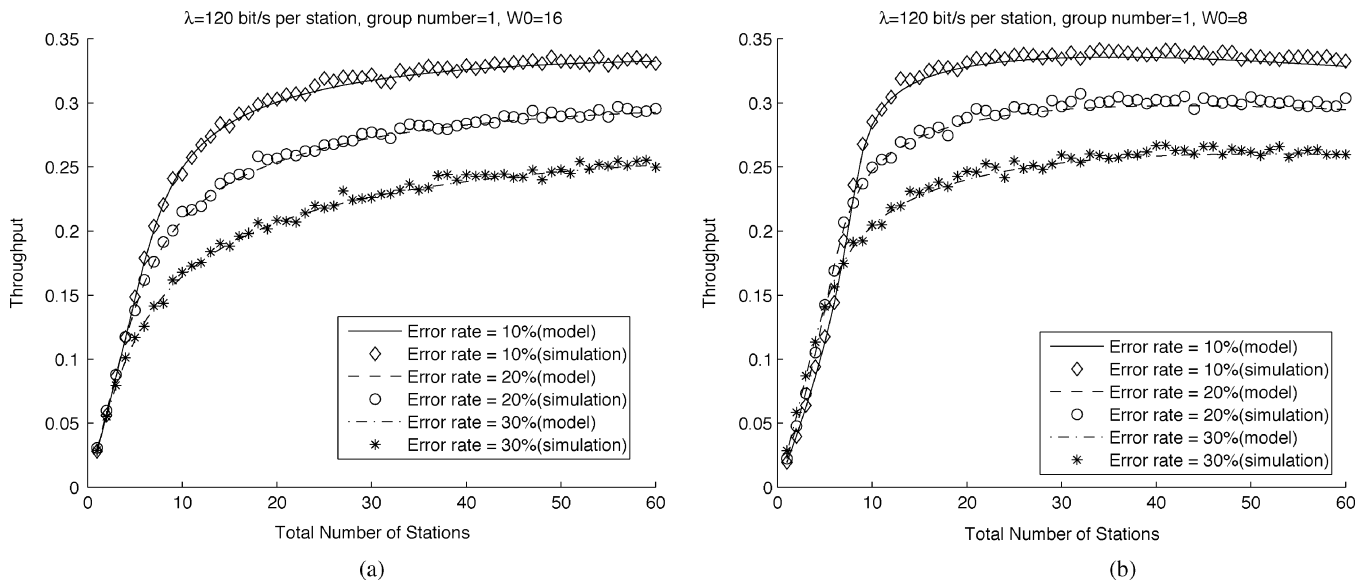

Fig. 6.   Normalized throughput (with channel noise). (a) $W_0 = 16$. (b) $W_0 = 8$.

1) when the initial contention window size (W_0) is 16. It demonstrates that, when the system is unsaturated (i.e., total number of stations in the system is small), no-grouping outperforms grouping. Furthermore, the more groups we divide the stations into, the lower the normalized throughput we can achieve. The benefit of grouping rises when the system becomes saturated. This benefit is shown more clearly in Fig. 3(b), where $W_0 = 8$. In such a case of no-grouping, the system gets saturated quickly and the throughput degrades after saturation. On the other hand, grouping keeps the system throughput stable even with a high number of stations (high load).

Fig. 4 confirms that, in all the cases, grouping reduces collision probability significantly. This is because more groups allow less stations to compete with each other which results in fewer collisions. However, by dividing stations into groups, and therefore having fewer stations per group, the model reduces the advantage of statistical multiplexing compared to no-grouping. Hence there is a tradeoff in using the grouping mode. When the system load becomes high, the risk of collisions also be-

comes high. The gains of reducing collisions are much higher than the losses of statistical multiplexing. Hence, grouping is recommended. On the other hand, when the system has a light load, grouping is not recommended.

