Energy-Efficient Congestion Control on Contention-Based BW-REQ in WiMAX System

Shih-Chin Shiu and Jen-Yi Pan,
Department of Communications Engineering, National Chung Cheng University Chia-Yi, Taiwan, Republic of China
E-mail: jokernov211@hotmail.com, jypan@comm.ccu.edu.tw

Abstract— **The IEEE 802.16 standard for broadband wire -less metropolitan area network supports diverse real time and non-real time services. The standard defined the contention-based bandwidth request mechanism for the non-real time (nrt-PS) and Best-Effort (BE) services. Before using the resource of base station (BS), the mobile station (MS) or subscriber station (SS) should send the bandwidth request (BW-REQ) in the allocated contention time interval. For avoiding collision, the standard supports truncated binary exponential backoff resolution. However this mechanism works inefficiently while the network traffic is overloading, and then the result will cause MSs waste power. In this paper we propose a new mechanism at MS, for BE service, to help the MS send the bandwidth request more efficiently and also avoid the traffic peak. By using the mechanism, the MS sends the bandwidth request in the low loading time and it reduces the power consumption of the MS. Additionally, the method balanced the burst of traffic loading and smoothed the traffic loading that brings the bandwidth request efficiency and higher power utilization.**

Keywords：Bandwidth Request, Congestion, Power Saving

I. Introduction

Fixed and mobile broadband wireless access network is the most important key technology of communication in the near future. For the reason, the IEEE that is the world's leading professional association for the advancement of technology defined many network communication standards such as IEEE 802.11 for wireless network, IEEE 802.3 for wired network and so on. The IEEE 802.16 standard is the novel wireless network resolution of last mile in the last few years. There are more and more different applications processed in Internet that brings more demand on Quality of Service (QoS). To support variable multimedia applications, the IEEE 802.16 standard [1] defines four specific types of service flows that meet different requirements of users. Each connection between SS (or MS) and BS is identified as one service flow by connection ID (CID). The Unsolicited Grant Service (UGS) is defined to support constant bit rate (CBR) flows such as Voice over IP (VoIP). The real-time Polling Service (rtPS) is designed for real-time applications which periodically send variable size data packets, e.g. video stream. The real-time application requires maximum delay and minimum assurance of bandwidth. To support non-real-time services, the IEEE 802.16 standard defined other two types of QoS named non-real-time Polling Service (nrtPS) and Best-Effort Service (BE). The nrtPS class is applied to no special delay guarantees services such as ftp. But the 802.16 standard provides minimum bandwidth to nrtPS. The last one type of QoS, the BE service is supporting the traffic generated by most Internet applications (web surfing, email, etc.). In this paper we will focus on BE service.

The standard has defined a Request and Grant mechanism for BE services to request bandwidth [1]. Before the BE services require the BS resources; SS or MS sends a bandwidth request and waits for the grant. In Section II we will have more detailed account of the mechanism. We review some related works about performance of the Request and Grant mechanism. A research analyzed the mechanism by mathematical model [3]. In [4], [5], the performance of the random access scheme specified in IEEE 802.16 is investigated. In these papers we can see that the traffic loading affects the mean delay of bandwidth request mechanism. When the traffic overloading, the mean delay grows up fast because the packet collision happens which means the MS will consume power in sending feckless requests. In this article we proposed a “BW-REQ Congestion Control” (BCC) scheme for abating the effect of burst traffic and balancing the traffic. The MS will send the bandwidth request in the adequate time interval and optimize the power utilization by using our BCC scheme.

The remainder of this paper is organized as follows. Section II describes the Request and Grant mechanism and the power saving mechanism of IEEE 802.16 standard. Section III introduces our scheme. Section IV presents the simulation environment. Simulation results are provided in Section V. Finally, we conclude this paper in Section VI.