Fig. 5 shows the mean delay performance for the same settings. It is shown that in most cases grouping induces a higher mean delay. This is again due to the loss of statistical multiplexing gain. That is, without grouping, each station can decrement its backoff counter in all K slots per frame since no carrier sensing and/or backoff freezing is needed in WiMAX, as explained in Section I-B. This is opposed to grouping, in which a station is not allowed to decrement its backoff counter until its group is polled. Hence, longer waiting time and higher mean delays are caused with grouping than without grouping. Again, when the system load is high, such a waiting in backoff can help reduce potential collisions. Hence, the loss of statistical multiplexing gains can be compensated by the gain from fewer collisions. This gain is more obvious with a smaller window size $(W_0 = 8)$. As a general remark, from the mean delay perfor-
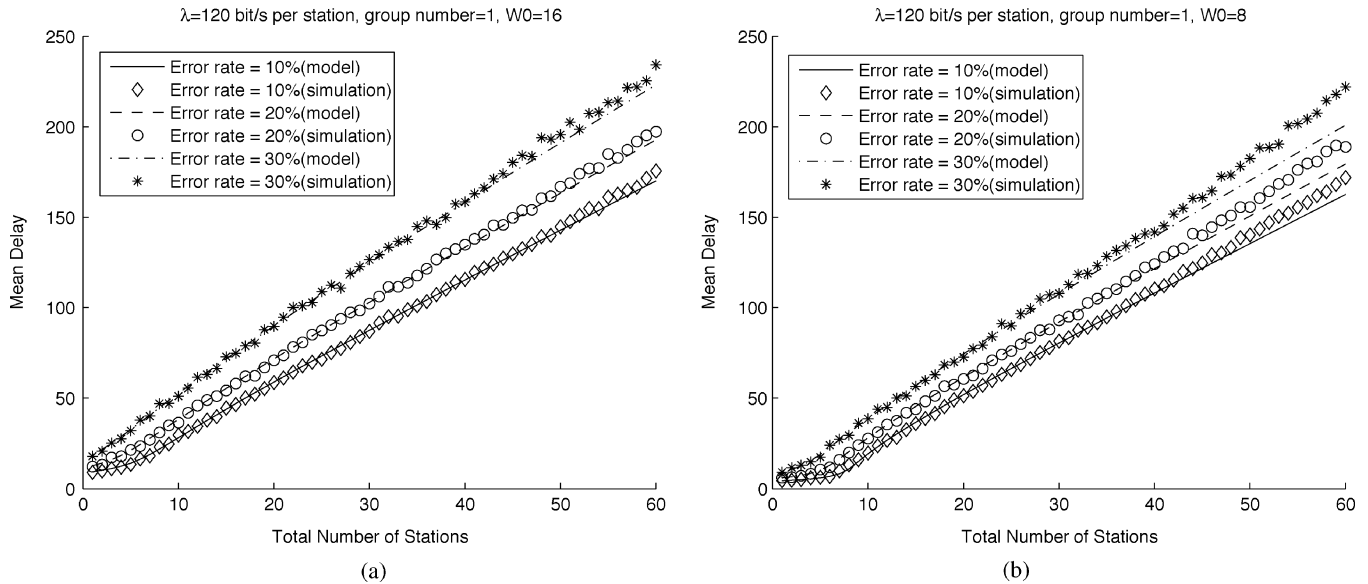
Fig. 7. Mean delay (with channel noise). (a) $W_0 = 16$. (b) $W_0 = 8$.

mance point of view, grouping is indeed not recommended when the system load is light.

Figs. 6 and 7 show the performances of the WiMAX system for the same settings in the presence of channel noise. To demonstrate the effect of channel noise, various error rates are chosen, i.e., 10%, 20%, and 30%. The initial contention window sizes (W_0) are set as 16 or 8 in order to compare the results to those without error rate. The group number is set at 1 (i.e., no-grouping) to clearly distinguish the influence from channel errors without from that of grouping.

Fig. 6 shows how the normalized throughput behaves under different error rates. It demonstrates clearly that the channel noise/error degrades the throughput performance. This is due to the fact that the channel noise/error increases the transmission failure probability of BW-REQ messages and hence reduces the throughput. The higher the error rate, the lower the throughput that is obtained. In order to improve the throughput performance, it is necessary to conduct further research on ameliorating the transmission channel environment or error-protection.

Fig. 7 shows the mean delay performance for the same settings. It is clear that the channel noise also degrades the delay performance, with a higher error rate inducing a higher mean delay. Again, this is due to the fact that the channel noise increases the transmission failure probability of BW-REQ messages, thus causing more retransmissions. Hence, longer waiting time period and mean delays result.

IV. CONCLUSION

In this paper, we analyzed the performance of contention based BW-REQ mechanisms as proposed in the WiMAX standard. We presented a new analytical model which can predict the performance of collision probability, throughput and mean delay of BW-REQs under various traffic and channel conditions (unsaturated, saturated, error-free and error-prone). The proposed analytical model is verified by extensive simulation results.

Both analytical and simulation results show that a grouping mechanism is helpful when the system load is high but not when the system load is light. It also demonstrates that the influence of channel noise on the BW-REQ mechanism is not negligible. The channel noise degrades the performance of both throughput and delay.

The proposed model can be utilized to facilitate the efficient operation of BW-REQ mechanisms and to optimize the system performance under varying network conditions. Our future work will include the performance analysis and optimization of the complete WiMAX MAC protocols, including bandwidth request, data packet scheduling and admission control algorithms under various traffic and channel conditions.

REFERENCES

[1] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: Current performance benchmarks and future potential," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 129–136, Feb. 2005.

[2] *IEEE Standard for Local and Metropolitan Area Networks*, IEEE Std 802.16-2005, Dec. 2005, Part16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems.

[3] WiMAX Forum; [Online]. Available: http://www.wimaxforum.org

[4] K. H. Teo, Z. Tao, and J. Zhang, "The mobile broadband WiMAX standard," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 144–148, Sep. 2007.

[5] B. Li, Y. Qin, C. P. Low, and C. L. Gwee, "A survey on mobile WiMAX," *IEEE Commun. Mag.*, vol. 45, no. 12, pp. 70–75, Dec. 2007.

[6] C. Cicconetti, A. Erta, L. Lenzini, and E. Mingozzi, "Performance evaluation of the IEEE 802.16 MAC for QoS support," *IEEE Trans. Mobile Comput.*, vol. 6, no. 1, pp. 26–38, Jan. 2007.