II. Overview of IEEE 802.16 Standard

A. IEEE 802.16 Frame Structure

We consider the network with a point-to-multipoint (PMP) architecture, which consists of one base station (BS) managing several MSs and SSs. Throughout this paper, we use the term “MS” to refer to “MS and SS”. Then let us consider the TDD (time division duplexing) mode in the PMP network. The frame structure is shown in Figure 1. It consists of two sub-frames to transmit that the subframe used for transmitting with the direction of BS to MS called down-link subframe, and the inverse one named up-link subframe. In the down-link subframe, the down-link map (DL-MAP) and up-link map (UL-MAP) are transmitted, in which the down-link and up-link

Proceedings of APCC2008 copyright © 2008 IEICE 08 SB 0083

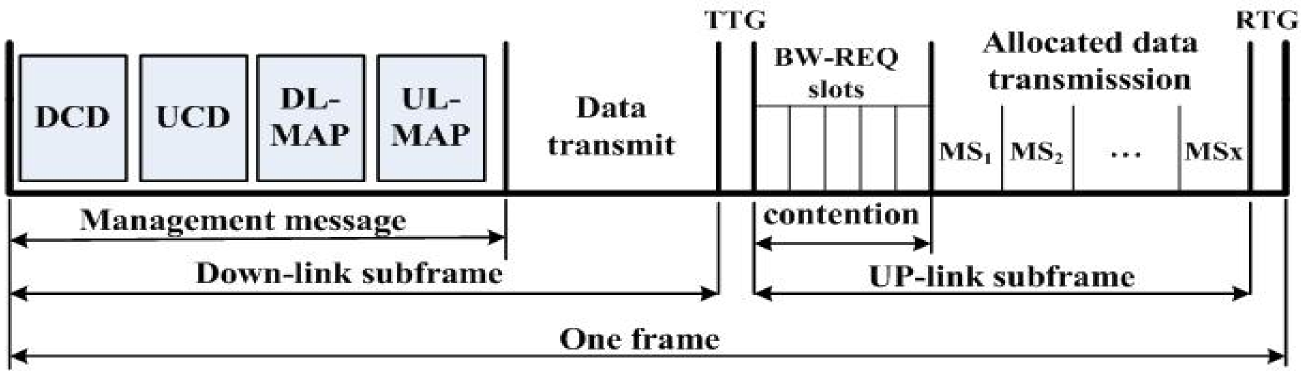

Figure 1. Frame structure in IEEE 802.16

bandwidth allocations for data transmission have been announced. In the up-link subframe, the contention time interval is allocated for MSs sending request messages to ask for the bandwidth. The up-link channel descriptor (UCD) and down-link channel descriptor (DCD) are also important information for management, which consists of some essential parameters like the value of minimum backoff window W_{min} and the value of maximum backoff window W_{max}. The Rx/Tx transmission gap (RTG/TTG) is specified between the up-link subframe and down-link subframe. The subframe duration can be dynamically tuned to accommodate the traffic in both directions. Besides the frame duration, the contention opportunities also can be flexibly arranged to adapt to different situations.

B. Request and Grant Mechanism

The MS supports two mandatory contention-based Bandwidth Request mechanisms, which are the Contention-based Bandwidth Requests and the Contention-based CDMA Bandwidth Requests [1], [2]. Our paper will focus on the Contention-based Bandwidth Requests mechanism. When an MS needs to ask for bandwidth on a connection of BE scheduling service, it sends a message to the BS containing the immediate requirements of the Demand Assigned Multiple Access (DAMA) connection. As the BS received the bandwidth request (BW-REQ) from the MS, BS will allocate resources for the MS in the next frame. In other words, before a connection with BE scheduling service using the bandwidth resource, the MS should send bandwidth request first. In the specification, a request may come as a stand-alone bandwidth request header or as a PiggyBack Request. The PiggyBack request is an optional feature and hence we will not discuss this, which is beyond the scope. Two main methods are defined in the standard for determining which MS is allowed to transmit their request, which are the contention-based random access and the contention-free polling. There is no acknowledgement message will be sent back to the MS to detect whether the BW-REQ successfully arrives at the BS or not in the both methods. For the reason, the timer T16 can determine whether the BW-REQ was collided and corrupt, and then triggers the contention resolution mechanism. The contention-free method is used for polling. In this paper we will focus on contention-based random access method.

With contention-based random access, the MS transmits a BW-REQ during the specific contention time interval, and uses random backoff mechanism to avoid the collision among different MSs. The standard specifies the truncated binary exponential backoff (BEB) as the mandatory random access contention resolution. Before MSs sending any BW-REQ, they randomly choose an integer number from $[0, 2^{W_{min}}]$ to be a backoff counter, where W_{min} is the backoff window size described in the UCD. As contention slots pass and the backoff counter counts down to zero, the MS transmits the BW-REQ requesting bandwidth for the amount of all data current in its buffers. Then the MS starts the T16 timer and waits for the bandwidth grant from the BS. If the timer expires, the MS will retry the process and choose a new backoff value in the interval of $[0, 2^{W_{min}+1}]$. In the other word, the MS will double the backoff window after each collision, until the window size goes up to 2 ^ W_{max}. The backoff method always efficiently reduces the collision occurrence. In [5], [6], the relation between traffic loading and BW-REQ transmission delay has been investigated that when the traffic loading is increasing, the delay of random access method is increasing fleetly. It means that if the MS transmits at the peak traffic load, the MS will waste power on feckless BW-REQ retransmission. In the following article we will introduce the power saving mechanism in the IEEE 802.16 standard.

C. Power Saving Mechanism

The IEEE 802.16 standard provides three power saving types for different service.

(1) Power-Saving Class of Type I

This class is recommended for the connections of the BE service and the non-real time service with variable data rate. In this mode, MS has a listen-window with fixed size and a variable sleep-window. In the sleep-window MS will turn off its receiver unit to save power and listen to the radio channel in the listen-window. After every sleeping cycle (which means a listen-window and a sleep-window), the sleep-window will double the length until being the maximum then keeping the max length to sleep.

(2) Power-Saving Class of Type II

Type II is recommended for the connections of the UGS service and the real-time service such as VoIP and video-streaming. In this mode, the MS uses the fixed listen window and fixed sleep window from cover to cover.

(3) Power-Saving Class of Type III

This type is used for management connections. It has no listening window and only involves a single sleep-window specified by BS. In the other word, the MS will sleep without listen-window until the appointed time.

In the Section III we will describe our proposed algorithm that makes the MSs send BW-REQ efficiently and we use the hybrid power saving scheme to reduce their power consumption.

III. PROPOSED ALGORITHM

In this article, we will introduce the proposed algorithm "BW-REQ Congestion Control" (BCC) in detail. For sending a BW-REQ efficiently, there are two parts of the BCC. One is how to determine whether the traffic load will be close to saturation and the other is how the MS can avoid the peak of traffic load.

A. Determine the Traffic Loading Peak

In the BCC we care about the traffic loading of the BW-REQ not the throughput loading so we determine the traffic loading by collecting statistics of the number of received collisions during a time interval. Because the contention time interval is modifiable, we consider this variable factor and define the ρ to determine the traffic loading. The traffic loading index can be obtained by

$$\rho = \frac{R_{total}}{S_{bw} \times T_{sta}}. \quad (1)$$

The R_{total} means the number of detected collisions in the statistic time interval, the S_{bw} is the contention slots of one frame defined by BS, and every T_{sta} frames we collect statistics of the BW-REQs. With the traffic loading analyzing function, the BS monitors the loading of traffic. When the value of ρ exceeds a threshold T_{max}, we can expect that after 100ms (in this paper we choose 100ms be the T16 value) the retrying MSs will make the traffic peak, so the BS will alert all the MSs after 100ms (T16) by a broadcast message. When the MS is aware of the alert of high loading, the MS will start the proposed algorithm.