[7] Q. Ni *et al.*, "Investigation of bandwidth request mechanisms under point-to-multipoint mode of WiMAX networks," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 132–138, May 2007.

[8] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[9] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1506–1515, Jul. 2005.

[10] I. Inan, F. Keceli, and E. Ayanoglu, "Analysis of the 802.11e enhanced distributed channel access function," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1753–1764, Jun. 2009.

[11] 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.

[12] G. R. Cantieni, Q. Ni, C. Barakat, and T. Turletti, "Performance analysis under finite load and improvement for multirate 802.11," *Comput. Commun.*, vol. 28, pp. 1095–1109, 2005.

[13] D. Malone, K. Duffy, and D. Leith, "Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions," *IEEE/ACM Trans. Netw.*, vol. 15, no. 1, Feb. 2007.

[14] A. Vinel, Y. Zhang, M. Lott, and A. Tiurlikov, *Performance Analysis of the Random Access in IEEE 802.16*. Berlin, Germany: IEEE PIMRC, Sep. 2005.

[15] J. He, K. Guild, K. Yang, and 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.

[16] H. L. Vu, S. Chan, and L. L. H. Andrew, "Performance analysis of best-effort service in saturated IEEE 802.16 networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 460–472, Jan. 2010.

[17] Y. P. Fallah, F. Agharebparast, M. R. Minhas, H. M. Alnuweiri, and C. M. Leung, "Analytical modeling of contention-based bandwidth request mechanism in IEEE 802.16 wireless networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 3094–3107, Sep. 2008.

[18] L. Kleinrock, *Queueing Systems, Volume 2: Computer Applications*. London, U.K.: Wiley, 1976.

Alexey Vinel (M'07) received the Bachelor's (Hons.) and Master's (Hons.) degrees in information systems from the State University of Aerospace Instrumentation, in 2003 and 2005, respectively, and the Ph.D. (kandidat nauk) degree in technical sciences from the Institute for Information Transmission Problems (Russian Academy of Sciences), Moscow, in 2007.

He is a Researcher at Tampere University of Technology, Tampere, Finland. He was an Intern with Siemens AG, Munich, Germany, during 2004–2005 and the University of Wuerzburg, Wuerzburg, Germany, during 2006–2007, and carried out numerous short-term visits with different European organizations. He has conducted scientific research in the areas of wireless network performance evaluation and development of random multiple access algorithms.

Dr. Vinel is a Fellow of the Alexander von Humboldt Foundation and a Member of the Organizing and Technical Committees of numerous international conferences.

Qiang Ni (M'04–SM'08) received the B.Sc., M.Sc., and Ph.D. degrees in engineering from Huazhong University of Science and Technology (HUST), Wuhan City, China.

He joined Brunel University, London, U.K., as a Lecturer and was promoted to Reader in August 2010, where he heads the Intelligent Wireless Communication Networking Group. Prior to that, he was a Senior Researcher with the Hamilton Institute, National University of Ireland Maynooth. He also previously worked with INRIA France as a Researcher (2001–2004). His research interests are wireless networking and mobile communications. He was an IEEE 802.11 wireless standard working group Voting Member and a contributor to the IEEE wireless standards.

Ling Hu received the M.Eng. degree in electronics engineering from Dublin City University, Dublin, Ireland, and the M.Sc. degree in system-on-chip design from Ecole Superieure DIngenieurs en Electrotechnique et Electronique(ESIEE) Paris, Paris, France.

She is a Research Assistant at Brunel University, London, U.K.

Yang Xiao (M'97–SM'04) received the B.S. and M.S. degrees from Jilin University, Jilin, China, and the M.S. and Ph.D. degrees in computer science and engineering from Wright State University, Dayton, OH.

He worked in industry as a Medium Access Control (MAC) Architect involving the IEEE 802.11 standard enhancement work before he joined academia. He is currently a Faculty Member (with tenure) with the Department of Computer Science , University of Alabama, Tuscaloosa. He was a voting member of the IEEE 802.11 Working Group from 2001 to 2004. His research areas are security, communications/networks, robotics, and telemedicine.

Mr. Xiao currently serves as Editor-in-Chief for the *International Journal of Security and Networks* (IJSN) and the *International Journal of Sensor Networks* (IJSNet).

Marios Hadjinicolaou (M'02) received the B.Sc. (Hons.) degree in electronics from the University of London, London, U.K., in 1979 and the M.Sc. and Ph.D. degrees in electronic and electrical engineering from Brunel University, London, U.K., in 1982 and 1986, respectively.

He is a Senior Lecturer at Brunel University, engaged in teaching wireless communications, and a Member of the Wireless Networks and Communications Centre. His research interests are in the field of multiple access systems, video-on-demand, telemedicine, teletraffic engineering, and QoS studies for multimedia applications. He has published more than 50 papers in refereed journals and conferences.

Dr. Hadjinicolaou is a regular reviewer for IEEE/IET journals. He is a Chartered Engineer and Member of the IET.