B. Avoiding the Traffic Loading Peak

The core concept of this algorithm is to disperse the crowd of BW-REQ in the peak load, and rearrange them in the low traffic load. The BE services require no delay guarantee and no real-time scheduling. For these reasons, we can postpone sending a BW-REQ to distribute the traffic loading. There are two parts of the proposed algorithm to cope with two cases respectively: (1) sending a new BW-REQ and (2) retransmitting a BW-REQ due to the collision or channel error.

(1) Sending a New BW-REQ

For an MS sending an new BW-REQ, the MS first determines whether the high loading flag of the broadcast message is true, if the result is in the traffic loading peak then MS will use the Power-Saving Class of Type III that introduced in section II and will enter the sleeping mode. We defined a parameter B_{max} to be the maximum number of sleeping frames and the parameter B_{min} to be the minimum one. The MS will choose a value of the interval between $[B_{min}, B_{max}]$ to be the sleeping window. Then the MS counts down the window by frame and retry the request method. We named this policy "Avoiding Peak" in the following.

(2) Retransmitting a BW-REQ

This policy is named "Collision Control". For the MS retransmitting the BW-REQ, we defined a parameter P_{sleep} to be a sleeping probability. Before a MS retransmitting a BW-REQ with high loading flag, it will choose a random value to determine whether postponing the retry process. If the MS does not pass the test, it will go to sleep and the sleeping window size will be 25ms (5 frames). After sleeping, the MS will do the retransmission process again.

By using our algorithm, we decentralize the new coming BW-REQ traffic and assure the better retransmission chance of MS. Because the starting time of listen window can be calculated by our algorithm, we can use the Power-Saving Class of Type III to reduce more power consumption and the MS can send BW-REQ with lower delay. In this policy, we reduce the number of retransmissions with a probability and in our simulation we set the probability to 0.5. The Figure 2 illustrates our algorithm in detail.

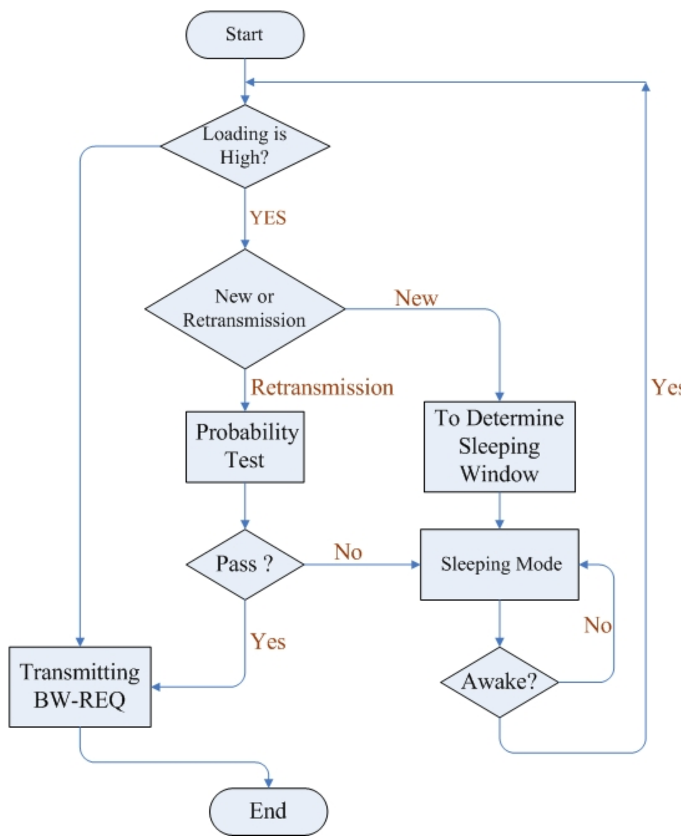

Figure 2. The proposed algorithm

IV. SIMULATION ENVIRONMENT

In this section we show the effect of our proposed scheme on the random access mechanism performance by the Network Simulator version 2 (NS2) [7] and use the IEEE 802.16 module of NIST [8] to simulate the random access mechanism. In this paper we will not consider with the collision that happens due to channel error; i.e., we assume the channel is perfect. In our simulation scenario the data traffic of BE service is modeled as Web source. We use the traffic named Web exponential [9], [10] and set the parameters interarrival time λ = 5/s and average rate =25Kb/s. The IEEE 802.16 network parameters are shown in table I.

Table I.
Simulation Parameters

Simulator parameter	Value
Channel bandwidth	11MHz
Frame duration	4 ms
Request backoff start	2
Request backoff stop	6
Contention bandwidth request collision detection timeout (T16)	100 ms
Number of nodes	26
T_{max} (T_{sta} = 1, S_{bw} = 5)	0.6
B_{min}, B_{max}	10, 20
P_{sleep}	0.5
Simulation time	40 s

V. Simulation Results and Performance Analysis

In the scenario, we compare the different policies of our scheme: the first is the original mechanism of IEEE 802.16 standard, the second is the policy "Avoiding Peak" that we send a new BW-REQ with lower traffic loading, the third is "Collision Control" that we try to reduce collisions and the last one is "Hybrid" that we use "Avoiding Peak" and "Collision control" in the same time. For these different situations, we compare the delay of BW-REQ and the number of collisions. Then we try to calculate the efficiency of power saving of our scheme.

Table II.
Statistic of BW-REQ

	Average delay of BW-REQ (frames)	Number of received BW-REQs	Number of received collisions
Origin	8.98 (100%)	2214 (100%)	538 (100%)
Avoiding Peak	6.70 (75%)	1737 (79%)	456 (85%)
Collision Control	5.25 (59%)	1724 (78%)	399 (75%)
Hybrid	6.64 (74%)	1731 (78%)	443 (83%)

The Table II shows the statistic of some important results that contain the number of received BW-REQs, total collisions and the average delay of BW-REQ. The average delay means the average time interval from MS sending a BW-REQ to the BS receive this packet. If the packet is collided, the delay will be the timeout T16 (100ms) and the time interval of packet retry will be added in one BW-REQ transmission delay. In other words, the delay means the time of a successful transmission or the time of T16 (100ms). The following percentages mean the ratio of comparing with original mechanism.

In the following we will describe other related results in detail.

The Figure 3 shows the number of collisions in four situations, when the traffic getting higher loading, the number of collisions is increasing and that in original mechanism is higher than others. That is because the truncated binary exponential backoff smooth over the traffic slowly. By using the policy "Collision Control", we stop the retry by a probability 50% and we can see that the number of collisions is reduced. The probability is a modifiable factor, in future work we will try to optimize it. The other results of the policy "Avoiding Peak" and the policy "Hybrid" also get fewer collisions than the original mechanism's. Comparing with the policy "Avoiding Peak" and the policy "Hybrid", we observe that the "Hybrid" gets higher number of collisions; we guess that the un-optimal parameters B_{min}, B_{max} and P_{sleep} make this result. To optimize the B_{min}, B_{max} and P_{sleep} is also the future work.

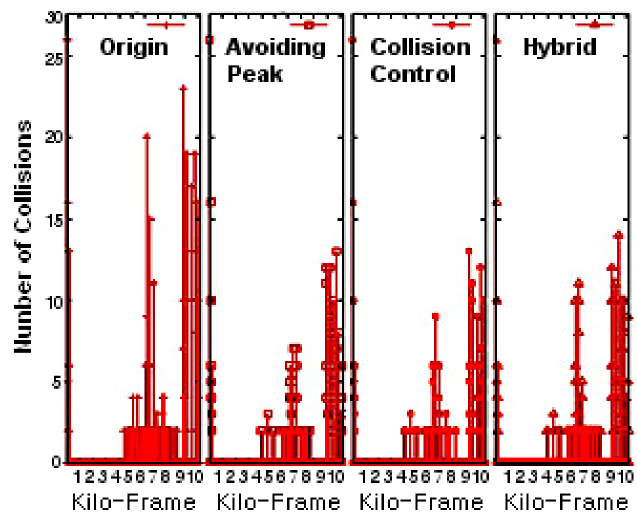

Figure 3. The number of collisions

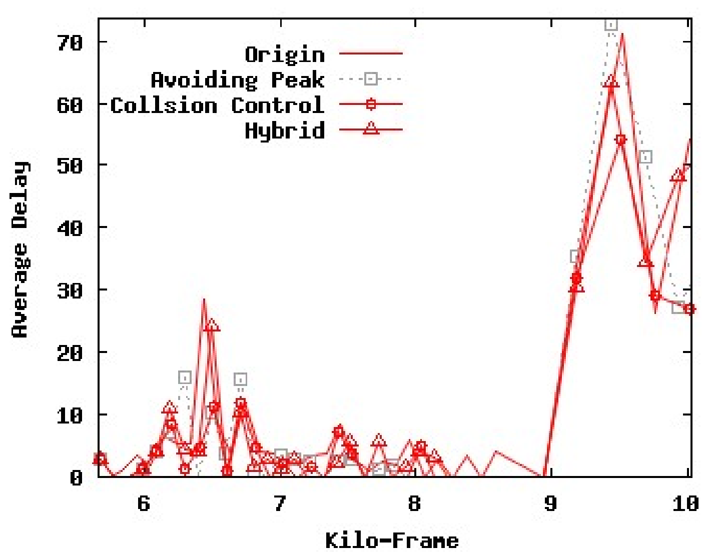

Figure. 4 Trend of average delay

The Figure 4 shows the trend of delay, and the result is similar to the Figure 3. The "Collision Control" reduces the delay efficiently. In the 9300th frame, the "Avoiding Peak" policy gets higher delay than others'. We conjecture that is because the burst of new BW-REQ causes frequent collisions, and the policy "Avoiding Peak" is not enough to improve that.

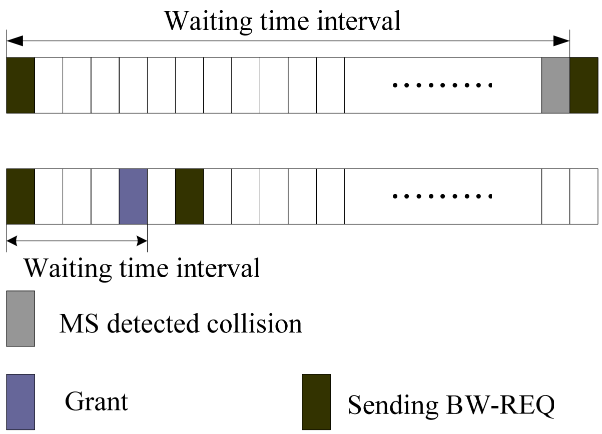

Figure. 5 Waiting time interval

Then we calculate the efficiency of power utilization. The Figure 5 illustrates the power consumption of sending a BW-REQ. After sending BW-REQ, MS should ceaselessly wait for the broadcast message with the grant or wait for the T16 timer being expired. In the proposed algorithm we try to send in low traffic load to avoid collisions and this policy not only saves the MS's power but also reduces the probability of colliding. The result makes other MSs reduce their power consumption. We can calculate the average power consumption "**δ**" of sending one successful BW-REQ by the average delay. In [11], the value of power consumption per frame, noted by **δ,** in active frame is defined as 1.5W. The power consumption of sending one BW-REQ can be defined by

$$C = \delta \times AverageDelay \,. \quad (2)$$

For different cases, we define C_{origin} as the power consumption calculated by using the average delay in original mechanism and the $C_{proposed}$ is of the proposed algorithm. We define the efficiency as

$$E = \frac{C_{origin} \times N_{origin} - C_{proposed} \times N_{proposed}}{C_{origin} \times N_{origin}} \,. \quad (3)$$

The N_{origin} means the number of outgoing BW-REQs with original mechanism in the simulation time, the $N_{proposed}$ is the number N with the proposed algorithm, and these values are shown in Table II. The Table III shows the results by percentage.

Table III.
The efficiency of power saving

	Avoiding Peak	**Collision Control**	**Hybrid**
E	**41.4%**	**54.4%**	**42.1%**

In Table III, the policy "Collision Control" gets the highest efficiency, but we think that if the parameters B_{min}, B_{max} and P_{sleep} being optimized the "Hybrid" will get better efficiency.

VI. Conclusion

In this paper we try to improve the efficiency of the power saving mechanism in the IEEE 802.16 standard. The Request and Grant mechanism in the traffic peak consumes much power due to the feckless BW-REQ retransmissions. By using the proposed algorithm, we enhance the power utilization by 40-55% when MSs sending BW-REQs. We also balance the traffic loading and decrease the number of feckless BW-REQs. With our algorithm, the MS sends BW-REQ more efficiently and gets more sleeping time. In this paper, we proposed a new algorithm to control the congestion of BW-REQ with the BE service and successfully smooth over the congestion.

ACKNOWLEDGEMENT

The authors would like to thank the Industrial Technology Research Institute and the Ministry of Economic Affairs (Project No. 7301XS2220) of the Republic of China, Taiwan, for financially supporting this research. This work was also supported partially by National Science Council of Republic of China under grant NSC 97-2221-E-194-025.

References

[1] IEEE Std. 802.16-2004, "Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems." Oct. 2004.

[2] IEEE Std. 802.16e-2005, "Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1." Feb 2006.

[3] Y. P. Fallah, F. Agharebparast, M. M. H. Alnuweiri, and V. C.M. Leung, "Analytical Modeling of Contention-Based Bandwidth Request Mechanism in IEEE 802.16 Wireless Networks," IEEE Transactions on Vehicular Technology, in press.

[4] A. Vine, Y. Zhang, M. Lott, and A. Tiurlikov, "Performance Analysis of the Random Access in IEEE 802.16," IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. vol. 3, 11-14 Sept. 2005, pp.1596-1600.

[5] Q. Ni, A. Vinel, Y. Xiao, A. Turlikov, and T. Jiang "WIRELESS BROADBAND ACCESS: WIMAX AND BEYOND - Investigation of Bandwidth Request Mechanisms under Point-to-Multipoint Mode of WiMAX Networks," IEEE Communications Magazine vol. 45, no. 5, May 2007, pp. 132 – 138.

[6] D. Staehle and R. Pries, "Comparative Study of the IEEE 802.16 Random Access Mechanisms," The International Conference on Next Generation Mobile Applications, Services and Technologies, 2007. NGMAST '07. 12-14 Sept. 2007, pp. 334 – 339.

[7] "The Network Simulator – ns-2," [Online]. Available: http://www.isi.edu/nsnam/ns/index.html

[8] "Seamless and secure Mobility," [Online]. Available: http://www.antd.nist.gov/seamlessandsecure/toolsuite.html

[9] C. Cicconetti, A. Erta, L. Lenzini, and E. Mingozzi, "Performance Evaluation of the IEEE 802.16 MAC for QoS Support," IEEE Transactions on Mobile Computing, vol. 6, no. 1, Jan. 2007, pp. 26 – 38.

[10] M. Molina, P. Catelli, and G. Foddis, "Web Traffic Modeling Exploiting TCP Connections' Temporal Clustering through HTML-REDUCE," IEEE Network, vol. 14, no. 3, May-June 2000, pp. 46 – 55.

[11] J. R. Lee, and D. H. Cho, "Performance Evaluation of Energy-Saving Mechanism Based on Probabilistic Sleep Interval Decision Algorithm in IEEE 802.16e," IEEE Transactions on Vehicular Technology, vol. 56, no. 4, Part 1, July 2007, pp.1773 – 1780.